

Mécanique Quantique

Table des Matières

1	Rappels des fondements et du formalisme de la Mécanique Quantique	9
1.1	L'équation de Schrödinger	9
1.2	Formalisme de Dirac	15
1.3	Opérateurs hermitiques, opérateurs unitaires	29
1.4	Postulats de la Mécanique Quantique	33
1.4.1	Notion d'état	33
1.4.2	Notion d'observable	35
1.4.3	Résultats possibles de la mesure d'une grandeur physique	38
1.4.4	La réduction du paquet d'ondes	41
1.4.5	Evolution des systèmes dans le temps	44
1.5	Evolution temporelle d'un système quantique	45
2	Symétrie et lois de conservation	55
2.1	Le principe euclidien de Relativité et le rôle de la symétrie en Physique	55
2.2	Symétrie spatiale continue : translations et rotations	59
2.2.1	Translations	59
2.2.2	Rotations	66
2.3	Invariance de jauge	71
2.4	Symétries discrètes	74
2.5	Symétrie et dégénérescence	79

3	Théorie du moment cinétique	81
3.1	Importance du moment cinétique	81
3.2	Propriétés générales des valeurs et vecteurs propres d'un moment cinétique	83
3.3	Etats propres du moment cinétique orbital	86
3.4	Cas particuliers des moments cinétiques $j = \frac{1}{2}, j = 1$	92
3.5	Addition de deux moments cinétiques. Coefficients de Clebsch-Gordan	94
3.5.1	Comment apparaît la somme des moments cinétiques	94
3.5.2	Somme de deux moments cinétiques	96
3.6	Théorème de Wigner-Eckart. Règles de sélection	105
4	Potentiel central et atome d'hydrogène	111
4.1	Définition du champ central et exemples	111
4.1.1	Hamiltonien du problème central	112
4.1.2	Comportements de la fonction radiale	117
4.1.3	La particule libre	119
4.1.4	Puits "carré" sphérique	122
4.2	Atome d'hydrogène : états propres liés	124
4.2.1	Résolution de l'équation radiale	124
4.2.2	Symétrie dynamique du potentiel Coulombien	132
5	Le spin	135
5.1	Insuffisances de la description par une seule fonction d'onde	135
5.2	Lagrangien d'une particule dans un champ électromagnétique	144
5.3	Magnétisme atomique : l'atome d'hydrogène	145
5.4	Rotation d'un spin	147
5.5	Retour sur le renversement du temps	149
5.6	Origine relativiste du spin	153
5.6.1	Construction de l'équation de Dirac	153
5.6.2	États stationnaires d'un électron libre	159

6	Méthodes d'approximation pour les états propres	165
6.1	Méthode variationnelle	165
6.1.1	Formulation variationnelle de l'équation aux valeurs et vecteurs propres	165
6.1.2	Calcul variationnel des états discrets (états liés)	168
6.2	Théorie des perturbations stationnaires	170
6.2.1	Cas d'un niveau non-dégénéré	172
6.2.2	Cas d'un niveau dégénéré	176
6.2.3	Développement systématique à l'aide de la résolvante	179
6.2.4	Exemples d'application : effet Stark et effet Zeeman pour l'hydrogène	181
7	Théorie des perturbations dépendant du temps	187
7.1	Présentation générale	187
7.2	Calcul systématique de l'opérateur d'évolution et approximations successives	191
7.3	Amplitudes de transition entre états discrets : exemples	195
7.3.1	Perturbation constante	195
7.3.2	Perturbation sinusoïdale	197
7.4	Amplitudes de transition entre un état discret et un continuum d'états finals. Règle d'or de Fermi	198
7.5	Applications exemplaires	201
7.5.1	Interaction d'un atome avec un champ électromagnétique (description semi-classique)	202
7.5.2	Théorie élémentaire de l'effet photoélectrique	204
7.6	Rudiments sur la description purement quantique de l'interaction champ - matière	210
8	Particules identiques	221
8.1	Indiscernabilité des particules identiques en Mécanique Quantique	221
8.2	Le postulat de symétrisation	224
8.3	Permutations. Opérateurs de symétrisation et d'antisymétrisation	227
8.4	Applications et exemples	232
8.4.1	Etats d'un système de particules indépendantes : différence fondamentale entre bosons et fermions	232
8.4.2	L'atome d'hélium	236
8.4.3	Modèle simpliste d'un solide. Conducteurs et isolants	241
8.5	Introduction à la Seconde Quantification	244

9	Introduction à la théorie de la diffusion	251
9.1	Présentation	251
9.2	Nature des états de diffusion	252
9.3	Méthode des déphasages	254

Avertissement

Tout en restant à l'intérieur du cadre fixé par le programme de la Maîtrise, le volume et le contenu de ces notes de cours dépassent – et de loin – ce qu'il est raisonnable et possible de traiter en deux amphes hebdomadaires semestriels. Cette disproportion est le résultat d'un choix délibéré : faire le tour d'horizon à propos 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.

Bien évidemment, ces notes n'ont aucun caractère exhaustif, compte tenu notamment de la richesse de la matière, et ne doivent surtout pas être considérées comme constituant un document de référence exclusif. Les avoir sous la main ne doit pas écarter l'envie de lire les ouvrages disponibles dans les bibliothèques, bien au contraire : il est souhaitable (et souhaité !) que les développements qui figurent dans ce polycopié et qui ne seront pas mentionnés oralement 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, étant entendu qu'il serait vain de vouloir épuiser les références bibliographiques tant sont nombreux les ouvrages de qualité exposant la Mécanique Quantique à des niveaux divers.

N. B. Le présent cours s'inscrit dans la continuité du cours de Physique Quantique dispensé en Licence de Physique. Pour cette raison, les rappels de connaissances acquises antérieurement seront strictement limités et forcément sommaires : pour plus de détails, se reporter au cours de Licence.

Chapitre 1

Rappels des fondements et du formalisme de la Mécanique Quantique

Le but de ce chapitre est de rappeler l'essentiel des fondements et du formalisme de la Mécanique Quantique. Les résultats mathématiques nécessaires pour la suite seront énoncés et justifiés d'une façon plus intuitive que rigoureuse, avant que soit donné le rappel des postulats.

1.1 L'équation de Schrödinger

L'équation de Schrödinger a été proposée de façon inductive par Schrödinger en 1926, un peu après la Mécanique des Matrices de Heisenberg (1925) et s'est développée d'abord dans le but de décrire les petits objets (atomes) constitués d'une seule particule située dans un certain champ de force (l'électron au sein de l'atome d'hydrogène, par exemple). L'objet central de la théorie de Schrödinger, nommée aussi Mécanique Ondulatoire, est une fonction $\Psi(\vec{r}, t)$ à valeurs *complexes*, appelée fonction d'onde. Cette fonction satisfait :

$$i\hbar \frac{\partial}{\partial t} \Psi(\vec{r}, t) = \left[-\frac{\hbar^2}{2m} \vec{\nabla}^2 + V(\vec{r}) \right] \Psi(\vec{r}, t) , \quad (1.1)$$

où \vec{r} est le rayon-vecteur repérant la particule dans l'espace ; $V(\vec{r})$ est l'énergie potentielle de la particule étudiée, $\vec{\nabla}$ est le vecteur gradient dont les trois composantes sont $(\partial/\partial x, \partial/\partial y, \partial/\partial z)$; i désigne le nombre imaginaire pur fondamental, m est la masse de la particule. Ceci étant admis, il apparaît que la Mécanique Quantique n'est formulable que pour des forces dérivant d'un potentiel ; dès lors, admettre que cette théorie s'applique à toute la Physique à l'échelle atomique ou subatomique, c'est admettre que toutes les interactions "fondamentales" dérivent d'un potentiel, au sens usuel ou en un sens généralisé¹.

L'équation (1.1) ne se démontre pas à proprement parler : construite sur la base d'arguments heuristiques, largement inspirée par l'analogie remarquée dès 1828 par Hamilton entre Optique et Mécanique², la validité de cette équation se mesure par la confrontation entre les résultats théoriques qu'elle engendre et les observations expérimentales. Ses succès sont immenses : à ce jour, aucune expérience ne permet de mettre en doute la théorie dont l'équation (1.1) est la pierre angulaire.

L'approche de Heisenberg³, appelée Mécanique des Matrices, procède d'une tout autre démarche, fondée sur l'analyse harmonique du mouvement classique revisité à la lumière des conditions de quantification de Bohr

¹Pour une particule chargée dans un champ électromagnétique, le potentiel ne dépend pas que des coordonnées mais est également fonction de la vitesse.

²Les 4 articles fondateurs de Schrödinger sont disponibles en français dans la référence [1].

³Si Heisenberg doit être considéré comme le Père-Fondateur de la Mécanique des Matrices, Born et Jordan ont participé très vite à ses premiers développements. L'usage des matrices en Physique avait été jusqu'alors exceptionnel, Born s'étant précisément "fait la main" avec ce nouvel outil pour la description des vibrations d'un solide.

- Sommerfeld - Wilson posées par l'Ancienne Théorie des Quanta. L'équation fondamentale porte sur des matrices (représentant des opérateurs dans un espace vectoriel rapporté à une base préalablement choisie), très précisément sur les matrices représentant les deux grandeurs fondamentales de la Mécanique Analytique de Lagrange et Hamilton : la coordonnée, notée q_j , et le moment conjugué p_j , reliés l'un à l'autre par :

$$p_j = \frac{\partial L}{\partial \dot{q}_j} . \quad (1.2)$$

L est le Lagrangien ; q_j désigne l'une des coordonnées (position, angle, ...) permettant de définir sans ambiguïté la position du système dans l'espace⁴. Pour des coordonnées *cartésiennes*, l'équation fondamentale de Heisenberg est⁵ :

$$[q_j, p_k] \equiv q_j p_k - p_k q_j = i\hbar \delta_{jk} , \quad (1.3)$$

où, au passage, le commutateur $[b, \#] \equiv b\# - \#b$ a été défini. Dans (1.3), q_j et p_k sont des *opérateurs* représentant les variables dynamiques fondamentales, coordonnées et moments conjugués. Par exemple, pour les trois coordonnées cartésiennes (x, y, z) , on a :

$$[u, p_v] = i\hbar \delta_{uv} \quad (u = x, y, z) . \quad (1.4)$$

La relation (1.3), obtenue précisément par Born et Jordan, introduit fondamentalement des objets dont l'algèbre est *non-commutative*. Ces objets sont les opérateurs associés aux grandeurs physiques, traditionnellement appelés observables. Il n'est pas exagéré de dire que toute la structure formelle de la Mécanique Quantique est construite sur – et contenue dans – la relation (1.3), étant entendu qu'il reste à en bâtir le contenu physique. Tout comme l'équation de Schrödinger, (1.3) ne se démontre pas à proprement parler⁶.

L'équivalence des descriptions de Heisenberg et de Schrödinger repose sur une remarque fondamentale de ce dernier, définissant d'ailleurs un mode de construction automatique de (1.1). On commence par former le Hamiltonien classique, $H_{cl}(\vec{r}, \vec{p})$, suivant les prescriptions de la Mécanique Analytique. Puis on remplace dans la fonction H_{cl} le vecteur \vec{r} par l'opérateur multiplication par \vec{r} , et le vecteur \vec{p} par $-i\hbar \vec{\nabla}$:

$$\vec{r} \rightarrow \times \vec{r} , \quad \vec{p} \rightarrow -i\hbar \vec{\nabla} . \quad (1.5)$$

La substitution pour \vec{p} ne vaut que pour les coordonnées cartésiennes. Ceci étant fait, on obtient un opérateur *différentiel* H , – c'est le Hamiltonien de la Mécanique Quantique – qui, agissant sur la fonction d'onde, constitue le second membre de l'équation de Schrödinger, à laquelle on peut dès lors donner la forme plus compacte :

$$i\hbar \frac{\partial}{\partial t} \Psi(\vec{r}, t) = H(\vec{r}, -i\hbar \vec{\nabla}) \Psi(\vec{r}, t) . \quad (1.6)$$

Par nature, cette méthode de construction fonctionne bien pour tous les degrés de liberté qui ont un équivalent classique. On verra cependant qu'il existe des degrés de liberté spécifiquement quantiques, comme le spin. En pareil cas (en théorie non-relativiste quand il s'agit du spin⁷), il convient d'ajouter "à la main" les termes d'interaction correspondants, sur la base d'arguments par analogie. Par exemple, à un spin \vec{S} (qui structurellement, comme on le verra, est un moment cinétique au sens usuel de la Mécanique Quantique) est associé un moment magnétique $\vec{\mu}$, par une simple relation de proportionnalité, tout comme un moment cinétique ordinaire (orbital) produit un moment magnétique par le biais du facteur gyromagnétique classique γ ; ainsi, dans le cas de l'électron, on doit poser⁸ :

$$\vec{\mu} = g_s \frac{e}{2m} \vec{S} , \quad (1.7)$$

où e est la charge de l'électron ($e = -1.6 \times 10^{-19}$ C) et où g_s est un facteur purement numérique très voisin de 2 ($g_s = 2.0023\dots$). Dès lors, en présence d'un champ magnétique \vec{B} , le Hamiltonien construit selon la prescription

⁴Pour un système à N particules dans \mathbb{R}^3 , il est nécessaire pour cela de définir $3N$ coordonnées ; on dit que le système possède $3N$ degrés de liberté.

⁵ δ_{jk} est le symbole de Kronecker, qui vaut 1 si $j = k$, 0 dans le cas contraire.

⁶Pour en savoir plus sur la démarche de Heisenberg, voir [8].

⁷La théorie de l'électron de Dirac introduit le spin naturellement.

⁸Noter que, la charge de l'électron étant négative, les deux vecteurs $\vec{\mu}$ et \vec{S} sont dirigés dans deux sens opposés.

de Schrödinger doit être complété par le terme représentant le couplage entre le moment magnétique lié au spin et le champ extérieur^{9,10}, soit $-\vec{\mu} \cdot \vec{B} \equiv -g_s(e/2m) \vec{S} \cdot \vec{B}$.

Pour établir dans le détail l'équivalence entre les deux approches de Heisenberg et de Schrödinger, il suffit maintenant de remarquer que les deux opérateurs \vec{r} et $-i\hbar\vec{\nabla}$ satisfont de fait la relation (1.4). En effet, pour le couple (x, p_x) écrit à la Schrödinger, on a, quelle que soit la fonction f différentiable :

$$\forall f(\vec{r}) : (xp_x - p_x x)f = -i\hbar \left[x \frac{\partial}{\partial x} f - \frac{\partial}{\partial x} (xf) \right] = -i\hbar \left(x \frac{\partial f}{\partial x} - f - x \frac{\partial f}{\partial x} \right) = +i\hbar f , \quad (1.8)$$

ce qui montre bien que :

$$[x, p_x] \equiv -i\hbar \left[x, \frac{\partial}{\partial x} \right] = +i\hbar \mathbf{1} , \quad (1.9)$$

où $\mathbf{1}$ désigne l'opérateur identité.

L'association $\vec{r} \rightarrow \times \vec{r}$, $\vec{p} \rightarrow -i\hbar\vec{\nabla}$ n'est qu'une façon parmi d'autres d'écrire la Mécanique Quantique (on dit qu'il s'agit de la représentation- q). Le rôle privilégié joué par la coordonnée \vec{r} semble même abusif, quand on se souvient que, en Mécanique Analytique, coordonnée et impulsion ordinaires sont traitées sur un strict pied d'égalité. De fait, il existe bien d'autres possibilités qui respectent la relation (1.3) fondamentale ; l'une d'entre elles, très utile pour certaines applications, constitue ce que l'on appelle la représentation- p , où la fonction d'onde est une fonction (toujours à valeurs complexes) d'un argument scalaire représentant précisément l'impulsion, soit $\Phi(\vec{p}, t)$. Cette représentation se définit par l'association suivante :

$$\vec{r} \rightarrow +i\hbar\vec{\nabla}_{\vec{p}} \quad \vec{p} \rightarrow \times \vec{p} , \quad (1.10)$$

où le vecteur gradient est maintenant :

$$\vec{\nabla}_{\vec{p}} = \left(\frac{\partial}{\partial p_x}, \frac{\partial}{\partial p_y}, \frac{\partial}{\partial p_z} \right) . \quad (1.11)$$

Il est clair que l'on a encore :

$$\forall g(\vec{p}) : [u, p_u] g = +i\hbar g \quad \iff \quad [u, p_v] = +i\hbar \delta_{uv} , \quad (1.12)$$

équation qui montre que les relations fondamentales (1.3) et (1.4) sont préservées. Le sens physique de $\Phi(\vec{p}, t)$ doit être précisé, tout comme celui de $\Psi(\vec{r}, t)$ d'ailleurs (voir ci-dessous). $\Psi(\vec{r}, t)$ et $\Phi(\vec{p}, t)$, toutes deux appelées fonctions d'onde, constituent deux *représentations* équivalentes du même état quantique et d'ailleurs se déduisent l'une de l'autre par une transformation de Fourier ; très précisément, on a¹¹ :

$$\Phi(\vec{p}, t) = (2\pi\hbar)^{-3/2} \int_{\mathbb{R}^3} d^3r e^{\frac{i}{\hbar} \vec{p} \cdot \vec{r}} \Psi(\vec{r}, t) \quad (1.13)$$

ainsi que la relation inverse :

$$\Psi(\vec{r}, t) = (2\pi\hbar)^{-3/2} \int_{\mathbb{R}^3} d^3p e^{-\frac{i}{\hbar} \vec{r} \cdot \vec{p}} \Phi(\vec{p}, t) . \quad (1.14)$$

Connaître une fonction et connaître sa transformée de Fourier constituent deux informations strictement équivalentes¹². Comme le terme cinétique introduit le terme $\vec{p}^2/(2m)$ – y compris en présence d'un champ magnétique –, l'opérateur différentiel en représentation- q contient toujours le Laplacien ; en revanche, en représentation- p , l'opérateur différentiel $V(+i\hbar\vec{\nabla}_{\vec{p}})$ ne se réduit pas au Laplacien et reflète la forme précise de V . Très explicitement, l'équation de Schrödinger prend les formes suivantes :

$$i\hbar \frac{\partial}{\partial t} \Psi(\vec{r}, t) = \left[-\frac{\hbar^2}{2m} \vec{\nabla}^2 + V(\vec{r}) \right] \Psi(\vec{r}, t) \quad (\text{représentation-}q) , \quad (1.15)$$

⁹C'est ce terme qui, dans le cas d'un spin demi-entier, donne lieu à l'effet Zeeman anormal, où les raies observées en champ nul éclatent en plusieurs raies voisines en présence du champ, le nombre de raies étant un nombre *pair*.

¹⁰Un tel procédé peut paraître assez artificiel. Il est totalement justifié lors de la construction de la version relativiste (théorie de l'électron de Dirac), dans laquelle le spin apparaît spontanément.

¹¹dans l'espace à trois dimensions.

¹²Quand on n'y prend pas garde, une transformation de Fourier (ou de Laplace) peut "effacer" des sauts finis ; on verra que la fonction d'onde doit être continue : cette perte d'information n'est donc pas à redouter dans le cadre de la Mécanique Quantique.

$$i\hbar \frac{\partial}{\partial t} \Phi(\vec{p}, t) = \left[\frac{\vec{p}^2}{2m} + V(+i\hbar \vec{\nabla}_{\vec{p}}) \right] \Phi(\vec{p}, t) \quad (\text{représentation-}p) . \quad (1.16)$$

À titre d'illustration, traitons le cas le plus simple, celui d'une particule libre dans \mathbb{R} . Après séparation espace-temps, $\Psi(x, t) = e^{\frac{i}{\hbar}Et} \psi(x)$, la fonction $\psi(x)$ en représentation- q doit satisfaire $H\psi = E\psi$, où $H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2}$ puisque $V \equiv 0$:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = E\psi(x) . \quad (1.17)$$

Toutes les solutions de cette équation sont des combinaisons linéaires quelconques de $e^{\pm ikx}$ avec $k = \hbar^{-1} \sqrt{2mE}$:

$$\psi(x) = A_+ e^{ikx} + A_- e^{-ikx} \quad (k = \hbar^{-1} \sqrt{2mE}) . \quad (1.18)$$

En ce qui concerne la représentation- p , on pose de façon analogue $\Phi(p, t) = e^{\frac{i}{\hbar}Et} \phi(p)$. $\phi(p)$ satisfait $H\phi = E\phi$ avec maintenant $H = \frac{p^2}{2m}$, soit :

$$\frac{p^2}{2m} \phi(p) = E\phi(p) . \quad (1.19)$$

Dans cette équation la fonction à trouver est au premier membre simplement *multipliée* par le facteur scalaire $\frac{p^2}{2m}$: ce dernier n'est plus un opérateur, au contraire de ce qui se passe en représentation- q . (1.19) se réécrit comme suit :

$$\left(\frac{p^2}{2m} - E \right) \phi(p) = 0 \quad (1.20)$$

et dit que ou bien la parenthèse est non-nulle, auquel cas $\phi(p) = 0$, ou bien la parenthèse est nulle et alors, pour les valeurs de p qui annulent cette dernière, la valeur de $\phi(p)$ est indéterminée. En dehors de la solution triviale $\phi(p) = 0$ partout, les solutions sont donc les fonctions de Dirac centrées là où $\frac{p^2}{2m} - E = 0$, soit $\delta(p \pm \sqrt{2mE})$. Comme l'équation (1.19) est *linéaire*, toute combinaison linéaire convient tout autant ; au total, la solution la plus générale de (1.19) est :

$$\phi(p) = B_+ \delta(p - \sqrt{2mE}) + B_- \delta(p + \sqrt{2mE}) . \quad (1.21)$$

Il est facile de vérifier que $\phi(p)$ donnée par (1.21) et $\psi(x)$ donnée par (1.18) sont bien reliées par la transformation de Fourier (1.14), écrite pour les seules fonctions spatiales (la dépendance en temps est ici un simple facteur, le même pour Ψ et Φ) :

$$\psi(x) = (2\pi\hbar)^{-1/2} \int_{\mathbb{R}} dp e^{-\frac{i}{\hbar}xp} \phi(p) . \quad (1.22)$$

On désigne habituellement par Mécanique Quantique la théorie unifiant les approches de Heisenberg - Born - Jordan et de Schrödinger ; les appellations Mécanique des Matrices et Mécanique Ondulatoire ont aujourd'hui un caractère un peu obsolète.

Fondamentalement, l'équation de Schrödinger est une équation aux dérivées partielles *linéaire* ; ceci veut dire que si on connaît deux solutions linéairement indépendantes Ψ_1 et Ψ_2 , toute combinaison linéaire est encore solution. Ceci permet de construire la solution générale en formant des combinaisons linéaires de solutions particulières ; les plus remarquables et les plus utiles – pour une équation aux dérivées partielles *linéaire* – sont celles dans lesquelles espace et temps sont séparés. Compte tenu de la forme de (1.1), et pour un Hamiltonien indépendant du temps (système isolé), on voit tout de suite qu'il existe une classe de solutions particulières de la forme :

$$\Psi_{\text{st}}(\vec{r}, t) = e^{\frac{i}{\hbar}Et} \psi(\vec{r}) . \quad (1.23)$$

On vérifie immédiatement que Ψ_{st} est une solution de (1.1) pourvu que la fonction spatiale ψ satisfasse l'équation plus simple :

$$H(\vec{r}, -i\hbar \vec{\nabla}) \psi(\vec{r}) = E \psi(\vec{r}) . \quad (1.24)$$

Dans ces écritures, E est une quantité homogène à une énergie, dont on sait qu'il convient de l'interpréter précisément comme l'énergie de la particule décrite par l'état ψ (ou Ψ_{st}). L'équation (1.24) est appelée équation aux valeurs (et fonctions propres) ; en effet, en bon français, elle dit que l'action de l'opérateur Hamiltonien

sur ψ produit la *même* fonction ψ , simplement multipliée par le scalaire E . Le couple (E, ψ) forme un couple propre, ou mode propre. Il est parfois utile de garder la trace de la valeur propre E dans le symbole représentant la fonction propre qui lui est associée ; s'il le faut, on écrit plus précisément :

$$H \psi_E = E \psi_E . \quad (1.25)$$

Les états $\Psi_{\text{st}}(\vec{r}, t)$ construits selon (1.23) – les ψ_E étant simplement multipliés par la phase $e^{\frac{i}{\hbar} Et}$ – sont remarquables à plus d'un titre ; en particulier, la moyenne quantique de toute observable calculée dans un tel état est toujours indépendante du temps ; pour cette raison, ces états sont appelés états stationnaires. On verra par la suite comment la connaissance de tous les modes propres de (1.24), pour un système donné, permet de fabriquer la solution la plus générale de l'équation de Schrödinger (1.1) en construisant la bonne (et unique) combinaison linéaire, qui, lorsque l'on y fait $t = t_0$, reproduit la condition initiale supposée donnée, $\Psi(\vec{r}, t_0)$.

Avant d'en venir au formalisme puis au rappel des postulats de la Mécanique Quantique, il importe de rappeler dès maintenant l'un des succès majeurs et précoces de la nouvelle théorie, à savoir la quantification de l'énergie des états liés d'un système. La fonction d'onde de Schrödinger est soumise à des conditions très particulières, liées au sens physique qu'ont proposé de lui attribuer Born et Jordan en 1927. Ces derniers postulent que, pour un état lié, la probabilité dP de trouver la particule au voisinage d'un point repéré par le vecteur \vec{r} est liée au module au carré de la fonction d'onde calculée en ce point. Très précisément, Born et Jordan posent, dans \mathbb{R}^3 :

$$dP(\vec{r}, t) = |\Psi(\vec{r}, t)|^2 d^3r , \quad (1.26)$$

où d^3r est le volume infinitésimal entourant l'extrémité de \vec{r} ; il en résulte que $\Psi(\vec{r}, t)$ a pour dimension physique¹³ [longueur]^{-3/2}. Pour que l'interprétation de Born et Jordan soit compatible avec l'équation de Schrödinger, il doit exister une équation de conservation, décrivant l'écoulement sans perte du "fluide" de probabilité. Une telle équation est en fait contenue dans l'équation (1.1) elle-même. En effet, en posant :

$$\rho(\vec{r}, t) = \Psi^*(\vec{r}, t) \Psi(\vec{r}, t) , \quad \vec{j}(\vec{r}, t) = \frac{\hbar}{2im} [\Psi^* \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^*] , \quad (1.27)$$

on vérifie facilement à partir de (1.1) que la densité ρ et le courant \vec{j} satisfont l'équation de conservation requise :

$$\frac{\partial}{\partial t} \rho(\vec{r}, t) + \vec{\nabla} \cdot \vec{j}(\vec{r}, t) = 0 . \quad (1.28)$$

La relation (1.26) doit bien être considérée comme un postulat, qui confère à la théorie une nature essentiellement *probabiliste*. *A priori*, il en résulte que l'intégrale du module carré doit être une quantité finie¹⁴ :

$$\int_{\mathbb{R}^3} dP(\vec{r}, t) \equiv \int_{\mathbb{R}^3} |\Psi(\vec{r}, t)|^2 d^3r < \infty . \quad (1.29)$$

En d'autres termes, la fonction d'onde doit être une fonction de carré sommable – on dit qu'elle est *normalisable* ; comme d'habitude, la probabilité de l'événement certain est posée conventionnellement égale à 1 – ce qui est assuré en multipliant Ψ par le bon coefficient (constante de normalisation), tel que :

$$\int_{\mathbb{R}^3} |\Psi(\vec{r}, t)|^2 d^3r = 1 . \quad (1.30)$$

Dans certaines applications, il peut sembler que ce n'est pas la condition de normalisabilité (1.30) qui joue le rôle-clé pour l'apparition spontanée de la quantification des états liés. Par exemple, en présence d'un saut fini de potentiel à une dimension, on doit écrire que la fonction d'onde et sa dérivée sont continues et ce sont ces conditions qui produisent la quantification lorsque l'on recherche des états liés. En réalité, il faut bien voir que

¹³ De la même façon, $|\Phi(\vec{p}, t)|^2 d^3p$ est la probabilité élémentaire de trouver la particule avec une impulsion \vec{p} dont l'extrémité est dans le volume d^3p de l'espace des impulsions. La dimension de $\Phi(\vec{p}, t)$ est [impulsion]^{-3/2}

¹⁴ Toutes les solutions de l'équation de Schrödinger représentant un état lié (c'est-à-dire localisé) satisfont la condition (1.29). Par la suite, on introduira également des solutions non-normalisables, par exemple des ondes planes, dont le sens physique devra être précisé ; ces solutions représentent les états stationnaires de diffusion où la particule ne reste pas dans le voisinage d'un centre d'attraction, aussi appelés états non-liés.

tout se tient : l'équation de Schrödinger contient aussi, comme on l'a vu, une équation de conservation pour la densité de probabilité et le courant qui lui est associé, (1.28), qui est l'expression locale de la conservation alors que (1.30) en est l'expression intégrale. Par ailleurs, c'est la nature même de (1.1) qui entraîne la continuité de la fonction d'onde et de sa dérivée, même quand V a un saut fini. D'une façon ou d'une autre, ce sont les conditions requises pour la fonction d'onde, et l'équation qu'elle satisfait, qui produisent quand il le faut la quantification spontanée de l'énergie et d'autres grandeurs (le moment cinétique orbital, par exemple).

Ceci étant, la solution $\Psi(\vec{r}, t)$ cherchée, descendante de l'état initial prescrit $\Psi(\vec{r}, t_0)$, s'écrit en combinaison linéaire des fonctions propres ψ_E ; la condition de normalisabilité se reporte bien évidemment sur ces dernières fonctions¹⁵, qui sont jusqu'à présent seulement contraintes à être solutions de l'équation aux valeurs propres. Il se trouve, fait remarquable, que la condition de normalisabilité vient faire le tri, sur des considérations purement physiques (le sens de la fonction d'onde) dans l'ensemble (en général infini non-dénombrable) de toutes les solutions de l'équation aux valeurs propres. En d'autres termes, la nécessité de normalisabilité impose de passer au crible (au peigne fin) toutes les solutions mathématiques du problème, pour ne retenir qu'un sous-ensemble discret (lui aussi le plus souvent infini, mais dénombrable). Cette propriété permet d'indicer les énergies possibles par un (ou plusieurs) entier(s), n ("nombre quantique(s)") et de désigner l'ensemble physiquement acceptable par $\{E_n\}$; en bonne logique, on pourrait alors désigner par ψ_{E_n} les fonctions propres correspondantes – dans un souci d'économie d'écriture, on les notera simplement ψ_n .

Ainsi, l'équation de Schrödinger agrémentée des conditions requises pour ses solutions, conditions dictées par des considérations physiques, produit ce qui avait jusqu'alors résisté à toutes les descriptions théoriques cohérentes : la quantification spontanée de l'énergie des états liés¹⁶. C'est certainement le plus spectaculaire succès précoce de la nouvelle théorie.

La théorie étant fondamentalement probabiliste, toutes les prévisions qui pourront en être tirées portent sur des valeurs *moyennes*, au sens usuel de l'espérance mathématique. Par exemple, la valeur moyenne de la position, \vec{r} , est, en représentation- q , donnée par l'intégrale :

$$\langle \vec{r} \rangle = \int_{\mathbb{R}^3} \vec{r} dP = \int_{\mathbb{R}^3} \vec{r} |\Psi(\vec{r}, t)|^2 d^3r \equiv \int_{\mathbb{R}^3} \Psi^*(\vec{r}, t) \vec{r} \Psi(\vec{r}, t) d^3r , \quad (1.31)$$

étant entendu que la fonction d'onde est supposée préalablement normalisée à l'unité. De la même façon, la valeur moyenne du moment conjugué est :

$$\langle \vec{p} \rangle = \int_{\mathbb{R}^3} \Psi^*(\vec{r}, t) (-i\hbar \vec{\nabla}) \Psi(\vec{r}, t) d^3r . \quad (1.32)$$

Plus généralement, la valeur moyenne d'une fonction $A(\vec{r}, \vec{p})$ est donnée par :

$$\langle A \rangle = \int_{\mathbb{R}^3} \Psi^*(\vec{r}, t) A(\vec{r}, -i\hbar \vec{\nabla}) \Psi(\vec{r}, t) d^3r . \quad (1.33)$$

La fonction $A(\vec{r}, -i\hbar \vec{\nabla})$ se forme à partir de son expression classique, en symétrisant le cas échéant les produits du genre up_u , qui contiennent deux opérateurs non-commutatifs. Comme d'habitude, l'opérateur différentiel sous l'intégrale de (1.33) agit sur ce qui est situé à sa droite. Ainsi, pour une particule de masse m située dans un champ de forces définissant une énergie potentielle¹⁷ $V(\vec{r})$, la valeur moyenne de l'énergie est :

$$\langle H \rangle = \int_{\mathbb{R}^3} \Psi^*(\vec{r}, t) \left[-\frac{\hbar^2}{2m} \Delta + V(\vec{r}) \right] \Psi(\vec{r}, t) d^3r . \quad (1.34)$$

Pour terminer, montrons pour la coordonnée comment les choses s'écrivent en représentation- p , en raisonnant à une dimension pour simplifier. La valeur moyenne est ici :

$$\langle x \rangle = \int_{\mathbb{R}} \Psi^*(x, t) x \Psi(x, t) dx . \quad (1.35)$$

¹⁵Pour que toute fonction Ψ développée sur les $\{\psi_E\}_E$ soit normalisable, il est évidemment *nécessaire* que chaque ψ_E le soit. En outre, on peut se poser la question de la complétude des $\{\psi_E\}_E$, permettant de représenter en série *n'importe quel* état initial ; la complétude est assurée par l'hermiticité du Hamiltonien.

¹⁶D'un autre côté, la quantification en tant que conséquence des conditions aux limites imposées aux solutions d'une équation aux dérivées partielles est en soi un phénomène banal : penser à la corde vibrante dont les extrémités sont fixes.

¹⁷En Mécanique Quantique, on commet presque toujours l'abus de langage consistant à appeler "potentiel" ce qui est en réalité l'énergie potentielle.

Maintenant, on exprime chaque Ψ en intégrale de Fourier, conformément à (1.14) :

$$\Psi^*(x, t) = (2\pi\hbar)^{-1/2} \int_{\mathbb{R}} dp e^{+\frac{1}{\hbar} xp} \Phi^*(p, t) , \quad \Psi(x, t) = (2\pi\hbar)^{-1/2} \int_{\mathbb{R}} dp' e^{-\frac{1}{\hbar} xp'} \Phi(p', t) , \quad (1.36)$$

de sorte que (1.35) devient :

$$\langle x \rangle = (2\pi\hbar)^{-1} \int_{\mathbb{R}} dp \int_{\mathbb{R}} dp' \int_{\mathbb{R}} dx e^{+\frac{1}{\hbar} xp} \Phi^*(p, t) x e^{-\frac{1}{\hbar} xp'} \Phi(p', t) . \quad (1.37)$$

Il suffit de remarquer que le produit des trois facteurs à droite s'écrit aussi :

$$x e^{-\frac{1}{\hbar} xp'} \Phi(p', t) = -i\hbar \frac{\partial}{\partial p'} \left[e^{-\frac{1}{\hbar} xp'} \Phi(p', t) \right] , \quad (1.38)$$

pour obtenir :

$$\langle x \rangle = (2\pi\hbar)^{-1} \int_{\mathbb{R}} dp \int_{\mathbb{R}} dp' \Phi^*(p, t) \Phi(p', t) (-i\hbar) \frac{\partial}{\partial p'} \int_{\mathbb{R}} dx e^{+\frac{1}{\hbar} x(p-p')} . \quad (1.39)$$

En utilisant $\int_{-\infty}^{+\infty} e^{ixy} dx = 2\pi\delta(y)$, $\delta(ax) = (1/|a|)\delta(x)$ et $\langle \delta', f \rangle = -f'(0)$, on trouve finalement :

$$\langle x \rangle = \int_{\mathbb{R}} dp \int_{\mathbb{R}} dp' \Phi^*(p, t) \Phi(p', t) (-i\hbar) \frac{\partial}{\partial p'} \delta(p-p') = \int_{\mathbb{R}} dp \Phi^*(p, t) (+i\hbar) \frac{\partial}{\partial p} \Phi(p, t) . \quad (1.40)$$

Cette expression confirme bien l'association $x \rightarrow +i\hbar\partial/\partial p$ en représentation- p .

1.2 Formalisme de Dirac

L'objet central de la théorie de Schrödinger, la fonction d'onde, est réputé contenir toute l'information possible sur un système. Cette connaissance est polymorphe et peut être représentée de plusieurs façons équivalentes ; par exemple, étant donnée la fonction d'onde au sens de Schrödinger, $\Psi(\vec{r}, t)$, il est possible d'introduire sa transformée de Fourier $\Phi(\vec{p}, t)$, conformément à (1.13), où \vec{p} est ici un vecteur au sens ordinaire et non pas l'opérateur $-i\hbar\vec{\nabla}$. Manifestement, cette nouvelle fonction contient exactement la même information que Ψ : quand on connaît l'une de ces fonctions, on peut en déduire l'autre. Elle est notée d'un symbole différent (Φ au lieu de Ψ) puisque, en tant que fonction au sens mathématique du terme, Φ ne coïncide pas avec Ψ . Il n'empêche que la fonction Ψ et la fonction Φ ne sont que deux écritures différentes d'une seule et même abstraction mathématique ; chacune d'entre elles constitue la donnée d'une infinité (non-dénombrable) de nombres (les valeurs successives de la fonction considérée aux différents points de l'espace des coordonnées ou des impulsions) ; l'ensemble de ces nombres constitue le champ quantique : c'est bien pourquoi on dit d'un champ qu'il a une infinité de degrés de liberté. Pour identifier la nature de cet objet, on peut aussi raisonner par analogie avec une autre situation plus usuelle et l'idée qui s'impose, au moins par sa simplicité, est celle de vecteur : le vecteur vitesse d'un point matériel est bien défini en soi ; selon que l'on prend un repère ou un autre, ce même vecteur sera représenté par l'une ou l'autre d'une suite de trois nombres (les composantes sur la base choisie), mais chacune de ces suites représente exactement la même réalité – le même vecteur.

Enfin, la notion de vecteur s'impose également au vu de la linéarité de l'équation de Schrödinger : toute combinaison linéaire quelconque de plusieurs solutions est encore une solution ; la combinaison linéaire est l'opération fondamentale permettant d'introduire la notion d'espace vectoriel – il est d'ailleurs bien connu que l'ensemble des solutions d'une équation différentielle du second ordre peut être muni d'une structure d'espace vectoriel, de dimension égale à 2.

Pour signaler explicitement la nature vectorielle sous-jacente, il est recommandé d'utiliser une notation particulière, afin de distinguer l'objet mathématique vecteur d'une part, Ψ , ses composantes d'autre part (l'ensemble des nombres $\{\Psi(\vec{r}, t)\}$, \vec{r} variant à t fixé¹⁸). L'usage, à la suite de Dirac, a consacré, non la flèche ordinaire de l'analyse vectorielle standard, mais le symbole $|\ \rangle$, appelé ket ; on notera donc comme suit l'état d'un

¹⁸On verra dans la suite pourquoi, de fait, on peut écrire : $\Psi(\vec{r}, t) = \langle \vec{r} | \Psi(t) \rangle$.

système à l'instant t : $|\Psi(t)\rangle$. Le corps sur lequel est défini l'espace vectoriel est celui des nombres complexes, \mathbb{C} ; ceci est bien évident puisque le nombre complexe fondamental i apparaît dans l'équation de Schrödinger et, comme on le verra par la suite, constitue une nécessité primordiale de la Mécanique Quantique.

Un vecteur étant donné, le choix d'une base permet de définir les composantes de ce vecteur et en constitue la *représentation*. Comme on le sait, il est parfois possible formellement d'effectuer toutes les opérations structurelles sur un espace vectoriel sans référence à une base donnée : ainsi, on peut souvent intégrer vectoriellement l'équation fondamentale de la Dynamique, sans projeter sur trois axes de coordonnées. D'un autre côté, certains calculs se font parfois plus commodément quand une base a été choisie, à condition qu'elle l'ait été judicieusement.

Une base est un ensemble de vecteurs linéairement indépendants en nombre égal à la dimension de l'espace vectoriel ; M vecteurs $\{|e_n\rangle\}_{1 \leq n \leq M}$, sont dits linéairement indépendants si l'équivalence suivante est vraie :

$$\sum_{n=1}^M \lambda_n |e_n\rangle = 0 \iff \lambda_n = 0 \quad \forall n \in [1, M] . \quad (1.41)$$

Si $M = N$ (la dimension de l'espace vectoriel \mathcal{E}), les $\{|e_n\rangle\}_{1 \leq n \leq N}$ forment alors une base de \mathcal{E} : il est en effet facile de montrer que tout vecteur $|\psi\rangle$ appartenant à \mathcal{E} peut se décomposer sur les $\{|e_n\rangle\}$:

$$\forall |\psi\rangle \in \mathcal{E} : \exists \{c_n\}_{1 \leq n \leq N} \quad |\psi\rangle = \sum_{n=1}^N c_n |e_n\rangle . \quad (1.42)$$

où les $\{c_n\}_n$ sont des nombres complexes. La décomposition suivant les $\{e_n\}$ est unique et constitue la suite des composantes de $|\psi\rangle$ sur la base $|e_n\rangle$; par nature, les composantes d'un vecteur donné dépendent de la base choisie. Comme il existe un nombre infini de bases (pour un espace vectoriel de dimension supérieure à 1), il convient de savoir exprimer les composantes d'un même vecteur relativement à une base en fonction de ses composantes relativement à une autre. Passer d'une base à une autre est ce que l'on appelle effectuer un changement de base, dont on donne maintenant les relations caractéristiques.

Soit deux bases distinctes d'un même espace vectoriel \mathcal{E} , $|e_n\rangle$ et $|f_p\rangle$, et un vecteur $|\psi\rangle$ de composantes respectives c_n et d_p . On a par définition :

$$|\psi\rangle = \sum_{n=1}^N c_n |e_n\rangle \quad |\psi\rangle = \sum_{p=1}^N d_p |f_p\rangle . \quad (1.43)$$

Il s'agit de trouver les relations entre les c_n et les d_p . Pour cela, il convient évidemment de dire comment les bases sont reliées entre elles ; on doit se donner, par exemple, l'expression de chaque $|e_n\rangle$ sur la base $|f_p\rangle$, ou l'inverse. Supposons donc connus les N développements donnant les "anciens" vecteurs de base en fonction des "nouveaux" :

$$|e_n\rangle = \sum_{p=1}^N T_{pn} |f_p\rangle . \quad (1.44)$$

Les N^2 nombres T_{np} forment ce que l'on appelle la matrice de passage de $|e_n\rangle$ à $|f_p\rangle$; on y trouve, dans la $n^{\text{ème}}$ colonne, les composantes du vecteur $|e_n\rangle$ exprimé sur la nouvelle base. En reportant ceci dans la première expression de $|\psi\rangle$, il vient :

$$|\psi\rangle = \sum_{n=1}^N \sum_{p=1}^N c_n T_{pn} |f_p\rangle \quad (1.45)$$

soit :

$$\sum_{n=1}^N \sum_{p=1}^N c_n T_{pn} |f_p\rangle = \sum_{p=1}^N d_p |f_p\rangle \iff \sum_{p=1}^N \left[\sum_{n=1}^N T_{pn} c_n - d_p \right] |f_p\rangle = 0 . \quad (1.46)$$

Comme $\{|f_p\rangle\}$ est une base, ses éléments sont linéairement indépendants. Selon (1.41), la dernière égalité ne peut être vraie que si chaque coefficient de la somme sur p est nul ; on en déduit la relation cherchée entre

anciennes et nouvelles composantes :

$$d_p = \sum_{n=1}^N T_{pn} c_n . \quad (1.47)$$

Il est aussi possible de procéder dans l'autre sens, en se donnant :

$$|f_p\rangle = \sum_{n=1}^N T'_{np} |e_n\rangle , \quad (1.48)$$

où les T'_{np} forment la matrice de la transformation inverse¹⁹ de T , que l'on peut donc aussi noter T^{-1} ; cette matrice à nouveau contient dans la $p^{\text{ème}}$ colonne les composantes du vecteur $|f_p\rangle$ exprimé sur les $\{|e_n\rangle\}$. La même démarche que ci-dessus conduit à :

$$c_n = \sum_{p=1}^N T'_{np} d_p \equiv \sum_{p=1}^N T_{np}^{-1} d_p . \quad (1.49)$$

D'ailleurs, en reportant l'expression des d_p , eq. (1.47), dans (1.49), il vient :

$$c_n = \sum_{p=1}^N \sum_{n'=1}^N T_{np}^{-1} T_{pn'} c_{n'} , \quad (1.50)$$

d'où la relation :

$$\sum_{p=1}^N T_{np}^{-1} T_{pn'} = \delta_{nn'} , \quad (1.51)$$

qui reproduit la règle de multiplication de deux matrices ; ici il s'agit d'une matrice T et de son inverse T^{-1} , dont le produit dans n'importe quel ordre est égal à l'unité. On remarque que les relations donnant les *nouvelles* composantes d_p en fonction des *anciennes* c_n impliquent les nombres T_{pn} exprimant les *anciens* vecteurs de base $|e_n\rangle$ en fonction des *nouveaux* $|f_p\rangle$; ceci fait penser à une opération blanche : au fond, ce qui est derrière, c'est l'invariance de l'objet $|\psi\rangle$ qui est défini intrinsèquement, indépendamment de toute base.

Un vecteur (ket) de l'espace des états peut donc ainsi être représenté par une suite de nombres (ses composantes) quand une base a été choisie ; ces nombres peuvent être rangés en colonne pour former un "vecteur-colonne" comme on le fait en analyse vectorielle ordinaire :

$$|\psi\rangle, \{|e_n\rangle\}_n \quad \longleftrightarrow \quad \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} , \quad (1.52)$$

et la relation de changement de base (1.47) s'écrit alors explicitement :

$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} & \dots & T_{1N} \\ T_{21} & T_{22} & \dots & T_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ T_{N1} & T_{N2} & \dots & T_{NN} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix} . \quad (1.53)$$

La représentation d'un vecteur ket par un vecteur-colonne est évidemment aussi valable pour chacun des vecteurs

¹⁹L'inverse de T existe puisque T transforme une base en une autre. Si T était singulière (une valeur propre nulle au moins), l'ensemble transformé par T agissant sur la base de départ contiendrait moins de N éléments non-nuls et ne saurait être une base de l'espace de dimension N .

de la base. On a ainsi :

$$|e_n\rangle, \{|e_n\rangle\}_n \longleftrightarrow \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (1.54)$$

où la seule composante non nulle (et égale à 1) est située sur la $n^{\text{ème}}$ ligne.

La notion de produit scalaire joue un rôle fondamental dans la théorie des espaces vectoriels. Ici comme ailleurs, il est fort utile d'en définir un, effectuant au passage une petite généralisation puisque le corps est ici celui des complexes \mathbb{C} et non pas \mathbb{R} , celui des réels. Le produit scalaire de deux vecteurs $|\psi\rangle$ et $|\phi\rangle$ est un nombre, ici c'est un nombre complexe, que l'on va noter provisoirement $(|\psi\rangle, |\phi\rangle)$. Cette opération est linéaire par rapport au second membre du couple :

$$(|\psi\rangle, \lambda_1|\phi_1\rangle + \lambda_2|\phi_2\rangle) = \lambda_1(|\psi\rangle, |\phi_1\rangle) + \lambda_2(|\psi\rangle, |\phi_2\rangle) . \quad (1.55)$$

Au contraire de ce que l'on sait du produit scalaire dans un espace vectoriel sur les réels, le produit scalaire est ici *non-commutatif*, parce que le corps est ici \mathbb{C} ; on pose de surcroît la relation suivante :

$$(|\psi\rangle, |\phi\rangle) = (|\phi\rangle, |\psi\rangle)^* , \quad (1.56)$$

où * note la conjugaison complexe ; il résulte en particulier :

$$(\lambda|\psi\rangle, \mu|\phi\rangle) = \lambda^*\mu(|\psi\rangle, |\phi\rangle) . \quad (1.57)$$

La raison du choix (1.56) se tient dans le souci de pouvoir définir une norme – et une distance – qui sont des quantités positives. D'ordinaire, le produit scalaire d'un vecteur par lui-même est le carré de sa norme ; ici, si l'on prend $\lambda = \mu$ et $\psi = \phi$, on a :

$$(\lambda|\psi\rangle, \lambda|\psi\rangle) = \lambda^*\lambda(|\psi\rangle, |\psi\rangle) = |\lambda|^2(|\psi\rangle, |\psi\rangle) \quad (1.58)$$

et ceci s'interprète en disant que le carré de la norme du vecteur $\lambda|\psi\rangle$ est égal à celui de $|\psi\rangle$ multiplié par le module carré du facteur d'échelle λ , ce qui est bien normal. On retiendra donc au passage la définition :

$$(|\psi\rangle, |\psi\rangle) = \text{carré de la norme de } |\psi\rangle \equiv \| |\psi\rangle \|^2 . \quad (1.59)$$

Deux vecteurs dont le produit scalaire est nul sont dits orthogonaux. En utilisant :

$$(|\psi\rangle, \mu|\phi\rangle) = \mu(|\psi\rangle, |\phi\rangle) \quad (1.60)$$

et en prenant $\mu = 0$ – alors quel que soit $|\phi\rangle$, $\mu|\phi\rangle$ est le vecteur nul – le second membre montre que le produit scalaire est nul, quel que soit l'autre vecteur $|\psi\rangle$ du couple. On en déduit que le vecteur nul est orthogonal à tous les vecteurs de l'espace vectoriel ; ce vecteur nul est le seul à posséder une telle propriété.

En définitive, les propriétés usuelles de linéarité d'un produit scalaire sont encore vraies, mais tiennent compte évidemment de la propriété de conjugaison caractéristique posée ci-dessus :

$$(|\psi\rangle, \mu_1|\phi_1\rangle + \mu_2|\phi_2\rangle) = \mu_1(|\psi\rangle, |\phi_1\rangle) + \mu_2(|\psi\rangle, |\phi_2\rangle) \quad (1.61)$$

$$(\lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle, |\phi\rangle) = \lambda_1^*(|\psi_1\rangle, |\phi\rangle) + \lambda_2^*(|\psi_2\rangle, |\phi\rangle) \quad (1.62)$$

Un tel produit scalaire est dit antilinéaire défini positif²⁰ ; le anti signifie qu'il faut appliquer les relations de conjugaison au premier vecteur du couple ($|\psi\rangle$, $|\phi\rangle$) ; le qualificatif défini positif rappelle que le produit scalaire

²⁰On dit aussi parfois : sesquilinéaire.

d'un vecteur par lui-même est positif (c'est le carré de sa norme), sauf pour le vecteur nul dont la norme est nulle.

Lorsque les deux vecteurs sont donnés par leurs composantes sur une base fixée, leur produit scalaire, ainsi défini, peut s'exprimer en fonction de ces composantes. Soit en effet :

$$|\psi\rangle = \sum_{n=1}^N x_n |e_n\rangle \quad |\phi\rangle = \sum_{m=1}^N y_m |e_m\rangle ; \quad (1.63)$$

alors, en vertu des équations (1.61) et (1.62) :

$$(|\psi\rangle, |\phi\rangle) = \sum_{n=1}^N \sum_{m=1}^N x_n^* y_m (|e_n\rangle, |e_m\rangle) . \quad (1.64)$$

Les produits scalaires des vecteurs de base deux à deux sont des caractéristiques de la base choisie et l'on introduit un objet à deux indices, g_{nm} , représentant ce que l'on appelle parfois la métrique de la base :

$$(|e_n\rangle, |e_m\rangle) = g_{nm} \quad g_{nm} = g_{mn}^* . \quad (1.65)$$

La deuxième relation caractérise par définition une métrique *hermitique*. Avec ceci, le produit scalaire (1.64) est :

$$(|\psi\rangle, |\phi\rangle) = \sum_{n=1}^N \sum_{m=1}^N x_n^* g_{nm} y_m . \quad (1.66)$$

Des bases particulières joueront un rôle important dans toute la suite ; il s'agit des bases orthonormées, dont tous les vecteurs sont de norme unité et deux à deux orthogonaux ; pour une telle base, par définition :

$$(|e_n\rangle, |e_m\rangle) = \delta_{nm} \quad (1.67)$$

où δ_{nm} est toujours le symbole de Kronecker. Sur une telle base, le produit scalaire de deux vecteurs prend une forme très simple :

$$(|\psi\rangle, |\phi\rangle) = \sum_{n=1}^N x_n^* y_n \quad (1.68)$$

et le carré de la norme d'un vecteur, $\| |\psi\rangle \|^2$, est :

$$\| |\psi\rangle \|^2 = \sum_{n=1}^N |x_n|^2 . \quad (1.69)$$

Par comparaison avec la norme dans un espace vectoriel sur le corps des réels, ceci apparaît bien comme en étant la généralisation la plus naturelle. On prendra garde à ne pas oublier que les expressions (1.68) et (1.69) ne sont correctes que si la base est orthonormée. *Dans toute la suite, on ne considèrera que des bases orthonormées*²¹. Sur la dernière écriture, on voit bien que le vecteur nul qui, par définition a toutes ses composantes nulles, est le seul et unique vecteur de norme nulle et orthogonal à tout autre vecteur de l'espace \mathcal{E} .

Toutes ces expressions peuvent recevoir une écriture sous forme matricielle. En effet, une base orthonormée étant donnée, l'expression (1.69) du produit scalaire montre qu'il convient d'introduire le vecteur-ligne :

$$\left[x_1^* \quad x_2^* \quad \dots \quad x_N^* \right] , \quad (1.70)$$

ce qui permet d'écrire :

$$(|\psi\rangle, |\phi\rangle) = \left[x_1^* \quad x_2^* \quad \dots \quad x_N^* \right] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} . \quad (1.71)$$

²¹Une telle hypothèse n'est évidemment pas nécessaire mais elle simplifie considérablement la présentation du formalisme. Dans le cas contraire, il conviendrait de distinguer entre composantes covariantes et contravariantes d'un vecteur, avec toutes les autres complications habituelles (voir [2] par exemple).

L'introduction du vecteur-ligne pour représenter le premier vecteur du couple alors que le second, représenté par un vecteur-colonne, est noté $|\phi\rangle$, donne l'idée d'introduire une notation symétrique pour désigner un vecteur-ligne ; il va en résulter une notation très commode pour le produit scalaire. Dirac a proposé le symbole $\langle\psi|$ appelé bra :

$$\left[x_1^* \quad x_2^* \quad \dots \quad x_N^* \right] \longleftrightarrow \langle\psi| . \quad (1.72)$$

À ce stade, il est utile de préciser un point. Soit le vecteur $\lambda|\psi\rangle$ où λ est un complexe. La matrice-colonne représentant ce vecteur est obtenue en multipliant toutes les composantes par le même facteur :

$$\lambda|\psi\rangle, \{ |e_n\rangle \}_n \longleftrightarrow \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_N \end{bmatrix} . \quad (1.73)$$

À ce vecteur ket correspond donc le bra :

$$\lambda|\psi\rangle \equiv |\lambda\psi\rangle \longleftrightarrow \left[\lambda^* x_1^* \quad \lambda^* x_2^* \quad \dots \quad \lambda^* x_N^* \right] \longleftrightarrow \langle\lambda\psi| , \quad (1.74)$$

et ce dernier vecteur ligne est associé au bra $\lambda^*\langle\psi|$. On retiendra donc le caractère *antilinéaire* de la correspondance entre ket et bra :

$$\lambda|\psi\rangle \longleftrightarrow \lambda^*\langle\psi| \equiv \langle\lambda\psi| . \quad (1.75)$$

Ainsi, quand on sort un scalaire d'un bra, il ne faut pas oublier d'en prendre le complexe conjugué. Cette correspondance étant précisée, le produit scalaire se note :

$$(|\psi\rangle, |\phi\rangle) = \langle\psi||\phi\rangle . \quad (1.76)$$

La double barre verticale n'est pas nécessaire et on notera désormais, conformément à l'usage :

$$(|\psi\rangle, |\phi\rangle) = \langle\psi|\phi\rangle , \quad (1.77)$$

sans oublier que la barre centrale est "double" et que l'on peut détacher les fragments si besoin est :

$$\langle\psi|\phi\rangle \equiv \langle\psi_{\leftarrow}||_{\rightarrow}\phi\rangle \rightarrow \langle\psi| \quad |\phi\rangle . \quad (1.78)$$

La correspondance ci-dessus introduisant le bra est définie pour tout vecteur de \mathcal{E} ; en particulier pour un vecteur $|e_n\rangle$ de la base introduite :

$$\left[0 \quad 0 \quad \dots \quad 0 \quad 1 \quad 0 \dots \quad 0 \right] \longleftrightarrow \langle e_n| . \quad (1.79)$$

Comme en géométrie vectorielle ordinaire, il est possible d'exprimer les composantes de tout vecteur en fonction des produits scalaires de ce vecteur avec les vecteurs de base et la métrique de cette base. En effet, soit :

$$|\psi\rangle = \sum_{n=1}^N x_n |e_n\rangle . \quad (1.80)$$

En multipliant scalairement par $|e_m\rangle$, il vient :

$$\langle e_m|\psi\rangle \equiv (|e_m\rangle, |\psi\rangle) = \sum_{n=1}^N x_n (|e_m\rangle, |e_n\rangle) \equiv \sum_{n=1}^N x_n g_{mn} . \quad (1.81)$$

En introduisant la matrice inverse²² de la matrice g_{mn} , d'éléments g_{pq}^{-1} , on a :

$$x_n = \sum_{m=1}^N g_{nm}^{-1} (|e_m\rangle, |\psi\rangle) , \quad (1.82)$$

²²Cette matrice existe nécessairement : en tant que matrice représentant la métrique hermitique, g_{mn} est définie positive et possède donc un inverse (toutes ses valeurs propres sont strictement positives).

d'où l'expression :

$$|\psi\rangle = \sum_{n=1}^N \sum_{m=1}^N g_{nm}^{-1} \langle e_m | \psi \rangle |e_n\rangle = \sum_{n=1}^N \sum_{m=1}^N |e_n\rangle g_{nm}^{-1} \langle e_m | \psi \rangle . \quad (1.83)$$

Dans le cas d'une base orthonormée, g est la matrice identité, ainsi que g^{-1} , et on obtient simplement :

$$|\psi\rangle = \sum_{n=1}^N \sum_{m=1}^N \delta_{nm} \langle e_m | \psi \rangle |e_n\rangle = \sum_{n=1}^N |e_n\rangle \langle e_n | \psi \rangle . \quad (1.84)$$

Cette dernière relation est vraie quel que soit le vecteur $|\psi\rangle$. En "détachant" les deux partenaires du produit scalaire :

$$|\psi\rangle = \sum_{n=1}^N |e_n\rangle \langle e_n | \psi \rangle = \left[\sum_{n=1}^N |e_n\rangle \langle e_n | \right] |\psi\rangle , \quad (1.85)$$

on en déduit la relation fondamentale :

$$\sum_{n=1}^N |e_n\rangle \langle e_n | = \mathbf{1} \quad (1.86)$$

qui constitue la relation de fermeture pour une base orthonormée. Elle sera utilisée à maintes reprises dans la suite. Le nom donné à cette relation rappelle le fait que cette base est complète, c'est-à-dire engendre bel et bien l'espace vectoriel dans son entier.

Le produit scalaire permet donc d'exprimer simplement les composantes d'un vecteur donné. Il permet aussi de trouver les éléments de la matrice d'un opérateur linéaire Ω donné, relativement à une base. En effet, une base étant choisie, Ω est complètement défini par sa matrice exprimant les transformés de vecteurs de la base :

$$\Omega |e_n\rangle = \sum_{m=1}^N \Omega_{mn} |e_m\rangle . \quad (1.87)$$

Connaître tous les $\Omega |e_n\rangle$, c'est pouvoir écrire le transformé par Ω de n'importe quel vecteur appartenant à \mathcal{E} . Compte tenu de l'orthogonalité toujours supposée de la base, en prenant le produit scalaire de (1.87) avec $|e_p\rangle$, il vient :

$$(|e_p\rangle, \Omega |e_n\rangle) = \sum_{m=1}^N \Omega_{mn} (|e_p\rangle, |e_m\rangle) = \Omega_{pn} \quad (1.88)$$

soit :

$$\Omega_{pn} = (|e_p\rangle, \Omega |e_n\rangle) \equiv \langle e_p | \Omega |e_n\rangle . \quad (1.89)$$

À nouveau, la relation simple entre Ω_{pn} et $\langle e_p | \Omega |e_n\rangle$ exprimée par (1.89) n'est vraie que pour une base orthonormée²³.

Une base orthonormée étant donnée, il s'en déduit une infinité d'autres par des transformations linéaires non-singulières ; la nouvelle base n'est pas en général orthonormée, sauf dans le cas particulier important des transformations dites unitaires : par définition, une transformation unitaire transforme une base orthonormée en une autre base orthonormée. Établissons maintenant les relations caractéristiques de ces transformations.

Soit une base orthonormée $\{|e_n\rangle\}$ se transformant en une autre base $\{|f_p\rangle\}$ orthonormée par une transformation U unitaire ; soit $|f_p\rangle$ le transformé de $|e_p\rangle$ par U :

$$|f_p\rangle = U |e_p\rangle = \sum_{n=1}^N U_{np} |e_n\rangle . \quad (1.90)$$

Compte tenu de l'unitarité de U :

$$\langle f_p | f_{p'} \rangle \equiv (|f_p\rangle, |f_{p'}\rangle) = \delta_{pp'} \iff \sum_{n=1}^N \sum_{n'=1}^N U_{np}^* U_{n'p'} (|e_n\rangle, |e_{n'}\rangle) = \delta_{pp'} . \quad (1.91)$$

²³On peut définir la matrice représentant un opérateur sans avoir pour autant préalablement introduit un produit scalaire.

En utilisant maintenant l'orthonormalisation de la base $\{|e_n\rangle\}$, (1.91) :

$$\sum_{n=1}^N \sum_{n'=1}^N U_{np}^* U_{n'p'} \delta_{nn'} = \delta_{pp'} \iff \sum_{n=1}^N U_{np}^* U_{np'} = \delta_{pp'} \quad (1.92)$$

et comme la matrice inverse de la transformation U , U^{-1} , possède des éléments satisfaisant par définition :

$$\sum_{n=1}^N U_{pn}^{-1} U_{np'} = \delta_{pp'} \quad , \quad (1.93)$$

on en déduit la propriété fondamentale de symétrie des éléments de la matrice (sur une base orthonormée) d'un opérateur unitaire :

$$U_{pn}^{-1} = U_{np}^* \quad . \quad (1.94)$$

Tout naturellement, une matrice dont l'inverse est égal à sa transposée conjuguée est appelée matrice unitaire. Une transformation unitaire est représentée par une matrice unitaire à condition que la base soit orthonormée et alors, la matrice de l'inverse s'obtient simplement en conjuguant la matrice transposée²⁴. Une transformation unitaire conserve les produits scalaires : elle laisse donc inchangés les angles et les distances, et bien sûr les normes. En particulier, si deux vecteurs sont orthogonaux, leurs transformés le sont aussi et le transformé de chacun a la même norme que son antécédent. En résumé, pour toute transformation unitaire, les deux propriétés suivantes sont vraies :

$$|\psi'\rangle = U|\psi\rangle \quad , \quad |\phi'\rangle = U|\phi\rangle \iff \langle\psi'|\phi'\rangle = \langle\psi|\phi\rangle \quad (1.95)$$

$$\| |\psi'\rangle \| = \| |\psi\rangle \| \quad \forall |\psi\rangle \quad . \quad (1.96)$$

Dans un espace vectoriel tel que \mathcal{E} , les transformations unitaires généralisent les rotations dans l'espace vectoriel ordinaire \mathbb{R}^3 , souvent appelées transformations orthogonales.

Remarque

Sous-jacente à la correspondance ket \leftrightarrow bra se trouve la notion mathématique d'espace dual, qui est l'espace vectoriel des formes linéaires d'un espace vectoriel sur son corps. Cette notion n'est pas absolument nécessaire et sera donc laissée de côté.

Les résultats ci-dessus sont maintenant généralisés dans deux directions ; la première consiste à considérer un espace vectoriel de dimension infinie dénombrable ($N \rightarrow +\infty$), la seconde à introduire formellement un espace de dimension infinie ayant la puissance du continu. Ces deux généralisations sont immédiates sur le plan formel, mais mériteraient d'être rigoureusement étayées sur un strict plan mathématique. On se bornera ici, dans l'esprit déjà annoncé, à les donner sans justifications, en restant à un niveau intuitif. Toutes les notions ainsi étendues sont parfaitement fondées : les mathématiciens se sont chargés d'édifier les théories mathématiques dont les physiciens avaient besoin²⁵.

Commençons d'abord par le plus simple, conceptuellement. Pour un espace de dimension infinie dénombrable, une base contient une infinité dénombrable d'éléments et on écrit :

$$|\psi\rangle = \sum_{n=1}^{+\infty} c_n |e_n\rangle \quad . \quad (1.97)$$

²⁴L'unitarité d'une transformation (d'un opérateur) est une qualité intrinsèque : elle est indépendante de la notion de base ; la matrice représentant cet opérateur ne possède elle-même cette propriété que si la base est orthonormée.

²⁵Un exemple : la Théorie des Distributions de L. Schwartz. Bien sûr, les théories physiques n'ont pas eu besoin de telles justifications formelles pour se développer : la théorie de Schwartz remonte au début des années 50 et, pour ne citer qu'un exemple, les bases de l'Electrostatique ont été jetées à la fin du XVIII^{ème} siècle... L'infini du physicien – petit ou grand – n'est pas celui du mathématicien. Pour le physicien, l'infini désigne la valeur d'une grandeur qui, comparée à toute autre grandeur de même dimension déjà présente dans le problème examiné, peut être choisie aussi grande que l'on veut – par exemple la dimension du récipient contenant un gaz parfait, comparée à la distance (moyenne) entre deux atomes du gaz. En Physique, le passage à la limite ne représente jamais qu'une commodité technique – et un saut conceptuel : l'écriture *différentielle* de l'équation fondamentale de la dynamique permet d'utiliser les techniques de résolution des équations différentielles alors qu'expérimentalement, on est seulement capable de la vérifier avec des intervalles de temps discrets, certes de plus en plus petits quand la technologie expérimentale progresse, mais toujours *finis*, jamais infiniment petits au sens du mathématicien.

Il est manifeste que, sous-jacentes, se trouvent impliquées des conditions de convergence topologique. Pour deux bases différentes, avec les mêmes notations, on écrira les relations de passage entre composantes d'un même vecteur :

$$d_p = \sum_{n=1}^{+\infty} T_{pn} c_n . \quad (1.98)$$

La matrice de passage est donc une matrice de dimension infinie²⁶. De même, un vecteur bra ou ket sera représenté par un vecteur-ligne ou vecteur-colonne d'extension infinie. Le produit scalaire s'écrit maintenant :

$$(|\psi\rangle, |\phi\rangle) = \sum_{n=1}^{+\infty} \sum_{m=1}^{+\infty} x_n^* g_{nm} y_m . \quad (1.99)$$

Il est clair qu'une telle définition suppose implicitement que des conditions de convergence sont satisfaites : on manipule maintenant des séries au lieu de sommes finies. Si une base orthonormée a été choisie, le produit scalaire s'exprime simplement comme :

$$(|\psi\rangle, |\phi\rangle) = \sum_{n=1}^{+\infty} x_n^* y_n \quad (1.100)$$

et le carré de la norme d'un vecteur, $\| |\psi\rangle \|^2$, est :

$$\| |\psi\rangle \|^2 = \sum_{n=1}^{+\infty} |x_n|^2 . \quad (1.101)$$

En raisonnant formellement comme auparavant, on obtient la relation de fermeture :

$$\sum_{n=1}^{+\infty} |e_n\rangle \langle e_n| = \mathbf{1} . \quad (1.102)$$

On admet ainsi finalement que toutes les propriétés établies pour N fini peuvent être étendues au cas où N est infini ; la base de la justification est un passage à la limite convenable : c'est plus simple à dire qu'à faire en détail. Un tel espace vectoriel sur les complexes, de dimension infinie dénombrable et muni du produit scalaire sesquilineaire défini comme ci-dessus porte le nom d'espace de Hilbert²⁷.

Ceci étant fait, le cas des représentations "discrètes" se trouve réglé. La deuxième généralisation indispensable (bases, ou représentations, continues) est beaucoup moins évidente et nécessiterait, elle, des justifications spécifiques plutôt peu familières. Elle est, elle aussi, présentée de façon très intuitive.

Revenons au point de départ, le développement d'un vecteur sur une base, orthonormée ou non :

$$|\psi\rangle = \sum_n c_n |e_n\rangle . \quad (1.103)$$

Les $\{c_n\}_n$ constituent une suite de nombres, que l'on peut considérer comme l'ensemble des valeurs d'une fonction *discrète* c , définie sur les nombres entiers. Quant aux $\{|e_n\rangle\}$, ce sont les vecteurs formant la base ; le symbole e a été utilisé pour pouvoir distinguer cette base d'une autre, étant entendu que chacun des éléments de deux telles bases dépend d'un indice discret ; quand cette distinction n'est pas nécessaire, on peut noter simplement $\{|n\rangle\}$ les éléments d'une base. Avec ces notations simplifiées, on a :

$$|\psi\rangle = \sum_n c_n |n\rangle . \quad (1.104)$$

Généralisons maintenant une telle expression au cas où la variable n , au lieu de prendre des valeurs discrètes dans l'ensemble des entiers naturels \mathbb{N} , prend des valeurs continues dans \mathbb{R} , notées ν . La variable c devient alors

²⁶ La théorie des matrices de dimension infinie est postérieure à l'édification de la Mécanique Quantique.

²⁷ David Hilbert (1862 - 1943), illustre mathématicien allemand, à qui l'on doit en particulier la notion de corps. En outre, il est l'auteur d'une célèbre liste de 23 problèmes à résoudre, énoncée en 1900 lors du Congrès International des Mathématiciens tenu à Paris (voir [5], p. 183). Le texte de la conférence de Hilbert est disponible sur le site [6].

une *fonction*, au sens ordinaire du terme ; selon l'usage, on la désignera en faisant figurer son argument entre parenthèses – il est naturel de la noter $\psi(\nu)$, pour rappeler qu'elle représente le vecteur $|\psi\rangle$ (on aurait pu tout autant noter ψ_n les composantes c_n de $|\psi\rangle$). La généralisation naturelle de la sommation discrète est l'intégrale de Riemann et on écrira ainsi :

$$|\psi\rangle = \sum_n \psi_n |n\rangle \quad \rightarrow \quad |\psi\rangle = \int d\nu \psi(\nu) |\nu\rangle . \quad (1.105)$$

Bien sûr, cette relation étant admise, on pourrait généraliser encore en introduisant l'intégrale sur une fonction analytique, le long d'une ligne quelconque du plan complexe ; ceci reviendrait à dire que la variable continue, ν , peut elle-même avoir des valeurs complexes ; ceci sera exclu dans toute la suite. Par ailleurs, il n'est pas évident que l'élément d'intégration soit simplement $d\nu$: on pourrait aussi introduire une "mesure" $d\rho(\nu)$; ceci ne sera pas non plus considéré dans la suite.

En tout état de cause, la décomposition (1.105), une fois acceptée, se prête à des manipulations similaires à celles effectuées sur les sommes discrètes, finies ou infinies, puisque tout repose en définitive sur le caractère *linéaire* du formalisme. Ainsi, le produit scalaire de deux vecteurs :

$$|\psi\rangle = \int d\nu \psi(\nu) |\nu\rangle , \quad |\phi\rangle = \int d\nu \phi(\nu) |\nu\rangle \quad (1.106)$$

peut s'écrire, en n'utilisant que ses propriétés de sesquilinearité :

$$\langle\psi|\phi\rangle \equiv (|\psi\rangle, |\phi\rangle) = \int d\nu \int d\nu' \psi^*(\nu) \phi(\nu') \langle\nu|\nu'\rangle . \quad (1.107)$$

Les bornes d'intégrations ne sont pas mentionnées, pour la simplicité des notations. Elles définissent un intervalle d'intégration, qui est manifestement le même pour les deux variables ν et ν' . Essayons maintenant de définir la propriété pour une base continue $\{|\nu\rangle\}$ d'être orthonormée. Si tel est le cas, l'expression du produit scalaire doit apparaître comme la généralisation de :

$$\langle\psi|\phi\rangle \equiv (|\psi\rangle, |\phi\rangle) = \sum_n x_n^* y_n , \quad (1.108)$$

c'est-à-dire que l'on doit avoir, sur une base orthonormée continue :

$$\langle\psi|\phi\rangle \equiv (|\psi\rangle, |\phi\rangle) = \int d\nu \psi^*(\nu) \phi(\nu) . \quad (1.109)$$

Ceci n'est possible que si le produit scalaire $\langle\nu|\nu'\rangle$ est tel que :

$$\langle\nu|\nu'\rangle = 0 \quad \text{si } \nu \neq \nu' \quad \langle\nu|\nu'\rangle \neq 0 \quad \text{si } \nu = \nu' , \quad (1.110)$$

et est l'analogue d'un symbole de Kronecker continu, que l'on peut noter provisoirement $\delta(\nu, \nu')$ et appeler, tout aussi provisoirement et par analogie, symbole de Dirac. Les dernières relations écrites forment une exigence incompatible avec ce que l'on sait de l'intégration ordinaire ; en effet, il est possible de supprimer un point dans une intégrale sans modifier la valeur de l'intégrale, pourvu que l'intégrand ait une valeur finie au point que l'on supprime – ou une discontinuité d'amplitude finie. De même, dans le plan \mathbb{R}^2 d'intégration pour le couple (ν, ν') , si on retire la première bissectrice $\nu = \nu'$ de la région d'intégration, la valeur de l'intégrale ne change pas. Comme le produit scalaire $\langle\nu|\nu'\rangle$ est supposé nul partout ailleurs, l'une des deux intégrales, celle sur ν' par exemple, est identiquement nulle : il semble que l'on soit dans une impasse. En réalité, il existe une issue : supprimer un point ne change rien si l'intégrand est *fini* ; mais que se passe-t-il si on ôte un point où l'intégrand est infini ? On sait bien que la combinaison du zéro et de l'infini peut donner tout et n'importe quoi, y compris un nombre fini. Il est donc raisonnable d'imaginer une procédure où enlever le "bon" point, celui où une "bonne" divergence se produit, n'est pas une opération neutre. On admet ainsi que :

$$\langle\nu|\nu'\rangle = \begin{cases} 0 & \text{si } \nu \neq \nu' \\ +\infty & \text{si } \nu = \nu' \end{cases} \quad (1.111)$$

ce qui définit un objet plutôt exotique appelé “fonction” de Dirac, usuellement notée $\delta(\nu - \nu')$:

$$\delta(\nu - \nu') \equiv \langle \nu | \nu' \rangle = \begin{cases} 0 & \text{si } \nu \neq \nu' \\ +\infty & \text{si } \nu = \nu' \end{cases} . \quad (1.112)$$

Clairement, il ne s’agit pas d’une fonction au sens ordinaire du terme, mais d’un objet se prêtant à un calcul symbolique dont les règles sont les suivantes. Revenons à l’expression du produit scalaire qui a motivé l’introduction de la fonction δ :

$$\int d\nu \int d\nu' \psi^*(\nu) \phi(\nu') \delta(\nu - \nu') = \int d\nu \psi^*(\nu) \phi(\nu) \quad (1.113)$$

et réécrivons les choses un peu différemment :

$$\int d\nu \psi^*(\nu) \left[\int d\nu' \phi(\nu') \delta(\nu - \nu') \right] = \int d\nu \psi^*(\nu) \phi(\nu) . \quad (1.114)$$

En comparant les deux membres, on voit que la règle fondamentale de calcul avec la fonction δ de Dirac est :

$$\int d\nu' \phi(\nu') \delta(\nu - \nu') = \phi(\nu) \quad \forall \text{ fonction } \phi , \quad (1.115)$$

où il est sous-entendu que ν' , en parcourant son intervalle d’intégration, rencontre, en dehors des bornes, la valeur ν . Ceci est une conséquence automatique du fait que ν et ν' sont définis sur le même intervalle ; si les intervalles en ν et ν' étaient disjoints, l’intégrale au premier membre de (1.115) serait nulle. En outre, pour que l’écriture ci-dessus – déjà symbolique puisque l’intégrale est une intégrale ordinaire au sens de Riemann – ait un sens, il faut évidemment que la valeur de la fonction ϕ au point ν existe. Il est donc prudent de préciser :

$$\int_a^b d\nu' \phi(\nu') \delta(\nu - \nu') = \phi(\nu) \quad \forall \text{ fonction } \phi \text{ définie en } \nu \text{ et si } a < \nu < b . \quad (1.116)$$

Ceci constitue la règle d’usage de δ . En particulier, si ϕ est la fonction qui vaut 1 partout :

$$\int_a^b d\nu' \delta(\nu - \nu') = 1 \quad (a < \nu < b) . \quad (1.117)$$

De cette dernière relation il résulte que si la variable ν a une dimension physique, δ a la dimension inverse ; par exemple, si ν est une longueur, x , $\delta(x)$ est homogène à l’inverse d’une longueur, si ν est une fréquence, alors $\delta(\nu)$ est homogène à un temps, etc. Les relations précédentes généralisent bien ce que l’on sait du symbole de Kronecker, δ_{nm} :

$$\sum_{m=1}^N \phi_m \delta_{nm} = \phi_n \quad \forall \text{ suite } \{\phi_m\} \text{ définie en } n \text{ et si } 1 \leq m \leq N , \quad (1.118)$$

et :

$$\sum_{m=1}^N \delta_{nm} = 1 \quad (1 \leq m \leq N) . \quad (1.119)$$

La relation (1.116) doit en réalité (comme toujours) être perçue comme résultant d’un processus de limite – et c’est ainsi que l’on peut se convaincre de la sagesse des relations symboliques ci-dessus. Le symbole de Kronecker vaut 0 partout sauf en un point “entier”, où il vaut 1 ; la fonction de Dirac doit donc lui ressembler : par analogie, s’il faut la prendre infinie au seul point où elle n’est pas nulle c’est parce qu’elle est non-nulle sur un intervalle de largeur nulle.

Pour ne pas être d’emblée dans un cas aussi singulier, on peut envisager dans un premier temps une fonction, une vraie cette fois, notée $\delta_\varepsilon(\nu' - \nu)$, dont le graphe est une courbe régulière mais très pointue, de largeur ε et dont l’aire vaut 1. Ce graphe a donc l’allure indiquée sur la figure 1.1 ; la valeur de la fonction au maximum est d’ordre $1/\varepsilon$ puisque la largeur à mi-hauteur est d’ordre ε et que la surface doit être égale à 1. ε paramétrise bien l’acuité de la fonction δ_ε .

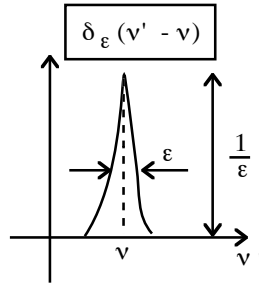
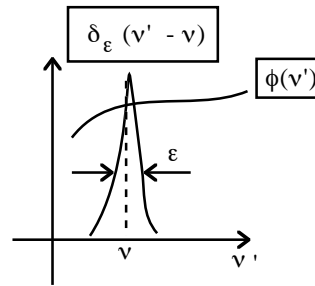


Figure 1.1: Précurseur de la “fonction” de Dirac.

Associons maintenant la fonction δ_ε avec une “bonne” fonction $\phi(\nu)$ dans une intégrale dont l’intervalle d’intégration non seulement contient le point d’abscisse ν mais encore englobe largement l’intervalle fini de largeur d’ordre ε où δ_ε n’est pas nulle. Soit donc l’intégrale I :

$$I = \int_a^b d\nu' \phi(\nu') \delta_\varepsilon(\nu' - \nu) . \quad (1.120)$$

Figure 1.2: Association du précurseur de la “fonction” de Dirac avec une bonne fonction ϕ .

Si la fonction $\phi(\nu')$ est régulière au voisinage de ν et si en particulier elle varie lentement²⁸ à l’échelle ε , il est évident géométriquement que :

$$I \simeq \phi(\nu) \int_a^b d\nu' \delta_\varepsilon(\nu' - \nu) = \phi(\nu) , \quad (1.121)$$

la dernière égalité venant du fait que δ_ε est de surface unité. Ainsi, pour toutes les fonctions bien régulières, aussi petit que soit ε fini, on a d’après (1.120) et (1.121) :

$$\int_a^b d\nu' \phi(\nu') \delta_\varepsilon(\nu' - \nu) \simeq \phi(\nu) , \quad (1.122)$$

et on a donc, en tant que limite :

$$\lim_{\varepsilon \rightarrow 0} \left[\int_a^b d\nu' \phi(\nu') \delta_\varepsilon(\nu' - \nu) \right] = \phi(\nu) , \quad (1.123)$$

où l’égalité est maintenant stricte, c’est-à-dire rigoureuse. Tout ce qui précède est parfaitement licite ; là où les choses se gâtent et où l’habitude conduit à une certaine désinvolture (sans conséquence si on reste lucide), c’est lorsque l’on échange l’ordre des opérations : limite et intégration. On obtient alors strictement la relation :

$$\int_a^b d\nu' \phi(\nu') \lim_{\varepsilon \rightarrow 0} [\delta_\varepsilon(\nu' - \nu)] = \phi(\nu) . \quad (1.124)$$

²⁸Pour une fonction différentiable ϕ donnée, la condition est toujours réalisable en prenant ε suffisamment petit.

qui est identique à (1.116) ci-dessus ; la fonction limite est bien la fonction nulle partout sauf en un point où elle est infinie.

Ce processus de limite a été présenté de façon un peu vague, mais il est facile de l'examiner en détail : il suffit de faire un choix explicite de fonctions reproduisant l'allure de δ_ε . Quelques choix courants sont les suivants :

$$\delta_\varepsilon(\nu' - \nu) = \frac{1}{\sqrt{2\pi\varepsilon}} e^{-[(\nu - \nu')^2]/(2\varepsilon^2)} \quad (\text{gaussienne}) , \quad (1.125)$$

$$\delta_\varepsilon(\nu' - \nu) = \frac{1}{\pi} \frac{\varepsilon}{(\nu - \nu')^2 + \varepsilon^2} \quad (\text{lorentzienne}) , \quad (1.126)$$

$$\delta_\varepsilon(\nu' - \nu) = \begin{cases} 0 & \text{si } |\nu' - \nu| > \frac{\varepsilon}{2} \\ 1/\varepsilon & \text{si } |\nu' - \nu| < \frac{\varepsilon}{2} \end{cases} \quad (\text{créneau}) . \quad (1.127)$$

Toutes ces (vraies) fonctions ont une largeur d'ordre ε et un maximum d'ordre ε^{-1} . Elles constituent quelques exemples des fonctions qui, à la limite, engendrent la fonction de Dirac. Le caractère de fonction se perd par le passage à la limite, processus assez usuel²⁹. Avec elles, on peut en toute tranquillité d'esprit, vérifier une à une toutes les relations établies plus haut formellement ; elles permettent aussi de se convaincre que la fonction de Dirac est une fonction paire :

$$\delta_\varepsilon(\nu' - \nu) = \delta_\varepsilon(\nu - \nu') , \quad (1.128)$$

et de se douter que, raisonnant à nouveau *avant* le passage à la limite, on peut introduire les dérivées des fonctions δ_ε ; elles permettront à leur tour de définir les “dérivées” de la fonction de Dirac. La règle, quand on manipule des objets un peu dangereux, est de se replacer au besoin dans un cadre sûr et d'effectuer au bon moment un passage à la limite ; en général, quand tout se passe bien, cette petite complication intermédiaire n'est pas nécessaire – mais il faut se souvenir que l'on peut toujours pouvoir l'appeler à la rescousse.

Ecrivons maintenant quelques relations typiques de l'espace vectoriel “continu” ainsi défini. Sur une base $|e_\nu\rangle \equiv |\nu\rangle$, un vecteur quelconque $|\psi\rangle$ se décompose donc comme suit :

$$|\psi\rangle = \int d\nu \psi(\nu) |\nu\rangle . \quad (1.129)$$

Un changement de base passant de $|e_\nu\rangle \equiv |\nu\rangle$ à $|f_\mu\rangle \equiv |\mu\rangle$, sur laquelle les composantes de $|\psi\rangle$ sont les fonctions notées $\phi(\mu)$, s'exprime comme :

$$\phi(\mu) = \int d\nu T_{\mu\nu} \psi(\nu) . \quad (1.130)$$

Il est d'usage de noter autrement la “matrice continue” $T_{\mu\nu}$, et de remonter les indices sous la forme d'arguments d'une fonction de deux variables, tout comme on l'a déjà fait pour les composantes ; cette matrice continue est d'ailleurs plutôt appelée noyau³⁰. On notera ainsi³¹ :

$$\phi(\mu) = \int d\nu T(\mu, \nu) \psi(\nu) . \quad (1.131)$$

Il existe également une relation de fermeture pour les bases continues. Partant de :

$$|\psi\rangle = \int d\nu' \psi(\nu') |\nu'\rangle , \quad (1.132)$$

on peut écrire :

$$\langle \nu | \psi \rangle = \int d\nu' \psi(\nu') \langle \nu | \nu' \rangle = \int d\nu' \psi(\nu') \delta(\nu - \nu') = \psi(\nu) . \quad (1.133)$$

²⁹Une somme de Fourier est une fonction continue partout ; si la somme devient série, on peut obtenir des sauts d'amplitude finie ou des points anguleux, bien que chaque composante soit continue et à dérivée continue. C'est aussi pour cela qu'un ampli hi-fi ne restituera jamais parfaitement un transitoire (un coup d'archet au violon par exemple) : il faudrait que sa bande passante soit infinie. Dans un tout autre domaine, c'est lorsque l'on passe explicitement à la limite thermodynamique qu'une fonction de partition peut (ce n'est pas systématique !) présenter des singularités associées à un point critique (transition de phase).

³⁰à ne pas confondre avec le noyau d'un opérateur, ensemble des vecteurs dont l'image par l'opérateur est le vecteur nul.

³¹Il y a ici un problème de notations. Comme on va le voir, $\psi(\nu)$ va s'identifier à $\langle \nu | \psi \rangle$, et donc on devrait noter les nouvelles composantes $\langle \mu | \psi \rangle$ tout autant ; mais ceci introduirait une confusion : les deux fonctions $\phi(\mu)$ et $\psi(\nu)$ sont deux fonctions différentes.

En injectant cette expression de $\psi(\nu)$ dans (1.132), il vient :

$$|\psi\rangle = \int d\nu \langle \nu | \psi \rangle |\nu\rangle = \int d\nu |\nu\rangle \langle \nu | \psi \rangle = \left[\int d\nu |\nu\rangle \langle \nu | \right] |\psi\rangle . \quad (1.134)$$

Vraie quel que soit $|\psi\rangle$, cette équation fournit la relation de fermeture sur une base continue orthonormée :

$$\int d\nu |\nu\rangle \langle \nu | = \mathbf{1} , \quad \langle \nu | \nu' \rangle = \delta(\nu - \nu') . \quad (1.135)$$

Compte tenu de ce qui précède, le produit scalaire de deux vecteurs, sur une base orthonormée au sens ci-dessus, s'écrit :

$$(|\psi\rangle, |\phi\rangle) \equiv \langle \psi | \phi \rangle = \int d\nu \psi^*(\nu) \phi(\nu) \quad \text{si } \langle \nu | \nu' \rangle = \delta(\nu - \nu') , \quad (1.136)$$

et le carré de la norme d'un vecteur, $\| |\psi\rangle \|^2$, est :

$$\| |\psi\rangle \|^2 = (|\psi\rangle, |\psi\rangle) \equiv \langle \psi | \psi \rangle = \int d\nu \psi^*(\nu) \psi(\nu) \equiv \int d\nu |\psi(\nu)|^2 . \quad (1.137)$$

Cette dernière intégrale doit rappeler quelque chose : si la variable ν est la position x d'une particule à une dimension, (1.137) est ce que l'on écrit pour prendre en compte le fait que la fonction d'onde de Schrödinger est, en module carré, normalisée à l'unité. En prenant comme base ces états notés naturellement $|x\rangle$, on voit que la fonction d'onde de Schrödinger peut être interprétée comme l'ensemble des composantes d'un vecteur $|\psi\rangle$:

$$|\psi\rangle = \int dx \psi(x) |x\rangle , \quad \psi(x) = \langle x | \psi \rangle , \quad \langle \psi | \phi \rangle = \int dx \psi^*(x) \phi(x) , \quad (1.138)$$

et c'est pourquoi la formulation initiale de Schrödinger s'appelle la "représentation- x " (ou représentation- q). Pour cette représentation, on a donc les relations de fermeture :

$$\int dx |x\rangle \langle x| = \mathbf{1} \quad (1.139)$$

et d'orthogonalité :

$$\langle x | x' \rangle = \delta(x - x') . \quad (1.140)$$

Au bout du compte, on retrouve la condition de normalisation sous sa forme familière :

$$\| |\psi\rangle \|^2 = (|\psi\rangle, |\psi\rangle) \equiv \langle \psi | \psi \rangle = \int dx \psi^*(x) \psi(x) \equiv \int dx |\psi(x)|^2 = 1 . \quad (1.141)$$

On montrera par la suite que le vecteur $|x_0\rangle$ est propre (et associé à la valeur propre x_0) de l'opérateur x (multiplication par x), qui est l'observable représentant la coordonnée en représentation- x .

Les relations de fermeture peuvent maintenant se combiner les unes avec les autres. Soit $\{a_n(x)\}$ un ensemble de fonctions formant une base complète orthonormalisée ; en notation de Dirac, on a donc :

$$\sum |a_n\rangle \langle a_n| = \mathbf{1} , \quad (1.142)$$

relation qui est une conséquence de :

$$\int dx a_n^*(x) a_m(x) \equiv \int dx \langle a_n | x \rangle \langle x | a_m \rangle = \delta_{nm} . \quad (1.143)$$

En multipliant (1.142) à gauche par $\langle x|$ et à droite par $|x'\rangle$, on fabrique la représentation- x de l'identité, qui est la "matrice" diagonale $\delta(x - x')$:

$$\sum_n \langle x | a_n \rangle \langle a_n | x' \rangle = \delta(x - x') , \quad (1.144)$$

qui peut aussi s'écrire en terme des fonctions d'onde $a_n(x)$:

$$\sum_n a_n^*(x) a_n(x') = \delta(x - x') . \quad (1.145)$$

Notons que ceci peut aussi se trouver en injectant $\mathbf{1} = \sum_n |a_n\rangle\langle a_n|$ au "milieu" du produit scalaire $\langle x|x'\rangle = \delta(x - x')$:

$$\langle x|x'\rangle \equiv \langle x|\mathbf{1}|x'\rangle = \delta(x - x') \iff \sum_n \langle x|a_n\rangle\langle a_n|x'\rangle = \delta(x - x') . \quad (1.146)$$

Cette dernière forme de la relation de fermeture prendra un sens particulièrement clair lors de l'étude de l'opérateur d'évolution $U(t)$; en effet, anticipant légèrement, celui s'écrit :

$$U(t) = \sum_n |\psi_n\rangle e^{-i\omega_n t} \langle \psi_n| , \quad (1.147)$$

où les $|\psi_n\rangle$ sont les états propres du Hamiltonien d'énergies $E_n = \hbar\omega_n$. Il en résulte que l'état à l'instant t , $|\Psi(t)\rangle$, s'écrit :

$$|\Psi(t)\rangle = U(t)|\Psi(t=0)\rangle = \sum_n |\psi_n\rangle e^{-i\omega_n t} \langle \psi_n|\Psi(t=0)\rangle . \quad (1.148)$$

Pour passer au langage des fonctions d'onde, il suffit d'injecter la relation de fermeture continue (1.139) au bon endroit, puis de multiplier à gauche par $\langle x|$:

$$\langle x|\Psi(t)\rangle = \sum_n \langle x|\psi_n\rangle e^{-i\omega_n t} \langle \psi_n|\int dx' |x'\rangle\langle x'|\Psi(t=0)\rangle . \quad (1.149)$$

Exprimé en terme de fonctions d'onde, ceci s'écrit :

$$\Psi(x, t) = \int dx' \sum_n \psi_n(x) e^{-i\omega_n t} \psi_n^*(x') \Psi(x', 0) \equiv \int dx' U(x, t; x', 0) \Psi(x', 0) , \quad (1.150)$$

et on voit que le propagateur a pour expression :

$$U(x, t; x', 0) = \sum_n \psi_n(x) e^{-i\omega_n t} \psi_n^*(x') . \quad (1.151)$$

Faisons maintenant $t = 0$ dans (1.151) ; on obtient, d'après la relation de fermeture (1.144) :

$$U(x, t = 0; x', 0) = \sum_n \psi_n(x) \psi_n^*(x') = \delta(x - x') . \quad (1.152)$$

Ceci exprime simplement le fait que la fonction d'onde n'évolue pas entre $t = 0$ et ... $t = 0$; dit autrement, $U(t)$ se réduit à l'opérateur identité en $t = 0$, comme il se doit.

1.3 Opérateurs hermitiques, opérateurs unitaires

Les opérateurs agissant dans l'espace vectoriel \mathcal{E} jouent un rôle fondamental en Mécanique Quantique. Selon l'algèbre linéaire élémentaire, un opérateur Ω est linéaire si et seulement si :

$$\Omega(\lambda|\psi\rangle + \mu|\phi\rangle) = \lambda\Omega|\psi\rangle + \mu\Omega|\phi\rangle , \quad (1.153)$$

où λ et μ sont des scalaires du corps. Ici, ce corps est celui des complexes et ceci donne un degré de liberté supplémentaire, permettant également de définir des opérateurs *antilinéaires*. Un opérateur K est antilinéaire si :

$$K(\lambda|\psi\rangle + \mu|\phi\rangle) = \lambda^* K|\psi\rangle + \mu^* K|\phi\rangle . \quad (1.154)$$

Il se révélera parfois nécessaire de considérer de tels opérateurs K (par exemple : pour la symétrie dans le renversement du temps), mais ceci constitue une situation plutôt exceptionnelle. Le cas usuel est celui où les opérateurs requis sont linéaires et c'est pourquoi on s'attarde sur eux dans ce qui suit. À la fin de cette section, seront données, pour les opérateurs antilinéaires, les quelques relations utiles pour la suite.

Donnons d'abord la définition de l'adjoint d'un opérateur (linéaire). Soit Ω un opérateur linéaire, mais quelconque par ailleurs ; l'adjoint de Ω , noté Ω^\dagger , est l'opérateur tel que :

$$(\Omega^\dagger|\psi\rangle, |\phi\rangle) = (|\psi\rangle, \Omega|\phi\rangle) . \quad (1.155)$$

En général, $\Omega^\dagger \neq \Omega$. Il existe cependant des opérateurs égaux à leur adjoint ; pour des raisons évidentes, ces opérateurs remarquables sont dits auto-adjoints, ou encore hermitiques, et on verra qu'ils jouent un rôle central en Mécanique Quantique. Pour un opérateur hermitique A , on a donc :

$$A^\dagger = A \iff (A|\psi\rangle, |\phi\rangle) = (|\psi\rangle, A|\phi\rangle) \quad \forall |\psi\rangle, |\phi\rangle . \quad (1.156)$$

Il faut se souvenir que la notion d'opérateur adjoint est totalement indépendante de toute référence à une base, quelle qu'elle soit. Par ailleurs, l'adjoint $(\Omega^\dagger)^\dagger$ de Ω^\dagger est l'opérateur Ω lui-même³². En effet :

$$((\Omega^\dagger)^\dagger|\phi\rangle, |\psi\rangle) = (|\phi\rangle, \Omega^\dagger|\psi\rangle) . \quad (1.157)$$

En prenant la complexe conjuguée de cette équation, ce qui revient à inverser les partenaires des deux produits scalaires :

$$(|\psi\rangle, (\Omega^\dagger)^\dagger|\phi\rangle) = (\Omega^\dagger|\psi\rangle, |\phi\rangle) . \quad (1.158)$$

Mais, par la définition de l'adjoint, le second membre vaut $(|\psi\rangle, \Omega|\phi\rangle)$; il vient ainsi :

$$(|\psi\rangle, (\Omega^\dagger)^\dagger|\phi\rangle) = (|\psi\rangle, \Omega|\phi\rangle) \quad (1.159)$$

d'où l'on déduit la relation suivante, valable au niveau des opérateurs puisque les vecteurs $|\psi\rangle$ et $|\phi\rangle$ sont absolument quelconques :

$$(\Omega^\dagger)^\dagger = \Omega . \quad (1.160)$$

Il est facile de voir que l'adjoint d'une combinaison linéaire d'opérateurs satisfait :

$$(\lambda_1\Omega_1 + \lambda_2\Omega_2)^\dagger = \lambda_1^*\Omega_1^\dagger + \lambda_2^*\Omega_2^\dagger . \quad (1.161)$$

Établissons maintenant la propriété essentielle suivante : les valeurs propres d'un opérateur hermitique A sont réelles et ses vecteurs propres sont orthogonaux. Soit $\{|a_n\rangle\}$ les vecteurs propres ; on a :

$$A|a_n\rangle = a_n|a_n\rangle . \quad (1.162)$$

A est supposé hermitique ; la relation de définition (1.155) avec $|\psi\rangle = |\phi\rangle = |a_n\rangle$, donne :

$$(A|a_n\rangle, |a_n\rangle) = (|a_n\rangle, A|a_n\rangle) . \quad (1.163)$$

A agit sur ses vecteurs propres : le résultat est donc la simple multiplication par a_n , conformément à (1.162) ; par la sesquilinearité (antilinéarité) du produit scalaire, (1.163) devient :

$$a_n^* (|a_n\rangle, |a_n\rangle) = a_n (|a_n\rangle, |a_n\rangle) . \quad (1.164)$$

Pour un vecteur propre non égal au vecteur nul, ceci donne bien a_n réel. Le spectre d'un opérateur hermitique est donc inclus dans \mathbb{R} .

³²L'espace de Hilbert est un espace à métrique hermitique où, par définition, il existe des bases orthonormées dont chacun des éléments a une norme positive et égale à 1 (il n'en va pas toujours ainsi : on connaît des espaces, appelés symplectiques, dont les éléments sont isotropes car leur carré est nul [2]). Dans un espace à métrique positive, l'adjoint de l'adjoint est de fait l'opérateur lui-même.

Montrons maintenant que deux vecteurs propres associés à deux valeurs propres *distinctes* sont forcément orthogonaux. En effet, partant de (1.155) appliqué à A avec deux vecteurs propres distincts³³ et tenant compte de la réalité des valeurs propres, on obtient :

$$a_n (|a_n\rangle, |a_m\rangle) = a_m (|a_n\rangle, |a_m\rangle) \iff (a_n - a_m) (|a_n\rangle, |a_m\rangle) = 0 . \quad (1.165)$$

Comme a_n est supposé différent de a_m , le produit scalaire est nul.

Les opérateurs hermitiques constituent une généralisation des opérateurs symétriques, comme on le verra par la suite. On se souviendra de l'implication (pour des vecteurs propres associés à deux valeurs propres distinctes) :

$$\text{opérateur hermitique} \implies \text{vecteurs propres orthogonaux} , \quad (1.166)$$

dont la réciproque est fautive en général³⁴. En revanche, si un opérateur a des vecteurs propres tous orthogonaux et des valeurs propres toutes réelles, alors il est hermitique.

Donnons maintenant la relation entre les éléments de matrice d'un opérateur Ω et de son adjoint Ω^\dagger . Par définition, on a :

$$(\Omega^\dagger |e_n\rangle, |e_m\rangle) = (|e_n\rangle, \Omega |e_m\rangle) = \Omega_{nm} . \quad (1.167)$$

Par l'antilinearité du produit scalaire, on a aussi :

$$(\Omega^\dagger |e_n\rangle, |e_m\rangle) = (|e_m\rangle, \Omega^\dagger |e_n\rangle)^* = (\Omega_{mn}^\dagger)^* . \quad (1.168)$$

Par comparaison, il vient :

$$(\Omega_{mn}^\dagger)^* = \Omega_{nm} \iff \Omega_{mn}^\dagger = \Omega_{nm}^* . \quad (1.169)$$

Ainsi, pour avoir la matrice de l'adjoint Ω^\dagger d'un opérateur Ω , sur une base orthonormée, il suffit de transposer et de prendre la complexe conjuguée de la matrice de cet opérateur.

Venons-en maintenant aux opérateurs unitaires, qui laissent invariant tout produit scalaire. Un opérateur unitaire possède nécessairement un inverse, puisqu'il conserve la norme d'un vecteur : il ne saurait transformer un vecteur de norme finie en un vecteur de norme nulle. L'adjoint d'un opérateur unitaire est égal à son inverse ; en effet, par l'unitarité :

$$(U|\psi\rangle, U|\phi\rangle) = (|\psi\rangle, |\phi\rangle) . \quad (1.170)$$

En utilisant à l'envers la définition de l'adjoint de U , U^\dagger , le premier membre s'écrit :

$$(U^\dagger U|\psi\rangle, |\phi\rangle) = (|\psi\rangle, |\phi\rangle) \quad \forall |\psi\rangle, |\phi\rangle . \quad (1.171)$$

de sorte que :

$$U^\dagger U = \mathbf{1} \iff U^\dagger = U^{-1} . \quad (1.172)$$

Il en résulte immédiatement que, sur une base orthonormée :

$$U_{np}^{-1} = U_{np}^\dagger = U_{pn}^* . \quad (1.173)$$

Des exemples d'opérateurs unitaires seront maintes fois rencontrés dans la suite (opérateur d'évolution $U(t)$, opérateurs associés aux translations ou rotations, *etc.*).

Remarques

1. La notation de Dirac permet d'écrire de façon quasi-automatique toutes les relations traduisant le jeu de l'algèbre vectorielle ; elle se révèle à l'usage d'une commodité extrême. Cette utilité est manifeste sur quelques exemples. Revenons à (1.80) et à (1.86) ; on peut écrire :

$$|\psi\rangle = \mathbf{1} |\psi\rangle = \sum_n |e_n\rangle \langle e_n | \psi \rangle . \quad (1.174)$$

³³Pour simplifier, on suppose qu'il n'y a pas de dégénérescence ; la généralisation est immédiate.

³⁴Par exemple, l'opérateur $\Omega = e^{iA}$ est unitaire quand A est hermitique ; en tant que fonction de A , il a les mêmes vecteurs propres que A et ceux-ci sont mutuellement orthogonaux puisque A est hermitique. En revanche, les valeurs propres de Ω ne sont pas en général réelles (ce sont des complexes de module 1). L'orthogonalité des vecteurs propres n'assure donc pas l'hermiticité.

La dernière égalité redonne (1.82) dans le cas d'une base orthonormée ($g_{nm} = \delta_{nm}$) :

$$x_n = \langle e_n | \psi \rangle . \quad (1.175)$$

Soit Ω un opérateur quelconque ; par (1.86) :

$$\Omega = \mathbf{1} \Omega \mathbf{1} = \sum_n |e_n\rangle \langle e_n| \Omega \sum_m |e_m\rangle \langle e_m| = \sum_n \sum_m \Omega_{nm} |e_n\rangle \langle e_m| . \quad (1.176)$$

En particulier, si on prend comme base la base propre de Ω , $\{|\omega_n\rangle\}$, (1.176) donne :

$$\Omega = \sum_n \omega_n |\omega_n\rangle \langle \omega_n| . \quad (1.177)$$

On donne le nom de *décomposition spectrale* au développement (1.177), puisqu'il fait apparaître³⁵ les valeurs propres de Ω . Si f est une certaine fonction, l'opérateur $f(\omega)$ a pour valeurs propres les valeurs de f pour les valeurs propres de Ω ; la décomposition spectrale de $f(\Omega)$ est ainsi :

$$f(\Omega) = \sum_n f(\omega_n) |\omega_n\rangle \langle \omega_n| . \quad (1.178)$$

Par exemple, si H a pour valeurs propres les énergies $E_n = \hbar\omega_n$ et les vecteurs propres $|\psi_n\rangle$, l'opérateur $e^{\frac{1}{\hbar}Ht}$ admet la décomposition :

$$e^{\frac{1}{\hbar}Ht} = \sum_n e^{-i\omega_n t} |\psi_n\rangle \langle \psi_n| . \quad (1.179)$$

La notation de Dirac permet aussi de multiplier les opérateurs et de retrouver au passage la règle usuelle de multiplication des matrices ; en injectant trois fois la relation de fermeture (1.102) aux bons endroits, il vient :

$$AB = \sum_{n,m,p} |e_n\rangle \langle e_n| A |e_m\rangle \langle e_m| B |e_p\rangle \langle e_p| = \sum_{n,p} |e_n\rangle \left[\sum_m A_{nm} B_{mp} \right] \langle e_p| \equiv \sum_n \sum_p |e_n\rangle (AB)_{np} \langle e_p| . \quad (1.180)$$

Sur cette dernière équation, on reconnaît la règle ordinaire donnant l'élément de matrice du produit de deux matrices.

2. Il convient de garder à l'esprit que pour deux vecteurs $|\psi\rangle$ et $|\phi\rangle$ quelconques (en particulier non orthogonaux), la quantité $\langle \psi | \Omega | \phi \rangle$ n'est pas un élément de la matrice représentant l'opérateur Ω . En effet, soit $\{|f_p\rangle\}$ une base *quelconque*, non orthonormée. Par définition de la matrice d'un opérateur, on a :

$$\Omega |f_p\rangle = \sum_n \Omega_{np} |f_n\rangle . \quad (1.181)$$

En multipliant scalairement par $|f_m\rangle$, on en déduit :

$$\langle f_m | \Omega | f_p \rangle = \sum_n \Omega_{np} \langle f_m | f_n \rangle \equiv \sum_n \Omega_{np} g_{mn} . \quad (1.182)$$

Trouver Ω_{np} passe par l'inversion du système linéaire (1.182), qui donne l'expression de l'élément de matrice en combinaison linéaire d'*a priori* tous les produits scalaires $\langle f_m | \Omega | f_p \rangle$; c'est seulement si $\langle f_m | f_n \rangle = \delta_{mn}$ que l'on retrouve la relation simple $\langle f_m | \Omega | f_n \rangle = \Omega_{mn}$. Ces distinctions disparaissent par usage de bases orthonormées, en principe toujours disponibles. En cas de nécessité, en présence d'une base non orthonormée, le plus sûr est de construire à partir de celle-ci une base orthonormée, par divers procédés dont le plus connu est celui dit de "l'orthonormalisation de Schmidt".

³⁵On appelle *spectre* d'un opérateur l'ensemble de ses valeurs propres.

Disons maintenant quelques mots des opérateurs antilinéaires³⁶, requis en Mécanique Quantique pour analyser l'invariance par renversement du temps. La définition d'un tel opérateur K est donnée en (1.154).

Attention ! De cette définition, il résulte immédiatement qu'un scalaire $\lambda \in \mathbb{C}$ ne commute pas avec un opérateur antilinéaire³⁷ :

$$K \lambda |\phi\rangle = \lambda^* K |\phi\rangle \iff [K, \lambda] = (\lambda^* - \lambda)K \iff [K, \lambda \mathbf{1}] \neq 0 . \quad (1.183)$$

L'adjoint d'un opérateur antilinéaire se définit comme suit :

$$(K^\dagger |\psi\rangle, |\phi\rangle) = (|\psi\rangle, K|\phi\rangle)^* ; \quad (1.184)$$

on notera, par rapport à la définition (1.155), la conjugaison complexe au second membre. La notation de Dirac doit être manipulée prudemment, quand il s'agit d'opérateurs antilinéaires. En effet, en ajoutant les parenthèses ici nécessaires :

$$(K^\dagger |\psi\rangle, |\phi\rangle) \equiv \langle K^\dagger \psi | \phi \rangle = \langle \psi | K | \phi \rangle . \quad (1.185)$$

D'un autre côté, suivant (1.184), le premier membre est égal à $[\langle \psi | (K|\phi) \rangle]^*$; d'où :

$$\langle \psi | K | \phi \rangle = \langle \psi | (K|\phi) \rangle^* , \quad (1.186)$$

où la nécessité des parenthèses est évidente³⁸. Pour comparaison, écrivons la même suite d'égalités pour un opérateur linéaire ; la seule différence est l'absence de conjugaison au niveau de la définition de l'adjoint et on a successivement :

$$(\Omega^\dagger |\psi\rangle, |\phi\rangle) \equiv \langle \Omega^\dagger \psi | \phi \rangle = \langle \psi | \Omega | \phi \rangle . \quad (1.187)$$

Suivant (1.155), le premier membre est égal à $\langle \psi | (\Omega|\phi) \rangle$; d'où :

$$\langle \psi | \Omega | \phi \rangle = \langle \psi | (\Omega|\phi) \rangle , \quad (1.188)$$

montrant que, pour un opérateur linéaire, les parenthèses ne servent à rien et peuvent donc être omises.

Pour terminer, donnons une dernière définition : un opérateur antiunitaire est un opérateur antilinéaire dont l'adjoint est égal à l'inverse :

$$U \text{ est antiunitaire} \iff U \text{ est antilinéaire et } U^\dagger = U^{-1} . \quad (1.189)$$

On rencontrera un opérateur antiunitaire lors de l'étude de l'invariance par renversement du temps. Le caractère "anti" vient du facteur imaginaire pur au premier membre de (1.1) : si on conjugue cette équation et que t est changé en $-t$, l'opérateur $i\hbar(\partial/\partial t)$ est invariant. La transformée de la fonction d'onde par renversement du temps est donc la fonction complexe conjuguée³⁹ ; l'opérateur renversement du temps est donc forcément antilinéaire – étant évident par ailleurs qu'il doit bien sûr conserver la norme de la fonction d'onde.

1.4 Postulats de la Mécanique Quantique

1.4.1 Notion d'état

Les arguments développés par Heisenberg, Schrödinger et de Broglie démontrent que la description du monde des particules doit être abordée de façon radicalement nouvelle et constitue une véritable révolution par rapport aux

³⁶Voir par exemple [20].

³⁷Il est tout à fait possible de définir la matrice M d'un opérateur antilinéaire sur une base donnée, selon la prescription habituelle : on porte successivement en colonne les transformés des vecteurs de la base. Toutefois, le vecteur transformé par l'opérateur d'un vecteur de composantes $\{c_n\}$ ne s'obtient pas simplement en multipliant la matrice M par la matrice colonne des c_n , car il faut en plus, conformément à (1.183), conjuguer ces derniers.

³⁸À tout prendre, la notation du produit scalaire sous forme d'un couple ordonné, $(|\psi\rangle, |\phi\rangle)$, n'est pas plus compliquée à écrire que la notation de Dirac. En cas de doute, mieux vaut y revenir.

³⁹Indépendamment de toute complication liée au spin.

conceptions classiques. Il y a quelque analogie entre la nécessité de renoncer à une vision purement mécaniste des particules et celle, maintenant idiosyncrasique, d'expliquer les propriétés d'un gaz parfait classique par des théories de nature statistique⁴⁰. Dans un cas comme dans l'autre, c'est la Nature elle-même qui impose sa loi, et ceci ne doit pas être ressenti comme une frustration : la raison d'être de la Physique est précisément de proposer une description du monde tel qu'il se révèle par l'expérience, autorisant la prévision des phénomènes à défaut d'en donner une explication dans l'absolu⁴¹.

La fonction d'onde de Schrödinger, ou ses généralisations⁴², constitue⁴³ notre actuelle connaissance ultime d'un système relevant de la Mécanique Quantique. Cette fonction d'onde a, fondamentalement, une interprétation statistique et ne prétend donc pas décrire l'état d'un système en particulier au sens classique. La description de l'état d'un système quantique consiste en l'énoncé des lois de probabilités permettant d'en décrire les propriétés et de fournir une explication – au sens usuel⁴⁴ – du monde à l'échelle atomique ou subatomique⁴⁵. Cet objectif, qui n'est pas réducteur, doit être bien compris, pour écarter le risque de malentendus ou de paradoxes. L'impossibilité d'observer avec les sens usuels la "réalité" physique à petite échelle ne doit pas être assimilée à la négation de la réalité : tout débat en la matière est du domaine de la Métaphysique, par définition. On ne doit pas davantage être tenté de croire que la Mécanique Quantique est une théorie incomplète⁴⁶ ou provisoire, pour deux raisons au moins.

La première des raisons est que, de toute façon, toute théorie physique constitue seulement, à un instant donné, une vérité relative toujours susceptible d'être bousculée par la suite. Affirmer que la Mécanique Quantique sera un jour remise en cause, même si à l'heure actuelle elle est à l'abri de toute contestation au vu des faits expérimentaux, est donc une remarque sans grand intérêt. La deuxième raison est que l'on ne saurait déclarer incomplète une théorie qui, à un instant donné, rend précisément compte de l'ensemble des expériences : il n'existe pas aujourd'hui d'expérience mettant en cause la Mécanique Quantique. L'incomplétude d'une théorie ne se mesure pas par référence aux présupposés mentaux applicables dans un autre cadre de pensée, même si l'habitude les avaient presque érigés en principes. Après tout, personne n'envisage de jeter aux orties la Relativité d'Einstein, sous prétexte qu'elle est en contradiction avec le Principe de Relativité de Galilée et bien qu'elle conduise à des "paradoxes" demeurés célèbres. Personne non plus ne songe à décrire un gaz, même parfait, autrement que par une approche statistique qui, pourtant, tire un trait sur la description détaillée du destin individuel d'une particule du gaz. La théorie statistique qui en découle permet d'expliquer et de comprendre tout ce que l'on sait, et que l'on peut vérifier expérimentalement, des systèmes composés d'un très grand nombre de particules (quel que soit leur "état", solide, liquide ou gazeux) et de ce fait constitue notre connaissance complète de ce que sont ces systèmes⁴⁷. Il en va de même, aujourd'hui, pour la Mécanique Quantique qui, en

⁴⁰Les déboires de Boltzmann – qui l'ont conduit au suicide – rappellent que cette révolution statistique, durant le dernier quart du XIX^{ème} siècle, ne s'est pas faite en douceur. Le dogme de Laplace

"Donnez-moi tous les x_0 et tous les p_{x_0} et je vous prédirai l'avenir du monde"

en tant que principe méthodologique *universel*, a eu la vie longue. . .

⁴¹Feynman écrit [7], p. 25 :

"J'espère que vous accepterez la Nature telle qu'elle est : absurde."

La lecture de ce petit livre est fortement recommandée.

⁴²Comme l'opérateur densité, ρ .

⁴³À température nulle.

⁴⁴En Physique, expliquer c'est proposer une théorie cohérente et non-contradictoire qui, rendant compte des phénomènes observés, permet d'en prédire d'autres. Le pouvoir explicatif de la Physique s'arrête là : écrire la loi de la Gravitation n'est pas fournir une explication ultime du pourquoi de l'attraction universelle. C'est seulement (!?) une rationalisation simple et magnifique de l'ensemble des observations ; en ce qui concerne la Mécanique Quantique :

"... il faut bien voir que personne n'est capable d'expliquer pourquoi la Nature se comporte de cette façon-là et pas d'une autre."

selon Feynman ([7], p. 24).

⁴⁵Étant entendu que c'est la description convenable des phénomènes survenant à cette échelle qui permet de comprendre les propriétés *macroscopiques* des objets de la vie courante (la conductivité des métaux, la rigidité d'une boule de billard, . . .).

⁴⁶On pourrait objecter que certains esprits, parmi les plus grands – dont Einstein – n'ont jamais accepté vraiment la Mécanique Quantique et surtout son interprétation. L'objection doit être relativisée par le contexte historique. Ainsi par exemple, lors de la formulation du célèbre paradoxe EPR [11], la Mécanique Quantique, quoique pratiquement achevée, avait eu finalement peu de temps pour faire ses preuves. Compte tenu de la révolution "culturelle" qu'elle imposait, il est normal qu'elle ait provoqué plus d'une réticence, y compris (surtout ?) dans les cerveaux les plus brillants. Rappelons d'ailleurs que des expériences récentes [12], [13], [14] ont démontré que la Mécanique Quantique est vérifiée jusque dans ses prévisions les plus surprenantes – sur lesquelles se sont appuyés Einstein, Podolski et Rosen pour formuler leur "paradoxe" ; les expériences, à l'époque impossibles technologiquement parlant, ne risquaient pas de leur donner tort.

⁴⁷Toutefois, l'analogie ne doit pas être poussée trop loin, pour écarter le risque de contresens. En Théorie cinétique des gaz,

dépôt d'un statut épistémologique unique⁴⁸, peut prétendre au même caractère complet que toute autre théorie rendant compte de l'ensemble des expériences connues.

Le tout premier postulat de la Mécanique Quantique est donc l'affirmation qu'il existe, pour tout système, un objet mathématique appelé fonction d'onde et représentant toute la connaissance possible de l'état de ce système. Comme on l'a rappelé, cette connaissance est multiforme et peut être représentée de plusieurs façons équivalentes (représentation- q , représentation- p , etc.) conduisant finalement à l'idée de *vecteur*. On en vient ainsi à devoir énoncer le premier postulat :

Postulat 1 : la connaissance de l'état d'un système à un instant t donné est complètement contenue dans un vecteur appelé vecteur d'état, noté génériquement $|\Psi(t)\rangle$ conformément à la notation de Dirac. L'espace vectoriel auquel appartient ce vecteur est appelé espace d'états, \mathcal{E} , défini sur le corps des complexes \mathbb{C} .

Dans les représentations- q et - p , l'argument des fonctions d'onde est une variable continue, car rien ne conduit à admettre que la position et l'impulsion d'une particule, confinée ou non, prennent des valeurs discrètes. En revanche, tout ce que l'on sait déjà de l'énergie d'un atome et de l'oscillateur harmonique – pour ne citer que deux exemples simples – montre au contraire qu'il existe des grandeurs dont la mesure⁴⁹ conduit à un ensemble de valeurs discrètes bien définies, spécifiques du système considéré. En la circonstance, il s'agit de l'énergie d'un état lié – mais l'expérience de Stern et Gerlach montre qu'il en va de même pour le moment cinétique d'un atome, qu'il soit orbital ou de spin. Il faut ainsi admettre l'idée générale suivant laquelle il existe des grandeurs quantifiées : les valeurs observées ne peuvent qu'appartenir à un ensemble *discret*, fini ou infini, mais dénombrable. La description de l'état d'un système en terme d'une grandeur A quantifiée, prenant N (N fini ou infini) valeurs distinctes a_n , $n = 1, 2, \dots, N$, repose maintenant sur la connaissance de N nombres, avec lesquels on doit pouvoir – comme avec une fonction d'onde au sens premier – calculer à l'avance la probabilité P_n d'obtenir la valeur a_n à la suite d'une opération de mesure. L'ensemble des a_n , $\{a_n\}_n$, constitue le spectre de la grandeur A . Si l'on désigne par c_n ces N nombres, on voit que ceux-ci constituent finalement une “fonction d'onde discrète”, une fonction définie sur l'ensemble des N premiers entiers. Il conviendra, en temps utile, d'établir explicitement le lien entre le vecteur d'état $|\Psi\rangle$ et l'ensemble de ses représentations en fonction d'onde (discrète ou continue).

Une dernière remarque : le Postulat 1 affirme que la fonction d'onde sous une forme ou sous une autre constitue la connaissance complète de l'état du système. Ce postulat nie donc l'existence de variables cachées.

1.4.2 Notion d'observable

En Mécanique Ondulatoire de Schrödinger, les grandeurs dynamiques q et p sont représentées par des opérateurs, la correspondance étant donnée par (1.5). C'est la généralisation de cette affirmation qui est érigée en postulat :

Postulat 2 : toute grandeur physique est représentée par un opérateur hermitique agissant dans l'espace des états \mathcal{E} .

L'usage a consacré l'adjectif substantivé “observable” pour cet opérateur ; on dit ainsi que l'opérateur $p_x = -i\hbar(\partial/\partial x)$ est, en représentation- q , l'observable associée à la composante suivant Ox de l'impulsion. Il est certain que toute observable, en tant que représentant quantique d'une grandeur physique, doit posséder des propriétés spécifiques⁵⁰. Par définition, une observable est un opérateur hermitique dont l'ensemble des vecteurs propres constitue une base complète de l'espace des états (un opérateur linéaire *quelconque* n'a pas le statut d'observable).

chaque atome est censé posséder à tout instant une position et une vitesse : on renonce d'emblée à les calculer mais on admet qu'elles existent en soi. Pour une particule – un électron par exemple –, il n'en va pas ainsi.

⁴⁸ On sait que la formulation précise de la Mécanique Quantique nécessite une référence constante à sa théorie-limite, la Mécanique Classique dans sa formulation lagrangienne ou hamiltonienne (voir les remarques dans la référence [15], p. 9.)

⁴⁹ Il s'agit bien des valeurs que *révèle* une mesure de la grandeur physique correspondante. Rien n'est affirmé sur la préexistence de ces valeurs *antérieurement* à la mesure.

⁵⁰ assurant notamment que toutes les valeurs moyennes sont des quantités réelles.

Parvenu à ce stade de l'énoncé des postulats, on mesure *a posteriori* encore mieux la distance⁵¹ qui sépare la description mécanique classique de celle que donne la Mécanique Quantique : celle-ci décrit l'état d'un système par un vecteur et les grandeurs dynamiques par des opérateurs (linéaires) agissant dans l'espace vectoriel des états. Au contraire, l'état d'un système en Mécanique Classique est complètement déterminé par des fonctions horaires, résultant de l'intégration de l'équation fondamentale à partir d'un jeu donné de conditions initiales.

En tant qu'opérateurs, les observables obéissent par nature à une algèbre en général non-commutative ; les relations fondamentales telles que (1.3) ont une image classique : c'est le crochet de Poisson de deux fonctions f et g ⁵², défini comme :

$$\{f(\dots q_j \dots, \dots p_j, \dots), g(\dots q_j \dots, \dots p_j, \dots)\} = \sum_k \left(\frac{\partial f}{\partial q_k} \frac{\partial g}{\partial p_k} - \frac{\partial g}{\partial q_k} \frac{\partial f}{\partial p_k} \right) . \quad (1.190)$$

La correspondance entre crochet de Poisson et commutateur est :

$$\{f, g\} \longleftrightarrow \frac{1}{i\hbar} [f, g] . \quad (1.191)$$

Par ce biais, on peut poser⁵³, par exemple, les relations de commutation caractéristiques d'un moment cinétique \vec{J} , de composantes J_u , $u = x, y, z$:

$$[J_x, J_y] = i\hbar J_z , \quad [J_y, J_z] = i\hbar J_x , \quad [J_z, J_x] = i\hbar J_y . \quad (1.192)$$

En remarquant que pour trois opérateurs quelconques A , B et C on a l'identité :

$$[A, BC] = [A, B]C + B[A, C] , \quad (1.193)$$

on déduit facilement de (1.192) que :

$$[\vec{J}^2, J_u] = 0 \quad (u = x, y, z) . \quad (1.194)$$

L'ensemble des relations de commutation du moment cinétique forme la base de la théorie du moment cinétique en Mécanique Quantique, qui sera développée au chapitre 3.

Il est bien connu qu'étant donné un opérateur linéaire A agissant dans un espace vectoriel, il existe des vecteurs de cet espace jouant un rôle éminent : ce sont les vecteurs propres de A ; l'action de A sur ceux-ci se traduit par la simple multiplication par un scalaire du corps sur lequel est défini l'espace vectoriel. Ainsi, en utilisant la notation de Dirac, si Ω est un opérateur quelconque, le couple propre de Ω ($\omega, |\omega\rangle$) satisfait par définition :

$$\Omega |\omega\rangle = \omega |\omega\rangle \quad (\omega \in \mathbb{C}) . \quad (1.195)$$

Un vecteur propre est défini à un facteur multiplicatif près, puisque (1.195) est une équation homogène. Pour un opérateur quelconque, la valeur propre ω est *a priori* un nombre complexe ; pour les observables, ω est toujours réel, ce qui résulte de la propriété d'hermiticité de l'opérateur lui-même. On sait par ailleurs que l'ensemble des valeurs propres est l'ensemble des solutions de l'équation algébrique :

$$\text{Det}(\Omega - \omega \mathbf{1}) = 0 \quad (1.196)$$

appelée équation caractéristique. Cette équation prend toujours la forme :

$$P(\omega) = 0 \quad (1.197)$$

où P est soit un polynôme de degré N , P_N , si l'espace vectoriel est de dimension N , soit une série, si l'espace est de dimension infinie dénombrable. En effet, pour que l'équation aux vecteurs propres – qui est homogène – admette une (des) solution(s) non triviale(s), il faut bien que le déterminant du système linéaire associé soit

⁵¹La distance est sidérale entre les descriptions classique et quantique – d'où l'incongruité des propriétés dynamiques des particules si on s'obstine à les percevoir dans une optique purement mécaniste héritée de la pensée classique.

⁵²Attention : la définition précise du crochet de Poisson varie d'un auteur à l'autre, conduisant à une fluctuation de signe.

⁵³On peut aussi les obtenir directement à partir de $[q_u, p_v] = i\hbar \delta_{uv}$, et vérifier ainsi la cohérence de la théorie sur ce point.

nul. Pour un espace vectoriel ayant une dimension infinie non dénombrable, l'équation aux valeurs propres est de nature intégrale :

$$\mathcal{P}(\omega) \equiv \int d\omega p(\omega) = 0 , \quad (1.198)$$

où $\mathcal{P}(\omega)$ est une fonction jouant le rôle de $P(\omega)$. Il est utile de se souvenir qu'un opérateur donné Ω satisfait sa propre équation caractéristique (théorème de Cayley-Hamilton), c'est-à-dire que partant de (1.197), on a en tant qu'équation opératorielle :

$$P(\Omega) = 0 . \quad (1.199)$$

Ceci étant rappelé, donnons un théorème qui jouera un rôle important dans la suite. A et B étant deux opérateurs dont le commutateur $[A, B]$ n'est pas nul, l'ensemble des vecteurs propres de A (resp. B) est – en tant qu'ensemble – distinct de l'ensemble des vecteurs propres de B (resp. A)⁵⁴. En effet, soit $|a\rangle$ l'un quelconque des vecteurs propres de A associé à la valeur propre a :

$$A|a\rangle = a|a\rangle \quad \forall a \in \text{spectre de } A . \quad (1.200)$$

Par hypothèse, on a :

$$AB|a\rangle \neq BA|a\rangle = aB|a\rangle . \quad (1.201)$$

Raisonnons maintenant par l'absurde : supposons que le vecteur $|a\rangle$, qui est l'un quelconque des vecteurs propres de A , est aussi vecteur propre de B ; c'est dire qu'il existe un scalaire b tel que :

$$B|a\rangle = b|a\rangle . \quad (1.202)$$

Dans cette hypothèse, la non-égalité (1.201) devient :

$$Ab|a\rangle = bA|a\rangle = ab|a\rangle \neq BA|a\rangle = aB|a\rangle = ab|a\rangle \iff ab|a\rangle \neq ab|a\rangle . \quad (1.203)$$

La dernière égalité est visiblement absurde, a et b étant des scalaires. À la réflexion, l'absence d'un ensemble de vecteurs propres communs pour deux opérateurs A et B ne commutant pas constitue une évidence. En effet, si l'on choisit comme base de l'espace vectoriel les vecteurs propres de l'un d'entre eux, A , par exemple, alors A est représenté par une matrice diagonale $D(A)$. Si, sur cette même base, B était représenté par une matrice elle aussi diagonale $D(B)$, alors on aurait :

$$D(A)D(B) = D(B)D(A) . \quad (1.204)$$

car deux matrices diagonales commutent. Comme les relations algébriques entre opérateurs linéaires se transposent fidèlement au niveau de leurs représentants (leurs matrices), cette dernière relation signifierait que $[A, B] = 0$, contrairement à l'hypothèse.

Ainsi, quand deux opérateurs ne commutent pas, l'ensemble des vecteurs propres de l'un ne peut être propre de l'autre : deux opérateurs dont le commutateur n'est pas nul n'ont pas un ensemble complet de vecteurs propres en commun. Deux telles observables sont dites *incompatibles*⁵⁵. Des exemples de tels couples sont fournis par (x, p_x) , (J_x, J_y) , etc.

À l'inverse, si deux opérateurs commutent, tout vecteur propre de l'un est aussi vecteur propre de l'autre. Pour simplifier, prenons le cas où il n'existe pas de dégénérescence (à chaque valeur propre ne correspond qu'un seul vecteur propre). Soit $|a\rangle$ un vecteur propre de A et B un autre opérateur tel que $[A, B] = 0$. On a :

$$A|a\rangle = a|a\rangle \quad (AB - BA)|a\rangle = AB|a\rangle - aB|a\rangle = 0 . \quad (1.205)$$

En ajoutant des parenthèses pour la clarté, la dernière équation s'écrit :

$$A(B|a\rangle) = a(B|a\rangle) . \quad (1.206)$$

⁵⁴ Deux tels opérateurs peuvent avoir certains vecteurs propres communs, mais ceci ne peut être vrai pour *tous* les vecteurs propres.

⁵⁵ Par le postulat 5 (réduction du paquet d'onde), il en résultera que la mesure *simultanée* de deux observables qui ne commutent pas est impossible – d'où la terminologie grandeurs *incompatibles*.

Ceci montre que le vecteur $B|a\rangle$, résultat de l'action de B sur le vecteur propre $|a\rangle$ de A , est encore vecteur propre de A et est associé à la même valeur propre a . Comme on a supposé qu'il n'y avait pas de dégénérescence, il en résulte que ce vecteur $B|a\rangle$ est simplement proportionnel à $|a\rangle$; en d'autres termes, $|a\rangle$ est aussi vecteur propre de B ; notant b le facteur de proportionnalité (la valeur propre de B), on a donc :

$$B|a\rangle = b|a\rangle . \quad (1.207)$$

En conséquence, pour rappeler ce fait et traiter A et B sur un pied d'égalité, on peut noter $|a, b\rangle$ le vecteur propre commun à A et B . La généralisation au cas où le spectre est dégénéré ne présente aucune difficulté. Si le spectre A est dégénéré, tout vecteur propre de A est loin d'être unique : toute combinaison linéaire de deux (ou plusieurs) vecteurs propres associés à une même valeur propre est encore vecteur propre ; si donc un vecteur propre ("tiré au hasard") de A n'est pas forcément propre de B , il est toujours possible, en formant les bonnes combinaisons linéaires de vecteurs propres associés à une même valeur propre (qui restent propres de A quoi qu'il arrive), de trouver précisément celles qui sont aussi propres de B .

On pourrait multiplier les exemples d'observables. La plus importante d'entre elles est sans doute l'énergie, associée au Hamiltonien H qui apparaît au cœur de l'équation fondamentale de Schrödinger : H permet de construire $d|\Psi\rangle/dt$ à partir de $|\Psi(t)\rangle$, c'est donc le générateur – infinitésimal – du mouvement. En Mécanique Classique, le crochet de Poisson d'une constante du mouvement avec H est nul ; en Mécanique Quantique, une observable sera une constante du mouvement si et seulement si son commutateur avec H est nul.

À l'instar d'un vecteur, représenté par ses composantes, le choix d'une base de l'espace vectoriel permet d'associer à tout opérateur A sa matrice⁵⁶ sur la base choisie. Plus particulièrement, un opérateur A étant donné, ses vecteurs propres $\{|a_m\rangle\}$ peuvent être utilisés pour engendrer effectivement l'espace vectoriel \mathcal{E} , en constituant une base. De la sorte, tout vecteur d'état $|\Psi\rangle$ peut s'écrire en combinaison linéaire :

$$|\Psi\rangle = \sum_m c_m |a_m\rangle . \quad (1.208)$$

Comme on l'a déjà noté, les coefficients $\{c_m\}$ jouent le rôle d'une "fonction d'onde discrète". En ce qui concerne la variable dynamique A , l'ensemble $\{|a_m\rangle\}$ doit en fait constituer une base de l'espace vectoriel des états. Si l'état considéré, $|\Psi\rangle$, dépend non trivialement du temps (état non stationnaire), les coefficients (les composantes de $|\Psi\rangle$) dépendent du temps autrement que par une simple phase.

On peut enfin considérer un cas plus compliqué où la description complète du système doit se faire simultanément en terme d'une grandeur continue, comme la position, et d'une grandeur discrète comme A (exemple : A est le moment cinétique (spin) d'un atome en voyage dans l'espace, dont la position est notée \vec{R}). Dans ce cas, l'état le plus général contiendra à la fois \vec{R} et l'indice m associé aux différentes valeurs possibles de A ; un tel état peut être noté :

$$|\Psi(\vec{R}; t)\rangle = \sum_m c_m(\vec{R}; t) |a_m\rangle . \quad (1.209)$$

En pareil cas, $|c_m(\vec{R}; t)|^2$ est la densité de probabilité de trouver l'atome près de \vec{R} avec la valeur a_m . Un exemple de cette décomposition est fourni par la description de l'expérience de Stern et Gerlach ; alors, les $\{|a_m\rangle\}$ sont les états de spin et \vec{R} désigne la position de l'atome. Le ; isole le paramètre temps qui, en Mécanique Quantique, est toujours un paramètre ne relevant d'aucune sorte de quantification et jamais rien de plus⁵⁷.

1.4.3 Résultats possibles de la mesure d'une grandeur physique

Les valeurs propres d'une observable jouent un rôle central ; elles sont au centre du 3^{ème} postulat qui s'énonce comme suit :

⁵⁶On ne sera pas surpris d'apprendre que les matrices des observables sur la base du Hamiltonien ne sont rien autre que les matrices introduites d'emblée par Heisenberg au sens de sa Mécanique des Matrices.

⁵⁷Il existe certes des relations du genre $\Delta E \tau \sim \hbar$, mais elles ont un statut très différent de celles (ex. : $\Delta x \Delta p_x \sim \hbar$) exprimant le principe d'incertitude de Heisenberg.

Postulat 3 : la mesure d'une grandeur physique représentée par l'observable A ne peut fournir que l'une des valeurs propres de A .

Le sens physique du 3^{ème} postulat impose aux opérateurs associés aux grandeurs physiques d'avoir toutes leurs valeurs propres réelles – ce qui est bien le cas puisque A est une observable, dont la définition vient d'être donnée (sous-section 1.4.2). En particulier, (section 1.3), les valeurs propres de A sont réelles et tous ses vecteurs propres sont orthogonaux.

Comme pour tout opérateur, le spectre d'un opérateur hermitique peut être discret, continu ou les deux à la fois⁵⁸. Les parties discrète et continue, si elles existent simultanément, peuvent être disjointes ou se recouvrir : une valeur propre discrète peut ainsi être "noyée" dans un *continuum* d'autres valeurs propres⁵⁹. Les valeurs propres discrètes sont celles qui expriment la quantification de la grandeur physique en question. Notons qu'une valeur propre discrète peut servir de standard métrologique puisqu'elle a, intrinsèquement, une valeur fixée par la Nature à une précision infiniment grande (exemple : une certaine transition de l'atome de ⁸⁶Kr a longtemps servi de référence pour la définition du mètre-étalon).

L'illustration la plus simple de ce postulat est sans doute le résultat de l'expérience de Stern et Gerlach ; pour le plus petit moment cinétique non nul ($J = 1/2$ en unités \hbar), on obtient deux taches symétriques, correspondant aux deux seules valeurs possibles en tant que résultat d'une mesure. Compte tenu de la symétrie de l'espace vis-à-vis du haut et du bas, ceci implique que, si $J = 1/2$ est la valeur du moment cinétique, les deux seules valeurs sont $-1/2$ et $-1/2 + 1 = +1/2$. Plus généralement, pour un moment cinétique de valeur J , on aura les $2J + 1$ valeurs $M = -J, -(J - 1), -(J - 2), \dots, +(J - 1), +J$, avec la valeur 0 au milieu si J est un nombre entier.

Une autre illustration est fournie par le processus d'absorption ou d'émission de photon(s) par un atome. Par l'équation de conservation de l'énergie⁶⁰ (en émission, pour l'exemple) :

$$E_i = E_f + h\nu . \quad (1.210)$$

L'observation de la fréquence ν constitue bien une mesure de l'énergie de l'atome, à partir d'une origine arbitraire. On ne trouve que les valeurs E_n appartenant au spectre du Hamiltonien du système considéré.

La nature probabiliste (statistique) essentielle de la Mécanique Quantique doit être maintenant bien admise. On vient de voir que les valeurs possibles des résultats d'une mesure appartiennent à un ensemble de nombres, discrets et/ou continus. La Mécanique Quantique doit donc, pour sa propre cohérence, énoncer une règle permettant la prédiction des résultats d'une mesure ; cette règle est une affirmation sur les probabilités d'observer une valeur ou une autre lors de la mesure d'une observable. Elle constitue le 4^{ème} Postulat.

Soit une grandeur physique représentée par l'opérateur (observable) A , dont les valeurs et vecteurs propres sont discrets et non dégénérés :

$$A|a_m\rangle = a_m |a_m\rangle . \quad (1.211)$$

L'absence de dégénérescence signifie qu'à une valeur propre il ne correspond qu'un seul vecteur propre ; ceci se traduit, dans les notations, par le fait que le ket contient uniquement le symbole correspondant à la valeur propre a_m . Les états propres constituent une base sur laquelle on peut décomposer l'état $|\Psi\rangle$:

$$|\Psi\rangle = \sum_m c_m |a_m\rangle . \quad (1.212)$$

⁵⁸Lorsqu'il s'agit du Hamiltonien, dont le spectre est l'ensemble des énergies possibles, la nature de ce spectre conditionne potentiellement toute la dynamique du système. Des valeurs propres discrètes donneront une suite discrète (finie ou infinie) de termes oscillants dans le temps. Si les fréquences sont commensurables (le rapport de deux d'entre elles étant un nombre rationnel), le mouvement est périodique ; l'oscillateur harmonique constitue un cas extrême où toutes les fréquences sont les multiples entiers d'une même fréquence fondamentale. Si ces rapports sont irrationnels (au moins certains d'entre eux), le mouvement est pseudo-périodique avec des temps de Poincaré éventuellement exponentiellement grands. Enfin, si les énergies forment un spectre continu, la dynamique est irréversible. On retiendra l'idée que le couplage à un *continuum* d'énergie est l'ingrédient nécessaire (et suffisant) pour l'apparition de l'irréversibilité dynamique. C'est ainsi que l'on peut rendre compte théoriquement de la durée de vie *finie* d'un état excité atomique, à condition d'incorporer le champ électromagnétique (les photons) dans la description quantique.

⁵⁹Cette situation où une valeur propre discrète est noyée dans un intervalle continu est assez rare.

⁶⁰et en négligeant les effets liés au recul de l'atome.

Les c_m sont les composantes de l'état $|\Psi\rangle$ sur la base choisie. On peut maintenant énoncer le 4^{ème} Postulat, aussi appelé "Principe de décomposition spectrale" :

Postulat 4 : la mesure de la grandeur physique représentée par l'observable A , effectuée sur un état quelconque (normalisé) $|\Psi\rangle$, donne le résultat a_m avec la probabilité P_m égale à $|c_m|^2$.

À nouveau, l'expérience de Stern et Gerlach fournit une illustration immédiate de ce postulat. Si $S = 1/2$, une base des états de spin d'un atome est constitué par deux vecteurs que l'on peut noter $|+\rangle$ et $|-\rangle$. L'état d'un atome, avant la traversée de l'appareil, est une combinaison linéaire :

$$|\Psi(\vec{R}; t = 0)\rangle = f(\vec{R}) (c_+|+\rangle + c_-|-\rangle) \equiv c_+(\vec{R}; t = 0)|+\rangle + c_-(\vec{R}; t = 0)|-\rangle , \quad (1.213)$$

où on ne connaît pas d'ailleurs la valeur des coefficients c_{\pm} : ceux-ci sont en effet complètement indéterminés puisque, à la sortie niveau du four (instant initial), il n'existe aucune direction privilégiée qui pourrait permettre à l'atome, par exemple, de préférer être plutôt dans un état que dans l'autre. Le module carré de la fonction f donne la densité de présence de l'atome au point \vec{R} : c'est le paquet d'ondes, à la sortie du four. On observe, sur l'écran situé à la sortie de l'appareil, deux taches de même intensité, ce qui signifie que, dans les conditions de l'expérience, les deux seules valeurs possibles, en tant que résultats d'une mesure, du moment cinétique intrinsèque de l'atome ont la même probabilité, soit $1/2$. Techniquement, ceci se traduit comme suit : après la traversée de l'aimant, l'état évolue comme un paquet d'ondes libre, formellement représenté par :

$$|\Psi(\vec{R}; t)\rangle = c_+(\vec{R}; t)|+\rangle + c_-(\vec{R}; t)|-\rangle . \quad (1.214)$$

Chacune des fonctions c_{\pm} présente un maximum pointu ; ces maxima suivent deux lignes ("trajectoires") aboutissant sur l'écran en deux points nettement séparés, pourvu que l'aimant soit assez long, condition qui exprime son caractère macroscopique. La biunivocité introduite par le dispositif entre position de l'atome (degré de liberté externe) et spin (degré de liberté interne) qualifie l'expérience en tant que mesure du spin.

On apprécie encore mieux l'importance des états propres d'une observable A donnée en notant ceci : si l'opération de mesure est faite sur un état $|\Psi\rangle$ qui coïncide avec l'un des états propres de A , soit $|a_{m_0}\rangle$, alors le résultat de la mesure est certainement la valeur propre a_{m_0} . En définitive, c'est le seul cas où la mesure de A donne un et un seul résultat avec certitude⁶¹.

Par ailleurs, deux vecteurs d'états ne différant que par une phase *globale* constante représentent le même état physique ; en effet, toutes les probabilités donnant les résultats d'une mesure coïncident dans un cas et dans l'autre, puisque le module au carré efface une phase globale. Une phase globale n'affectant aucune des prévisions physiques, il est naturel d'affirmer que l'on est en présence d'un seul et même état⁶².

Le postulat 4 se généralise au cas où le spectre de l'opérateur A présente une dégénérescence, c'est-à-dire lorsqu'il existe un certain nombre entier, $g_m > 1$, de vecteurs propres linéairement indépendants associés à la même valeur propre. Dans ce cas, chaque vecteur propre doit être repéré par la valeur propre, a_m , et par un autre label, indice r , permettant de distinguer entre eux les différents vecteurs propres tous associés à la même valeur propre a_m . L'ensemble de ces g_m vecteurs engendre le sous-espace propre de dimension g_m associé à a_m . Ainsi, on note :

$$A|a_m, r\rangle = a_m|a_m, r\rangle , \quad r = 1, 2, \dots, g_m . \quad (1.215)$$

Dans ce cas, l'état $|\Psi\rangle$ se décompose en sommant d'une part sur les différentes valeurs propres et d'autre part, pour une valeur propre, sur toutes les "directions" dans le sous-espace dégénéré :

$$|\Psi\rangle = \sum_m \sum_r c_{m r} |a_m, r\rangle . \quad (1.216)$$

La probabilité d'obtenir la valeur a_m par une mesure effectuée sur l'état $|\Psi\rangle$ s'obtient alors à partir du postulat 4 étendu :

⁶¹Le résultat d'une mesure n'est donc pas toujours entaché d'aléatoire au sens des postulats ; une opération de mesure ne constitue donc pas *toujours* une perturbation incontrôlable de l'état du système : si *juste avant* la mesure le système est dans un état propre de l'observable mesurée, l'opération de mesure ne modifie pas l'état.

⁶²D'ailleurs, la condition de normalisation de la fonction d'onde ne permet jamais de trouver la constante de normalisation à mieux qu'un facteur de phase près.

Postulat 4 étendu : la mesure de la grandeur physique représentée par l'observable A , effectuée sur l'état $|\Psi\rangle$ normalisé donne le résultat a_m avec la probabilité $\sum_r |c_{mr}|^2$.

Il est clair que cette affirmation contient la première formulation du 4^{ème} postulat en tant que cas particulier : il suffit de faire $g_m = 1$. Par ailleurs, pour que cette généralisation ait bien un sens physique, il faut que la probabilité P_m ne dépende pas de la base choisie, c'est-à-dire que n'importe quel changement de base effectué à l'intérieur du sous-espace propre de a_m ne modifie pas P_m . C'est bien le cas, puisqu'un tel changement de base est représenté par un opérateur unitaire, qui ne change pas la norme et donc ne change pas la longueur de la projection de $|\Psi\rangle$ dans le sous-espace dégénéré.

Remarque

En Mécanique Quantique, il est essentiel de bien faire la distinction entre somme des modules élevés au carré et module de la somme élevée au carré. Avec ceci en tête, l'affirmation du Postulat 4 étendu mérite d'être commentée.

L'existence d'une dégénérescence pour A signifie que la connaissance d'un vecteur propre recouvre une "réalité" multiple, puisque, en pareil cas, ce vecteur propre peut être n'importe lequel des vecteurs du sous-espace propre associé à la valeur propre considérée (en quelque sorte, la dégénérescence traduit une sorte d'ignorance tant qu'aucune autre observable n'est prise en considération) ; plus précisément, ce qui a un sens du point de vue de l'invariance physique, c'est bien ce sous-espace et non pas l'un quelconque de ces vecteurs. Ce sous-espace, lui, est défini sans aucune ambiguïté ; dans les notations précédentes, sa dimension est égale à g_m .

Pour donner un sens physique direct à chacun de ces vecteurs, il suffit de trouver d'autres observables qui commutent avec A et de déterminer leurs vecteurs propres. Soit par exemple le vecteur $|a_m, b_n\rangle$, propre à la fois à A et à B qui commutent. Si tous les couples de valeurs propres (a_m, b_n) sont distincts (en tant que couples), il n'y a plus de dégénérescence à proprement parler. Comme A et B commutent, il s'agit de deux grandeurs compatibles. Le contenu du postulat 4 étendu consiste à dire que lors de la mesure de A donnant la valeur propre a_m dégénérée, l'observable B a ou bien la valeur b_1 , ou bien la valeur b_2 , etc. Ceci renvoie à l'addition des probabilités relatives à des événements mutuellement exclusifs, comme l'affirme d'ordinaire la théorie des probabilités.

Bien évidemment, si le couple (A, B) présente encore de la dégénérescence, il suffit de trouver une troisième observable, C , compatible avec A et B , et de déterminer les vecteurs propres communs à A , B et C , $|a_m, b_n, c_p\rangle$, et ainsi de suite. Un ensemble d'observables dont chaque vecteur propre est unique (tous les n -plets (a_m, b_n, c_p) sont distincts les uns des autres) porte le nom d'ECOC (ensemble complet d'observables qui commutent). On reviendra sur ce point dans la suite.

Il reste à dire un mot du cas où l'observable A possède un spectre continu, qui englobe en fait le postulat (1.26) à la base de l'interprétation de Born et Jordan de la fonction d'onde de Schrödinger. Soit $|a\rangle$ un état propre associé à la valeur propre a , appartenant à une partie continue du spectre ; l'état $|\Psi\rangle$ se décompose comme :

$$|\Psi\rangle = \int da c(a) |a\rangle . \quad (1.217)$$

Alors, la probabilité élémentaire d'obtenir comme résultat de mesure la valeur située dans l'intervalle $[a, a + da]$ est :

$$dP(a) = |c(a)|^2 da . \quad (1.218)$$

Si a désigne la position x , $c(a)$ n'est autre que la fonction d'onde $\Psi(x)$ déjà introduite et, en effet, selon Born et Jordan, $|\Psi(x)|^2 dx$ désigne la probabilité élémentaire $dP(x)$ d'observer la particule dans l'intervalle $[x, x + dx]$.

1.4.4 La réduction du paquet d'ondes

On désigne ainsi un phénomène⁶³ spécifique de la Mécanique Quantique qui apparaît, à ce stade de l'énoncé des principes, presque comme une nécessité compte tenu des postulats déjà énoncés, mais qu'il faut toutefois inclure

⁶³Aussi appelé *collapse* de la fonction d'onde.

dans la construction axiomatique. Ce phénomène, assez déconcertant au premier abord, a été (et est encore) l'objet de débats passionnés sur la notion de réalité physique et de la connaissance scientifique en général que l'on peut souhaiter en avoir. Pour Schrödinger (d'après [9]), la réduction du paquet d'ondes était de la magie et, pour marquer son opposition, énonça le fameux paradoxe connu sous le nom du "Paradoxe du Chat de Schrödinger"⁶⁴. La nécessité d'accepter ce phénomène se comprend aisément. En effet, ayant procédé à une mesure à l'instant t et ayant trouvé la valeur a_{n_0} , on ne peut que retrouver la même valeur a_{n_0} , et ce avec certitude, si, à l'instant immédiatement postérieur $t + dt$, on fait une nouvelle mesure de A . Autrement dit : toute mesure de A donne une certaine valeur appartenant au spectre de A . Si une première mesure ayant donné a_{n_0} est immédiatement suivie d'une autre, le résultat doit être certain (sans dispersion des résultats) et la valeur trouvée la deuxième fois ne peut être qu'identique à celle qui vient d'être trouvée lors de la première mesure. Il ne serait pas acceptable, physiquement parlant, que deux mesures infiniment proches dans le temps l'une de l'autre, produisent deux résultats arbitrairement différents. Cela signifierait qu'entre deux instants arbitrairement voisins le système a changé d'état "autant" que l'on veut, ce qui n'aurait pas de sens⁶⁵. *Force* est donc d'admettre le 5^{ème} Postulat :

Postulat 5 : si la mesure de l'observable A donne le résultat a_m , et si la valeur propre a_m est non-dégénérée, alors, immédiatement après la mesure, le système est dans l'état propre $|a_m\rangle$.

Ce postulat est clairement lié à l'aspect probabiliste des prédictions de la Mécanique Quantique : celle-ci permet de calculer des probabilités pour que tel ou tel événement se produise. Bien évidemment, juste après que l'événement s'est produit, il n'est plus question de raisonner, à cet instant précis, en terme de probabilités ; quand on joue à pile ou face avec une pièce symétrique, la probabilité d'obtenir pile ou face vaut exactement $1/2$, mais une fois obtenu pile ou face à la suite d'un jet de la pièce, la notion de probabilité disparaît en tant que telle, sauf si on décide de recommencer une expérience ... ou si la table sur laquelle est tombée la pièce se met en mouvement de façon erratique susceptible de provoquer un nouveau jet de la pièce. Dans ce dernier cas, c'est par l'évolution propre du système que la notion de probabilité reprend le devant de la scène. Dit autrement, lorsqu'une suite d'événements est seulement prévisible en terme de probabilités, l'occurrence (la réalisation) de l'un d'entre eux modifie radicalement la perception du phénomène. La conséquence de ce postulat, allié au Postulat 4, est bien d'assurer qu'une deuxième mesure de A redonnera la même valeur que la mesure immédiatement antérieure et déjà effectuée. Le processus de réduction du paquet d'ondes lors d'une mesure de A peut s'illustrer comme montré sur la figure 1.3. Cette évolution spécifique ne relève pas de l'équation de Schrödinger ; ceci sera rediscuté dans la suite.

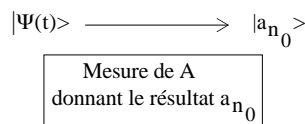


Figure 1.3: Illustration de la réduction du paquet d'ondes Ψ .

Afin d'inclure le cas où le spectre de l'observable mesurée est dégénéré, il faut énoncer la généralisation :

⁶⁴Les résultats des expériences d'A. Aspect *et al.* conduisent certains à affirmer que l'on doit renoncer soit au postulat de la Relativité affirmant que la vitesse de la lumière c est une vitesse limite, soit à l'idée que deux systèmes ayant interagi sont séparés, totalement indépendants l'un de l'autre. En effet, si on dit que deux tels systèmes sont séparés, alors, selon la Mécanique Quantique, toute mesure effectuée sur l'un donne instantanément une information sur l'autre : ceci viole le postulat relativiste. Si on refuse cette violation, alors il faut admettre que les systèmes ne sont pas séparés... même s'il s'agit de deux particules distantes de dizaines de mètres, de kilomètres ou d'années-lumière. En fait, la Mécanique Quantique oblige à une révision complète de la notion de séparabilité de deux systèmes.

Cette analyse présuppose – ce qui n'est pas toujours assuré dans les discussions à ce sujet – la définition précise de ce que l'on appelle un échange d'information ; s'il requiert la propagation d'un signal, il est certain que surgit alors la contrainte imposée par la Relativité. D'un autre côté, on peut acquérir de l'"information" à distance, de façon instantanée : deux personnes sur le point de se séparer, tirent chacune (sans la regarder) une boule dans un sac qui en contient deux ; les deux personnes savent que le sac contient une boule blanche et une boule noire. Une fois rendue à destination, chaque personne, prenant connaissance de la couleur de la boule qu'elle a emportée, en déduit immédiatement la couleur de la boule tirée par l'autre ; il n'y a là rien de choquant, ni de contradictoire avec quoi que ce soit. Il reste cependant une différence fondamentale avec la situation quantique : la couleur des boules est définie intrinsèquement, sans référence à l'intervention d'un observateur, préalablement à toute opération de mesure.

⁶⁵Ceci impliquerait que la vitesse d'évolution du système dans l'espace des états serait *infinie*, ou, en termes plus imagés, que sa "trajectoire" ne serait pas différentiable. Or l'équation (1.1) – qui gouverne l'évolution temporelle – assure que la dérivée en temps existe.

Postulat 5 étendu : si la mesure de l'observable A donne la valeur a_m et si cette valeur propre est dégénérée, l'état du système juste après la mesure est la projection du vecteur juste avant la mesure dans le sous-espace correspondant.

Analytiquement, ceci s'écrit comme suit. Le système étant dans l'état (1.216), supposons que la mesure de A donne la valeur a_m ; alors, juste après la mesure, le système se trouve dans l'état (défini à un facteur près) :

$$\Pi_m |\Psi\rangle = \sum_r c_{m,r} |a_m, r\rangle \quad (1.219)$$

où Π_m désigne l'opération de projection sur le sous-espace dégénéré $\{|a_m, r\rangle\}_r$. Cet état n'est plus normé, mais on peut le normaliser et obtenir $|\Psi_{a_m}\rangle$, état normalisé après avoir trouvé a_m :

$$|\Psi_{a_m}\rangle = \frac{1}{\sum_r |c_{m,r}|^2} \sum_r c_{m,r} |a_m, r\rangle . \quad (1.220)$$

Dans le cas où la valeur propre a_m est dégénérée, l'état immédiatement après la mesure est donc la projection, dans le sous-espace dégénéré, de l'état sur lequel la mesure est faite ; les coefficients du développement ne sont donc en aucune façon arbitraires et possèdent au contraire une relation mutuelle de *phase* parfaitement déterminée, celle qui était présente juste avant la mesure. Une telle phase joue un rôle essentiel dans l'apparition des interférences quantiques.

Ce postulat permet de mieux saisir l'impossibilité de mesurer simultanément deux observables incompatibles A et B ($[A, B] \neq 0$) ; en effet, à l'issue d'une telle mesure hypothétique, le vecteur d'état devrait être l'un des vecteurs propres communs à la fois à A et à B , et on sait que deux opérateurs qui ne commutent pas n'ont pas un système de vecteurs propres en communs. Autrement dit, vouloir mesurer simultanément deux grandeurs incompatibles, ce serait essayer de mettre simultanément le système dans deux états distincts, ce qui est manifestement impossible.

Un exemple, dû à Bohm ([3], p. 126) montre que le changement abrupt des objets mathématiques représentant l'évolution d'un système se présente aussi, par nature, en théorie usuelle des Probabilités. Soit une personne dont on sait qu'elle est âgée de 20 ans ; compte tenu de ceci, on peut lui estimer une durée de vie de l'ordre de 50 ans environ. Si, à un moment donné, on apprend que cette personne atteinte d'une grave maladie, on peut alors affirmer, presque instantanément, que le nombre d'années qui lui restent à vivre est beaucoup plus faible. Il y a bien, au moment où l'information plus précise sur l'état de la personne est donnée, une variation très rapide de la prévision sur l'avenir, que l'on peut être tenté de comparer à la réduction du paquet d'ondes, et qui se produit strictement au niveau de la connaissance de l'observateur : apprendre que la personne est malade ne modifie pas l'âge de cette personne. La brusque modification porte donc uniquement sur la connaissance que l'on a de la situation, non sur l'état de l'objet analysé. Il existe toutefois une différence essentielle entre l'exemple de Bohm et la situation quantique ; en effet, l'état de la personne (qu'il soit connu ou inconnu de l'interrogateur) est une donnée en soi, existant préalablement à toute opération de "mesure" ; au contraire, la position de l'atome dans l'appareil de Stern et Gerlach est décrite comme un ensemble de potentialités et il ne faut pas dire que, avant localisation sur l'écran, l'atome se trouve au voisinage de l'une des deux bosses de probabilité du paquet d'ondes (le dire, ce serait – comme pour les fentes d'Young – commencer à reconstruire une trajectoire à l'envers).

La notion de probabilité en Mécanique Quantique est bien à prendre au sens usuel de la théorie classique des probabilités, et c'est en effet la fonction d'onde qui permet de prévoir et de calculer les probabilités quantiques. De ce point de vue, le cadre probabiliste inévitable est donc, en ce sens, "classique". On ne peut toutefois prétendre que $|\Psi\rangle$ décrit seulement la connaissance que l'on a d'un système et non pas l'état du système lui-même : quand un électron se matérialise sur un écran, il s'agit bien d'un processus physique qui se produit intrinsèquement, qu'il y ait ou non un observateur pour regarder le spot lumineux. Autrement dit, la réduction du paquet d'ondes n'est pas un processus impliquant la seule perception mentale de l'observateur, un grand champ de possibilités venant se réduire à l'un d'entre eux sous l'effet d'une mesure apportant une précieuse information. La matérialisation d'une particule est le résultat d'une interaction physique entre cette particule et un autre système : que l'état juste après la matérialisation soit très différent de l'état qui serait survenu en l'absence de cette matérialisation ne doit donc ni étonner ni surprendre. On retiendra l'idée que la réduction du

paquet d'ondes est la contrepartie physique naturelle d'un processus se produisant réellement⁶⁶. La description détaillée de ce processus implique, en bonne logique, de décrire quantiquement l'appareil de mesure et son interaction avec l'objet de la mesure : personne ne sait faire ceci actuellement. D'ailleurs, tôt ou tard on retomberait sur des problèmes de nature conceptuelle ; si l'appareil est décrit quantiquement, le statut de l'aiguille est le même que celui des électrons dans l'expérience d'Young : l'aiguille n'est nulle part, mais potentiellement partout, le chat de Schrödinger est dans une superposition d'états mort et vivant, *etc.*

Ce problème de la réduction du paquet d'ondes est discuté dans une abondante littérature où les présupposés philosophiques se heurtent parfois de front avec les réflexes hérités d'une démarche scientifique dure. En tout état de cause, il serait quelque peu abusif d'affirmer que les problèmes de fond soulevés dans ces débats sont tous complètement résolus à l'heure actuelle⁶⁷. Leur discussion est ici de toute façon totalement hors de propos. Quoi qu'il en soit, il est clair que ce postulat se révèle nécessaire pour la cohérence interne de la théorie et assure qu'elle est sensée physiquement. Il est intéressant de remarquer que cette nécessité, admise sur des considérations physiques, a des conséquences pour le moins "paradoxaes", au sens usuel.

Enfin, il est important de remarquer que l'opération de mesure n'est pas définie en soi, préalablement à l'énoncé du postulat. En définitive, l'opération de mesure n'est pas décrite en détail en termes physiques : elle est seulement définie par ses conséquences (juste après l'issue d'une mesure, le système est dans l'état propre *etc.*) Notons pour terminer qu'une opération de mesure permet de préparer un système dans un état bien particulier ; dans l'expérience de Stern et Gerlach, immédiatement après sa matérialisation sur l'écran, chaque atome est bien localisé spatialement et se trouve dans l'état de spin, $|+\rangle$ ou $|-\rangle$, en corrélation parfaite et totale avec l'endroit où il est arrivé. Après une mesure ayant donné la valeur propre non-dégénérée a_{n_0} , on dispose d'un système préparé dans l'état $|a_{n_0}\rangle$. Le cas du spectre continu mérite une discussion plus poussée (voir la discussion dans [4], III. E 2. b, sur les appareils de mesure insuffisamment sélectifs).

1.4.5 Evolution des systèmes dans le temps

Le postulat énonçant le mode d'évolution temporelle a déjà été formulé, puisque l'équation de Schrödinger, ne se démontrant pas, a le statut d'un postulat ; on a vu comment elle peut s'induire par des arguments par analogie et de nature heuristique. Il faut bien ériger la conviction correspondante sous forme de postulat - c'est le dernier pour l'instant⁶⁸, exprimé à propos du vecteur d'état $|\Psi(t)\rangle$:

Postulat 6 : l'évolution d'un système dans le temps est gouvernée par l'équation :

$$i\hbar \frac{d}{dt} |\Psi(t)\rangle = H(t) |\Psi(t)\rangle \quad (1.221)$$

où $H(t)$ est l'opérateur associé à l'énergie du système.

Ainsi formulé, ce postulat englobe le cas où le système décrit par $|\Psi(t)\rangle$ ne constitue pas un système isolé. On pourra donc partir de cette équation pour décrire, par exemple, l'évolution d'un atome couplé à un champ électromagnétique variable dans le temps et supposé classique⁶⁹.

Comme déjà mentionné à plusieurs reprises, cette équation suppose la donnée supplémentaire d'un état initial, $|\Psi(t_0)\rangle$; à défaut, aucune prévision n'est possible. À l'inverse, un état initial étant donné, la solution qui

⁶⁶Bohm ([3], p. 127) récuse l'idée d'un *collapse* soudain de la fonction d'onde et invoque une perte graduelle de la cohérence de phase entre le système quantique et l'appareil de mesure. Cette position ne va pas à l'encontre de l'interprétation usuelle : la réduction du paquet d'ondes est en effet le résultat d'un processus physique modifiant l'état du système et ce processus prend en effet un certain temps. Le 5^{ème} Postulat n'affirme pas que ce processus est instantané.

⁶⁷Exemple de question que l'on peut légitimement se poser, par exemple à propos des fentes d'Young : juste avant d'être matérialisé sur l'écran en un certain point, l'électron est-il "partout" ou à proximité du point où il va se localiser ? Il peut paraître raisonnable d'admettre qu'il est à proximité ; mais alors, de proche en proche, on peut reconstruire une trace dans l'espace, qui n'est pas vraiment une trajectoire, et "remonter" ainsi au trou où est passé l'électron. Dès lors, on en vient à admettre que l'électron passe par un trou ou par l'autre ; mais alors, occulter l'un des deux trous ne devrait pas qualitativement changer la répartition des impacts ; or il n'en est rien... De la même façon, dans l'expérience de Stern et Gerlach, juste avant de se matérialiser près de la tache indiquant qu'il a le spin +, où est l'atome ??? Mystère...

⁶⁸Il faudra en poser un autre, à propos des particules identiques.

⁶⁹Bien sûr, si on englobe le champ (les photons) dans la description quantique et si l'atome et les photons constituent un système isolé, alors le Hamiltonien global ne dépend pas du temps.

en est issue à un instant ultérieur est unique. C'est heureux : les probabilités des événements ultérieurs sont calculables de façon unique et ceci constitue l'expression essentielle du déterminisme quantique. L'indéterminisme ne réapparaît que dans la mesure où l'opération de mesure elle-même n'est pas, en pratique, descriptible à l'aide de cette équation. On lit parfois que cette équation est valable "en dehors de toute opération de mesure". Il n'y a pas lieu de donner cette restriction : si le système interagit avec un autre, il faut élargir la description et considérer le "super-système", auquel on peut appliquer l'équation ci-dessus – dans le cas d'une mesure à proprement parler – qui implique forcément un système macroscopique –, cette tâche est hors de portée.

1.5 Evolution temporelle d'un système quantique

Pour un système isolé, le Hamiltonien ne dépend pas du temps, par définition. Ceci ne veut évidemment pas dire que $|\Psi(t)\rangle$ ne dépend pas du temps, ni, *a fortiori*, que toutes les moyennes de grandeurs physiques calculées avec $|\Psi(t)\rangle$ sont constantes. Toutes les valeurs moyennes sont indépendantes du temps seulement dans le cas d'états très particuliers appelés états stationnaires définis par les équations (1.23) et (1.24). Un tel état peut être préparé par une opération de mesure de l'énergie, donnant l'énergie E_n ⁷⁰. À l'issue de cette opération de mesure, effectuée à $t = t_0$, le système, abandonné à lui-même, évolue par l'équation de Schrödinger à partir de cet état initial $|\Psi_n\rangle$; comme c'est le Hamiltonien qui pilote l'évolution dans le temps, l'état à l'instant $t > t_0$ est :

$$\text{Mesure de } E \text{ à } t_0 \text{ donnant } E_n \implies \forall t > t_0 : |\Psi(t)\rangle = e^{\frac{i}{\hbar} E_n(t-t_0)} |\Psi_n(t_0)\rangle . \quad (1.222)$$

Si on effectue à nouveau une mesure de l'énergie, on retrouve évidemment la même valeur E_n : ceci est l'expression la plus simple, en Mécanique Quantique, de la conservation de l'énergie pour un système isolé. La mesure de l'énergie d'un système situé dans un état stationnaire ne modifie donc pas l'énergie de celui-ci ; toute autre mesure ultérieure redonnera encore la même énergie⁷¹. En outre, la dépendance en temps d'un tel état stationnaire est remarquablement simple : c'est un simple facteur de phase. En conséquence, les probabilités des résultats d'une première mesure de n'importe quelle observable A , à partir d'un état stationnaire, sont des nombres indépendants du temps⁷², que cette observable soit ou non une constante du mouvement. Si A est une constante du mouvement, $[A, H] = 0$ et on peut mesurer A et H simultanément ; toute nouvelle mesure ultérieure redonnera la même valeur pour A et pour H . En revanche, si $[A, H] \neq 0$, les probabilités des résultats d'une deuxième mesure de A seront toujours des constantes mais la mesure modifiera l'état du système (qui ne sera plus stationnaire après la deuxième mesure) ; alors les probabilités des résultats d'une troisième mesure deviendront des fonctions explicites du temps auquel est effectuée celle-ci. En toute généralité, quel que soit l'état (stationnaire ou non) sur lequel est effectuée la mesure, les probabilités des résultats de mesure de l'énergie sont toujours des constantes, puisque H commute avec H : ainsi s'exprime la conservation de l'énergie en Mécanique Quantique.

La description de la dynamique quantique procède usuellement de deux méthodes : soit en introduisant explicitement le propagateur permettant de calculer le vecteur d'état $|\Psi(t)\rangle$ à tout instant (c'est le "point de vue" de Schrödinger), soit par l'emploi des équations de Heisenberg ("point de vue" de Heisenberg).

La description de Heisenberg propose une méthode directe : $|\Psi(t)\rangle$ contient souvent trop d'information et l'intégration explicite de l'évolution de $|\Psi(t)\rangle$ est un effort parfois inutilement excessif⁷³. Bien sûr, si on veut se réserver la possibilité de calculer les valeurs moyennes de toutes les observables, c'est $|\Psi(t)\rangle$ qu'il convient de trouver, en préalable. Cette situation, sans être rare, n'est pas la plus fréquente en pratique : en effet, on est intéressé très souvent à obtenir les valeurs moyennes à l'instant t de certaines observables et non pas de toutes. En pareil cas, il est bien plus direct de se poser la question de trouver les équations du mouvement pour les observables elles-mêmes ; conceptuellement, ces équations présentent un intérêt majeur : elles sont les pendants quantiques des équations de la Mécanique Classique. En définitive, si la description de Schrödinger porte sur le

⁷⁰Ceci est vrai même si l'énergie correspondante est dégénérée ; il reste en effet que, même dans ce cas, la dépendance temporelle se réduit à un facteur de phase global qui disparaît de toute opération de moyenne.

⁷¹Ceci est un exemple où une opération de mesure ne modifie en rien l'état du système. Il peut paraître surprenant que la mesure effectuée dans ces conditions, (c'est un processus physique d'interaction) ne modifie pas l'énergie du système, mais c'est ainsi...

⁷²Le facteur de phase temporel disparaît quand on prend les modules au carré.

⁷³Le cas du paquet d'ondes libre qui s'étale est un exemple simple montrant que si l'on cherche seulement l'écart quadratique de la coordonnée, l'usage des équations du mouvement de Heisenberg conduit au résultat de façon très économique (voir plus loin).

vecteur d'état lui-même, celle de Heisenberg décrit directement l'évolution des valeurs moyennes et propose des équations (équations du mouvement de Heisenberg) pour obtenir commodément ces dernières.

La description de Schrödinger est fondée formellement sur l'emploi de l'opérateur d'évolution U . L'équation fondamentale (1.1) peut être réécrite en introduisant un opérateur $U(t, t_0)$, dit opérateur d'évolution, engendrant l'état à l'instant t à partir de l'état à l'instant t_0 . En effet, compte tenu de l'unicité de la solution quand on connaît l'état de départ et de la linéarité de l'équation de Schrödinger, on peut affirmer que $|\Psi(t)\rangle$ s'obtient à partir de $|\Psi(t_0)\rangle$ par l'action d'un opérateur linéaire agissant dans l'espace des états ; on pose ainsi :

$$|\Psi(t)\rangle = U(t, t_0) |\Psi(t_0)\rangle . \quad (1.223)$$

En reportant dans (1.1), il vient :

$$i\hbar \frac{\partial}{\partial t} U(t, t_0) |\Psi(t_0)\rangle = H(t) U(t, t_0) |\Psi(t_0)\rangle . \quad (1.224)$$

Maintenant, l'état initial peut être choisi absolument quelconque ; en particulier, on peut le prendre successivement comme étant l'un des éléments d'une base de l'espace des états ; l'équation pour U doit donc être valide au niveau des opérateurs eux-mêmes. On obtient ainsi l'équation d'évolution pour l'opérateur U lui-même :

$$i\hbar \frac{\partial}{\partial t} U(t, t_0) = H(t) U(t, t_0) . \quad (1.225)$$

Bien évidemment, cette équation est aussi du premier ordre ; par la définition de U , on doit évidemment avoir:

$$U(t_0, t_0) = \mathbf{1} . \quad (1.226)$$

Par ailleurs, comme la norme de $|\Psi(t)\rangle$ est conservée au cours du temps (puisque H est hermitique), U est un opérateur unitaire :

$$[U(t, t_0)]^{-1} = U^\dagger(t, t_0) . \quad (1.227)$$

L'évolution de t_0 à t peut évidemment être décomposée en deux étapes : évolution de t_0 à t_1 puis évolution de t_1 à t (il n'est d'ailleurs pas nécessaire que t_1 se situe entre t_0 et t). Par la définition de l'opérateur d'évolution, il en résulte :

$$U(t, t_0) = U(t, t_1) U(t_1, t_0) \quad (1.228)$$

et, prenant $t = t_0$, on trouve ainsi :

$$\mathbf{1} = U(t_0, t_1) U(t_1, t_0) \iff U(t_1, t_0) = [U(t_0, t_1)]^{-1} = U^\dagger(t_0, t_1) . \quad (1.229)$$

L'inverse de U s'obtient en permutant les deux arguments – ce qui est bien naturel. Les diverses propriétés qui viennent d'être établies montrent que l'ensemble des opérateurs d'évolution d'un même système, pris à des instants différents, a une structure de groupe.

Remarque

Avant de continuer, une mise en garde est nécessaire. Il faut éviter l'erreur consistant à passer par mégarde sur le fait que l'équation (1.225) porte sur des *opérateurs*, dont l'algèbre est *a priori* non-commutative. Il en résulte que la solution de cette équation n'est pas :

$$U(t, t_0) = \exp \left[\frac{1}{i\hbar} \int_{t_0}^t dt' H(t') \right] \quad \mathbf{FAUX !} \quad (1.230)$$

Cette écriture n'est correcte que dans le cas où $H(t)$ commute avec $H(t')$, quels que soient les deux instants t et t' ⁷⁴. L'obtention explicite de l'opérateur U dans le cas où le Hamiltonien dépend du temps est en général très difficile et on ne sait résoudre ce problème (explicitement, à l'aide éventuellement de fonctions spéciales) que dans quelques cas assez simples, qui jouent un rôle exemplaire et servent de modèles.

⁷⁴C'est par exemple le cas lorsque $H(t)$ est de la forme $f(t)V$ où $f(t)$ est une fonction numérique (scalaire !) et V un opérateur indépendant du temps.

D'une façon générale, lorsque le problème ne peut être résolu exactement, on recourt aux méthodes de perturbation.

En revanche, lorsque H ne dépend pas du temps, la difficulté précédente tombe complètement : quand H est constant dans le temps, H commute bien sûr avec lui-même à tout temps et la solution formelle de (1.225) est :

$$U(t, t_0) = \exp \left[\frac{t - t_0}{i\hbar} H \right] \quad \left(\frac{\partial H}{\partial t} \equiv 0 \right) \quad (1.231)$$

ce que l'on écrit d'habitude :

$$U(t, t_0) = e^{\frac{1}{i\hbar} H (t - t_0)} \quad (1.232)$$

quand on est sûr de voir dans les parenthèses de l'exponentielle un facteur multiplicatif et non pas une dépendance fonctionnelle⁷⁵. Bien évidemment, en tant que fonction de H , $U(t, t_0)$ commute avec H :

$$[U(t, t_0), H] = 0 \quad \forall t, t_0 . \quad (1.233)$$

En désignant par E_n les énergies (valeurs propres de H) et par $|\psi_n\rangle$ les vecteurs propres correspondants, de (1.232) on déduit que $U(t)$ admet la décomposition :

$$U(t, t_0) = \sum_n |\psi_n\rangle e^{\frac{1}{i\hbar} E_n t} \langle \psi_n | . \quad (1.234)$$

Un exemple simple permet de mémoriser le piège à éviter quand H dépend du temps. Imaginons que le Hamiltonien d'un système change soudainement à un certain instant (pris comme origine), passant de la valeur constante H_0 à $t < 0$ à la valeur H_1 , également constante, à $t > 0$. Ecrivons maintenant l'évolution de $t_0 < 0$ à $t_1 > 0$. On a évidemment :

$$U(t_1, t_0) = U(t_1, 0+) U(0+, 0-) U(0-, t_0) . \quad (1.235)$$

L'introduction du U intermédiaire est pour la forme : il vaut l'identité $\mathbf{1}$ comme on va le voir. Par ailleurs, dans les deux autres intervalles, le Hamiltonien prend une valeur constante, indépendante du temps ; on peut donc utiliser les résultats précédents et écrire :

$$U(t_1, t_0) = e^{\frac{1}{i\hbar} H_1 (t_1 - 0)} \mathbf{1} e^{\frac{1}{i\hbar} H_0 (0 - t_0)} = e^{\frac{1}{i\hbar} H_1 t_1} e^{-\frac{1}{i\hbar} H_0 t_0} . \quad (1.236)$$

Comme H_1 et H_0 n'ont aucune raison fondamentale de commuter entre eux, on n'a pas le droit de regrouper les arguments dans une même exponentielle – qui en effet serait l'intégrale de $H(t)$:

$$U(t_1, t_0) = e^{\frac{1}{i\hbar} H_1 t_1} e^{\frac{1}{i\hbar} H_0 t_0} \neq e^{\frac{1}{i\hbar} (H_1 t_1 - H_0 t_0)} \equiv e^{\frac{1}{i\hbar} \int_{t_0}^{t_1} dt' H(t')} . \quad (1.237)$$

L'exemple qui vient d'être traité donne l'occasion de préciser ce qui se passe lors d'un brusque changement du Hamiltonien, et d'établir le résultat annoncé ci-dessus par anticipation : lors d'un saut fini de H , le vecteur d'état est continu. En effet, avec les mêmes notations, soit une variation instantanée de H passant de H_0 à H_1 à l'instant $t = 0$. L'intégration formelle de l'équation de Schrödinger entre deux instants $\pm \delta t$ de part et d'autre de $t = 0$ donne :

$$|\Psi(+\delta t)\rangle - |\Psi(-\delta t)\rangle = \frac{1}{i\hbar} \int_{-\delta t}^{+\delta t} dt' H(t') |\Psi(t')\rangle . \quad (1.238)$$

Le vecteur $|\Psi(t)\rangle$ est borné à tout instant puisqu'il est de norme finie ; par ailleurs, $H(t)$ est borné si l'on ne considère que des sauts d'amplitude finie (aucune des valeurs propres de H_0 et de H_1 n'est infinie). Au total, l'intégrand est borné : quand $\delta t \rightarrow 0$, l'intégrale tend aussi vers zéro et on en déduit :

$$\lim_{\delta t \rightarrow 0} [|\Psi(+\delta t)\rangle - |\Psi(-\delta t)\rangle] = 0 . \quad (1.239)$$

ce qui établit la continuité du vecteur d'état ; on peut raffiner l'argument en imaginant que H passe graduellement d'une valeur à l'autre. Par exemple, on peut envisager une "montée" linéaire [20] entre deux instants finis 0 et T , ce qui revient à poser explicitement :

$$H(t) = (H_1 - H_0) \frac{t}{T} + H_0 . \quad (1.240)$$

⁷⁵D'ailleurs, il faut bien que l'argument de l'exponentielle soit sans dimension.

L'intégration de l'équation de Schrödinger entre $t = 0$ et $t = T$ donne maintenant :

$$|\Psi(T)\rangle - |\Psi(0)\rangle = \frac{1}{i\hbar} (H_1 - H_0) \int_0^T dt' \frac{t'}{T} |\Psi(t')\rangle + H_0 \int_0^T dt' |\Psi(t')\rangle . \quad (1.241)$$

À nouveau, tous les intégrands sont bornés et si le facteur $(H_1 - H_0)$ l'est aussi, la limite $T \rightarrow 0$ du second membre est nulle. Avec l'opérateur d'évolution, on voit bien que, pour un saut fini du Hamiltonien $H(t)$ autour de $t = 0$:

$$\lim_{\delta t \rightarrow 0} U(+\delta t, -\delta t) = \mathbf{1} . \quad (1.242)$$

comme on l'a utilisé explicitement ci-dessus. On retiendra l'idée que juste après un tel changement, le vecteur d'état est inchangé et n'a pas le temps d'évoluer. À l'inverse, il est possible montrer que pour un changement infiniment lent (modification "adiabatique" de $H(t)$, d'où le nom de ce résultat : *théorème adiabatique*), l'état du système évolue de façon à être, à tout instant, un état propre du Hamiltonien instantané qui varie infiniment lentement. Ainsi, lorsqu'un système est dans l'état fondamental d'un Hamiltonien H_0 à $t = -\infty$ et si ce Hamiltonien varie infiniment lentement pour arriver à la nouvelle valeur H_1 (à $t = 0$ par exemple), le système se retrouve dans un état propre de H_1 qui est souvent l'état fondamental – mais ce n'est pas obligatoire.

Ceci étant, il est possible de trouver la moyenne à l'instant t de n'importe quelle observable A , qu'elle dépende ou non du temps ; cette moyenne est la moyenne d'un ensemble de mesures effectuées dans les mêmes conditions, au même instant t , à partir d'un même état initial $|\Psi(t_0)\rangle$ et se calcule suivant les principes énoncés antérieurement ; dans le cas le plus général où A dépend du temps, on a ($|\Psi\rangle$ étant préalablement normalisé à l'unité) :

$$\langle A(t) \rangle(t) = \langle \Psi(t) | A(t) | \Psi(t) \rangle = \langle U(t, t_0) \Psi(t_0) | A(t) | U(t, t_0) \Psi(t_0) \rangle \quad (\langle \Psi(t) | \Psi(t) \rangle = 1) . \quad (1.243)$$

Même si A ne dépend pas du temps, sa moyenne en dépend en général, sauf si $|\Psi(t)\rangle$ est un état stationnaire ; en général, pour un état quelconque, il vient :

$$\langle A \rangle(t) = \langle \Psi(t) | A | \Psi(t) \rangle = \langle U(t, t_0) \Psi(t_0) | A | U(t, t_0) \Psi(t_0) \rangle . \quad (1.244)$$

Parmi toutes les observables, l'énergie joue évidemment un rôle déterminant. Dans le cas le plus général, H dépend du temps et la moyenne de l'énergie s'obtient par :

$$E(t) \equiv \langle H(t) \rangle(t) = \langle \Psi(t) | H(t) | \Psi(t) \rangle = \langle U(t, t_0) \Psi(t_0) | H(t) | U(t, t_0) \Psi(t_0) \rangle . \quad (1.245)$$

La dérivée de $E(t)$ est :

$$\frac{d}{dt} E(t) = \left\langle \frac{d}{dt} \Psi(t) \right| H(t) | \Psi(t) \rangle + \langle \Psi(t) | \frac{d}{dt} H(t) | \Psi(t) \rangle + \langle \Psi(t) | H(t) | \frac{d}{dt} \Psi(t) \rangle . \quad (1.246)$$

En utilisant (1.1), l'antilinearité du produit scalaire et l'hermiticité de $H(t)$, on voit que les termes extrêmes se compensent et il reste :

$$\frac{d}{dt} E(t) = \langle \Psi(t) | \frac{d}{dt} H(t) | \Psi(t) \rangle \equiv \left\langle \frac{d}{dt} H(t) \right\rangle . \quad (1.247)$$

Ceci permet de réaliser comment s'exprime la conservation de l'énergie en Mécanique Quantique : si H est indépendant du temps, la moyenne de l'énergie dans tout état est constante dans le temps puisque cette dérivée s'exprime comme la valeur moyenne d'un opérateur identiquement nul. Dans le cas général, la dérivée de $E(t)$ s'obtient simplement en calculant la valeur moyenne de la dérivée de H . À titre d'exemple, revenons au cas où $H(t)$ varie instantanément de H_0 à H_1 . Comme $H(t)$ est en réalité constant par morceaux, dans chacun des demi-intervalles $t < 0$ et $t > 0$, on a :

$$E(t) = C^{\text{ste}} \equiv E_+ \quad \forall t < 0 \quad E(t) = C^{\text{ste}} \equiv E_- \quad \forall t > 0 \quad (1.248)$$

mais il n'y a *a priori* aucune raison pour que E_- soit égal E_+ . Pour calculer ces deux constantes, il suffit de se placer soit à $t = 0-$, soit à $t = 0+$; on a ainsi :

$$E_- = \langle \Psi(0-) | H_0 | \Psi(0-) \rangle \quad E_+ = \langle \Psi(0+) | H_1 | \Psi(0+) \rangle \quad (1.249)$$

et comme $|\Psi(0-)\rangle = |\Psi(0+)\rangle \equiv |\Psi(0)\rangle$, c'est toujours avec le même état que les moyennes sont prises. En ce qui concerne les valeurs de E de part et d'autre du saut de $H(t)$, tout est possible : elles peuvent être différentes ou égales, selon la nature du saut de H et aussi de la symétrie de l'état $|\Psi(0)\rangle$. Par exemple, si on applique un champ électrique à un oscillateur initialement dans un état propre de H_0 , l'énergie ne change pas⁷⁶. Autre exemple : si la fréquence propre de l'oscillateur est soudainement réduite à une nouvelle valeur plus faible, l'énergie change. À l'inverse, une particule confinée dans un puits infini qui se dilate soudainement a la même énergie en moyenne avant et après l'expansion. Il en va évidemment tout autrement pour la variance de l'énergie : si le système est initialement dans un état stationnaire de H_0 , alors ΔE est nulle ; à $t > 0$, sauf changement trivial de $H(t)$, le système n'est plus dans un état stationnaire de H_1 et la variance devient strictement positive⁷⁷.

Remarquons que l'on peut aussi calculer la variation éventuelle de l'énergie en partant de la relation (1.247). Dans le cas où $H(t)$ a seulement un saut en $t = 0$, on a⁷⁸ :

$$\frac{d}{dt} H(t) = (H_1 - H_0) \delta(t) \quad (1.251)$$

où $\delta(t)$ est la fonction de Dirac. Il en résulte :

$$\frac{d}{dt} E(t) = \langle \Psi(0) | (H_1 - H_0) | \Psi(0) \rangle \delta(t) \quad (1.252)$$

et $E(t)$ aura un saut si effectivement l'écart $(H_1 - H_0)$ a une valeur moyenne non nulle dans l'état impliqué. La dernière relation est finalement évidente : en l'intégrant de part et d'autre de $t = 0$, le premier membre donne $E(0+) - E(0-)$; quant au second, il représente bien la variation d'énergie puisqu'il vaut $\langle \Psi(0) | H_1 | \Psi(0) \rangle - \langle \Psi(0) | H_0 | \Psi(0) \rangle$.

En résumé, dans la description de Schrödinger, la dépendance en temps apparaît explicitement dans le vecteur représentant l'état du système à l'instant t . On appelle propagateur⁷⁹ l'objet qui permet d'écrire la fonction d'onde à l'instant t en fonction de l'état initial donné à l'instant t_0 . Seul sera ici considéré le cas où le Hamiltonien est indépendant du temps : comme on peut s'en douter, le traitement d'un problème où H n'est pas statique est très difficile. La relation engendrant l'état développé à partir d'un état initial⁸⁰ s'écrit :

$$\Psi(x, t) = \int dx_0 U(x, t; x_0, t_0) \Psi(x_0, t_0) \quad (1.253)$$

où $U(x, t; x_0, t_0) = \langle x | U(t, t_0) | x_0 \rangle$. Le noyau $U(x, t; x_0, t_0)$ est en quelque sorte une "matrice continue". En représentation- q , le noyau U satisfait :

$$i\hbar \frac{\partial}{\partial t} U(x, t; x_0, t_0) = H \left(x, p = -i\hbar \frac{\partial}{\partial x} \right) U(x, t; x_0, t_0) \quad (1.254)$$

avec la condition initiale :

$$U(x, t = t_0; x_0, t_0) = \delta(x - x_0) \quad (1.255)$$

exprimant que $U(x, t_0; x_0, t_0)$ est le noyau de l'opérateur identité $\mathbf{1}$. H étant supposé indépendant du temps, U ne dépend que de la *différence* des temps et on peut toujours prendre $t_0 = 0$; dans la suite, on pose $U(x, t; x_0, t_0 = 0) \equiv U(x, t; x_0)$. U est unitaire (il doit conserver la norme) ; en vertu de (1.229), on a :

$$U^{-1}(x, t; x_0) = U^\dagger(x, t; x_0) = U(x, -t; x_0) . \quad (1.256)$$

⁷⁶ H_0 et H_1 diffèrent par un terme linéaire en position et l'état $|\Psi(0)\rangle$, propre de H_0 est soit pair, soit impair en x : l'intégrale $\langle \Psi(0) | x | \Psi(0) \rangle$ est nulle et $E(t < 0) = E(t > 0)$.

⁷⁷ Dans le cas du puits infini soudainement dilaté, cette variance est même infinie.

⁷⁸ On utilise le fait que pour la fonction de Heaviside Y , on a :

$$\frac{d}{dt} Y(t) = \delta(t) \quad (1.250)$$

⁷⁹ La terminologie est fluctuante : le propagateur s'appelle aussi fonction de Green ; par ailleurs, le nom propagateur est utilisé aussi dans une acception un peu différente. Dans tous les cas, il s'agit d'une représentation ou d'une autre de ce qui permet de relier explicitement un "ancêtre" à sa descendance. Il doit être clair que la notion de propagateur n'est pas spécifiquement quantique... D'une façon générale, le propagateur est l'objet qui fait évoluer le système dans l'espace-temps.

⁸⁰ En raisonnant toujours à une dimension pour simplifier les écritures.

Il suffit donc de déterminer une fonction U_+ égale à U si $t > 0$ et nulle si $t < 0$; le noyau U à $t < 0$ s'en déduit par (1.256). U_+ est appelé propagateur causal, ou avancé.

L'équation (1.254), écrite pour U_+ , se résout aisément en utilisant une transformation de Laplace, en posant⁸¹ :

$$K(x, z; x_0) = \int_0^{+\infty} dt U_+(x, t; x_0) e^{-zt} \quad (\Re z > 0) . \quad (1.257)$$

La restriction⁸² $\Re z > 0$ est suffisante⁸³ ; il est possible de s'en convaincre intuitivement comme suit. U_+ est directement relié à un opérateur unitaire ($U(t)$) dont toutes les valeurs propres, ayant un module égal à 1, sont bornées. L'intégrale (1.257) converge dès que z a une partie réelle finie, aussi petite soit-elle ; $K(x, z; x_0)$ n'a donc aucune singularité dans le demi-plan de droite $\Re z > 0$.

La transformée de Laplace de l'équation (1.254) pour $U_+(x, t; x_0)$ se construit suivant la règle bien connue et s'écrit :

$$i\hbar [z K(x, z; x_0) - \delta(x - x_0)] = H(x, p) K(x, z; x_0) , \quad (1.258)$$

de sorte que l'équation de K reste différentielle en x en représentation- q ou différentielle en p en représentation- p . Pour une particule libre, $H = -[\hbar^2/(2m)](\partial^2/\partial x^2)$ et (1.258) donne :

$$z K(x, z; x_0) - \frac{i\hbar}{2m} \frac{\partial^2}{\partial x^2} K(x, z; x_0) = \delta(x - x_0) . \quad (1.259)$$

Pour résoudre (1.259), il est commode de faire maintenant une transformation de Fourier en x et de poser :

$$K(x, z; x_0) = \int_{-\infty}^{+\infty} \frac{dk}{2\pi} \mathcal{U}(k, z; x_0) e^{+ikx} , \quad \mathcal{U}(k, z; x_0) = \int_{-\infty}^{+\infty} dx K(x, z; x_0) e^{-ikx} . \quad (1.260)$$

La transformée de Fourier de $\delta(x - x_0)$ est e^{-ikx_0} ; la transformée de Fourier de l'équation (1.259) est :

$$z \mathcal{U}(k, z; x_0) + \frac{i\hbar}{2m} k^2 \mathcal{U}(k, z; x_0) = e^{-ikx_0} , \quad (1.261)$$

d'où l'on déduit immédiatement :

$$\mathcal{U}(k, z; x_0) = \frac{e^{-ikx_0}}{z + \frac{i\hbar k^2}{2m}} . \quad (1.262)$$

Soit $\mathcal{K}(k, t; x_0)$ la fonction de Laplace inverse de $\mathcal{U}(k, z; x_0)$; un calcul élémentaire utilisant le théorème des résidus fournit alors :

$$\mathcal{K}(k, t; x_0) = e^{-ikx_0} e^{-i \frac{\hbar k^2}{2m} t} . \quad (1.263)$$

On note au passage la relation :

$$[\mathcal{K}(k, t; x_0)]^* = \mathcal{K}(-k, -t; x_0) , \quad (1.264)$$

qui exprime la symétrie dans le renversement du temps : en renversant le temps et la vitesse (k est changé en son opposé), on obtient le complexe conjugué de \mathcal{K} . Ceci est en accord avec le fait que si $\Psi(x, t)$ est solution de l'équation de Schrödinger, $\Psi(x, -t)^*$ l'est aussi. Pour avoir enfin $U_+(x, t; x_0)$ il suffit maintenant de prendre la transformée de Fourier inverse de \mathcal{K} ; compte tenu de (1.263) :

$$U_+(x, t; x_0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk e^{ikx} e^{-ikx_0} e^{-i \frac{\hbar k^2}{2m} t} . \quad (1.265)$$

Ceci est une intégrale gaussienne qui se calcule aisément, par un changement de variable élémentaire. On trouve ainsi finalement :

$$U_+(x, t; x_0) = \sqrt{\frac{m}{2i\pi\hbar t}} e^{i \frac{m}{2\hbar t} (x-x_0)^2} \equiv U(x, t; x_0) \quad \forall t > 0 \quad (1.266)$$

⁸¹Pour être tout à fait explicite, il conviendrait de noter K_+ cette transformée de Laplace. On se souviendra qu'elle est reliée au propagateur avancé.

⁸² \Re désigne la partie réelle.

⁸³Et non pas $\Re z > s_0 > 0$.

qui est donc le propagateur U pour $t > 0$. On reconnaît dans l'exponentielle donnant U_+ l'action classique $S_{\text{class}} = m(x - x_0)^2/(2t)$ de la particule libre partie de x_0 à $t = 0$ et arrivée en x à l'instant t .

Pour avoir U pour $t < 0$, il suffit d'utiliser $U(t) = U^\dagger(-t)$ en prenant $t < 0$ – auquel cas $U^\dagger(-t)$ se déduit de (1.266) puisque son argument est positif. Il vient ainsi :

$$U(x, t < 0; x_0) = [U_+(x, -t; x_0)]^* = \left[\sqrt{\frac{m}{2i\pi\hbar t}} e^{i\frac{m}{2\hbar t}(x-x_0)^2} \right]^* . \quad (1.267)$$

La branche de la racine carrée est celle qui prend des valeurs réelles positives sur l'axe réel positif ; il en résulte que :

$$(\sqrt{z})^* = \sqrt{z^*} \quad (1.268)$$

et finalement :

$$U(x, t < 0; x_0) = \sqrt{\frac{m}{2(-i)\pi\hbar(-t)}} e^{-i\frac{m}{2\hbar(-t)}(x-x_0)^2} = \sqrt{\frac{m}{2i\pi\hbar t}} e^{i\frac{m}{2\hbar t}(x-x_0)^2} . \quad (1.269)$$

Ainsi, l'expression de U est la même, que t soit positif ou négatif⁸⁴ :

$$U(x, t; x_0) = \sqrt{\frac{m}{2i\pi\hbar t}} e^{i\frac{m(x-x_0)^2}{2\hbar t}} \equiv \sqrt{\frac{m}{2i\pi\hbar t}} e^{i\frac{1}{\hbar} S_{\text{class}}(x, x_0; t)} \quad \forall t \quad (1.270)$$

Il est recommandé de s'exercer au maniement de U en choisissant une forme spécifique pour l'état initial (gaussienne, lorentzienne, etc.). Une fois connu le propagateur pour un problème donné, le calcul de $\Psi(x, t)$ se réduit à une simple intégration, conformément à (1.253).

Passons maintenant à la description du mouvement selon Heisenberg. Soit un système isolé (son Hamiltonien ne dépend donc pas du temps⁸⁵) ; dans ces conditions, l'opérateur d'évolution est donné par (1.232) : en raison de l'invariance par translation dans le temps, U ne dépend que de la différence des temps $t - t_0$ et, pour simplifier les notations, on pose désormais :

$$U(t - t_0) \equiv \exp \left[\frac{1}{i\hbar} H(t - t_0) \right] \quad (1.271)$$

En outre, comme seul compte l'intervalle de temps écoulé entre l'instant initial et l'instant considéré, on peut toujours poser conventionnellement $t_0 = 0$. Soit maintenant une observable A indépendante du temps. Sa valeur moyenne dans l'état $|\Psi(t)\rangle$ issu de $|\Psi(0)\rangle$ est :

$$\langle A \rangle(t) = \langle U(t)\Psi(0) | A | U(t)\Psi(0) \rangle = \langle \Psi(0) | U^\dagger(t) A U(t) | \Psi(0) \rangle , \quad (1.272)$$

de sorte que $\langle A \rangle(t)$ peut tout autant se calculer en prenant la valeur moyenne de l'association $U^\dagger(t) A U(t)$ sur l'état initial prescrit. Ceci conduit à donner un nom à ce groupement remarquable ; on note ainsi :

$$A_H(t) = U^\dagger(t) A U(t) , \quad (1.273)$$

d'où la règle de calcul de la valeur moyenne de A à l'instant t :

$$\langle A \rangle(t) = \langle \Psi(0) | A_H(t) | \Psi(0) \rangle ; \quad (1.274)$$

ainsi, la valeur moyenne de A à l'instant t s'obtient en prenant la valeur moyenne de $A_H(t)$ sur l'état *initial*. $A_H(t)$ est le représentant selon Heisenberg⁸⁶ de l'opérateur A , ce dernier étant alors appelé, par opposition, représentant de Schrödinger. Comme $U(0) = U^\dagger(0) = \mathbf{1}$, on a :

$$A_H(0) = A . \quad (1.275)$$

⁸⁴Indépendamment du préfacteur contenant la constante de Planck, il est remarquable que la fonction action *classique* S_{class} apparaisse aussi simplement dans l'expression du propagateur. Il n'en va pas toujours ainsi, mais c'est vrai pour tous les Lagrangiens au plus quadratiques en \vec{p} et \vec{r} (par exemple : oscillateur harmonique, particule accélérée par un champ constant).

⁸⁵C'est seulement dans ce cas que l'approche de Heisenberg est utile.

⁸⁶On dit aussi que $A_H(t)$ est l'observable A dans le "point de vue" de Heisenberg, A étant la même observable dans le "point de vue" de Schrödinger ; les deux représentants coïncident à $t = 0$: $A_H(t = 0) = A$.

Ainsi, à l'instant zéro, les deux représentants coïncident. Par ailleurs, pour une fonction quelconque développable en série entière, on a :

$$[f(A)]_{\text{H}} = f(A_{\text{H}}) . \quad (1.276)$$

Le passage de A à A_{H} réalise ce que l'on appelle une transformation canonique qui, par définition, conserve la forme des équations fondamentales. En effet, soit par exemple une relation de commutation :

$$[X, Y] = C . \quad (1.277)$$

Il est facile de voir que les représentants de Heisenberg ont entre eux la même relation :

$$[X, Y]_{\text{H}} = [X_{\text{H}}, Y_{\text{H}}] = C_{\text{H}} . \quad (1.278)$$

Ainsi, toute relation de commutation s'exprime à l'aide des opérateurs transformés comme la relation de commutation avant transformation : cette relation est donc invariante en forme.

Ceci étant établi, il est maintenant possible d'écrire les équations du mouvement pour une observable quelconque A , toujours supposée indépendante du temps, dans la "représentation" de Heisenberg. Par la définition (1.273) :

$$\frac{d}{dt} A_{\text{H}}(t) = \left[\frac{d}{dt} U^{\dagger}(t) \right] A U(t) + U^{\dagger}(t) A \frac{d}{dt} U(t) . \quad (1.279)$$

D'après (1.225), U et U^{\dagger} satisfont :

$$i\hbar \frac{\partial}{\partial t} U(t) = H U(t) , \quad -i\hbar \frac{\partial}{\partial t} U^{\dagger}(t) = U^{\dagger}(t) H^{\dagger} = H^{\dagger} U^{\dagger}(t) . \quad (1.280)$$

où a été utilisé le fait que H , U et U^{\dagger} commutent tous deux à deux ; l'expression (1.279) de la dérivée devient :

$$\frac{d}{dt} A_{\text{H}}(t) = \frac{1}{i\hbar} [A_{\text{H}}(t), H] . \quad (1.281)$$

Comme H et U commutent :

$$H_{\text{H}}(t) = H \quad (1.282)$$

et l'on obtient finalement l'équation de Heisenberg pour l'observable A :

$$\frac{d}{dt} A_{\text{H}} = \frac{1}{i\hbar} [A, H]_{\text{H}} . \quad (1.283)$$

En pratique l'équation de Heisenberg pour une observable quelconque s'obtient rapidement en calculant le commutateur de A avec H (en représentation de Schrödinger) et en affectant l'indice H au résultat. Usuellement, une fois les idées bien mises en place, on omet l'indice H pour alléger l'écriture.

À ce stade, cela vaut la peine de comparer (1.283) à l'équation correspondante que l'on obtient en Mécanique Classique. Soit a une grandeur dynamique⁸⁷ ; sa dérivée totale est :

$$\frac{d}{dt} a(q(t), p(t)) = \frac{\partial a}{\partial q} \dot{q} + \frac{\partial a}{\partial p} \dot{p} . \quad (1.284)$$

Le long d'une trajectoire réelle, les équations de Hamilton sont satisfaites ; il en résulte alors, utilisant la notation en crochet de Poisson :

$$\frac{d}{dt} a(q(t), p(t)) = \{a, H\} . \quad (1.285)$$

La comparaison de cette équation avec l'équation du mouvement de Heisenberg (1.283) fait à nouveau émerger la correspondance (1.191) entre crochet de Poisson classique et commutateur quantique.

Dans le cas où l'observable A dépend du temps avant le passage en Heisenberg, la dérivation effectuée comme en (1.279) introduit la moyenne sur l'état initial de $U^{\dagger}(t)(\partial A/\partial t)U(t)$, que l'on note $(\partial A_{\text{H}}/\partial t)$. L'équation du mouvement de Heisenberg prend alors la forme :

$$\frac{d}{dt} A_{\text{H}} = \frac{\partial A_{\text{H}}}{\partial t} + \frac{1}{i\hbar} [A, H]_{\text{H}} , \quad \text{avec} : \quad \frac{\partial A_{\text{H}}}{\partial t} \equiv U^{\dagger}(t) \frac{\partial A}{\partial t} U(t) . \quad (1.286)$$

⁸⁷dont on suppose d'abord, pour simplifier et pour faciliter la comparaison, qu'elle ne dépend du temps que *via* $q(t)$ et $p(t)$.

L'utilisation de la notation dérivée partielle rappelle que cette dérivée est prise sans tenir compte du "mouvement" décrit par $U(t)$. Une fois écrites les équations de Heisenberg, on peut en principe les intégrer et donc obtenir par elles l'expression explicite des quantités intéressantes. En prenant alors la valeur moyenne sur l'état *initial* prescrit, on en déduit immédiatement les valeurs moyennes à l'instant t des observables cherchées.

Les équations de Heisenberg conduisent naturellement à s'intéresser aux observables qui commutent avec H , notées généralement C . Pour une telle observable, par définition :

$$[C, H] = 0 \quad (1.287)$$

de sorte que C et son image de Heisenberg, C_H , coïncident ; l'équation du mouvement de C_H est donc :

$$\frac{d}{dt} C_H = 0 \iff \frac{d}{dt} \langle C \rangle = 0 \quad \forall |\Psi(t)\rangle . \quad (1.288)$$

En d'autres termes, la valeur moyenne de C dans n'importe quel état ne varie pas au cours du temps : c'est bien une constante du mouvement. Comme on s'en doute, les constantes du mouvement jouent un rôle essentiel : comme elles commutent avec H , on peut trouver un jeu de vecteurs propres communs à ces constantes et à H . En outre, à toute constante du mouvement est associée une propriété de symétrie – d'où inversement le rôle majeur de la symétrie pour l'analyse d'un problème donné.

Des exemples permettent de montrer explicitement la puissance et l'efficacité des équations de Heisenberg. Même dans le cas simple d'une particule libre à une dimension, on sait que le calcul explicite de $|\Psi(t)\rangle$ est un peu laborieux ; une fois ceci fait, il reste encore à calculer les valeurs moyennes avec $|\Psi(t)\rangle$ ainsi obtenu. En revanche, ce problème se résout facilement à l'aide des équations de Heisenberg. On obtient, en rétablissant t_0 comme instant initial :

$$\langle x \rangle(t) = \frac{t - t_0}{m} \langle p \rangle(t_0) + \langle x \rangle(t_0) . \quad (1.289)$$

Pour ces valeurs moyennes, on retrouve formellement la même loi du mouvement qu'en Mécanique Classique. Si la fonction d'onde initiale est réelle, l'écart quadratique $\Delta x^2(t)$ est :

$$\Delta x^2(t) \equiv \langle x^2 \rangle(t) - [\langle x \rangle(t)]^2 = \frac{(t - t_0)^2}{m^2} \Delta p^2(t_0) + \Delta x^2(t_0) . \quad (1.290)$$

Comme le produit de $\Delta x^2 \Delta p^2$ est toujours borné inférieurement par $\hbar^2/4$, l'étalement du paquet d'ondes est d'autant plus rapide qu'il est initialement bien localisé spatialement.

L'oscillateur harmonique est un autre exemple montrant l'intérêt des équations du mouvement de Heisenberg⁸⁸ ; après intégration en temps de ces équations, on trouve ($t_0 = 0$) :

$$x_H(t) = x \cos \omega t + \frac{p}{m\omega} \sin \omega t , \quad (1.291)$$

$$p_H(t) = p \cos \omega t + m\omega x \sin \omega t . \quad (1.292)$$

Les moyennes à l'instant t sont donc :

$$\langle x \rangle(t) = \langle x \rangle \cos \omega t + \left\langle \frac{p}{m\omega} \right\rangle \sin \omega t , \quad (1.293)$$

$$\langle p \rangle(t) = \langle p \rangle \cos \omega t + \langle m\omega x \rangle \sin \omega t , \quad (1.294)$$

toutes les valeurs moyennes à droite étant calculées sur l'état initial à $t = 0$. Le mouvement décrit par ces moyennes est strictement identique à celui des grandeurs classiques correspondantes ; la raison de ceci est la linéarité des équations du mouvement, linéarité liée au fait que l'énergie potentielle est quadratique par rapport à la coordonnée. En général, la moyenne de $\vec{\nabla}V(\vec{r})$, $\langle \vec{\nabla}V(\vec{r}) \rangle$, étant différente de $\vec{\nabla}V(\langle \vec{r} \rangle)$, les valeurs moyennes ne suivent pas les lois classiques : le théorème d'Ehrenfest ([4], p. 242) n'affirme pas que le centre d'un paquet d'ondes suit strictement la trajectoire classique ; c'est le cas – approximativement – seulement si le potentiel V varie *lentement* à l'échelle du paquet d'ondes.

⁸⁸Le calcul explicite du propagateur, qui serait nécessaire dans l'image de Schrödinger, est encore plus laborieux que pour la particule libre, compte tenu de la complexité des fonctions propres (polynômes de Hermite tempérés par des gaussiennes).

Un dernier exemple est intéressant : c'est celui d'une particule de charge q dans un champ électromagnétique \vec{E}, \vec{B} . On sait que si \vec{A} et ϕ désignent les potentiels vecteur et scalaire, le Hamiltonien classique est :

$$H_{\text{class}}(\vec{r}, \vec{p}, t) = \frac{1}{2m} [p - q\vec{A}(\vec{r}, t)]^2 + q\phi(\vec{r}) . \quad (1.295)$$

Comme le potentiel-vecteur dépend du rayon-vecteur \vec{r} , il faut veiller à respecter l'ordre des produits où l'impulsion figure également. Un calcul un peu long⁸⁹ permet d'écrire alors d'écrire l'équation du mouvement pour position et impulsion et d'en déduire la version quantique de l'équation de Newton :

$$m \frac{d^2 \tilde{r}_H}{dt^2} = q\vec{E} + \frac{q}{2} \left(\frac{d\tilde{r}_H}{dt} \times \vec{B} - \vec{B} \times \frac{d\tilde{r}_H}{dt} \right) . \quad (1.296)$$

où l'on reconnaît à droite l'expression symétrisée⁹⁰ de la force de Lorentz.

Terminons cette discussion sur l'évolution en temps par une remarque. Il est important de noter que les deux postulats 5 (réduction du paquet d'ondes) et 6 (évolution temporelle) sont distincts et non réductibles l'un à l'autre. En effet, l'évolution par l'équation de Schrödinger est une évolution unitaire qui conserve la norme du vecteur d'état ; c'est une nécessité, requise pour la conservation de la probabilité. À l'inverse, la réduction du paquet d'ondes est non-unitaire, puisque l'état après la mesure est la projection de l'état juste avant la mesure, opération qui manifestement modifie la norme du vecteur. Penrose [10] discute longuement ces deux types d'évolution, qu'il appelle respectivement **U** et **R**, pour des raisons évidentes.

Fondamentalement, on peut dire que l'évolution **U** est complètement déterministe ; c'est au contraire par l'opération **R** qu'éclate l'imprévisibilité du comportement des particules, ou plutôt sa seule prévisibilité en terme de probabilités. Le non-déterminisme de la description quantique – il est visiblement nécessaire puisque les expériences cruciales mettent en évidence la non-reproductibilité d'un événement élémentaire (l'arrivée d'un électron sur l'écran dans l'expérience des fentes d'Young, *etc*) – ce non-déterminisme est situé exclusivement dans le processus de réduction du paquet d'ondes.

La distinction entre **U** et **R** mériterait d'ailleurs d'être approfondie. En effet, l'unitarité est une conséquence de la description globale du système considéré et, au moins sur le plan conceptuel, on ne voit aucune raison de ne pas traiter comme un "super-système" le système objet de la mesure et l'appareil de mesure lui-même. Dans ces conditions, la non-unitarité provient du fait que l'on examine un sous-système du système total, tout comme, si l'énergie totale d'un système isolé est constante, l'énergie d'une partie de ce système ne l'est pas en général⁹¹.

⁸⁹ Voir [18], p. 178.

⁹⁰ Classiquement, les deux produits vectoriels $(d\vec{r})(dt) \times \vec{B}$ et $-\vec{B} \times (d\vec{r})(dt)$ sont bien égaux ; l'expression quantique s'obtient en faisant la *somme* des termes classiques tous égaux entre eux puisque ne différant que par l'ordre de facteurs commutant tous deux à deux. Dans l'expression quantique symétrisée, c'est bien au total un signe – qui ressort.

⁹¹ Un autre exemple : l'équation de Liouville est une équation invariante par renversement du temps puisqu'elle repose sur des équations *mécaniques*. La densité $\rho(q, p, t)$ est une fonction, définie dans l'espace des phases, dont le "mouvement" est réversible. Il n'empêche que si l'on s'intéresse seulement à certains degrés de liberté – la coordonnée par exemple – il apparaît spontanément une évolution irréversible. C'est le cas lorsqu'il existe une incertitude sur les conditions initiales : l'incertitude qui en résulte ne fait que croître au cours du temps, signature d'une évolution non-réversible. Un autre exemple est fourni par l'instabilité de Landau dans les plasmas, dont la description repose sur l'équation de Vlasov, équation pourtant réversible (absence d'intégrale de collisions).

Chapitre 2

Symétrie et lois de conservation

Le but de ce chapitre est de montrer explicitement le lien qui existe entre la symétrie, le plus souvent géométrique, et les lois de conservation qui en découlent. Ce lien sera établi d'abord en Mécanique Classique, puis en Mécanique Quantique afin de bien mettre en évidence les parallèles entre ces deux théories.

2.1 Le principe euclidien de Relativité et le rôle de la symétrie en Physique

L'étude des phénomènes physiques requiert l'usage d'un système de référence (repère). D'un repère à l'autre, les lois déterminant le mouvement n'ont pas en général la même forme (penser à un repère accéléré, rectilignement ou en rotation).

L'hypothèse fondamentale de Galilée¹ est l'existence de repères vis-à-vis desquels l'espace ordinaire (physique) est isotrope et homogène et où le temps est le même : un système isolé et fermé étant défini, ses propriétés ne changent pas si on le déplace d'un point à l'autre² ou si on le fait tourner sur lui-même ; si les propriétés ne changent pas, aucun résultat d'expérience ne peut être affecté par une telle opération. Une théorie physique visant à décrire les propriétés des systèmes doit donc posséder des propriétés d'invariance : les lois qu'elle énonce doivent être les mêmes dans tout repère galiléen. Le principe de Relativité nie l'existence d'une position ou d'une orientation privilégiée de l'espace dans lequel sont plongés les objets que la Physique prétend étudier et décrire, et affirme l'universalité du temps. Un corollaire de ce principe est qu'il n'existe pas de repère "absolu" à privilégier, et que toutes les lois de la Nature sont les mêmes dans tous les référentiels galiléens.

Bien évidemment, ce principe s'étend au cas d'un système non isolé : par exemple, pour un système soumis à un champ de forces variable ou non dans le temps, les propriétés ne changent pas si le champ de forces est déplacé ou tourné en même temps que le système pré-défini. En réalité, ce dernier et le seul champ de forces supposé exister constituent un système isolé et fermé auquel le principe de relativité s'applique.

En Physique, on nomme habituellement "le système" un ensemble d'objets sur lesquels le "reste du monde" (environnement) peut agir ; à l'inverse, cet ensemble n'a pas de rétroaction sur l'environnement. Exemple : quand une particule est dite soumise à un champ (électrique, magnétique, électromagnétique), ceci veut dire que les sources de ce champ (condensateur, aimant, courants variables dans le temps) sont supposées avoir leur dynamique propre définie en soi, quoi que fasse la particule – qui constitue alors le système physique d'étude. Il est clair que ceci, dans l'absolu, est toujours une approximation que l'on doit justifier.

¹Voir [17], p. 371. Merzbacher emploie l'expression *Euclidean Principle of Relativity*.

²Cette affirmation contient implicitement l'hypothèse selon laquelle l'espace est infini dans toutes les directions.

Par exemple, construire la Mécanique Quantique dans le cadre du principe de Relativité a un sens, même vis-à-vis des expériences faites sur Terre ; en effet, il y existe évidemment une direction privilégiée (le sens du champ de force de gravitation), mais ses effets, comparés à ceux qui sont pertinents pour la dynamique des petits objets (particules, atomes, molécules, etc.), sont totalement négligeables, sauf cas très exceptionnel³.

L'affirmation contenue dans le principe de relativité de Galilée ne requiert par nature aucune démonstration : ériger en principe une affirmation, c'est s'en remettre pour la preuve de sa véracité à la confrontation entre les résultats théoriques qui en découlent et les observations expérimentales. Aucune violation n'ayant été observée jusqu'à présent, ce principe peut (et doit) être invoqué dans la construction de toute théorie physique non-relativiste au sens d'Einstein.

À titre d'exemple, montrons comment le principe de Relativité permet de former le Lagrangien L d'une particule libre⁴. L'uniformité (homogénéité de l'espace) permet d'affirmer que L ne dépend d'aucune coordonnée spatiale, puisque toutes les positions sont équivalentes. De la même façon, L ne peut dépendre du temps puisque tous les instants se valent. Il en résulte que L est au plus une fonction de la vitesse. Utilisons maintenant l'isotropie de l'espace : comme toutes les directions sont équivalentes, L ne peut dépendre de l'orientation du vecteur vitesse : seul importe le module de ce vecteur. Celui-ci pouvant toujours s'exprimer en fonction du carré v^2 , on en arrive à la conclusion que le Lagrangien d'une particule libre, tel qu'il résulte du principe de Galilée, est nécessairement une certaine fonction du carré de la vitesse :

$$L = f(v^2) , \quad (2.1)$$

où la fonction f est pour l'instant inconnue et doit être trouvée.

Le Lagrangien permet d'écrire l'intégrale donnant l'action, S , qui est extrémale (minimale, c'est le Principe de Moindre Action) pour la trajectoire réellement suivie :

$$S = \int_{t_1}^{t_2} L dt , \quad \delta S = 0 \text{ pour la trajectoire réellement suivie .} \quad (2.2)$$

Le principe de Galilée permet de trouver la fonction L , au moins pour une particule libre. Soit un point matériel (particule) dont les coordonnées dans deux repères galiléens R et R' sont \vec{r} et \vec{r}' , R' étant animé d'une vitesse \vec{V} par rapport à R ; on a :

$$\vec{r}' = \vec{r} - \vec{V}t , \quad \vec{v}' = \vec{v} - \vec{V} . \quad (2.3)$$

Dans chacun de ces repères, le mouvement est décrit par les équations de Lagrange déduites du principe de moindre action. Soit L et L' les deux Lagrangiens manipulés par deux observateurs appartenant l'un à R , l'autre à R' . Comme pour ces deux observateurs les lois du mouvement doivent être les mêmes, les deux actions S et S' ne doivent différer que d'une constante ; ce sera le cas si $L - L'$ est une dérivée totale en temps, dont l'intégration conduisant à l'action donnera un terme constant. Plus précisément, les deux observateurs vont respectivement écrire :

$$\delta \int_{t_1}^{t_2} L dt = 0 , \quad \delta \int_{t_1}^{t_2} L' dt = 0 . \quad (2.4)$$

Le premier observateur en déduira une loi exprimée par :

$$\vec{r} = g(t; \vec{r}_0, \vec{v}_0) , \quad (2.5)$$

où g est une fonction du temps paramétrée par les conditions initiales. Le deuxième observateur trouvera la *même* loi, exprimée par la *même* fonction g mais paramétrée par les conditions initiales mesurées dans R' :

$$\vec{r}' = g(t; \vec{r}'_0, \vec{v}'_0) . \quad (2.6)$$

Le temps est le même dans les deux référentiels. Compte tenu du fait que les équations du mouvement s'obtiennent par variation de L , qui est une intégrale sur le temps, les lois du mouvement (traduites par la

³La description très précise de la dynamique de particules ultra-froides (neutrons, atomes, etc) peut exiger l'incorporation semi-classique de la gravitation terrestre.

⁴voir [19], p. 12.

fonction g) seront les mêmes dans R et R' pourvu que L et L' diffèrent seulement par une dérivée totale en temps.

Partant de :

$$L = f(v^2) , \quad L' = f(v'^2) \quad (2.7)$$

et prenant en compte (2.3), L' s'écrit :

$$L' = f(v^2 - 2\vec{v} \cdot \vec{V} + V^2) . \quad (2.8)$$

Supposons que la vitesse \vec{V} est un infiniment petit. En développant le second membre, on obtient L' à des infiniment petits d'ordre 2 près :

$$L' = f(v^2) - 2\vec{v} \cdot \vec{V} \frac{\partial f}{\partial v^2} \equiv L - 2\vec{v} \cdot \vec{V} \frac{\partial f}{\partial v^2} . \quad (2.9)$$

Le terme additionnel doit être de la forme $(d/dt)\bullet$ pour que les lois trouvées dans R et R' , déduites de L et de L' , coïncident. Comme $\vec{v} \equiv (d/dt)\vec{r}$ est déjà une dérivée totale, ceci est assuré si $\partial f/\partial v^2$ est une constante ; il en résulte que f est un monôme du genre av^2 où a désigne une constante :

$$f(v^2) = av^2 \iff L = av^2 . \quad (2.10)$$

Comme on le sait, la constante a n'est autre que la moitié de la masse m de la particule – ceci afin de retrouver l'équation de Newton. Une fois obtenue cette dépendance pour une transformation infinitésimale, il est facile de voir que le Lagrangien $L = av^2$ est également invariant⁵ pour une transformation finie :

$$L' = av'^2 = a(\vec{v} - \vec{V})^2 = a(v^2 - 2\vec{v} \cdot \vec{V} + V^2) . \quad (2.11)$$

Ceci peut s'écrire :

$$L' = av^2 + a \frac{d}{dt} (-2\vec{r} \cdot \vec{V} + V^2 t) \equiv L + \frac{d}{dt} \phi . \quad (2.12)$$

L et L' donnent les mêmes équations du mouvement dans les deux repères.

Pour une particule soumise à un champ (de gravitation, électrostatique, *etc.*) dérivant d'un potentiel U , le Lagrangien s'écrit :

$$L = T - V , \quad (2.13)$$

où V est l'énergie potentielle, différant simplement du potentiel U par un scalaire multiplicatif (masse pour le champ de gravitation, charge pour le champ électrostatique, *etc.*)⁶. C'est ce choix qui assure que les équations de Lagrange déduites du principe variationnel $\delta S = 0$ reproduisent les équations de la Mécanique dans sa formulation élémentaire.

L'énergie potentielle V est *a priori* une fonction des trois coordonnées d'espace x , y et z , ou de tout autre jeu de coordonnées fixant sans ambiguïté la position de la particule dans l'espace. V est la représentation désincarnée physiquement de l'action de l'extérieur sur la particule (celle-ci étant le système d'étude) : on a remplacé tous les autres systèmes en interaction avec la particule par un champ de forces "éthérées" ; ces forces agissent sur la particule, qui n'a en retour aucune influence sur elles. C'est ainsi que s'introduit la notion de champ en Physique.

La symétrie proclamée par le principe de Galilée sur l'espace où plongent les systèmes physiques est une symétrie universelle. De surcroît, un système donné, dans un environnement donné, peut posséder des symétries internes propres. Il convient de bien distinguer l'invariance universelle résultant des hypothèses d'isotropie et d'homogénéité de l'espace, de l'invariance, pour un système donné, résultant de sa symétrie propre. L'invariance euclidienne affirme que les propriétés de tout système sont indifférentes à la position ou à l'orientation du système, en conséquence de l'homogénéité et de l'isotropie de l'espace où il est plongé. Maintenant, pour un système

⁵Au sens : produit des équations du mouvement invariante.

⁶La formulation de la Mécanique Quantique repose sur le formalisme hamiltonien habituel, où il est possible de définir une énergie potentielle V . On peut toujours se trouver dans cette situation en élargissant suffisamment le système à quantifier, quitte à prendre des traces partielles par la suite si seule une partie des degrés de liberté est digne d'intérêt. En l'absence d'énergie potentielle, on ne sait pas quantifier un système.

donné, certains de ses éléments peuvent être indifférents à telle ou telle opération et c'est cette indifférence qui constitue la symétrie interne *propre*. Affirmer que les propriétés ne changent pas, c'est prendre en compte à la fois le principe euclidien et la symétrie interne propre : la seconde repose sur le premier.

Par exemple, au voisinage de la surface terrestre, comptant positivement vers le haut l'altitude z , le champ de gravitation est $U = gz$, l'énergie potentielle est $V = mgz$. La fonction $V(z)$ a une propriété remarquable : elle ne dépend que de la variable verticale z et non des deux coordonnées horizontales x et y (physiquement : tout déplacement horizontal ne donne lieu à aucun travail de la force de pesanteur). Il en résulte que les propriétés d'équilibre ne dépendent que de z : l'équilibre dans un plan perpendiculaire à Oz est indifférent aux valeurs de x et de y . Ces deux coordonnées absentes de l'énergie potentielle sont appelées variables *cycliques*⁷ et entraînent la conservation de l'impulsion dans un plan horizontal (p_x et p_y sont des constantes du mouvement). La symétrie correspondante est une symétrie de translation parallèlement au plan xOy : l'image de tout point d'équilibre dans une telle translation est aussi un point d'équilibre. Ceci constitue un exemple de propriété invariante dans une opération géométrique : c'est l'une des symétries du système. La situation est identique pour une particule de charge q plongée dans un champ électrique extérieur $\vec{\mathcal{E}}$. Le Lagrangien s'écrit alors :

$$L = \frac{1}{2} m\vec{v}^2 - (-q\vec{\mathcal{E}} \cdot \vec{r}) . \quad (2.14)$$

Le champ définit une direction remarquable : il est tout indiqué, pour la commodité, de choisir l'un des axes du repère (cartésien) parallèle à ce champ, soit Oz pour fixer les idées. Il vient alors :

$$L = \frac{1}{2} m\vec{v}^2 + q\mathcal{E}z \quad (2.15)$$

et on se retrouve exactement dans la même situation que pour la pesanteur, étudiée localement à la surface terrestre – ou plus généralement au voisinage de tout point dont la distance au centre de la Terre est donnée une fois pour toutes.

Autre exemple : l'atome d'hydrogène dans l'approximation du noyau infiniment massif. Le système physique d'étude se réduit à une seule particule (l'électron, de masse m et de charge $e < 0$) soumise à la force de Coulomb. Le Lagrangien est alors :

$$L = \frac{1}{2} m\vec{v}^2 - \left(-\frac{e'^2}{r}\right) \quad \left(e'^2 = \frac{e^2}{4\pi\epsilon_0}\right) . \quad (2.16)$$

Ici, $V(r) = -e'^2/r$ a la propriété remarquable de ne dépendre que de la distance électron – noyau, en aucune façon de la *direction* du rayon vecteur joignant les deux particules. Il en résulte une symétrie interne sphérique : toute rotation autour d'un axe passant par le noyau est une opération blanche, qui ne change en rien l'état du système. L'électron se moque de savoir ce que le physicien définit comme étant les trois axes d'un repère cartésien. Ici d'ailleurs, les coordonnées sphériques s'imposent, toujours pour des raisons de commodité ; il en faut trois : r et les deux angles θ (latitude) et ϕ (azimut) définis comme d'habitude. Maintenant, les deux variables cycliques sont ces deux angles : ils n'apparaissent pas dans le Lagrangien. Il s'agit encore d'une symétrie qui, comme toujours, produira une loi de conservation : ici, celle du moment cinétique orbital de l'électron.

Supposons maintenant que l'on impose de l'extérieur un champ électrique, définissant toujours l'axe Oz . Le Lagrangien contient maintenant un terme de plus :

$$L = \frac{1}{2} m\vec{v}^2 + \frac{e'^2}{r} + e\mathcal{E}z \quad (2.17)$$

et, visiblement, la symétrie est abaissée par l'adjonction du champ externe : il n'y a maintenant invariance que par rapport aux rotations autour de l'axe Oz – et également par rapport aux réflexions par rapport à tout plan contenant ce même axe. On dit que le champ a *brisé* la symétrie sphérique, tout en préservant la symétrie cylindrique et la symétrie miroir ; au passage, on notera que le phénomène de brisure de symétrie fonctionne par tout ou rien : un champ infinitésimal fait changer *qualitativement* la symétrie. Il y a maintenant moins de

⁷L'origine de cette terminologie tient au fait que, très souvent, les variables absentes de L (ou de H) sont des angles, définis à 2π près.

variables cycliques que précédemment. Avec les coordonnées cylindriques ρ , ϕ et z , qui sont appropriées à ce problème, on a :

$$L = \frac{1}{2} m \vec{v}^2 + \frac{e'^2}{\sqrt{\rho^2 + z^2}} + e \mathcal{E} z \quad (2.18)$$

et il n'y a plus qu'une seule variable cyclique, l'angle azimutal ϕ . À cette variable absente de L sera associée la conservation de la projection le long de Oz du moment cinétique. Notons que, grâce à la symétrie de réflexion, les deux mouvements qui se correspondent dans cette opération sont physiquement équivalents ; pour deux tels mouvements, l'électron tourne dans un sens ou dans le sens contraire autour de Oz . Les propriétés physiques ne dépendront donc que du module de cette composante, pas de son signe⁸.

Dernier exemple, important : un système isolé, possédant par nature la symétrie de translation dans le temps, ni L ni H ne contiennent explicitement le temps. Alors, c'est l'énergie qui est une constante du mouvement.

En résumé, toute propriété de symétrie interne d'un système⁹ se traduit techniquement par l'absence dans le Lagrangien de certaines coordonnées, pourtant nécessaires par ailleurs pour fixer la position du système dans l'espace : ce sont ces variables absentes qui sont appelées variables cycliques et c'est cette absence qui produit automatiquement les lois de conservation des moments conjugués correspondants. En effet, dans les notations traditionnelles, les équations de Lagrange sont :

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad \iff \quad \frac{d}{dt} p_i = \frac{\partial L}{\partial q_i} \quad (2.19)$$

Imaginons que l'une des coordonnées, q_{i_0} , n'apparaît pas dans L ; la dérivée correspondante $\partial L / \partial q_{i_0}$ est donc identiquement nulle ; par (2.19), il en résulte immédiatement que $p_{i_0} = C^{\text{ste}}$.

2.2 Symétrie spatiale continue : translations et rotations

Une symétrie spatiale *continue* se caractérise par le fait que l'on peut lui associer des transformations *infinitésimales*, arbitrairement proches de la transformation identité. Sur un strict plan géométrique, deux types de transformations jouent un rôle de tout premier plan en Mécanique (classique ou quantique)¹⁰, les translations et les rotations. À ces symétries continues s'opposent les symétries discrètes (la réflexion, le renversement du temps, ...), pour lesquelles on ne peut définir d'opérations infinitésimales¹¹. Les translations et les rotations appartiennent à la classe des *déplacements*, opérations géométriques qui ne changent ni les longueurs, ni les angles, ni la chiralité.

2.2.1 Translations

Un système étant donné, la symétrie d'invariance par translation dans l'espace se traduit par le fait que certaines coordonnées rectangulaires (x , y ou z) sont absentes dans l'expression du Lagrangien et donc du Hamiltonien. Pour fixer les idées, le système de référence pour l'exemple sera une particule dans un champ constant (gravité, champ électrique, ...), possédant l'énergie potentielle $V(z) = -Kz$. Le Lagrangien est :

$$L = \frac{1}{2} m \vec{v}^2 - V(z) . \quad (2.20)$$

⁸Il en va tout autrement avec un champ *magnétique* ; c'est pourquoi la levée de dégénérescence est seulement partielle pour l'effet Stark (champ électrique), alors qu'elle est totale pour l'effet Zeeman (champ magnétique). Physiquement la différence vient du fait que le couplage magnétique implique une *vitesse* (la brique élémentaire du magnétisme est une boucle de courant), qui change de signe quand on effectue la symétrie miroir.

⁹On doit à Emmy Noether d'avoir reconnu pour la première fois le lien entre symétrie et constantes du mouvement, dont l'énoncé constitue le *Théorème de Noether*.

¹⁰Récemment, une autre symétrie a révélé son importance : la similitude, ou symétrie affine (dilatation ou contraction des longueurs). Cette symétrie donne lieu à l'invariance d'échelle sur laquelle est construite la zoologie des structures fractales.

¹¹En outre, la réflexion ne respecte pas la chiralité, au contraire de toute rotation, même discrète.

Les moments conjugués $p_u = \partial L / \partial \dot{q}_u$, ($u = x, y, z$) sont égaux à mv_u . Les équations de Lagrange donnent :

$$\frac{d}{dt}(mv_u) = \frac{\partial L}{\partial u} = -\frac{\partial V}{\partial u} \equiv F_u . \quad (2.21)$$

Ces équations représentent l'équation fondamentale de la dynamique projetée sur les trois axes. Comme V ne dépend que de z :

$$\frac{d}{dt}(mv_x) = 0 \quad \frac{d}{dt}(mv_y) = 0 . \quad (2.22)$$

Ainsi, le caractère cyclique des deux coordonnées x et y assure *ipso facto* la conservation des projections correspondantes de la quantité de mouvement.

Évidemment, tout ceci se retrouve dans la formulation hamiltonienne. Le Hamiltonien se construit comme d'habitude. Après avoir trouvé les moments conjugués p_i , on fabrique la combinaison¹² :

$$H = \sum_i p_i \dot{q}_i - L \quad (2.23)$$

que l'on exprime enfin à l'aide des p_i et des q_i . On trouve ainsi facilement :

$$H = \frac{\vec{p}^2}{2m} + V(z) . \quad (2.24)$$

H est la somme de l'énergie cinétique et de l'énergie potentielle : c'est l'énergie totale du système dans le champ considéré, qui est indépendante du temps pour un système isolé (conservation de l'énergie assurant l'existence de l'intégrale première $H(q_i, p_i) = E$, constante dont la valeur est fixée une fois pour toutes par les conditions initiales prescrites). La conservation de p_x et de p_y découle immédiatement du fait que L et H ne dépendent ni de x ni de y ; soit par (2.22), soit à partir des équations de Hamilton :

$$\dot{q}_i = \frac{\partial H}{\partial p_i} , \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} ; \quad (2.25)$$

on déduit alors de (2.24) que p_x et p_y sont des constantes le long de la trajectoire réellement suivie.

La notion de variables cycliques se transpose en Mécanique Quantique. En effet, il suffit de revenir aux équations de Heisenberg :

$$i\hbar \frac{dA_H}{dt} = [A_H, H] \equiv [A, H]_H \quad (2.26)$$

et de choisir pour A les coordonnées et les moments conjugués. Pour une particule dans \mathbb{R}^3 on obtient ainsi les six équations :

$$i\hbar \frac{du_H}{dt} = [u, \frac{p_u^2}{2m}]_H , \quad i\hbar \frac{dp_{u,H}}{dt} = [p_u, V(\vec{r})]_H \quad (u = x, y, z) \quad (2.27)$$

qui, après calcul, donnent :

$$\dot{u}_H = \frac{p_{u,H}}{m} , \quad \dot{p}_{u,H} = -\left(\frac{\partial V}{\partial u}\right)_H . \quad (u = x, y, z) \quad (2.28)$$

Ces équations du mouvement quantique sont en tout point analogues formellement à leurs homologues classiques, mais, bien sûr, impliquent des opérateurs. Supposons maintenant que certaines variables soient absentes de H ; dans l'exemple ci-dessus, mais maintenant traité quantiquement – donc décrit par le Hamiltonien (2.24) où les variables classiques ont été remplacées par leurs opérateurs –, ceci signifie que les opérateurs x et y ne figurent pas dans H ; les commutateurs de x et de y avec H sont alors nuls, et on obtient à nouveau $\dot{p}_{x,H} = \dot{p}_{y,H} = 0$: $p_{x,H}$ et $p_{y,H}$ sont des constantes du mouvement au sens quantique : les opérateurs $p_{x,H}$ et $p_{y,H}$ coïncident avec p_x et p_y à tout instant, leurs valeurs moyennes calculées avec n'importe quel état $|\Psi(t)\rangle$ ne dépendent pas du temps.

Bien évidemment, l'invariance euclidienne a aussi des conséquences sur le formalisme quantique, qui vont se manifester de façon plus abstraite : toute prévision quantique doit satisfaire le principe euclidien,

¹²C'est une transformation de Legendre.

ce qui impose un mode de transformation bien déterminé pour la fonction d'onde et les observables. En toute généralité, l'invariance euclidienne contraint très fortement la structure géométrique de la Mécanique Quantique. De surcroît, toute symétrie interne propre se traduit sur les attributs quantiques d'un système donné ; pour l'exemple cité plus haut, on va voir que le Hamiltonien lui-même est invariant dans toute translation parallèle au plan xOy .

Le système quantique décrit par (2.24) possède intrinsèquement une symétrie d'invariance par translation perpendiculairement au champ. Ceci signifie que toute translation du système physique¹³ n'affecte en rien les prévisions effectuées à son propos : comme l'état du système ne change pas dans une telle translation, celui-ci se moque du choix qui est fait pour des raisons pratiques de calcul ; trouver des résultats différents avant et après translation serait une absurdité puisque cela reviendrait à dire qu'un système et le même système ont des propriétés différentes.

Considérons d'abord une translation *quelconque* du système, $\mathcal{T}_{\vec{a}}$, caractérisée par le vecteur \vec{a} . Dans cette hypothèse, la condition d'invariance explorée et utilisée sera celle qui découle du principe euclidien – au contraire, quand \vec{a} est perpendiculaire à l'axe Oz , c'est alors la symétrie propre du système qui doit ressortir, toujours assise sur l'invariance galiléenne universelle.

Dans une translation quelconque, un vecteur \vec{r} se transforme en un vecteur \vec{r}' :

$$\vec{r}' \equiv \mathcal{T}_{\vec{a}} \vec{r} = \vec{r} + \vec{a} . \quad (2.29)$$

Avant translation, la fonction d'onde est $\Psi(\vec{r}, t)$; après translation, c'est *a priori* une nouvelle fonction, $\Psi'(\vec{r}, t)$, le paramètre temps étant évidemment le même. Ces deux fonctions d'onde n'ont évidemment aucune raison de coïncider par elles-mêmes : en revanche, ce qui importe c'est l'identité des prévisions faites avec l'une et avec l'autre. Si elles ne coïncident pas, il doit toutefois exister un opérateur agissant dans l'espace des fonctions et transformant l'une en l'autre. Ceci étant admis, on peut donc écrire formellement, en termes de vecteurs, pour deux états quelconques $|\Psi\rangle$ et $|\Phi\rangle$:

$$|\Psi'\rangle = \mathcal{T}_{\vec{a}} |\Psi\rangle , \quad |\Phi'\rangle = \mathcal{T}_{\vec{a}} |\Phi\rangle , \quad (2.30)$$

$\mathcal{T}_{\vec{a}}$ étant l'opérateur agissant dans l'espace des états \mathcal{E} et représentant la translation $\mathcal{T}_{\vec{a}}$ effectuée dans l'espace physique. En Mécanique Quantique, toute prévision s'énonce en terme de probabilités, lesquelles s'expriment d'une façon générale par le module au carré du produit scalaire approprié. L'invariance requise par la symétrie euclidienne de translation impose que l'on ait, $\forall \vec{a}$:

$$(|\Psi'\rangle, |\Phi'\rangle)|^2 = (|\Psi\rangle, |\Phi\rangle)|^2 \iff (|\Psi'\rangle, |\Phi'\rangle) = (|\Psi\rangle, |\Phi\rangle) \quad (2.31)$$

et rien de plus. Notons que l'invariance ne demande pas nécessairement l'égalité des produits scalaires :

$$(|\Psi'\rangle, |\Phi'\rangle) = (|\Psi\rangle, |\Phi\rangle) \quad (2.32)$$

auquel cas l'opérateur $\mathcal{T}_{\vec{a}}$ serait forcément unitaire, mais seulement, d'après (2.31), l'égalité de leurs *modules*.

Il est possible de montrer¹⁴ que la condition (2.31) impose en général à tout opérateur représentant une opération de symétrie d'être soit unitaire (c'est de loin le cas le plus répandu) – tout produit scalaire est alors invariant –, soit antiunitaire – auquel cas tout produit scalaire est transformé en son complexe conjugué. Pour une symétrie continue, l'opérateur est forcément unitaire. Ce théorème peut d'ailleurs se démontrer au cas par cas en utilisant des arguments physiques.

¹³Lorsqu'il s'agit d'étudier l'invariance par symétrie, il y a toujours deux points de vue possibles :

- soit on effectue un changement de repère, sans toucher au système, le second repère se déduisant du premier par une transformation géométrique (point de vue "passif"),
- soit on effectue la transformation inverse sur le système lui-même (point de vue "actif").

S'agissant de rotations, par exemple, tourner le repère dans un sens ou tourner le système dans l'autre sens revient au même, précisément grâce au principe de relativité. Dans la suite, sauf mention du contraire, on adopte le point de vue actif : c'est le système que l'on translate, que l'on fait tourner, *etc.*

¹⁴C'est le théorème de Wigner ([17] p. 372, [20] p. 540, théorème III).

Comme exemple de prévision, considérons la probabilité d'observer la particule au voisinage du point \vec{r} ; la densité de probabilité correspondante est donnée par les modules au carré des fonctions d'onde. Une fois que l'on a translaté le système, c'est donc $|\Psi'(\vec{r}, t)|^2$ qui donne la densité de probabilité de présence au point \vec{r} ; l'invariance (euclidienne) exige que ce nombre soit le même que celui calculé, pour le système avant translation, au point qui est l'antécédent de \vec{r} dans la translation. Il faut donc¹⁵ :

$$|\Psi'(\vec{r}, t)|^2 \equiv |T_{\vec{a}} \Psi(\vec{r}, t)|^2 = |\Psi(\mathcal{T}_{\vec{a}}^{-1} \vec{r}, t)|^2 . \quad (2.34)$$

Cette égalité entraîne :

$$\Psi'(\vec{r}, t) = e^{i\alpha} \Psi(\mathcal{T}_{\vec{a}}^{-1} \vec{r}, t) \quad (2.35)$$

où α est une phase dépendant *a priori* de l'espace (et éventuellement du temps). Il est facile de se convaincre que, puisque la translation dans l'espace n'implique pas la vitesse (donc pas le moment conjugué), α est en réalité une constante (dans le cas contraire, la moyenne de l'impulsion calculée avec Ψ' ne serait pas égale à la moyenne calculée avec Ψ , ce qui serait stupide – voir Remarque ci-après) ; α étant une constante, on peut toujours s'en affranchir et prendre celle-ci égale à zéro¹⁶, d'où :

$$\Psi'(\vec{r}, t) = \Psi(\mathcal{T}_{\vec{a}}^{-1} \vec{r}, t) . \quad (2.36)$$

Dans ces conditions, l'opérateur $T_{\vec{a}}$ est alors clairement unitaire. En effet, pour deux fonctions Ψ et Φ , on a :

$$\int_{\mathbb{R}^3} d^3r \Psi'^*(\vec{r}, t) \Phi'(\vec{r}, t) = \int_{\mathbb{R}^3} d^3r \Psi^*(\mathcal{T}_{\vec{a}}^{-1} \vec{r}, t) \Phi(\mathcal{T}_{\vec{a}}^{-1} \vec{r}, t) . \quad (2.37)$$

Dans le second membre, il suffit de changer de variable d'intégration en posant $\rho = \mathcal{T}_{\vec{a}}^{-1} \vec{r} = \vec{r} - \vec{a}$ pour arriver à¹⁷ :

$$\int_{\mathbb{R}^3} d^3r \Psi'^*(\vec{r}, t) \Phi'(\vec{r}, t) = \int_{\mathbb{R}^3} d^3\rho \Psi^*(\rho, t) \Phi(\rho, t) \quad (2.38)$$

qui exprime bien l'invariance du produit scalaire. L'unitarité est d'ailleurs évidente – par opposition à l'antiunitarité – quand on réalise qu'il existe des translations infiniment petites qui tendent par continuité vers l'opération identité ; celle-ci est trivialement un opérateur unitaire et il ne peut qu'en aller de même des translations, infinitésimales ou finies. L'antiunitarité ne peut être associée qu'à une opération ne pouvant *continûment* se réduire à ne rien faire du tout (caractère "discret", "quantifié", de l'opération considérée – comme le renversement du temps^{18, 19}). Des résultats précédents, on déduit la valeur moyenne de la coordonnée dans l'état translaté, notée²⁰ $\langle \vec{r} \rangle'$; par définition de la valeur moyenne :

$$\langle \vec{r} \rangle' = \int_{\mathbb{R}^3} d^3r \Psi'^*(\vec{r}, t) \vec{r} \Psi'(\vec{r}, t) . \quad (2.39)$$

D'après(2.29) et (2.36), il vient :

$$\langle \vec{r} \rangle' = \int_{\mathbb{R}^3} d^3\rho \Psi^*(\vec{\rho}, t) (\vec{\rho} + \vec{a}) \Psi(\vec{\rho}, t) \equiv \langle \vec{r} \rangle + \vec{a} , \quad (2.40)$$

où $\langle \vec{r} \rangle$ désigne la valeur moyenne dans l'état avant translation ; ce résultat, trivial, est bien ce que l'on attend. En ce qui concerne l'impulsion – représentée par un opérateur de dérivation d'espace, insensible à une translation donnée – on a bien sûr $\langle \vec{p} \rangle' = \langle \vec{p} \rangle$ comme il se doit.

¹⁵L'équation (2.34) est bien la généralisation formelle de la transformée d'une fonction. En termes plus simples, soit une fonction d'une seule coordonnée, $f(x)$, représentant un certain attribut d'un système. Si le système est translaté de x_0 , $x' \equiv T_{x_0} x = x + x_0$ et le système est alors décrit par la fonction $f(x - x_0)$ (on prend en bloc $f(x)$ et on la déplace de x_0). On a bien :

$$f'(x) \equiv T_{x_0} f(x) = f(x - x_0) = f(\mathcal{T}_{x_0}^{-1} x) . \quad (2.33)$$

¹⁶Les fonctions Ψ' et $e^{i\alpha}\Psi'$ donnent exactement les mêmes prévisions physiques.

¹⁷On suppose bien évidemment l'espace infini dans toutes les directions.

¹⁸Renverser ou ne pas renverser le temps est visiblement une opération par tout ou rien, tout comme une symétrie-miroir.

¹⁹Ne pas en déduire que toute symétrie discrète est associée à un opérateur antiunitaire. La parité ou la symétrie-miroir sont représentées par des opérateurs unitaires.

²⁰Bien remarquer que le ' est à l'extérieur du signe $\langle \dots \rangle$ représentant la moyenne.

Remarque

Cela vaut la peine de revenir sur l'absence de phase α dans la relation entre Ψ' et Ψ (comparer (2.35) et (2.36)). Imaginons qu'il y en ait une ; on aurait alors :

$$\langle \vec{p} \rangle' = \int_{\mathbb{R}^3} d^3r e^{-i\alpha} \Psi^*(\mathcal{T}_{\vec{a}}^{-1} \vec{r}, t) \vec{p} e^{+i\alpha} \Psi(\mathcal{T}_{\vec{a}}^{-1} \vec{r}, t) . \quad (2.41)$$

Comme :

$$e^{-i\alpha} \vec{p} e^{+i\alpha} = \vec{p} - i[\alpha, \vec{p}] + \frac{i^2}{2} [\alpha, [\alpha, \vec{p}]] + \dots = \vec{p} - \hbar [\alpha, \vec{\nabla}] = \vec{p} + \hbar \vec{\nabla} \alpha , \quad (2.42)$$

une phase α non constante dans l'espace donnerait :

$$\langle \vec{p} \rangle' = \langle \vec{p} \rangle + \hbar \langle \alpha \vec{\nabla} \alpha \rangle . \quad (2.43)$$

Si une phase spatiale apparaissait, on trouverait cette absurdité : traduire un système dans l'espace modifie la valeur moyenne de sa vitesse !

Il est maintenant facile d'écrire explicitement l'expression de l'opérateur de translation $T_{\vec{a}}$. Pour cela, commençons par considérer une translation infinitésimale d'amplitude $\delta\vec{a}$. Par définition de l'opérateur de translation (2.30) et d'après (2.36), on a :

$$\Psi'(\vec{r}) \equiv T_{\delta\vec{a}} \Psi(\vec{r}) = \Psi(\mathcal{T}_{\vec{a}}^{-1} \vec{r}) = \Psi(\vec{r} - \delta\vec{a}) . \quad (2.44)$$

En développant au premier ordre, on obtient :

$$\forall \Psi : T_{\delta\vec{a}} \Psi(\vec{r}) = \Psi(\vec{r}) - \delta\vec{a} \cdot \vec{\nabla} \Psi(\vec{r}) + \dots \equiv \left(\mathbf{1} + \frac{\vec{p}}{i\hbar} \cdot \delta\vec{a} + \dots \right) \Psi(\vec{r}) . \quad (2.45)$$

Ceci montre que l'opérateur $T_{\delta\vec{a}}$ est :

$$T_{\delta\vec{a}} = \mathbf{1} + \frac{\vec{p}}{i\hbar} \delta\vec{a} + \dots . \quad (2.46)$$

En raison de ce résultat, on dit que \vec{p} est la *génératrice* des translations. Pour avoir l'opérateur d'une translation finie, il suffit maintenant de multiplier une suite de translations, toutes parallèles entre elles en posant :

$$\delta\vec{a} = \frac{1}{N} \vec{a} \quad (2.47)$$

et en écrivant :

$$T_{\vec{a}} = (T_{\delta\vec{a}})^N = \left(\mathbf{1} + \frac{\vec{p}}{i\hbar} \frac{\vec{a}}{N} + \dots \right)^N . \quad (2.48)$$

Dans la limite $N \rightarrow \infty$, compte tenu de $\lim_{N \rightarrow \infty} [1 + (x/N)]^N = e^x$ on trouve :

$$T_{\vec{a}} = e^{\frac{1}{i\hbar} \vec{p} \cdot \vec{a}} . \quad (2.49)$$

Cet opérateur est visiblement unitaire, puisque \vec{p} est hermitique et que \vec{a} est un vecteur réel. Comme on le sait, l'ordre des translations dans un produit n'importe pas, que les translations considérées soient parallèles entre elles ou non (les translations forment un groupe abélien) ; l'expression (2.49) assure bien ceci puisque les différentes composantes de \vec{p} commutent également entre elles, ce qui permet d'écrire :

$$T_{\vec{a}} = e^{\frac{1}{i\hbar} \sum_{u=x,y,z} p_u a_u} = \prod_{u=x,y,z} e^{\frac{1}{i\hbar} p_u a_u} . \quad (2.50)$$

La commutation entre elles des composantes de \vec{p} permet d'écrire que l'exponentielle de la somme est le produit des exponentielles, et que, dans ce produit, l'ordre des facteurs est sans importance.

Faisons à nouveau le lien avec les éventuelles variables cycliques de H . Tout comme le vecteur d'état change quand on translate le système, le Hamiltonien H , lui aussi, change, et se transforme en H'^{21} . La recette définissant le mode de calcul des valeurs moyennes étant définie sans référence à un repère quelconque, l'invariance euclidienne impose l'égalité des moyennes :

$$(|\Psi'\rangle, H'|\Psi'\rangle) = (|\Psi\rangle, H|\Psi\rangle) \iff \langle \Psi | T_{\vec{a}}^\dagger H' T_{\vec{a}} | \Psi \rangle = \langle \Psi | H | \Psi \rangle . \quad (2.51)$$

soit²², ceci étant vrai $\forall |\Psi\rangle$ (par exemple pour l'ensemble des vecteurs propres de H , qui forment une base complète) :

$$H = T_{\vec{a}}^\dagger H' T_{\vec{a}} \quad \forall \vec{a} \iff H' = T_{\vec{a}} H T_{\vec{a}}^\dagger \quad (\text{invariance euclidienne}) . \quad (2.52)$$

Le vecteur de translation étant quelconque, ceci est l'expression analytique du principe euclidien²³. On retiendra l'idée suivante : dans une transformation le vecteur d'état change ; puisque toutes les prévisions doivent être invariantes au sens euclidien, il faut bien que les observables changent simultanément, de façon à compenser la variation du vecteur d'état.

Revenons maintenant au cas précis considéré au début, servant d'exemple à la symétrie (interne, propre) de translation parallèlement à Ox et à Oy ; H est invariant dans toute translation perpendiculaire à Oz , puisque celle-ci ne change que x et y , les deux coordonnées dont H ne dépend pas :

$$\forall \vec{a} \perp Oz : H' = H \iff T_{\vec{a} \perp Oz}^\dagger H T_{\vec{a} \perp Oz} = H . \quad (2.53)$$

Comme $T_{\vec{a}}$ est unitaire, $T_{\vec{a}}^\dagger = T_{\vec{a}}^{-1}$; finalement on obtient :

$$\forall \vec{a} \perp Oz : H T_{\vec{a} \perp Oz} = T_{\vec{a} \perp Oz} H \iff [H, T_{\vec{a} \perp Oz}] = 0 . \quad (2.54)$$

D'où le résultat majeur : le Hamiltonien commute avec les opérateurs de symétrie interne du système, une affirmation qui est quasi-évidente après coup : par nature, le système est totalement insensible à toute opération de symétrie. De fait, si $\vec{a} \perp Oz$ ($\vec{a} = a_x \vec{e}_x + a_y \vec{e}_y$), l'opérateur de translation est :

$$T_{\vec{a} \perp Oz} = e^{\frac{i}{\hbar}(a_x p_x + a_y p_y)} . \quad (2.55)$$

Comme H ne contient ni x ni y , H et $T_{\vec{a}}$, donné par (2.55), commutent.

Ainsi, le Hamiltonien commute avec tout opérateur associé à une opération de symétrie du système. Le Hamiltonien étant le générateur du mouvement, toute propriété de symétrie (la parité, la façon de se transformer par une rotation, etc.) est une "constante du mouvement"²⁴. À titre d'exemple, montrons explicitement qu'un état pair ne peut se transformer en état impair par évolution libre²⁵. Soit la valeur moyenne $\langle \Pi \rangle = \langle \Psi | \Pi | \Psi \rangle$ calculée avec un certain état. On a :

$$\frac{d}{dt} \langle \Pi \rangle = -\frac{1}{i\hbar} \langle \Psi | H \Pi | \Psi \rangle + \frac{1}{i\hbar} \langle \Psi | \Pi H | \Psi \rangle \equiv \frac{1}{i\hbar} \langle [\Pi, H] \rangle . \quad (2.56)$$

Si H est invariant par parité, le commutateur est nul et la valeur moyenne $\langle \Pi \rangle$ est une constante dans le temps, quel que soit l'état $|\Psi\rangle$: l'opérateur Π – qui est linéaire – est une constante du mouvement. Pour un état de parité déterminée $|\Psi_\varepsilon\rangle$ ($\Pi |\Psi_\varepsilon\rangle = \varepsilon |\Psi_\varepsilon\rangle$, $\varepsilon = \pm 1$), $\langle \Pi \rangle = \varepsilon$; pour un état quelconque $|\Psi\rangle = \cos \theta |\Psi_{+1}\rangle + e^{i\alpha} \sin \theta |\Psi_{-1}\rangle$, la moyenne vaut $\langle \Pi \rangle = \cos 2\theta$.

En outre et en toute généralité, H et les opérateurs liés à la symétrie interne possèdent des vecteurs propres en commun, ce qui pourra grandement faciliter la résolution effective de l'équation aux valeurs et vecteurs propres.

²¹Par exemple, soit un oscillateur harmonique dont le point d'équilibre est pris à l'origine ; le terme d'énergie potentielle est donc $(m\omega^2/2)x^2$; si on translate le système de x_0 , le point d'équilibre a maintenant l'abscisse x_0 et l'énergie potentielle devient $(m\omega^2/2)(x - x_0)^2$.

²²voir [20], p. 549.

²³Pour l'oscillateur harmonique pris ci-dessus comme exemple, on vérifiera que l'équation (2.52) est bien satisfaite avec l'opérateur $T_{x_0} = e^{-x_0(d/dx)}$

²⁴Rien n'interdit formellement de définir la moyenne quantique d'un opérateur qui n'est pas hermitique, en particulier celle de tout opérateur représentant une opération de symétrie ; la moyenne en question ne sera pas forcément réelle, puisque l'opérateur considéré est unitaire ou, exceptionnellement, antiunitaire, mais peu importe. De même, on peut écrire des équations de Heisenberg pour un opérateur non hermitique ; pour les opérateurs décrivant la symétrie, ces équations montrent qu'il s'agit de constantes du mouvement, précisément (*attention !* Ceci ne vaut que pour les opérateurs linéaires, pas pour les opérateurs antilinéaires, voir plus loin la discussion sur le renversement du temps).

²⁵On dit souvent de cette impossibilité qu'elle résulte d'une règle de "supersélection", pour l'opposer aux règles de sélection qui sont, souvent, définies à l'intérieur d'une approximation donnée (au 1^{er} ordre, au 2^{ème} ordre, etc.). Comme toujours, la symétrie fonctionne par tout ou rien et verrouille absolument.

Remarques

1. Il convient de distinguer les deux types de conditions imposées par l'invariance vis-à-vis d'une opération de symétrie S , qu'elle soit euclidienne ou de symétrie propre. S'agissant des probabilités, il faut bien :

$$|(|\Psi'\rangle, |\Phi'\rangle)|^2 = |(|\Psi\rangle, |\Phi\rangle)|^2, \quad (2.57)$$

quels que soient les états $|\Psi\rangle$ et $|\Phi\rangle$; c'est bien le *module carré* qu'il faut considérer et, suivant le théorème de Wigner, ceci ouvre une double possibilité : S est soit unitaire, soit antiunitaire.

D'un autre côté, il est bien évident que toutes les valeurs moyennes doivent être elles aussi invariantes : si A est une observable quelconque, les valeurs moyennes de A doivent être les mêmes (pas seulement leurs modules !) pour deux observateurs, liés à deux repères différents pour la symétrie euclidienne, ou acteur et aveugle en ce qui concerne la symétrie interne²⁶ ; ceci impose :

$$\langle \Psi' | A' | \Psi' \rangle = \langle \Psi | A | \Psi \rangle \quad \forall |\Psi\rangle, \quad (2.58)$$

d'où il résulte²⁷ que les deux opérateurs $S^\dagger A' S$ et A sont égaux ; le transformé d'un opérateur linéaire A représentant une observable est donc :

$$A' = S A S^\dagger. \quad (2.59)$$

La relation (2.59) étant établie, il en résulte, cette fois pour les éléments non-diagonaux (pas seulement pour les valeurs moyennes, qui sont des éléments diagonaux) :

$$(|\Psi\rangle, A|\Phi\rangle) = (|\Psi\rangle, S^\dagger A' S|\Phi\rangle) \quad \forall |\Psi\rangle, |\Phi\rangle. \quad (2.60)$$

Maintenant de deux choses l'une :

- ou bien S est unitaire, auquel cas le second membre de (2.60) est $(S|\Psi\rangle, A' S|\Phi\rangle)$, soit $(|\Psi'\rangle, A'|\Phi'\rangle)$ et on retrouve, même pour les éléments non-diagonaux, une relation du genre (2.51) :

$$(|\Psi'\rangle, A'|\Phi'\rangle) = (|\Psi\rangle, A|\Phi\rangle) \quad \forall |\Psi\rangle, |\Phi\rangle. \quad (2.61)$$

- ou bien S est antiunitaire ; alors le second membre de (2.60), compte tenu de la définition de l'adjoint d'un opérateur antiunitaire, est $(S|\Psi\rangle, A' S|\Phi\rangle)^*$ et maintenant :

$$(|\Psi'\rangle, A'|\Phi'\rangle) = (|\Psi\rangle, A|\Phi\rangle)^* \quad \forall |\Psi\rangle, |\Phi\rangle. \quad (2.62)$$

2. À titre d'exemple, revenons à l'exemple de la translation de vecteur \vec{a} , conduisant à la relation (2.40) entre les valeurs moyennes. Imaginons maintenant deux observateurs du même système : l'un ne bouge pas, l'autre suit le système dans sa translation ; ce dernier trouve la valeur moyenne de la coordonnée en manipulant l'opérateur transformé \vec{r}' défini comme :

$$\vec{r}' = T_{\vec{a}} \vec{r} T_{\vec{a}}^\dagger = e^{\frac{i}{\hbar} \vec{a} \cdot \vec{p}} \vec{r} \left(e^{\frac{i}{\hbar} \vec{a} \cdot \vec{p}} \right)^\dagger = e^{\frac{i}{\hbar} \vec{a} \cdot \vec{p}} \vec{r} e^{-\frac{i}{\hbar} \vec{a} \cdot \vec{p}}, \quad (2.63)$$

conformément à (2.49) et (2.59). Le développement de chaque exponentielle fait apparaître une série de commutateurs multiples, série dont seuls les deux premiers termes sont non-nuls :

$$\vec{r}' = \vec{r} + \frac{1}{i\hbar} [\vec{a} \cdot \vec{p}, \vec{r}] + \dots = \vec{r} + \frac{1}{i\hbar} [\vec{a} \cdot \vec{p}, \vec{r}] = \vec{r} - \vec{a}. \quad (2.64)$$

L'observateur déplacé avec le système calcule la valeur moyenne selon²⁸ $\langle \Psi' | \vec{r}' | \Psi' \rangle \equiv \langle \vec{r}' \rangle'$:

$$\langle \vec{r}' \rangle' = \int_{\mathbb{R}^3} d^3 r \Psi^*(T_{\vec{a}}^{-1} \vec{r}, t) (\vec{r} - \vec{a}) \Psi(T_{\vec{a}}^{-1} \vec{r}, t) = \int_{\mathbb{R}^3} d^3 \rho \Psi^*(\rho, t) \rho \Psi(\rho, t) \equiv \langle \vec{r} \rangle, \quad (2.65)$$

et au total :

$$\langle \vec{r}' \rangle' = \langle \vec{r} \rangle. \quad (2.66)$$

²⁶ L'observateur acteur est celui qui translate le système, l'autre est celui qui se ferme les yeux pendant la translation.

²⁷ voir [20], p. 549.

²⁸ Bien noter le ' supplémentaire à l'intérieur de $\langle \dots \rangle$ qui représente la moyenne.

Cette identité exprime que les deux observateurs trouvent le même résultat chacun dans son repère, comme le requiert l'invariance euclidienne.

On peut dire que dans ce type d'analyse, il y a en réalité trois observateurs ; deux d'entre eux sont liés chacun à un repère R ou R' (ici R' est translaté de \vec{a} par rapport à R). Le troisième joue le rôle de “juge de paix” et compare les résultats des deux autres. Le premier observateur mesure $\langle \vec{r} \rangle$, le second mesure $\langle \vec{r}' \rangle$ qui est identique à ce qu'a trouvé le premier. Le troisième observateur constate avec soulagement que l'invariance euclidienne imposée donne la bonne relation $\langle \vec{r}' \rangle = \langle \vec{r} \rangle + \vec{a}$ (voir (2.40)).

3. Les deux conditions (2.31) et (2.58) ne sont pas réductibles l'une à l'autre. En effet, même si les probabilités sont identiques (assurées par la première condition), les valeurs moyennes sont différentes si les deux opérateurs A et A' n'ont pas le même spectre de valeurs propres. Il est donc bien nécessaire d'imposer la deuxième condition ; c'est suffisant car la relation (2.59) est une relation d'équivalence assurant que A et A' ont bel et bien le même spectre. Avec l'identité des probabilités et l'identité des spectres, l'invariance est assurée.
4. Examinons ce que deviennent les relations de commutation dans une transformation S , en raisonnant simplement avec x et p_x . Partant de $[x, p_x] = i\hbar$, multiplions à gauche par S et à droite par S^\dagger :

$$S [x, p_x] S^\dagger = S i\hbar S^\dagger . \quad (2.67)$$

Dans tous les cas, on a $SS^\dagger = S^\dagger S = \mathbf{1}$. En insérant le produit SS^\dagger là où il faut, on voit tout de suite que le premier membre est de fait toujours égal à $[x', p'_x]$ ($x' = SxS^\dagger$, $p'_x = Sp_xS^\dagger$). Au contraire, le second membre dépend du caractère unitaire ou antiunitaire de S . Si S est unitaire, on a : $S(i\hbar)S^\dagger = i\hbar\mathbf{1}$ et, si S est antiunitaire : $S(i\hbar)S^\dagger = -i\hbar\mathbf{1}$.

Plus généralement, soit A et B deux observables ; étant hermitiques, leur commutateur peut être noté iC où C est aussi hermitique :

$$[A, B] = iC . \quad (2.68)$$

Entre opérateurs transformés, on a la relation :

$$S [A, B] S^\dagger = S(iC)S^\dagger , \quad (2.69)$$

soit :

$$[A', B'] = \begin{cases} +iSCS^\dagger = +iC' & \text{si } S \text{ est unitaire} \\ -iSCS^\dagger = -iC' & \text{si } S \text{ est antiunitaire} \end{cases} . \quad (2.70)$$

Autrement dit, les relations de commutation fondamentales sont invariantes si la symétrie est unitaire et changent de signe si la symétrie est antiunitaire. Comme on le verra par la suite, le renversement du temps, représenté par la transformation notée K , est antiunitaire. Il en résultera les relations, nécessaires (et évidentes) physiquement²⁹ :

$$\vec{r}' = K \vec{r} K^\dagger = \vec{r} \quad \vec{p}' = K \vec{p} K^\dagger = -\vec{p} , \quad (2.71)$$

de sorte que la relation de commutation fondamentale $[u, p_v] = i\hbar\delta_{uv}$ devient :

$$[u, p'_v] = -i\hbar\delta_{uv} . \quad (2.72)$$

2.2.2 Rotations

Tout ce qui vient d'être dit à propos des translations dans l'espace tient aussi pour les rotations. L'invariance par translation implique l'homogénéité présupposée de l'espace (supposé illimité), telle que l'affirme le principe euclidien (pas d'origine privilégiée). En ce qui concerne les rotations, c'est l'isotropie de l'espace qui est à l'œuvre (pas de direction privilégiée). Les rotations constituent également des déplacements : il existe donc des transformations infinitésimales, arbitrairement voisines de la transformation identité. L'analyse qui suit jouera un rôle de tout premier plan dans l'étude des systèmes possédant une symétrie de rotation ; le champ central

²⁹Changer t en $-t$ ne modifie pas les coordonnées mais inverse toutes les vitesses (et donc aussi les moments cinétiques).

(par exemple : l'atome d'hydrogène) en fournira un exemple très important et ces arguments d'invariance par rotation seront aussi utilisés à propos des effets Zeeman et Stark.

Considérons une rotation du système d'un angle θ autour d'un axe défini par le vecteur unitaire vecteur \vec{u} . Par définition, dans cette transformation $\mathcal{R}_{\vec{u},\theta}$, \vec{r} devient \vec{r}' :

$$\vec{r}' = \mathcal{R}_{\vec{u},\theta} \vec{r} . \quad (2.73)$$

Soit $\Psi(\vec{r}, t)$ une fonction d'onde du système ; lorsque ce dernier subit la rotation $\mathcal{R}_{\vec{u},\theta}$, la fonction d'onde devient $\Psi'(\vec{r}, t)$, qui doit pouvoir se déduire de $\Psi(\vec{r}, t)$ par l'action d'un opérateur $R_{\vec{u},\theta}$; à nouveau, l'invariance euclidienne permet d'écrire :

$$\Psi'(\vec{r}, t) = R_{\vec{u},\theta} \Psi(\vec{r}, t) = \Psi(\mathcal{R}_{\vec{u},\theta}^{-1} \vec{r}, t) \quad (2.74)$$

qui est l'analogie de (2.36). L'argument conduisant à cette relation est le même que pour les translations : l'invariance des probabilités requiert l'invariance des modules au carré des produits scalaires, ce qui autorise toutefois une phase additionnelle non-triviale entre Ψ' et Ψ . Cependant, cette phase (variable dans l'espace) donnerait à nouveau une absurdité : le moment cinétique du système – pendant de l'impulsion quand il s'agit de rotation – serait changé quand on fait tourner le système. Par ailleurs, l'unitarité de $R_{\vec{u},\theta}$ se démontre de la même façon que pour une translation (il suffit à nouveau de changer les variables muettes d'intégration). De même, on montre facilement que la valeur moyenne de la coordonnée dans l'état tourné est le transformé de la valeur moyenne dans l'état avant rotation par la rotation $\mathcal{R}_{\vec{u},\theta}$.

L'opérateur $R_{\vec{u},\theta}$ est facile à trouver, en suivant une procédure du même type que pour les translations. L'écart $\delta\vec{r} = \mathcal{R}_{\vec{u},\delta\theta} \vec{r} - \vec{r}$ est perpendiculaire à la fois à \vec{u} et à \vec{r} ; il est donc colinéaire à $\vec{u} \times \vec{r}$, son module est $r_{\perp} \delta\theta$, avec $r_{\perp} = r |\sin(\vec{r}, \vec{u})|$, d'où $\delta\vec{r} = \vec{u} \times \vec{r} \delta\theta$. Pour une rotation infinitésimale d'angle $\delta\theta$, le vecteur transformé est donc :

$$\mathcal{R}_{\vec{u},\delta\theta} \vec{r} = \vec{r} + \delta\theta \vec{u} \times \vec{r} . \quad (2.75)$$

Selon (2.74), il en résulte³⁰ :

$$\Psi'(\vec{r}, t) = \Psi(\vec{r} - \delta\theta \vec{u} \times \vec{r}, t) . \quad (2.76)$$

À des infiniment petits d'ordre supérieur :

$$\Psi'(\vec{r}, t) = \Psi(\vec{r}) - \delta\theta \vec{u} \times \vec{r} \cdot \vec{\nabla} \Psi(\vec{r}, t) = (\mathbf{1} - \delta\theta \vec{u} \cdot \vec{r} \times \vec{\nabla}) \Psi(\vec{r}, t) . \quad (2.77)$$

En remplaçant $\vec{\nabla}$ par $-\vec{p}/(i\hbar)$, on fait apparaître le moment cinétique orbital $\vec{L} = \vec{r} \times \vec{p}$:

$$\Psi'(\vec{r}, t) = R_{\vec{u},\delta\theta} \Psi(\vec{r}, t) = \left(\mathbf{1} + \frac{\delta\theta}{i\hbar} \vec{u} \cdot \vec{L} \right) \Psi(\vec{r}, t) , \quad (2.78)$$

d'où par identification :

$$R_{\vec{u},\delta\theta} = \mathbf{1} + \frac{\delta\theta}{i\hbar} \vec{u} \cdot \vec{L} . \quad (2.79)$$

Visiblement, \vec{p} et \vec{L} jouent des rôles analogues pour les translations et les rotations respectivement : \vec{L} est aussi appelé le générateur infinitésimal des rotations, comme le montre (2.79), équation qui est l'analogie de (2.46) pour les translations.

Pour obtenir l'opérateur associé à une rotation d'angle fini θ , on combine un grand nombre de rotations élémentaires : posant $\delta\theta = \theta/N$ et on fait le produit de N rotations d'angle $\delta\theta$ autour du même axe de rotation³¹. On obtient ainsi, après passage à la limite $N \rightarrow \infty$, $\delta\theta \rightarrow 0$, et usage de $\lim_{N \rightarrow \infty} [1 + (x/N)]^N = e^x$:

$$R_{\vec{u},\theta} = e^{\frac{1}{i\hbar} \theta \vec{u} \cdot \vec{L}} . \quad (2.80)$$

³⁰L'inverse d'une rotation d'angle θ est la rotation d'angle $-\theta$.

³¹Attention : au contraire des translations, les rotations ne commutent pas en général ; seules commutent entre elles les rotations autour d'un *même* axe – c'est pourquoi les composantes du moment cinétique ne commutent pas entre elles, voir ci-dessous. Il est intéressant de noter que translations et rotations sont *topologiquement* inéquivalentes – ce qui n'empêche pas, quand le terrain est bien balisé (non-pertinence des effets de bord), d'assimiler souvent localement rotations et translations, en particulier pour un "grand" système où la transformation peut être considérée comme infinitésimale.

Cet opérateur est visiblement unitaire puisque le moment cinétique est hermitique. En outre, cette expression corrobore le fait que deux rotations autour de deux axes différents (Ox et Oy , par exemple) ne commutent pas ; les composantes L_x , L_y et L_z , en effet, ne commutent pas entre elles. En conséquence (comparer avec (2.50)) :

$$R_{\vec{u},\theta} = e^{\frac{1}{i\hbar}\theta \sum_{v=x,y,z} L_v u_v} \neq \prod_{v=x,y,z} e^{\frac{1}{i\hbar}\theta L_v u_v} . \quad (2.81)$$

Quand on fait tourner le système, le Hamiltonien se transforme en H' ; l'invariance euclidienne de la Mécanique Quantique, cette fois par rapport aux rotations (comp. (2.51)) requiert à nouveau l'invariance des valeurs moyennes – et aussi des éléments de matrice puisque $R_{\vec{u},\theta}$ est unitaire, voir Remarque 1 ci-dessus ; en particulier, pour H :

$$(|\Psi'\rangle, H'|\Phi'\rangle) = (|\Psi\rangle, H|\Phi\rangle) \iff \langle \Psi | R_{\vec{u},\theta}^\dagger H' R_{\vec{u},\theta} | \Phi \rangle = \langle \Psi | H | \Phi \rangle \quad (2.82)$$

quels que soient $|\Psi\rangle$ et $|\Phi\rangle$, d'où :

$$R_{\vec{u},\theta}^\dagger H' R_{\vec{u},\theta} = H . \quad (2.83)$$

Compte tenu de l'unitarité de $R_{\vec{u},\theta}$, H' est finalement donné par³² :

$$H' = R_{\vec{u},\theta} H R_{\vec{u},\theta}^\dagger = R_{\vec{u},\theta} H R_{\vec{u},\theta}^{-1} \quad (\text{invariance euclidienne}) . \quad (2.84)$$

Maintenant, si le système possède la symétrie de rotation, son Hamiltonien est invariant :

$$H' = H \quad (\text{invariance par symétrie propre}) \quad (2.85)$$

Et en utilisant l'unitarité de $R_{\vec{u},\theta}$:

$$H = R_{\vec{u},\theta} H R_{\vec{u},\theta}^\dagger \iff [H, R_{\vec{u},\theta}] = 0 . \quad (2.86)$$

Cette relation est vraie, quels que soient \vec{u} et θ lorsque H ne dépend que du module carré du moment cinétique et du module du rayon vecteur, r (champ central) ; H étant à symétrie sphérique, les trois composantes du moment cinétique orbital commutent chacune avec H et c'est bien le *vecteur* moment cinétique \vec{L} qui est une constante du mouvement : tout comme pour les translations ressort le lien indissoluble entre symétrie et conservation dans le temps. À l'inverse, en présence d'un champ (dirigé le long de Oz), si le terme d'énergie cinétique continue à n'être fonction que du module carré du moment cinétique, le terme potentiel sera seulement invariant par rotation autour de Oz ³³ : le champ appliqué, par l'intermédiaire de V , brise la symétrie sphérique.

La relation (2.86) conduit naturellement à la notion d'observable *scalaire*, généralisant la notion usuelle de scalaire en géométrie ordinaire³⁴. On dit qu'une observable A est un opérateur scalaire si elle est invariante dans toute rotation ; cette propriété se traduit par :

$$A = R_{\vec{u},\theta} A R_{\vec{u},\theta}^\dagger \iff [A, R_{\vec{u},\theta}] = 0 \quad \forall \vec{u} . \quad (2.87)$$

Selon (2.80), ceci veut dire qu'un opérateur scalaire commute avec chacune des trois composantes du moment cinétique orbital :

$$[A, L_x] = 0 , \quad [A, L_y] = 0 , \quad [A, L_z] = 0 . \quad (2.88)$$

Par exemple, pour un système invariant par rotation, le Hamiltonien est un opérateur scalaire.

Dans le même ordre d'idée et par généralisation, on introduit les opérateurs dits *vectoriels* ; un opérateur vectoriel \vec{A} est défini par la propriété que les valeurs moyennes de ses composantes se comportent par rotation comme les composantes d'un vecteur de la géométrie ordinaire ; soit R une rotation dans \mathbb{R}^3 induisant la transformation (unitaire) \mathcal{R} dans l'espace des états : $|\Psi'\rangle = \mathcal{R}|\Psi\rangle$. Comme précédemment, on note $\langle A_u \rangle$

³²Soit une charge q couplée à un champ électrique \mathcal{E} dirigé le long de Oz ; l'interaction est $-q\mathcal{E}z$. Faisons la rotation de $+\pi/2$ autour de Oy : l'axe Oz se transforme en Ox . On vérifie bien que (2.82) et $R_{Oy, \pi/2} = e^{(1/i\hbar)(\pi/2)L_y}$ conduisent à $H' = -q\mathcal{E}x$ comme il se doit.

³³et seule la composante du moment cinétique le long de Oz sera une constante du mouvement.

³⁴Le produit scalaire de deux vecteurs ordinaires est un invariant dans toute rotation des deux vecteurs.

($u = x, y, z$) la valeur moyenne dans l'état $|\Psi\rangle$ avant rotation et $\langle A_u \rangle'$ la valeur moyenne dans l'état $|\Psi'\rangle$ après rotation. Compte tenu de la définition adoptée, on doit avoir :

$$\begin{bmatrix} \langle A_x \rangle' \\ \langle A_y \rangle' \\ \langle A_z \rangle' \end{bmatrix} = \begin{bmatrix} R_{xx} & R_{xy} & R_{xz} \\ R_{yx} & R_{yy} & R_{yz} \\ R_{zx} & R_{zy} & R_{zz} \end{bmatrix} \begin{bmatrix} \langle A_x \rangle \\ \langle A_y \rangle \\ \langle A_z \rangle \end{bmatrix} \quad (2.89)$$

où les R_{uv} sont les éléments de la matrice de la rotation en question dans l'espace à trois dimensions³⁵. L'expression de l'opérateur de rotation est déjà connue, donnée par (2.80), de sorte que tout est calculable dans l'équation (2.89) – qui de ce fait est plutôt une identité. Le calcul explicite mettrait donc en évidence un ensemble de relations caractéristiques entre les composantes d'un opérateur réputé vectoriel et les composantes du moment cinétique orbital \vec{L} (énoncées sous forme de commutateurs).

On peut aussi utiliser l'argument qui suit, exploitant directement la définition *physique* des opérateurs vectoriels adoptée ici – traduite par (2.89) – afin notamment de montrer comment le moment cinétique “sort” spontanément dès qu'il s'agit de rotations. Ce faisant, on établit ces relations de commutation caractéristiques, et d'autre part on retrouve l'expression de $R_{\vec{u},\theta}$ obtenue par d'autres moyens.

Considérons plus précisément une transformation infinitésimale, par exemple une rotation de $\delta\theta$ autour de Oz ; dans ce cas, (2.89) est explicitement :

$$\begin{bmatrix} \langle A_x \rangle' \\ \langle A_y \rangle' \\ \langle A_z \rangle' \end{bmatrix} = \begin{bmatrix} 1 & -\delta\theta & 0 \\ +\delta\theta & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \langle A_x \rangle \\ \langle A_y \rangle \\ \langle A_z \rangle \end{bmatrix} . \quad (2.92)$$

L'opérateur de rotation étant unitaire, on peut *a priori* l'écrire :

$$R_{z,\delta\theta} = \mathbf{1} - i\Omega_z \delta\theta , \quad (2.93)$$

où Ω_z est un opérateur hermitique et où l'indice z rappelle qu'il s'agit d'une rotation autour de Oz . Avec ceci, les valeurs moyennes deviennent :

$$\langle A_x \rangle' \equiv \langle \Psi' | A_x | \Psi' \rangle = \langle \Psi | (\mathbf{1} + i\Omega_z \delta\theta) A_x (\mathbf{1} - i\Omega_z \delta\theta) | \Psi \rangle = \langle A_x \rangle + i\delta\theta \langle \Psi | [\Omega_z, A_x] | \Psi \rangle . \quad (2.94)$$

D'après (2.92), ceci doit être égal à $\langle A_x \rangle - \delta\theta \langle A_y \rangle$, $\forall |\Psi\rangle$, d'où par identification des termes en $\delta\theta$:

$$i[\Omega_z, A_x] = -A_y \iff [\Omega_z, A_x] = iA_y . \quad (2.95)$$

En procédant de même avec les composantes suivant Oy et Oz , on trouve :

$$[\Omega_z, A_y] = -iA_x , \quad [\Omega_z, A_z] = 0 . \quad (2.96)$$

En faisant le même travail avec des rotations infinitésimales autour de Ox et Oy , on trouve :

$$[\Omega_x, A_y] = iA_z , \quad [\Omega_x, A_z] = -iA_y , \quad [\Omega_x, A_x] = 0 , \quad (2.97)$$

$$[\Omega_y, A_z] = iA_x , \quad [\Omega_y, A_x] = -iA_z , \quad [\Omega_y, A_y] = 0 . \quad (2.98)$$

Ces relations sont vraies pour n'importe quel opérateur vectoriel \vec{A} ; il reste à trouver le triplet $(\Omega_x, \Omega_y, \Omega_z)$. Physiquement, le moment cinétique \vec{L} doit être lui-même un opérateur vectoriel : on attend des valeurs moyennes de ses composantes qu'elles se transforment comme les composantes cartésiennes d'un vecteur. Il en résulte que dans toutes les relations ci-dessus, on peut remplacer \vec{A} par \vec{L} . L'écriture explicite des neuf relations distinctes

³⁵On sait écrire cette matrice indépendamment du contexte présent ; par exemple, dans la rotation autour de l'axe Oz d'un angle θ , les composantes d'un vecteur \vec{V} se transforment en :

$$V'_x = V_x \cos \theta - V_y \sin \theta , \quad V'_y = V_x \sin \theta + V_y \cos \theta , \quad V'_z = V_z , \quad (2.90)$$

ce qui s'écrit sous forme matricielle :

$$\begin{bmatrix} V'_x \\ V'_y \\ V'_z \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} . \quad (2.91)$$

(2.95) - (2.98) permet alors de constater que le vecteur $\vec{\Omega}$ n'est autre que le moment cinétique \vec{L} lui-même – au facteur \hbar près :

$$\vec{\Omega} = \hbar^{-1} \vec{L} . \quad (2.99)$$

Cette identification permet de reformuler comme suit les relations (2.95) - (2.98) :

$$[V_x, L_y] = i\hbar V_z , \quad [V_y, L_y] = 0 , \quad [V_z, L_y] = -i\hbar V_x , \quad (2.100)$$

$$[V_y, L_z] = i\hbar V_x , \quad [V_z, L_z] = 0 , \quad [V_x, L_z] = -i\hbar V_y , \quad (2.101)$$

$$[V_z, L_x] = i\hbar V_y , \quad [V_x, L_x] = 0 , \quad [V_y, L_x] = -i\hbar V_z . \quad (2.102)$$

Il est facile de vérifier que le produit scalaire $\vec{V}_1 \cdot \vec{V}_2$ de deux opérateurs vectoriels est un opérateur scalaire, satisfaisant (2.88).

En conclusion, et en vue de la généralisation aux opérateurs tensoriels irréductibles, on peut affirmer qu'il existe donc une totale équivalence entre les deux définitions :

- un opérateur vectoriel est un ensemble de trois opérateurs dont les valeurs moyennes des composantes se transforment par rotation comme les trois composantes d'un vecteur de la géométrie ordinaire.
- un opérateur vectoriel est un ensemble de trois opérateurs satisfaisant avec les composantes du moment cinétique les neuf équations caractéristiques (2.100) - (2.102), ou les relations équivalentes³⁶ :

$$\begin{aligned} [L_+, V_x] &= +\hbar V_z , & [L_+, V_y] &= +i\hbar V_z , & [L_+, V_z] &= -\hbar (V_x + iV_y) \\ [L_-, V_x] &= -\hbar V_z , & [L_-, V_y] &= +i\hbar V_z , & [L_-, V_z] &= +\hbar (V_x - iV_y) \\ [L_z, V_x] &= +i\hbar V_y , & [L_z, V_y] &= -i\hbar V_x , & [L_z, V_z] &= 0 . \end{aligned} \quad (2.103)$$

Ces relations peuvent d'ailleurs s'écrire de façon plus systématique en introduisant les composantes standard $V_q^{(1)}$ de l'opérateur vectoriel $\vec{V} \equiv V^{(1)}$, définies comme :

$$V_1^{(1)} = -\frac{1}{\sqrt{2}} (V_x + iV_y) , \quad V_0^{(1)} = V_z , \quad V_{-1}^{(1)} = +\frac{1}{\sqrt{2}} (V_x - iV_y) . \quad (2.104)$$

Avec cette définition, et en notant généralement \vec{J} le moment cinétique, on a :

$$[J_{\pm}, V_q^{(1)}] = \hbar \sqrt{2 - q(q \pm 1)} V_{q \pm 1}^{(1)} , \quad [J_z, V_q^{(1)}] = \hbar q V_q^{(1)} . \quad (2.105)$$

Ce sont ces dernières relations que l'on généralise en définissant les opérateurs tensoriels irréductibles (voir [20], ch. XIII.)

Remarques

1. S'agissant du spin, on argumentera exactement de la même façon en s'appuyant sur le fait physique que le moment magnétique $\vec{\mu}$ doit être un opérateur vectoriel au sens défini comme ci-dessus (les valeurs moyennes de ses composantes se transforment comme les composantes cartésiennes d'un vecteur ordinaire). Comme $\vec{\mu}$ et le spin \vec{S} sont proportionnels, cette réquisition physique pour $\vec{\mu}$ fait *ipso facto* de \vec{S} un opérateur vectoriel. Quand on fait "tourner l'aimant" qui permet de mesurer l'une ou l'autre des composantes de $\vec{\mu}$ ou de \vec{S} , on induit une transformation R_{spin} agissant dans l'espace des états de spin. Les calculs précédents peuvent être recommencés pour obtenir l'expression de R_{spin} ; en conformité avec (2.80), on trouve :

$$R_{\text{spin } \vec{u}, \theta} = e^{\frac{i}{\hbar} \theta \vec{u} \cdot \vec{S}} . \quad (2.106)$$

D'autres arguments seront d'ailleurs développés dans le chapitre 5, qui permettent de se convaincre que \vec{S} est un moment cinétique.

2. Aucun des arguments précédents ne repose sur la dimension de l'espace des états. Il s'ensuit que tous les résultats obtenus sont vrais pour toutes les valeurs du moment cinétique, entières s'il s'agit d'un moment orbital, demi-entières quand intervient des spins demi-entiers en nombre impair.

³⁶ $L_{\pm} = L_x \pm iL_y$.

2.3 Invariance de jauge

Les éléments de symétrie d'un système ne sont pas toujours de nature géométrique. De fait, la notion d'invariance par rapport à certaines transformations peut aussi résulter de la formalisation d'un problème donné en termes d'objets qui n'ont pas en eux-mêmes une signification physique directe, si toutefois les éléments de la réalité physique peuvent en être déduits.

C'est le cas de l'invariance de jauge pour le champ électromagnétique. Pour une particule de charge q soumise à un champ (\vec{E}, \vec{B}) , on sait poser d'emblée les équations du mouvement à l'aide des champs, en écrivant l'équation fondamentale de la Dynamique où figure l'expression de la force de Lorentz $\vec{F}_L = q(\vec{E} + \vec{v} \times \vec{B})$. D'un autre côté, il se révèle commode pour les calculs d'exprimer les champs eux-mêmes à l'aide des potentiels. Pour un même champ, on peut définir une infinité de potentiels se déduisant les uns des autres par ce que l'on appelle une transformation de jauge, et qui tous donnent les mêmes équations du mouvement en Mécanique Classique : les potentiels dépendent de la jauge choisie mais les champs eux-mêmes – et donc les forces – n'en dépendent pas. En d'autres termes, les équations de Maxwell sont invariantes quand on passe d'un certain potentiel (U, \vec{A}) à un autre, (U', \vec{A}') ³⁷. En effet, comme $\vec{B} = \vec{\nabla} \times \vec{A}$, on peut, sans changer le champ magnétique, ajouter à \vec{A} le gradient de n'importe quelle fonction de l'espace et du temps, $\phi(\vec{r}, t)$:

$$\vec{B} = \vec{\nabla} \times \vec{A} = \vec{\nabla} \times [\vec{A} + \vec{\nabla}\phi(\vec{r}, t)] \equiv \vec{\nabla} \times \vec{A}' . \quad (2.107)$$

Le champ électrique est :

$$\vec{E} = -\vec{\nabla}U - \frac{\partial \vec{A}}{\partial t} ; \quad (2.108)$$

on veut qu'il ne doit pas non plus altéré quand on remplace \vec{A} par \vec{A}' : il faut donc simultanément changer U en $U' = U - (\partial\phi/\partial t)$. Pour toute fonction ϕ , le changement des potentiels :

$$U'(\vec{r}, t) = U(\vec{r}, t) - \frac{\partial\phi}{\partial t} , \quad \vec{A}'(\vec{r}, t) = \vec{A}(\vec{r}, t) + \vec{\nabla}\phi(\vec{r}, t) \quad (2.109)$$

porte le nom de transformation de jauge et, par construction, laisse les champs \vec{E} et \vec{B} inaltérés.

On sait par ailleurs que le Lagrangien L pour une particule de charge q située dans un champ associé à (U, \vec{A}) est :

$$L = \frac{1}{2} m\vec{v}^2 - qU(\vec{r}, t) + q\vec{A}(\vec{r}, t) \cdot \vec{v} . \quad (2.110)$$

Pour un autre choix de jauge, le Lagrangien sera donc L' , différent de L :

$$L' = \frac{1}{2} m\vec{v}^2 - qU'(\vec{r}, t) + q\vec{A}'(\vec{r}, t) \cdot \vec{v} . \quad (2.111)$$

D'un autre côté, la vitesse et la position de la particule sont indépendantes de la jauge en Mécanique Classique, puisqu'elles se déduisent de l'équation fondamentale $m\dot{\vec{v}} = \vec{F}_L$, où les potentiels n'apparaissent pas : on dit que vitesse et position sont des quantités invariantes de jauge ; si on désigne par des primes les grandeurs dynamiques pertinentes quand on travaille dans la deuxième jauge, l'invariance de jauge de la position et de la vitesse en Mécanique Classique s'exprime suivant :

$$\vec{r}' = \vec{r} , \quad m\vec{v}' = m\vec{v} \quad (\text{classique}) . \quad (2.112)$$

Il en résulte immédiatement que le moment conjugué, lui, n'est pas invariant de jauge en Mécanique Classique :

$$\vec{p} \equiv \frac{\partial L}{\partial \vec{v}} = m\vec{v} + q\vec{A}(\vec{r}, t) , \quad (2.113)$$

$$\vec{p}' \equiv \frac{\partial L'}{\partial \vec{v}'} = m\vec{v}' + q\vec{A}'(\vec{r}, t) = m\vec{v} + q\vec{A}'(\vec{r}, t) , \quad (2.114)$$

³⁷Tout comme l'équation fondamentale de la dynamique ne change pas si on ajoute une constante additive à l'énergie potentielle.

d'où :

$$\vec{p}' = \vec{p} + q[\vec{A}'(\vec{r}, t) - \vec{A}(\vec{r}, t)] \neq \vec{p} \quad (\text{classique}) . \quad (2.115)$$

L'écart entre les moments conjugués dans les deux jauges n'implique que le potentiel-vecteur.

Examinons maintenant comment se traduit l'invariance de jauge en Mécanique Quantique. Les grandeurs dynamiques (\vec{r} , \vec{v} , \vec{p} , *etc.*) sont maintenant représentées par des opérateurs et l'état dynamique du système l'est par le vecteur d'état $|\Psi\rangle$. La transformation de jauge est une opération de symétrie, il existe donc un opérateur T reliant les deux vecteurs d'états représentant le système dans une jauge ou une autre :

$$|\Psi'\rangle = T|\Psi\rangle . \quad (2.116)$$

T est sûrement unitaire (pas antiunitaire), car il existe évidemment des transformations de jauge infiniment proches de la transformation identité. Continuant à mettre un prime (') pour les grandeurs associées à la jauge (U' , \vec{A}'), l'invariance requise par la symétrie de jauge s'exprime par l'égalité des valeurs moyennes de la position et de la vitesse (ou, de façon équivalente, de la quantité de mouvement). *A priori*, on doit donc écrire :

$$\langle \vec{r}' \rangle' = \langle \vec{r} \rangle \iff \int_{\mathbb{R}^3} d^3r \Psi'^*(\vec{r}, t) \vec{r}' \Psi'(\vec{r}, t) = \int_{\mathbb{R}^3} d^3r \Psi^*(\vec{r}, t) \vec{r} \Psi(\vec{r}, t) , \quad (2.117)$$

$$\langle m\vec{v}' \rangle' = \langle m\vec{v} \rangle \iff \int_{\mathbb{R}^3} d^3r \Psi'^*(\vec{r}, t) m\vec{v}' \Psi'(\vec{r}, t) = \int_{\mathbb{R}^3} d^3r \Psi^*(\vec{r}, t) m\vec{v} \Psi(\vec{r}, t) . \quad (2.118)$$

C'est bien ainsi que doit s'exprimer la symétrie de jauge en Mécanique Quantique, et non pas par l'invariance des opérateurs \vec{r} et \vec{v} qui n'ont pas *a priori* un sens physique direct – seules les *valeurs moyennes* quantiques en ont un. D'ailleurs et de toute façon, il est clair que les conditions (2.117) et (2.118) doivent être vraies dans la limite classique – laquelle s'exprime en termes de *valeurs moyennes* quantiques, comparées aux grandeurs mécaniques classiques.

À l'inverse, les relations canoniques de quantification ont été énoncées en toute généralité, indépendamment de toute référence à une jauge particulière. Il en résulte que, en représentation- q , on doit *dans toutes les jauges* utiliser la correspondance universelle de Schrödinger :

$$\vec{r} \longrightarrow \times \vec{r} , \quad \vec{p} \longrightarrow -i\hbar\vec{\nabla} ; \quad (2.119)$$

ce qui assure le maintien de la relation fondamentale $[q, p] = i\hbar$, et ceci quelle que soit la jauge. Il en résulte que \vec{r} et \vec{p} sont invariants de jauge par construction de la théorie ; par conséquent les opérateurs primés coïncident avec les opérateurs non primés (comparer avec (2.112)) :

$$\vec{r}' = \vec{r} , \quad \vec{p}' = \vec{p} \quad (\text{quantique}) . \quad (2.120)$$

En revanche, pour l'opérateur associé à la quantité de mouvement, on a (compte tenu de (2.120)) :

$$m\vec{v} = \vec{p} - q\vec{A}(\vec{r}, t) , \quad m\vec{v}' = \vec{p}' - q\vec{A}'(\vec{r}, t) = \vec{p} - q\vec{A}'(\vec{r}, t) . \quad (2.121)$$

soit (comparer avec (2.115)) :

$$m\vec{v}' = m\vec{v} + q[\vec{A}(\vec{r}, t) - \vec{A}'(\vec{r}, t)] \neq m\vec{v} \quad (\text{quantique}) . \quad (2.122)$$

En conséquence, l'opérateur associé à la quantité de mouvement $m\vec{v}$ n'est pas invariant de jauge ; ce n'est pas grave : en Mécanique Quantique, les prévisions ne s'expriment pas à l'aide des seuls opérateurs³⁸ mais impliquent simultanément le vecteur d'état.

En vertu de (2.120), et (2.121), les équations d'invariance (2.117) et (2.118) s'expriment alors comme :

$$\langle \Psi' | \vec{r}' | \Psi' \rangle = \langle \Psi | \vec{r} | \Psi \rangle , \quad \langle \Psi' | \vec{p}' - q\vec{A}' | \Psi' \rangle = \langle \Psi | \vec{p} - q\vec{A} | \Psi \rangle . \quad (2.123)$$

³⁸Le spectre d'un opérateur donne les valeurs possibles des résultats d'une mesure : il doit donc être invariant de jauge. C'est bien le cas puisque la relation $\Omega' = T\Omega T^\dagger$ est une relation d'équivalence assurant que les deux opérateurs Ω' et Ω ont les mêmes valeurs propres.

Avec (2.116) et puisque que ces relations sont vraies $\forall |\Psi\rangle$, elles peuvent être élevées au rang de relations entre opérateurs ; il en résulte :

$$T^\dagger \vec{r} T = \vec{r} , \quad (2.124)$$

et :

$$T^\dagger (\vec{p} - q\vec{A}') T = \vec{p} - q\vec{A} . \quad (2.125)$$

De (2.124), on déduit que la transformation T ne dépend que de \vec{r} (elle ne peut dépendre de \vec{p} , qui changerait forcément \vec{r}). T commute donc avec les potentiels-vecteurs, ce qui permet de transformer (2.125) comme suit (et tenant compte de (2.109)) :

$$T^\dagger \vec{p} T - q\vec{A}' = \vec{p} - q\vec{A} \iff T^\dagger \vec{p} T = \vec{p} + q\vec{\nabla}\phi . \quad (2.126)$$

T est une fonction de \vec{r} seul et est un opérateur unitaire ; on peut donc toujours l'écrire :

$$T = e^{iF(\vec{r}, t)} , \quad (2.127)$$

où F est une fonction à valeurs réelles. Dans ces conditions, la deuxième écriture de (2.126) devient :

$$e^{-iF(\vec{r}, t)} \vec{p} e^{iF(\vec{r}, t)} = \vec{p} + q\vec{\nabla}\phi . \quad (2.128)$$

Le premier membre est :

$$e^{-iF(\vec{r}, t)} \vec{p} e^{iF(\vec{r}, t)} = \vec{p} - i[F(\vec{r}, t), \vec{p}] + \frac{i^2}{2} [F(\vec{r}, t), [F(\vec{r}, t), \vec{p}]] + \dots = \vec{p} + \hbar \vec{\nabla}F(\vec{r}, t) , \quad (2.129)$$

où a été utilisé le fait que tous les commutateurs d'ordre supérieur à 1 sont nuls. Il faut donc :

$$\vec{p} + \hbar \vec{\nabla}F = \vec{p} + q\vec{\nabla}\phi \iff \vec{\nabla}F = \hbar^{-1} q\vec{\nabla}\phi \iff F(\vec{r}, t) = \frac{q}{\hbar} \phi(\vec{r}, t) + F_0(t) . \quad (2.130)$$

La fonction $F_0(t)$ donne un facteur de phase global additionnel à l'état transformé et peut donc être omise. Finalement, la fonction F à retenir est :

$$F(\vec{r}, t) = \frac{q}{\hbar} \phi(\vec{r}, t) , \quad (2.131)$$

et, selon (2.116), à l'état transformé $|\Psi'\rangle$ correspond la fonction d'onde :

$$\Psi'(\vec{r}, t) = e^{\frac{i}{\hbar} q \phi(\vec{r}, t)} \Psi(\vec{r}, t) . \quad (2.132)$$

Bien noter que le facteur de phase, parce qu'il dépend de l'espace, n'est pas trivial et ne peut de ce fait être ignoré. S'il y a non pas une mais plusieurs particules de charges q_1, q_2, \dots, q_N , la transformation de jauge entre fonctions d'onde s'écrit :

$$\Psi'(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t) = e^{\frac{i}{\hbar} \sum_n q_n \phi(\vec{r}_n, t)} \Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t) . \quad (2.133)$$

Montrons pour terminer que l'équation de Schrödinger est invariante de jauge, ce qui est bien la moindre des choses. Pour cela, il faut que les deux équations suivantes soient vérifiées :

$$i\hbar \frac{\partial}{\partial t} |\Psi\rangle = H|\Psi\rangle \quad (S) , \quad i\hbar \frac{\partial}{\partial t} |\Psi'\rangle = H'|\Psi'\rangle \quad (S') , \quad (2.134)$$

avec :

$$|\Psi'\rangle = e^{iF} |\Psi\rangle , \quad H = \frac{1}{2m} (\vec{p} - q\vec{A})^2 + qU , \quad H' = \frac{1}{2m} (\vec{p} - q\vec{A}')^2 + qU' . \quad (2.135)$$

Le premier membre de (S') est :

$$i\hbar \frac{\partial}{\partial t} e^{iF} |\Psi(t)\rangle = i\hbar \left(i \frac{\partial F}{\partial t} e^{iF} |\Psi(t)\rangle + e^{iF} \frac{\partial}{\partial t} |\Psi(t)\rangle \right) . \quad (2.136)$$

L'équation (S') devient donc :

$$i\hbar \frac{\partial}{\partial t} |\Psi\rangle = e^{-iF} \left(H' + \hbar \frac{\partial F}{\partial t} \right) e^{iF} |\Psi\rangle . \quad (2.137)$$

Il faut que ceci soit identique à l'équation (S) ; compte tenu de (2.131), on doit donc avoir :

$$e^{-iF} \left(H' + q \frac{\partial \phi}{\partial t} \right) e^{iF} = H ; \quad (2.138)$$

or, d'après (2.129) et (2.109) :

$$e^{-iF} (\vec{p} - q\vec{A}') e^{iF} = e^{-iF} \vec{p} e^{iF} - q\vec{A}' = \vec{p} + q\vec{\nabla}\phi - q\vec{A}' = \vec{p} - q\vec{A} ; \quad (2.139)$$

par ailleurs :

$$e^{-iF} \left(qU' + q \frac{\partial \phi}{\partial t} \right) e^{iF} = qU' + q \frac{\partial \phi}{\partial t} = qU ; \quad (2.140)$$

donc l'équation (2.138) est vérifiée, assurant que (S') est équivalente à (S), c'est-à-dire que les deux équations (S) et (S') sont simultanément vraies.

Pour terminer, notons que la densité $\rho = \Psi^* \Psi$ et le courant associé, \vec{j} , sont invariants de jauge, ce qui n'est pas surprenant puisque l'équation de Schrödinger elle-même contient l'équation de conservation. En ce qui concerne ρ , c'est évident puisque la phase e^{iF} disparaît dans le module au carré. Quant au courant, il suffit pour établir l'invariance de prendre soin du fait qu'il s'exprime à l'aide de la *vitesses*³⁹ $\vec{v} = \vec{p} - q\vec{A}$; le vecteur \vec{j} a donc pour expression :

$$\vec{j} = \frac{1}{2} [\Psi^* \vec{v} \Psi + \Psi \vec{v} \Psi^*] \equiv \Re [\Psi^* \vec{v} \Psi] = \Re \frac{1}{m} [\Psi^* (\vec{p} - q\vec{A}) \Psi] = \frac{1}{m} \Re [\Psi^* (-i\hbar\vec{\nabla} - q\vec{A}) \Psi] . \quad (2.141)$$

Dans l'autre jauge, on a :

$$\vec{j}' = \frac{1}{m} \Re [\Psi'^* (-i\hbar\vec{\nabla} - q\vec{A}') \Psi'] = \frac{1}{m} \Re [e^{-iF} \Psi^* (-i\hbar\vec{\nabla} - q\vec{A} - q\vec{\nabla}\phi) e^{+iF} \Psi] . \quad (2.142)$$

Le terme entre crochets à droite vaut :

$$\Psi^* [-i\hbar(\text{ie}^{-iF} \vec{\nabla} F + \vec{\nabla} \Psi) e^{+iF} - q\vec{A} - q\vec{\nabla}\phi] \Psi \quad (2.143)$$

d'où, compte tenu de (2.130) :

$$\vec{j}' = \frac{1}{m} \Re \left\{ \Psi^* \left[-i\hbar(\text{ie}^{-iF} \hbar^{-1} q\vec{\nabla}\phi + \vec{\nabla} \Psi) e^{+iF} - q\vec{A} - q\vec{\nabla}\phi \right] \Psi \right\} = \frac{1}{m} \Re [\Psi^* (\vec{p} - q\vec{A}) \Psi] \equiv \vec{j} . \quad (2.144)$$

2.4 Symétries discrètes

Toutes les symétries rencontrées jusqu'à présent sont associées à des transformations continues, au sens où celles-ci dépendent continûment d'un paramètre (angle de rotation, amplitude de la translation, charge électrique (formellement !) pour la transformation de jauge, *etc.*) En particulier, il existe des transformations infinitésimales, aussi proches que l'on veut de la transformation identité ; cette constatation permet de se convaincre que les opérateurs associés ne sauraient être antiunitaires : ils sont donc forcément unitaires.

Les opérations de symétrie ne sont pas toutes continues (elles peuvent être *discrètes*). Par exemple, pour un polygone régulier à N côtés, les symétries⁴⁰ sont les rotations de $2\pi/N$ autour d'un axe perpendiculaire au plan et passant par le centre du polygone ; le groupe correspondant est visiblement un sous-groupe (fini, lui) des rotations quelconques autour du même axe et est constitué d'opérateurs unitaires, quoique relatifs à des transformations discrètes. Pour un réseau régulier infini de points, il existe de même des translations

³⁹Tout comme l'énergie cinétique, qui vaut $(\vec{p} - q\vec{A})^2 / (2m)$.

⁴⁰Ce ne sont pas les seules.

finies superposant le réseau à lui-même, sous-groupe (infini dénombrable) du groupe (continu) des translations quelconques.

Il en va également ainsi de la réflexion par rapport à un plan (symétrie-miroir) qui, avec l'identité, forme un groupe à deux éléments. Il n'existe pas alors de transformation infinitésimale, ni de générateurs. Clairement, la réflexion change la chiralité et il est manifestement impossible de passer d'une forme chirale à une autre par continuité, sauf à inclure d'autres types de transformations, non considérées ici⁴¹. Un autre exemple est l'inversion d'espace, traditionnellement appelée parité en Mécanique Quantique.

Ces transformations discrètes sont encore représentées par des opérateurs unitaires. Par exemple, pour la parité (inversion d'espace), représentée par l'opérateur noté Π agissant dans l'espace des états, la fonction d'onde se transforme suivant :

$$\Psi'(\vec{r}, t) \equiv \Pi\Psi(\vec{r}, t) = \Psi(-\vec{r}, t) \quad (2.145)$$

et à la fois position et impulsion changent de signe :

$$\vec{r}' \equiv \Pi\vec{r}\Pi^\dagger = -\vec{r}, \quad \vec{p}' \equiv \Pi\vec{p}\Pi^\dagger = -\vec{p} \quad (2.146)$$

de sorte que les relations de commutation sont inchangées et Π est unitaire – ce que l'on peut tout autant vérifier en constatant l'invariance du produit scalaire (il suffit de faire un changement de variable muette dans les intégrales). Dans cette opération, le moment cinétique orbital ne change pas de signe non plus.

Les mêmes observations valent pour la symétrie-miroir⁴², représentée par M ; si \vec{n} est le vecteur unitaire normal au plan de réflexion, on a :

$$\vec{r}' \equiv M\vec{r}M^\dagger = \vec{r} - 2(\vec{n}\cdot\vec{r})\vec{n}, \quad \vec{p}' \equiv M\vec{p}M^\dagger = \vec{p} - 2(\vec{n}\cdot\vec{p})\vec{n} \quad (2.147)$$

et il est à nouveau facile de voir que les relations de commutation sont toujours inchangées, d'où l'on déduit que M est unitaire. Dans la symétrie-miroir par rapport au plan xOz , le report des transformés \vec{r}' et \vec{p}' (2.147) dans $\vec{L}' = \vec{r}' \times \vec{p}'$ donne :

$$L'_x = -L_x, \quad L'_y = L_y, \quad L'_z = -L_z. \quad (2.148)$$

Physiquement, il est clair que si le point tourne autour de Oz , la symétrie-miroir par rapport au plan xOz inverse le sens de rotation (et de même si la rotation a lieu autour de Ox) ; en revanche, la rotation autour de Oy (normal au plan) est inchangée. Au total, quoique discrètes, ces dernières symétries sont encore représentées par des opérateurs unitaires.

Il en va tout autrement pour une symétrie importante, la symétrie dans le renversement du temps. Cette opération, qui n'est pas propre à la Mécanique Quantique, revient à laisser invariantes les coordonnées mais à inverser les vitesses. Soit un système classique conservatif dont le Hamiltonien (par conséquent indépendant du temps) est, de surcroît, invariant dans le renversement des moments conjugués ; pour fixer les idées, ce peut être une particule de masse m dans un champ de forces usuel. En pareil cas, on a :

$$H(\vec{r}, \vec{p}) = \frac{\vec{p}^2}{2m} + V(\vec{r}). \quad (2.149)$$

Visiblement, puisque \vec{p} n'apparaît qu'au carré, la fonction H ne change pas : $H(\vec{r}, \vec{p}) = H(\vec{r}, -\vec{p})$. Il en résulte que si la fonction $\vec{r}(t)$ est solution des équations du mouvement classique, la fonction définie par :

$$\vec{r}_{\text{renv}}(t) = \vec{r}(-t) \quad (2.150)$$

est aussi solution des équations du mouvement dont les conditions initiales ont été elles aussi renversées (position inchangée, vitesse renversée). Schématiquement, l'invariance de l'équation fondamentale de la Dynamique (EFD) par renversement du temps résulte du fait qu'il s'agit d'une équation du second ordre par rapport au temps :

⁴¹La pyramide NH_3 peut évidemment passer d'une forme à l'autre en s'aplatissant, mais ceci fait appel à des déformations (continues).

⁴²La symétrie-miroir est aussi le produit d'une inversion et d'une rotation de π autour d'un axe perpendiculaire au miroir.

comme le temps figure “au carré”, changer t en $-t$ ne fait donc rien. De la relation précédente pour la coordonnée, on déduit, pour la vitesse et le moment conjugués du mouvement renversé :

$$\vec{v}_{\text{renv}}(t) = \frac{d}{dt}\vec{r}(-t) = (-1) \left(\frac{d\vec{r}(u)}{du} \right)_{u=-t} = -\vec{v}(-t) \iff \vec{p}_{\text{renv}}(t) = -\vec{p}(-t) . \quad (2.151)$$

Une deuxième dérivation en temps rétablit le signe + :

$$\frac{d}{dt}\vec{v}_{\text{renv}}(t) = -\frac{dv(-t)}{dt} = \frac{dv(t)}{dt} \quad (2.152)$$

et on retrouve bien l’EFD :

$$\frac{d}{dt}\vec{v}_{\text{renv}}(t) = \frac{dv(t)}{dt} = \vec{F}(\vec{r}) = \vec{F}(\vec{r}_{\text{renv}}) . \quad (2.153)$$

L’invariance par renversement du temps peut s’exprimer également d’une autre façon ; \vec{r} est une certaine fonction, f , paramétrée par les conditions initiales :

$$\vec{r} = g(t, t_0; \vec{r}_0, \vec{v}_0) . \quad (2.154)$$

Si on inverse les temps et les vitesses – en laissant évidemment les coordonnées inchangées – on retrouve la même coordonnée \vec{r} ; d’où :

$$g(-t, -t_0; \vec{r}_0, -\vec{v}_0) = g(t, t_0; \vec{r}_0, \vec{v}_0) , \quad (2.155)$$

égalité que l’on vérifie facilement dans des cas simples. Si l’on part d’un point A à l’instant $-t_1$ avec la vitesse \vec{v}_A , on arrive à $+t_1$ en un certain point B avec la vitesse \vec{v}_B . L’invariance dit que si l’on part à $-t_1$ du point B avec $-\vec{v}_B$, on arrive à $+t_1$ en A avec la vitesse $-\vec{v}_A$.

Considérons maintenant la situation quantique. En changeant t en $-t$ dans l’équation de Schrödinger, il vient :

$$i\hbar \frac{\partial}{\partial(-t)} \Psi(\vec{r}, -t) = \left[-\frac{\hbar^2}{2m} \Delta + V(\vec{r}) \right] \Psi(\vec{r}, -t) . \quad (2.156)$$

Pour récupérer le signe + au premier membre, il suffit d’en prendre le complexe conjugué ; par ailleurs, V est une fonction à valeurs réelles⁴³, et l’opérateur laplacien est aussi réel. Au total, en prenant les complexes conjugués des deux membres, on trouve :

$$i\hbar \frac{\partial}{\partial t} \Psi^*(\vec{r}, -t) = \left[-\frac{\hbar^2}{2m} \Delta + V(\vec{r}) \right] \Psi^*(\vec{r}, -t) . \quad (2.157)$$

Ainsi, avec le Hamiltonien apparaissant au second membre de (2.157), si $\Psi(\vec{r}, t)$ est solution, la fonction $\Psi^*(\vec{r}, -t)$ est aussi solution. Ceci conduit naturellement à définir d’une façon générale la fonction transformée de Ψ par renversement du temps $\Psi_{\text{renv}}(\vec{r}, t)$:

$$\Psi_{\text{renv}}(\vec{r}, t) = \Psi^*(\vec{r}, -t) . \quad (2.158)$$

En particulier, pour un état stationnaire, on a⁴⁴ :

$$\Psi_{n \text{ st, renv}}(\vec{r}, t) = \psi_n^*(\vec{r}) e^{-\frac{i}{\hbar} E_n(-t)} = \psi_n^*(\vec{r}) e^{\frac{i}{\hbar} E_n t} = \Psi_{n \text{ st}}^*(\vec{r}, -t) \quad (2.159)$$

et, pour une fonction indépendante du temps :

$$\psi_{n \text{ renv}}(\vec{r}) = \psi_n^*(\vec{r}) . \quad (2.160)$$

Dans la suite, on désignera par K l’opérateur effectuant le renversement du temps dans l’espace des états. Dans ces notations, par définition :

$$|\Psi_{\text{renv}}(t)\rangle = K|\Psi(t)\rangle . \quad (2.161)$$

Il est bien clair que K ne peut être un opérateur unitaire, à cause de la conjugaison complexe qui apparaît dans la définition (2.158) (voir aussi [17], p. 407). De fait, raisonnant par exemple sur la base propre de H , on voit

⁴³Dans le cas contraire, H ne serait pas hermitique, l’évolution par $U(t)$ ne serait pas unitaire, on perdrait de la probabilité, ..., etc.

⁴⁴Les énergies sont réelles !

que l'opérateur K change chaque état propre en son complexe conjugué (selon (2.160)) et transforme donc tout produit scalaire en son complexe conjugué : c'est par conséquent un opérateur *antiunitaire*.

Une fois définie la transformation K , il convient de se poser la question de la transformation des observables. Comme précédemment, il existe une symétrie universelle permettant de prendre en compte le fait que le sens du temps est, dans une théorie *mécanique*, totalement arbitraire⁴⁵. Toutes les valeurs moyennes obtenues quand un choix a été fait doivent coïncider avec celles trouvées avec le choix contraire⁴⁶. Il convient donc d'écrire, pour toute observable⁴⁷ A :

$$\langle A' \rangle = \langle A \rangle \iff (|\Psi'\rangle, A'|\Psi'\rangle) = (|\Psi\rangle, A|\Psi\rangle) \quad \forall |\Psi\rangle, \quad (2.162)$$

soit, en utilisant (2.161) et en tenant compte de la définition de l'opérateur adjoint d'un opérateur antiunitaire :

$$(|\Psi\rangle, A'K|\Psi\rangle) = (|\Psi\rangle, A|\Psi\rangle) \iff (|\Psi\rangle, K^\dagger A' K |\Psi\rangle)^* = (|\Psi\rangle, A|\Psi\rangle). \quad (2.163)$$

A étant une observable, sa valeur moyenne est réelle ; la dernière égalité donne donc après conjugaison complexe des deux membres :

$$(|\Psi\rangle, K^\dagger A' K |\Psi\rangle) = (|\Psi\rangle, A|\Psi\rangle) \quad \forall |\Psi\rangle \iff K^\dagger A' K = A. \quad (2.164)$$

Puisque K est antiunitaire, $K^\dagger = K^{-1}$ et l'on retrouve, même pour le cas antiunitaire, la relation habituelle pour la transformation des observables :

$$A' = K A K^\dagger. \quad (2.165)$$

Soit maintenant à trouver les transformés des grandeurs fondamentales \vec{r} et \vec{p} ; physiquement, on s'attend à obtenir que, tout comme dans le cas classique, ces variables dynamiques se transforment comme suit :

$$\vec{r}' \equiv K \vec{r} K^\dagger = \vec{r}, \quad \vec{p}' \equiv K \vec{p} K^\dagger = -\vec{p}. \quad (2.166)$$

(2.166) entraîne que les relations de commutation pour les variables transformées *changent de signe*, signature de l'antiunitarité de l'opérateur K .

Pour trouver \vec{r}' et \vec{p}' , plaçons-nous en représentation coordonnée, là où la transformation de renversement du temps a été définie en premier. Cette transformation ne touche pas aux coordonnées⁴⁸, \vec{r} est donc manifestement invariant d'où $\vec{r}' = \vec{r}$. En ce qui concerne l'impulsion, on a :

$$\vec{p}' = K (-i\hbar\vec{\nabla}) K^\dagger = +i\hbar K \vec{\nabla} K^\dagger. \quad (2.167)$$

Noter le changement de signe venant de la commutation de i avec K antiunitaire. Maintenant :

$$\forall \Psi(\vec{r}, t) : K \vec{\nabla} \Psi(\vec{r}, t) = K \frac{\partial \Psi(\vec{r}, t)}{\partial \vec{r}} = \left(\frac{\partial \Psi(\vec{r}, -t)}{\partial \vec{r}} \right)^* = \vec{\nabla} \Psi^*(\vec{r}, -t) = \vec{\nabla} K \Psi(\vec{r}, t), \quad (2.168)$$

autrement dit $[K, \vec{\nabla}] = 0$ de sorte que (2.167) devient :

$$\vec{p}' = +i\hbar \vec{\nabla} K K^\dagger = +i\hbar \vec{\nabla} \equiv -\vec{p} \quad (2.169)$$

qui donne le résultat attendu⁴⁹. En conséquence, le moment cinétique orbital est lui aussi changé en son opposé :

$$\vec{L}' = \vec{r}' \times \vec{p}' = -\vec{r} \times \vec{p} = -\vec{L}. \quad (2.170)$$

⁴⁵Le Second Principe est d'essence *statistique* et ne survient qu'après avoir renoncé, parce que c'est impossible et parce que ce serait absurde méthodologiquement parlant au vu de l'expérience, à décrire *tous* les degrés de liberté d'un système macroscopique.

⁴⁶Tout comme le sens positif sur l'axe Ox , choisi conventionnellement pour faire des calculs, n'a évidemment aucun sens physique par lui-même dans l'espace supposé homogène, isotrope, pair, *etc.*

⁴⁷Il est facile de voir que l'antiunitarité de K ne modifie pas le fait que les moyennes d'un opérateur hermitique transformé par K sont toujours des nombres réels.

⁴⁸Dans la définition (2.158), c'est le même argument \vec{r} qui figure dans les deux membres.

⁴⁹Une autre façon de se convaincre de la nécessité pour K d'être antiunitaire.

Ceci étant, par construction de K et conformément à (2.162), les moyennes sont invariantes :

$$\langle \vec{r}' \rangle' = \langle \vec{r} \rangle, \quad \langle \vec{p}' \rangle' = \langle \vec{p} \rangle, \quad \langle \vec{L}' \rangle' = \langle \vec{L} \rangle, \dots \quad (2.171)$$

H est une observable parmi d'autres et se transforme en $H' = KHK^\dagger$. Compte tenu de (2.169), H est invariant dès qu'il ne contient que des puissances paires de \vec{p} . Ce n'est pas le cas en présence d'un champ magnétique, où des termes du genre $\vec{A} \cdot \vec{p}$ apparaissent – ce sont d'ailleurs eux qui produisent le couplage sous la forme $-\vec{\mu} \cdot \vec{B}$, c'est-à-dire $\propto -\vec{L} \cdot \vec{B}$ par l'intermédiaire du facteur gyromagnétique ; or, précisément, le moment cinétique change de signe par K , donc H n'est pas invariant par renversement du temps. En revanche, en présence d'un champ électrique, H est invariant : un champ électrique lèvera donc *a priori* moins de dégénérescence qu'un champ magnétique. Physiquement, ceci est bien évident : un champ magnétique implique d'une façon ou d'une autre des boucles de courant, où interviennent des vitesses, qui changent de signe par renversement du temps.

Quand H est invariant par renversement du temps⁵⁰, $H' = H$ et la relation de commutation $[H, K] = 0$ se trouve *ipso facto* vérifiée ; toutefois, en raison du caractère antiunitaire de K , cela vaut la peine de s'en convaincre explicitement. Soit un état stationnaire d'énergie E_n , $\Psi_{n,\text{st}}(\vec{r}, t) = e^{E_n t / (\hbar)} \psi_n(\vec{r})$, satisfaisant :

$$H\Psi_{n,\text{st}}(\vec{r}, t) = E_n \psi_n(\vec{r}) e^{\frac{1}{\hbar} E_n t} \equiv E_n \Psi_{n,\text{st}}(\vec{r}, t) \quad (2.172)$$

On a d'abord :

$$HK\Psi_{n,\text{st}} = H\psi_n^* e^{-\frac{1}{\hbar} E_n (-t)} = H\psi_n^* e^{\frac{1}{\hbar} E_n t} \quad (2.173)$$

Comme H est réel – c'est pour ceci qu'il est invariant par K –, si on a $H\psi_n = E_n\psi_n$, on a également⁵¹, en prenant les complexes conjugués des deux membres, $H\psi_n^* = E_n\psi_n^*$. Il vient donc :

$$HK\Psi = E_n \psi_n^* e^{\frac{1}{\hbar} E_n t} \quad (2.174)$$

Ensuite, pour le produit KH :

$$KH\Psi = KE_n \psi_n e^{\frac{1}{\hbar} E_n t} = E_n \psi_n^* e^{-\frac{1}{\hbar} E_n (-t)} = E_n \psi_n^* e^{\frac{1}{\hbar} E_n t} \quad (2.175)$$

(2.174) et (2.175) montrent que HK et KH donnent le même résultat. La commutation de H avec K vaut tout autant pour une combinaison linéaire d'états stationnaires. Remarquons toutefois que la commutation de K avec H ne signifie pas pour autant que K est une “constante du mouvement” au sens défini antérieurement ; en effet, le lien bien connu entre commutation avec H et constance dans le temps ne tient que pour les opérateurs *linéaires*, alors que K est antilinéaire. De fait, on a :

$$\begin{aligned} \frac{d}{dt} \langle K \rangle &= \frac{d}{dt} (|\Psi\rangle, K|\Psi\rangle) = \left(\frac{1}{i\hbar} H|\Psi\rangle, K|\Psi\rangle \right) + (|\Psi\rangle, K \frac{1}{i\hbar} H|\Psi\rangle) = \frac{1}{-i\hbar} (H|\Psi\rangle, K|\Psi\rangle) + \frac{1}{-i\hbar} (|\Psi\rangle, KH|\Psi\rangle) \\ &= \frac{i}{\hbar} (|\Psi\rangle, (HK + KH)|\Psi\rangle) \equiv \frac{i}{\hbar} (|\Psi\rangle, \{H, K\}|\Psi\rangle) \end{aligned} \quad (2.176)$$

C'est donc l'*anticommutateur* $\{, \}$ qui apparaît dans la dérivée en temps de la valeur moyenne : le fait que le commutateur soit nul n'entraîne donc pas $d\langle K \rangle / dt = 0$.

Bien évidemment, appliquer deux fois consécutivement le renversement du temps doit n'avoir aucune conséquence observable. Ceci entraîne que l'on doit avoir :

$$K^2\Psi = e^{i\alpha}\Psi \quad (2.177)$$

où α une phase qui est la même pour tous les états. On peut alors écrire successivement, pour deux états quelconques $|\Phi\rangle$ et $|\Psi\rangle$:

$$(K\Phi, \Psi) = (K\Phi, K^\dagger K\Psi) \quad \text{puisque } K^\dagger K = \mathbf{1} \quad (2.178)$$

$$(K\Phi, K^\dagger K\Psi) = (K K\Phi, K\Psi)^* \quad \text{par définition de l'adjoint d'un opérateur antilinéaire} \quad (2.179)$$

⁵⁰Ceci est vrai si H est un opérateur réel. Comme $\vec{p} = -i\hbar\vec{\nabla}$, c'est le cas lorsque H ne dépend que du carré de \vec{p} .

⁵¹On note au passage que pour un Hamiltonien invariant par renversement du temps, ψ_n et ψ_n^* correspondent à la même énergie. Dès lors, si le niveau E_n est non-dégénéré, on peut toujours supposer que la fonction ψ_n est réelle ; si le niveau est dégénéré, on peut toujours, si besoin est, former deux combinaisons linéaires réelles $2^{-1/2}(\psi_n + \psi_n^*)$ et $i^{-1}2^{-1/2}(\psi_n - \psi_n^*)$.

d'où :

$$(K\Phi, \Psi) = (K^2\Phi, K\Psi)^* = (K\Psi, K^2\Phi) = (K\Psi, e^{i\alpha}\Phi) = e^{i\alpha}(K\Psi, \Phi) , \quad (2.180)$$

où (2.177) a été prise en compte. Les deux vecteurs $|\Phi\rangle$ et $|\Psi\rangle$ sont quelconques : on peut récrire (2.180) en les échangeant :

$$(K\Psi, \Phi) = e^{i\alpha}(K\Phi, \Psi) ; \quad (2.181)$$

En rapprochant (2.180) et (2.181), il vient :

$$(K\Phi, \Psi) = e^{i\alpha}(K\Psi, \Phi) = e^{i\alpha} e^{i\alpha}(K\Phi, \Psi) , \quad (2.182)$$

d'où $e^{2i\alpha} = 1$, soit $\alpha = 0$ ou π . En d'autres termes, pour un système donné (et selon la *nature* de ce système), on a :

$$K^2\Psi = \Psi \quad \text{ou} \quad K^2\Psi = -\Psi . \quad (2.183)$$

Faisant maintenant $\Phi = \Psi$ dans (2.180), on trouve l'égalité :

$$(K\Psi, \Psi) = (K^2\Psi, K\Psi)^* = (K\Psi, K^2\Psi) = e^{i\alpha}(K\Psi, \Psi) , \quad (2.184)$$

de sorte que dans le cas $\alpha = \pi$:

$$(K\Psi, \Psi) = -(K\Psi, \Psi) \iff (K\Psi, \Psi) = 0 . \quad (2.185)$$

Prenons pour Ψ l'un des états propres du système, ψ_n ; l'égalité (2.185) dit que ψ_n et $K\psi_n$ sont orthogonaux ; comme, pour un Hamiltonien invariant par renversement du temps, un état propre et son renversé correspondent à la *même* valeur propre, il en résulte que lorsque $\alpha = \pi$, les énergies sont dégénérées au moins deux fois, chaque doublet étant formé des deux états (orthogonaux) qui se correspondent l'un l'autre par renversement du temps. Cette association par paires dégénérées globalement invariantes par renversement du temps porte le nom de dégénérescence de Kramers et se rencontre pour les systèmes à nombre impair de fermions.

Remarque

Quand on revient à la définition première de K , (2.158) et (2.161), on ne voit pas bien comment K^2 peut être autre chose que l'identité, puisque conjuguer et/ou changer t en $-t$ sont des opérations binaires. En fait, cette définition du renversement du temps doit être généralisée afin de tenir compte du spin⁵². On verra de fait que la phase α dépend du nombre $N_{1/2}$ de particules de spin demi-entier et vaut précisément $N_{1/2}\pi$ de sorte que :

$$K^2 = e^{iN_{1/2}\pi} = (-1)^{N_{1/2}} . \quad (2.186)$$

2.5 Symétrie et dégénérescence

Soit un système possédant une symétrie (invariance) vis-à-vis de transformations \mathcal{G}_i , géométriques ou non, représentées par des opérateurs unitaires ou antiunitaires notés G_i . D'après ci-dessus, on a par hypothèse :

$$H' = G_i H G_i^\dagger = H \iff [H, G_i] = 0 . \quad (2.187)$$

Maintenant, soit $|\psi_n\rangle$ un état propre de H associé à l'énergie E_n . Quand on applique \mathcal{G}_i à un système situé dans l'état $|\psi_n\rangle$, ce vecteur se transforme en un vecteur $|\psi'_n\rangle$, *a priori* différent de $|\psi_n\rangle$ et tel que :

$$|\psi'_n\rangle = G_i |\psi_n\rangle . \quad (2.188)$$

On va montrer que $|\psi'_n\rangle$ est encore propre de H , associé à la même énergie E_n – heureusement d'ailleurs : les deux états $|\psi'_n\rangle$ et $|\psi_n\rangle$ étant image et objet l'un de l'autre dans une opération de symétrie, il faut bien que l'énergie ne change pas⁵³. En effet :

$$[H, G_i]|\psi_n\rangle = 0 \iff H G_i |\psi_n\rangle = G_i H |\psi_n\rangle \iff H |\psi'_n\rangle = E_n |\psi'_n\rangle . \quad (2.189)$$

⁵²Tout ce qui a été dit en considérant une seule fonction d'onde – c'est-à-dire un état à une seule composante – reste donc valide pour une particule de spin nul.

⁵³Sinon les deux "positions" du système ne seraient pas équivalentes, et on pourrait les distinguer justement par leur énergie.

Remarquons que cette chaîne d'équations est vraie même si G_i est antiunitaire, puisque de toute façon E_n est une quantité réelle.

En toute hypothèse, il y a donc deux cas possibles :

- ou bien $|\psi'_n\rangle$ coïncide en fait avec $|\psi_n\rangle$ et alors le niveau E_n n'est pas dégénéré. Ce cas doit être considéré comme relativement exceptionnel, surtout pour les problèmes dans plus d'une dimension d'espace.
- ou bien $|\psi'_n\rangle$ est différent de $|\psi_n\rangle$ et alors, nécessairement, le niveau E_n présente une dégénérescence. Ceci est le cas le plus fréquent et implique l'existence de différents états associés à la même énergie ; pour distinguer ces états, on peut les noter $|\psi_{n,\alpha}\rangle$:

$$H|\psi_{n,\alpha}\rangle = E_n |\psi_{n,\alpha}\rangle . \quad (2.190)$$

Ce qui précède permet d'affirmer que toute opération de symétrie transforme les $|\psi_{n,\alpha}\rangle$ les uns dans les autres puisque, précisément, $G_i|\psi_{n,\alpha}\rangle$ produit un vecteur propre d'énergie E_n , qui est la même que celle de $|\psi_{n,\alpha}\rangle$; autrement dit, on doit pouvoir écrire l'un d'entre eux en combinaison linéaire des autres :

$$G_i|\psi_{n,\alpha}\rangle = \sum_{\beta} G_{i,\beta\alpha} |\psi_{n,\beta}\rangle , \quad (2.191)$$

où les $G_{i,\beta\alpha}$ sont des scalaires. Cette observation est le point de départ de la théorie de la représentation linéaire des groupes.

Chapitre 3

Théorie du moment cinétique

Ce chapitre est consacré à la théorie générale du moment cinétique en Mécanique Quantique.

3.1 Importance du moment cinétique

On sait l'importance du moment cinétique \vec{L} en Mécanique Classique, exprimée par le théorème dit du moment cinétique :

$$\frac{d}{dt} \vec{L} = \vec{r} \times \vec{F} . \quad (3.1)$$

Ce théorème est une conséquence immédiate de l'équation fondamentale de la dynamique. En particulier, pour une force centrale, associée à une énergie potentielle V ne dépendant que de la distance au centre, $V(r)$, le produit vectoriel est nul et \vec{L} est une constante du mouvement. C'est pourquoi la trajectoire est plane, dans un plan perpendiculaire à la valeur constante de \vec{L} , fixée une fois pour toutes par les conditions initiales ; c'est aussi la raison pour laquelle la trajectoire obéit à la loi des aires : la surface Σ balayée par le rayon-vecteur \vec{r} entre deux instants t_0 et t est proportionnelle au temps ($\Sigma = \|\vec{L}\| (t - t_0)/m$).

Bien évidemment, le moment cinétique joue également un rôle fondamental en Mécanique Quantique, et ce d'autant plus que des expériences cruciales ont montré qu'il est quantifié (par exemple : expériences de Stern et Gerlach), tout comme l'énergie des états liés d'un système. En outre, c'est le moment cinétique (orbital ou de spin) qui permet de comprendre la nature des sources du magnétisme, ce que la Physique classique est à nouveau dans l'incapacité de décrire.

Par ailleurs, le moment cinétique joue un rôle central en tant que générateur des rotations spatiales, comme on l'a vu dans le chapitre précédent. L'importance de la symétrie en Mécanique Quantique n'est plus à démontrer.

Enfin, force est d'admettre (au vu notamment de certaines complexités des spectres atomiques) que les "particules" possèdent un degré de liberté intrinsèque, en tout point analogue à un moment cinétique, mais sans équivalent classique ; ce degré de liberté a été appelé *spin*¹ et est noté traditionnellement \vec{S} . C'est d'ailleurs ce degré de liberté qui conduit à définir deux grandes classes de particules, les fermions et les bosons, dont les comportements, radicalement différents, produisent des effets spectaculaires². Pour rappeler sa nature "conventionnelle", \vec{L} sera souvent plus précisément appelé moment cinétique orbital, évoquant le fait qu'il est

¹À l'époque où cette terminologie a été choisie, on a imaginé que ce moment cinétique représentait une rotation de la particule sur elle-même, ce que les astronomes désignent par spin. En réalité, le spin d'une particule n'est susceptible d'aucune représentation de cette sorte : c'est un effet spécifiquement quantique, sans contrepartie classique.

²C'est parce que les électrons sont des fermions que les métaux sont conducteurs de l'électricité. C'est aussi parce qu'ils peuvent se condenser par paires que certains métaux deviennent supraconducteurs à (très) basse température.

l'analogie quantique d'un moment cinétique d'une particule qui, classiquement, serait en orbite autour d'un centre fixe.

\vec{L} est un opérateur ayant trois composantes L_x , L_y et L_z ($L_x = yp_z - zp_y$, etc.) dont la définition à partir de son expression classique ne pose aucun problème puisque dans chaque produit, les deux opérateurs commutent. Il en résulte immédiatement que ces composantes sont, telles quelles, des opérateurs hermitiques. En utilisant la relation fondamentale $[u, p_v] = i\hbar \delta_{uv}$, il est très facile de montrer que L_x , L_y et L_z satisfont les relations de commutation suivantes, qui se déduisent les unes des autres par permutations circulaires :

$$[L_x, L_y] = i\hbar L_z, \quad [L_y, L_z] = i\hbar L_x, \quad [L_z, L_x] = i\hbar L_y. \quad (3.2)$$

Le \hbar au second membre est bien nécessaire pour l'homogénéité ; quant au i , il ne doit pas surprendre : tous les opérateurs sont hermitiques, mais le commutateur de deux opérateurs hermitiques est antihermitique³.

Par contraste, on désignera habituellement par \vec{S} un moment cinétique de spin⁴, qui n'a pas d'équivalent classique. Les trois composantes de \vec{S} satisfont les mêmes relations de commutation que celles de \vec{L} : c'est pourquoi il est possible de faire une théorie générale du moment cinétique en Mécanique Quantique ; quand la nature (orbitale ou de spin) est non-pertinente, l'usage est de désigner par \vec{J} un tel moment cinétique "abstrait". Ses trois composantes satisfont les relations caractéristiques :

$$[J_x, J_y] = i\hbar J_z, \quad [J_y, J_z] = i\hbar J_x, \quad [J_z, J_x] = i\hbar J_y. \quad (3.3)$$

À partir de (3.3), il est facile d'établir que :

$$[J_u, \vec{J}^2] = 0 \quad (u = x, y, z; \vec{J}^2 = J_x^2 + J_y^2 + J_z^2) \iff [\vec{J}, \vec{J}^2] = 0. \quad (3.4)$$

Il existe donc des vecteurs propres communs à \vec{J}^2 et à l'une de ses autres composantes ; l'habitude est de choisir le couple (\vec{J}^2, J_z) , mais il est clair que ceci est purement conventionnel. En outre, il se révèle très commode d'introduire les deux combinaisons remarquables :

$$J_+ = J_x + iJ_y, \quad J_- = J_x - iJ_y, \quad (3.5)$$

J_+ et J_- sont hermitiques conjugués l'un de l'autre, mais leur produit est hermitique :

$$J_+^\dagger = J_-^\dagger, \quad (J_- J_+)^\dagger = J_- J_+, \quad (J_+ J_-)^\dagger = J_+ J_-. \quad (3.6)$$

Les J_\pm satisfont des relations déduites facilement de (3.3) :

$$[J_z, J_+] = +\hbar J_+, \quad [J_z, J_-] = -\hbar J_-, \quad [J_+, J_-] = 2\hbar J_z, \quad [J_\pm, \vec{J}^2] = 0. \quad (3.7)$$

On a aussi :

$$J_x^2 + J_y^2 = \frac{1}{2}(J_+ J_- + J_- J_+), \quad \vec{J}^2 = J_z^2 + \frac{1}{2}(J_+ J_- + J_- J_+) \quad (3.8)$$

Avec $J_+ J_- = J_- J_+ + 2\hbar J_z$ et (3.8), on trouve :

$$\vec{J}^2 = J_z^2 + \hbar J_z + J_- J_+ = J_z^2 - \hbar J_z + J_+ J_- \iff J_\pm J_\mp = \vec{J}^2 - J_z^2 \pm \hbar J_z \quad (3.9)$$

Les relations de commutation permettent aussi d'écrire des inégalités portant sur les écarts quadratiques des composantes du moment cinétique. Avec la définition conventionnelle $\Delta A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2}$ et compte tenu du théorème général :

$$[A, B] = iC \implies \Delta A \Delta B \geq \frac{1}{2} |\langle C \rangle|, \quad (3.10)$$

on a :

$$\Delta J_x \Delta J_y \geq \frac{1}{2} |\langle J_z \rangle| \quad (\text{et permutations circulaires}). \quad (3.11)$$

³au sens où $\Omega^\dagger = -\Omega$.

⁴On verra au ch. 5 pourquoi il est nécessaire physiquement de considérer que \vec{S} est effectivement un moment cinétique.

On ne peut, en général, trouver des vecteurs propres communs à J_x , J_y et J_z . Il existe toutefois des états exceptionnels $|\psi\rangle$ où les écarts quadratiques des *trois* composantes sont simultanément nuls. D'après (3.11), une condition nécessaire pour ceci est $\langle \vec{J} \rangle = 0$; pour avoir également $\Delta \vec{J} = 0$, il faut de surcroît $\langle J_x^2 \rangle = \langle J_y^2 \rangle = \langle J_z^2 \rangle = 0$. Pour un tel état $|\psi\rangle$, on a :

$$\vec{J}|\psi\rangle = 0 . \quad (3.12)$$

Ces états possèdent la symétrie sphérique.

Enfin, une autre relation remarquable est l'inégalité :

$$\langle \vec{J}^2 \rangle \geq \langle J_u^2 \rangle \geq \langle J_u \rangle^2 , \quad (3.13)$$

qui résulte de $\vec{J}^2 - J_z^2 = J_x^2 + J_y^2$; l'opérateur au second membre étant positif, le premier l'est tout autant et la valeur moyenne d'un opérateur positif est un nombre positif. En fait, l'égalité (3.13) n'est obtenue que si les deux autres composantes sont de carré moyen nul, correspondant à l'état exceptionnel $|\psi\rangle$ considéré ci-dessus. Dans tous les autres cas, on a :

$$\langle \vec{J}^2 \rangle > \langle J_u \rangle^2 \iff \sqrt{\langle \vec{J}^2 \rangle} > \langle J_z \rangle , \quad (3.14)$$

et les relations analogues avec J_x et J_y : la valeur moyenne du module d'un moment cinétique est donc toujours plus grande que la valeur moyenne de n'importe laquelle de ses composantes. Le vecteur \vec{J} n'est jamais en moyenne colinéaire à l'un des axes de coordonnées.

3.2 Propriétés générales des valeurs et vecteurs propres d'un moment cinétique

On va montrer, de façon purement algébrique⁵, en quoi consiste le spectre d'un moment cinétique. Ce qui va être établi vaut autant pour le moment orbital que pour le moment de spin : seules les relations de commutation caractéristiques (3.3) vont être utilisées, avec, sous-jacente, l'hypothèse que les vecteurs propres sont tous des vecteurs de norme finie – et comme d'habitude fixée conventionnellement à l'unité. La méthode utilisée est une méthode d'échelle (ou d'escalier) en bien des points similaire à celle qui a été utilisée pour la quantification de l'oscillateur harmonique et introduisant naturellement les opérateurs de création et d'annihilation.

Comme \vec{J}^2 et J_z commutent, il est possible de leur trouver des vecteurs propres communs ; désignons ceux-ci par $|\lambda, m\rangle$: λ est un nombre pur qui donne la valeur propre associée à l'opérateur \vec{J}^2 , m donne de même celle associée à J_z :

$$\vec{J}^2 |\lambda, m\rangle = \lambda \hbar^2 |\lambda, m\rangle , \quad J_z |\lambda, m\rangle = m \hbar |\lambda, m\rangle . \quad (3.15)$$

Les nombres λ et m sont réels (les opérateurs sont hermitiques) mais à part ceci quelconques ; toutefois, en conséquence de l'inégalité (3.13), ils satisfont :

$$\lambda \geq m^2 , \quad (3.16)$$

ce qui montre (sans surprise) que λ est positif. Par ailleurs, le sens positif le long de l'axe Oz est purement conventionnel : la symétrie galiléenne exige donc que, pour λ fixé, si la valeur m est possible, la valeur $-m$ est également possible.

Mettons maintenant en œuvre la procédure d'échelle annoncée ([17], p. 375 *et sq.*) ; partons de la relation $J_z J_+ = J_+ J_z + \hbar J_+$ (voir (3.7)) appliquée à $|\lambda, m\rangle$; compte tenu de la deuxième relation (3.15) elle donne :

$$J_z J_+ |\lambda, m\rangle = (J_+ J_z + \hbar J_+) |\lambda, m\rangle = (m + 1) \hbar J_+ |\lambda, m\rangle . \quad (3.17)$$

⁵ par opposition à une méthode purement analytique pour le moment orbital, conduisant aux harmoniques sphériques (voir section 3.3).

De la même façon, en partant de $J_z J_- = J_- J_z - \hbar J_-$, on obtient :

$$J_z J_- |\lambda, m\rangle = (m - 1)\hbar J_- |\lambda, m\rangle . \quad (3.18)$$

Enfin, avec $[\vec{J}^2, J_{\pm}] = 0$, on trouve :

$$J^2 J_{\pm} |\lambda, m\rangle = \lambda \hbar^2 J_{\pm} |\lambda, m\rangle . \quad (3.19)$$

Ainsi, $|\lambda, m\rangle$ étant un vecteur propre de (\vec{J}^2, J_z) avec les valeurs propres $\lambda \hbar^2$ et $m \hbar$, $J_{\pm} |\lambda, m\rangle$ est également un vecteur propre de ces deux opérateurs, mais associé aux valeurs propres $\lambda \hbar^2$ et $(m \pm 1)\hbar$. Les deux relations suivantes sont donc vraies :

$$J_{\pm} |\lambda, m\rangle = C_{\pm}(\lambda, m)\hbar |\lambda, m \pm 1\rangle , \quad (3.20)$$

où les $C_{\pm}(\lambda, m)$ sont des fonctions numériques de λ et m à déterminer.

Précisons maintenant le lien entre λ et m . Compte tenu de (3.16), la valeur absolue de m est bornée quand λ est donné. Désignons par j la valeur maximale de m et par $|\lambda, j\rangle$ le vecteur propre correspondant – qui n'est pas le vecteur nul ; à ce stade on sait que (voir (3.16)) :

$$m_{\max} \equiv j \leq \sqrt{\lambda} . \quad (3.21)$$

Comme J_+ augmente le nombre m d'une unité, l'application de J_+ à ce dernier vecteur ne peut produire un nouveau vecteur propre ; on doit donc avoir :

$$J_+ |\lambda, j\rangle = 0 . \quad (3.22)$$

Appliquons maintenant J_- à cette équation et prenons en compte (3.9) :

$$J_- J_+ |\lambda, j\rangle = 0 \iff (\vec{J}^2 - J_z^2 - \hbar J_z) |\lambda, j\rangle = 0 , \quad (3.23)$$

soit :

$$(\lambda - j^2 - j)\hbar^2 |\lambda, j\rangle = 0 \iff \lambda = j^2 + j = j(j + 1) . \quad (3.24)$$

Ceci donne la relation précise entre la valeur maximale de m , j , et le nombre λ donnant la valeur propre de \vec{J}^2 ; j étant positif ou nul, l'inégalité (3.21) est visiblement satisfaite.

De la même façon, soit m_{\min} la plus petite valeur prise par m ; notons-la j' et par définition $-\sqrt{\lambda} \leq j' \leq j$; une argumentation analogue montre que :

$$J_- |\lambda, j'\rangle = 0 \implies J_+ J_- |\lambda, j'\rangle = 0 \iff (\vec{J}^2 - J_z^2 + \hbar J_z) |\lambda, j'\rangle = 0 , \quad (3.25)$$

soit :

$$(\lambda - j'^2 + j')\hbar^2 |\lambda, j'\rangle = 0 \iff \lambda = j'^2 - j' = j'(j' - 1) , \quad (3.26)$$

ce qui fournit une relation précise entre j' et λ . Les deux expressions (3.24) et (3.26) de λ ne sont compatibles que si $j(j + 1) = j'(j' - 1)$, soit $j' = -j$ ou si $j' = j + 1$. Cette dernière possibilité est à exclure, car elle dit que j' est plus grand que j . Ainsi arrivent les résultats importants :

$$\lambda = j(j + 1) , \quad (3.27)$$

$$-j \leq m \leq j . \quad (3.28)$$

Il est ainsi possible d'atteindre $|\lambda, -j\rangle$ – le vecteur propre ayant la plus petite valeur de m – à partir de $|\lambda, j\rangle$ en appliquant l'opérateur J_- un nombre suffisant de fois. Comme à chaque application de J_- , la valeur propre m décroît d'une unité, la différence $j - j' = 2j$ est nécessairement un entier positif ou nul. Par ailleurs, comme deux valeurs m et $-m$ doivent être simultanément présentes, les valeurs de m sont réparties symétriquement autour de zéro. D'où le résultat majeur : j ne peut être qu'un nombre *entier ou demi-entier* :

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \dots \iff \lambda \equiv j(j + 1) = 0, \frac{3}{4}, 2, \frac{15}{4}, 6, \frac{30}{4}, \dots . \quad (3.29)$$

Pour un j donné, les valeurs propres possibles de J_z sont en nombre égal à $2j + 1$ et sont données par :

$$j\hbar, (j-1)\hbar, (j-2)\hbar, \dots, (-j+1)\hbar, -j\hbar . \quad (3.30)$$

Ainsi, le nombre quantique⁶ m prend des valeurs entières si j est entier, demi-entières si j est demi-entier :

$$m = j, j-1, j-2, \dots, -j+1, -j . \quad (3.31)$$

À j fixé, m varie donc exactement d'une unité quand on passe d'un vecteur propre au suivant.

Trouvons maintenant les scalaires C_{\pm} introduits en (3.20). Le carré scalaire de (3.20) est :

$$|C_+(\lambda, m)|^2 \hbar^2 (|m+1\rangle, |m+1\rangle) = (J_+|\lambda, m\rangle, J_+|\lambda, m\rangle) = \langle \lambda, m | J_- J_+ | \lambda, m \rangle . \quad (3.32)$$

En supposant tous les vecteurs propres préalablement normalisés à l'unité, il vient :

$$|C_+(\lambda, m)|^2 \hbar^2 (|m+1\rangle, |m+1\rangle) = \langle \lambda, m | J_- J_+ | \lambda, m \rangle = [j(j+1) - m^2 - m] \hbar^2 . \quad (3.33)$$

De ceci, on ne peut pas déduire la phase de C_+ , qui est sans importance. Par usage, on adopte le choix le plus simple en posant :

$$C_+(\lambda, m) = \sqrt{j(j+1) - m(m+1)} . \quad (3.34)$$

Par des moyens analogues, on trouve :

$$C_-(\lambda, m) = \sqrt{j(j+1) - m(m-1)} . \quad (3.35)$$

Maintenant que la relation entre λ et j est précisée, changeons légèrement les notations afin de se conformer à l'usage et notons désormais $|j, m\rangle$ les états propres $|\lambda, m\rangle$ (bien que $\lambda \neq j$, si $j \neq 0$) :

$$J_z |j, m\rangle = m\hbar |j, m\rangle , \quad \vec{J}^2 |j, m\rangle = j(j+1)\hbar^2 |j, m\rangle . \quad (3.36)$$

Dans ces nouvelles notations, les relations fondamentales à retenir sont :

$$J_{\pm} |j, m\rangle = \hbar \sqrt{j(j+1) - m(m \pm 1)} |j, m \pm 1\rangle . \quad (3.37)$$

Elles permettent d'engendrer les $2j + 1$ vecteurs propres de (\vec{J}^2, J_z) , pour un j donné, à partir de l'un d'entre eux. Notons que ces dernières relations s'écrivent aussi :

$$J_{\pm} |j, m\rangle = \hbar \sqrt{(j \mp m)(j \pm m + 1)} |j, m \pm 1\rangle . \quad (3.38)$$

Il en sera fait un usage intensif dans la suite et dans les applications.

Il est clair que, à j fixé, les $2j + 1$ vecteurs forment un sous-espace dégénéré pour la valeur propre de \vec{J}^2 . Toujours à j fixé, on voit que ce sous-espace est stable par l'application de J_x et J_y puisque $J_{\pm} |j, m\rangle$ est un vecteur du même sous-espace. En l'absence d'autres degrés de liberté, \vec{J}^2 et J_z forment un ECO⁷, puisque la donnée des valeurs propres (représentées par j et m) spécifie complètement leur état propre commun.

Remarques

La relation (3.14) prend maintenant la forme explicite :

$$\sqrt{\langle \vec{J}^2 \rangle} = \sqrt{j(j+1)} \hbar > j\hbar = \hbar \max_{-j \leq m \leq +j} m = \langle J_z \rangle . \quad (3.39)$$

Le rapport $\langle J_z \rangle / \sqrt{\langle \vec{J}^2 \rangle}$ qui évoque le cosinus de l'angle entre le moment cinétique et sa projection sur Oz est donc toujours strictement plus petit que 1 : sa plus grande valeur est $\sqrt{\frac{j}{j+1}}$ (elle tend vers 1 à la limite classique caractérisée par les grands nombres quantiques).

⁶appelé nombre quantique magnétique pour des raisons claires par la suite.

⁷Suivant la définition de [4], II. D, p. 144.

J_x et J_y ont toujours une moyenne nulle dans un état $|\lambda, m\rangle$ (ce sont des combinaisons linéaires des opérateurs J_{\pm} qui changent m d'une unité). Le petit dessin que l'on rencontre souvent (un vecteur \vec{J} avec une orientation quantifiée) peut être trompeur, sauf à imaginer que c'est un instantané du vecteur tournant à toute vitesse autour de Oz .

En revanche, l'écart-type de J_x et J_y est fini. Dans ce même type d'état, on a :

$$\Delta J_x = \Delta J_y = \frac{\hbar}{\sqrt{2}} \sqrt{j(j+1) - m^2} > \hbar \sqrt{\frac{j}{2}} . \quad (3.40)$$

3.3 Etats propres du moment cinétique orbital

Le moment cinétique orbital, noté \vec{L} , est un moment cinétique particulier. On peut donc d'ores et déjà affirmer que ses états propres, notés $|l, m\rangle$ selon l'usage, satisfont :

$$\vec{L}^2 |l, m\rangle = l(l+1)\hbar^2 |l, m\rangle , \quad L_z |l, m\rangle = m\hbar |l, m\rangle , \quad (3.41)$$

où l est un entier (ou un demi-entier, pour l'instant on n'en sait pas plus) positif ou nul, et m un entier (ou un demi-entier) assujéti aux inégalités $-l \leq m \leq +l$. Toutefois, on va montrer que, pour un moment cinétique *orbital* – qui est égal à un produit $\vec{r} \times \vec{p}$ –, seules les valeurs *entières* de l (et donc de m) sont possibles.

\vec{L}^2 et L_z sont deux opérateurs dont on connaît l'expression en représentation- q , par exemple ; il est donc facile d'écrire les équations différentielles exprimant les équations aux valeurs propres et de les résoudre.

Le moment cinétique orbital joue un rôle fondamental dans tous les problèmes à force centrale, qui possèdent de ce fait⁸ une symétrie sphérique. En pareil cas, les coordonnées adaptées au problème sont évidemment les coordonnées sphériques, définies comme d'habitude par r, θ et ϕ ($r \geq 0, 0 \leq \phi \leq 2\pi, 0 \leq \theta \leq \pi$) :

$$x = r \sin \theta \cos \phi , \quad y = r \sin \theta \sin \phi , \quad z = r \cos \theta , \quad (3.42)$$

l'élément de volume $d^3r = dx dy dz$ étant égal à :

$$d^3r = dx dy dz = r^2 \sin \theta dr d\theta d\phi . \quad (3.43)$$

On verra en effet que \vec{L} n'agit que sur les angles θ et ϕ , ce qui est *a priori* évident quand on se souvient que \vec{L} entre dans l'expression des opérateurs de rotation et qu'aucune rotation ne modifie la distance r au centre : \vec{L} ne saurait donc altérer r .

On connaît au départ les expressions des composantes cartésiennes de \vec{L} , que l'on sait exprimer à l'aide des coordonnées cartésiennes et de leurs moments conjugués. Par exemple :

$$L_x = yp_z - zp_y = -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \quad (3.44)$$

et de même pour L_y et L_z . En procédant comme d'habitude quand on passe d'un jeu de variables à un autre, il est facile, mais un peu pénible, d'obtenir L_x, L_y et L_z exprimés en coordonnées sphériques. On trouve :

$$L_x = i\hbar \left(\sin \phi \frac{\partial}{\partial \theta} + \cos \phi \cot \theta \frac{\partial}{\partial \phi} \right) , \quad L_y = i\hbar \left(-\cos \phi \frac{\partial}{\partial \theta} + \sin \phi \cot \theta \frac{\partial}{\partial \phi} \right) . \quad (3.45)$$

$$L_z = -i\hbar \frac{\partial}{\partial \phi} . \quad (3.46)$$

L'expression de L_z ne doit pas surprendre, et peut d'ailleurs s'obtenir très vite en raisonnant comme suit. On sait du ch. 2 que l'opérateur de la rotation d'un angle ϕ autour de Oz est $e^{\frac{i}{\hbar} \phi L_z}$; pour une rotation infinitésimale

⁸sauf adjonction éventuelle de champs extérieurs.

d'angle $\delta\phi$, l'opérateur de rotation est donc $\mathbf{1} + \frac{\delta\phi}{i\hbar} L_z$. Par ailleurs, pour toute fonction $f(\phi)$ sa transformée par cette rotation est $f(\phi - \delta\phi) = f(\phi) - \delta\phi f'(\phi) + \dots$. D'où l'égalité :

$$[\mathbf{1} + \frac{\delta\phi}{i\hbar} L_z + \dots]f(\phi) = f(\phi) - \delta\phi f'(\phi) + \dots, \quad (3.47)$$

d'où, par identification, l'expression (3.46) de L_z .

Il est facile de vérifier que les expressions (3.45) et (3.46) satisfont les relations de commutation (3.2) ; elles permettent d'obtenir :

$$\vec{L}^2 = -\hbar^2 \left(\frac{\partial^2}{\partial\theta^2} + \cot\theta \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right). \quad (3.48)$$

$$L_{\pm} = \pm\hbar e^{\pm i\phi} \left(\frac{\partial}{\partial\theta} \pm i \cot\theta \frac{\partial}{\partial\phi} \right). \quad (3.49)$$

En outre, on vérifie aussi, par des intégrations par parties, que les L_u sont hermitiques dans l'espace des fonctions 2π -périodiques, et que $(L_-)^\dagger = L_+$, comme il se doit.

Une fois ces expressions obtenues, on peut écrire explicitement les équations aux valeurs et fonctions propres pour \vec{L}^2 et L_z , qui commutent entre eux. Leurs fonctions propres communes, indicées par les nombres quantiques l et m , associées aux valeurs propres selon (3.41), sont traditionnellement désignées par la lettre Y et portent le nom d'harmoniques sphériques. Ainsi, les fonctions cherchées, $Y_{lm}(\theta, \phi)$, satisfont les deux équations aux dérivées partielles :

$$-\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right) Y_{lm}(\theta, \phi) = l(l+1) Y_{lm}(\theta, \phi), \quad (3.50)$$

$$-i \frac{\partial}{\partial\phi} Y_{lm}(\theta, \phi) = m Y_{lm}(\theta, \phi). \quad (3.51)$$

Dans la théorie générale du moment cinétique, notamment lors de la mise en œuvre de la procédure d'échelle, il a toujours été supposé que les vecteurs propres sont de norme finie. Dans la représentation en cours, ceci impose aux fonctions Y_{lm} d'être de module carré sommable. Il en résulte que seules les solutions des équations précédentes possédant cette propriété doivent être retenues physiquement. On les normalise à l'unité en posant :

$$\int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi |Y_{lm}(\theta, \phi)|^2 = 1. \quad (3.52)$$

Les équations aux dérivées partielles (3.50) et (3.51) ont par elles-mêmes un grand nombre de solutions, qui ne conviennent pas toutes physiquement. En effet, la fonction d'onde d'une particule (sans spin) doit avoir une et une seule valeur en un point donné de l'espace, elle doit être monovaluée⁹ et continue. Ceci impose notamment qu'elle prenne la *même* valeur en ϕ et en $\phi + 2\pi$:

$$Y_{lm}(\theta, 0) = Y_{lm}(\theta, 2\pi). \quad (3.53)$$

D'ailleurs, on sait du ch. 2 que pour toute rotation \mathcal{R} du système dans l'espace \mathbb{R}^3 , la fonction d'onde $\Psi(\vec{r}, t)$ se transforme en $\Psi(\mathcal{R}^{-1}\vec{r}, t)$, sans phase additionnelle qui violerait l'invariance galiléenne (une rotation du système changerait son moment cinétique orbital !).

La solution de (3.51) est du type :

$$Y_{lm}(\theta, \phi) = \Theta_{lm}(\theta) e^{im\phi}, \quad (3.54)$$

où Θ_{lm} est pour l'instant une fonction quelconque de θ . La fonction Y_{lm} n'est monovaluée que si m est un entier ($m = 0, \pm 1, \pm 2, \dots$). De la théorie générale, on savait que m pouvait être entier ou demi-entier et que l

⁹Si on s'autorisait une phase entre $\Phi(0)$ et $\Phi(2\pi)$, les prévisions physiques resteraient encore inaltérées, mais la fonction d'onde ne serait pas continue. Sur le caractère monovalué de la fonction d'onde, voir [21]. Au ch. 5, on verra que la fonction d'onde (à deux composantes) d'un spin 1/2 acquiert une phase $e^{i\pi} = -1$ lors d'une rotation de 2π : avec $S = \frac{1}{2}$, $e^{\frac{i}{\hbar}(2\pi)(\frac{\hbar}{2})} = -1$.

et m sont solidaires dans cette double possibilité. En fait, compte tenu du lien entre les rotations spatiales et le moment cinétique orbital, m ne peut être ici qu'un entier, et il en va alors de même pour le nombre quantique l . On retrouve à nouveau le rôle décisif des conditions que doit satisfaire la fonction d'onde sur l'apparition précise de la quantification.

Ainsi, pour un moment cinétique *orbital*, seules les valeurs *entières* de m et de l sont donc autorisées, les valeurs demi-entières se trouvant *ipso facto* exclues. Ce résultat étant acquis, on peut poursuivre soit en se branchant sur la littérature de Mathématiques appliquées (équation de Legendre, voir Remarque dans la suite), soit utiliser directement les résultats établis ci-dessus, se contentant d'une description *a minima* – ce qui va être fait maintenant – en réitérant avec les Y_{lm} la procédure d'échelle développée antérieurement sur les kets $|j, m\rangle$.

De la théorie générale pour \vec{J} , et d'après (3.22), L_+ agissant sur $Y_{l, m=l}$ doit donner zéro :

$$L_+ Y_{ll}(\theta, \phi) = 0 . \quad (3.55)$$

Compte tenu de l'expression (3.49) de l'opérateur L_+ , cette dernière équation s'écrit :

$$L_+ Y_{ll}(\theta, \phi) \equiv \hbar e^{+i\phi} \left(\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right) \Theta_{ll}(\theta) e^{il\phi} = 0 , \quad (3.56)$$

soit :

$$\left(\frac{\partial}{\partial \theta} - l \cot \theta \right) Y_{ll}(\theta, \phi) = 0 \iff \left(\frac{d}{d\theta} - l \cot \theta \right) \Theta_{ll}(\theta) = 0 ; \quad (3.57)$$

$\cot \theta$ étant la dérivée de $\ln(\sin \theta)$, ceci s'écrit :

$$\frac{1}{\Theta_{ll}} \frac{d\Theta_{ll}}{d\theta} = l \frac{d}{d\theta} \ln(\sin \theta) \iff \Theta_{ll}(\theta) = C_l (\sin \theta)^l , \quad (3.58)$$

où C_l est une constante arbitraire, qui sera fixée par normalisation. Au total :

$$Y_{ll}(\theta, \phi) = C_l (\sin \theta)^l e^{il\phi} . \quad (3.59)$$

Il est maintenant possible de construire les $2l$ autres fonctions Y_{lm} , $m = l-1, l-2, \dots, -l$, par application répétée de L_- et usage de la relation générale (3.37) :

$$L_- Y_{lm} = \hbar \sqrt{l(l+1) - m(m-1)} Y_{l, m-1} \iff Y_{l, m-1} = \frac{1}{\hbar \sqrt{l(l+1) - m(m-1)}} L_- Y_{lm} \quad (3.60)$$

En représentation- q , compte tenu de (3.49), ceci s'écrit :

$$Y_{l, m-1} = \frac{1}{\sqrt{l(l+1) - m(m-1)}} e^{-i\phi} \left(-\frac{\partial}{\partial \theta} - m \cot \theta \right) Y_{lm} . \quad (3.61)$$

Ainsi par exemple, on a, pour $m = l$ et selon (3.59) :

$$Y_{l, l-1} = \frac{1}{\sqrt{2l}} e^{-i\phi} \left(-\frac{\partial}{\partial \theta} - l \cot \theta \right) Y_{ll} = -C_l \sqrt{2l} e^{i(l-1)\phi} \cos \theta (\sin \theta)^{l-1} . \quad (3.62)$$

Connaissant $Y_{l, l-1}$, on peut de même engendrer $Y_{l, l-2}$ et ainsi de suite. Il n'est pas difficile d'obtenir des formules générales pour $Y_{l, l-\mu}$ ¹⁰ ; il apparaît de fait des fonctions bien connues appelées polynômes associés de Legendre, dont la définition et certaines propriétés seront données en temps utile.

Les fonctions Y_{lm} sont orthogonales entre elles puisqu'elles sont fonctions propres du couple d'opérateurs hermitiques (\vec{L}^2, L_z) ; après normalisation à l'unité des Y_{lm} , on a donc¹¹ :

$$\langle Y_{lm} | Y_{l' m'} \rangle \equiv \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi Y_{lm}^*(\theta, \phi) Y_{l' m'}(\theta, \phi) = \delta_{ll'} \delta_{mm'} . \quad (3.63)$$

¹⁰ voir [20], p. 446, [4], complément A_{VI}.

¹¹ bien noter que l'élément différentiel n'est pas simplement $d\theta d\phi$!

Comme d'habitude, une telle contrainte laisse encore un facteur de phase indéterminé, sans aucune incidence physique. La convention d'usage consiste à prendre $Y_{l0}(0,0)$ réel et positif¹² ; les phases de toutes les autres Y_{lm} de même l s'en trouvent automatiquement fixées grâce aux relations (3.60).

En outre, les Y_{lm} constituent une base complète orthonormée pour les fonctions de carré sommable définies de la sphère unité dans \mathbb{C} . À ce titre, elles satisfont la relation de fermeture exprimant cette complétude. Comme l'intégration sur θ se fait avec $\sin \theta d\theta = -d(\cos \theta)$, la relation de fermeture s'écrit explicitement¹³ :

$$\sum_{l=0}^{+\infty} \sum_{m=-l}^{+l} Y_{lm}(\theta, \phi) Y_{lm}^*(\theta', \phi') = \delta(\cos \theta - \cos \theta') \delta(\phi - \phi') \equiv \frac{1}{\sin \theta} \delta(\theta - \theta') \delta(\phi - \phi') , \quad (3.65)$$

où il serait superflu d'écrire $|\sin \theta|$ puisque $\theta \in [0, \pi]$. Ainsi toute fonction de carré sommable $f(\theta, \phi)$ peut être développée en série des Y_{lm} :

$$f(\theta, \phi) = \sum_{l=0}^{+\infty} \sum_{m=-l}^{+l} f_{lm} Y_{lm}(\theta, \phi) , \quad (3.66)$$

les coefficients f_{lm} s'exprimant simplement comme suit, grâce à l'orthonormalisation (3.63) des Y_{lm} ; l'intégrale représentant toujours le produit scalaire, on a :

$$f_{lm} = \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi Y_{lm}^*(\theta, \phi) f(\theta, \phi) \equiv \langle Y_{lm} | f \rangle . \quad (3.67)$$

La parité (ou inversion d'espace), Π , change θ en $\pi - \theta$ et ϕ en $\phi + \pi$, puisqu'elle est égale au produit d'une symétrie-miroir par rapport à xOy et d'une rotation de 180 degrés autour de Oz . Par une telle transformation les Y_{lm} au plus changent de signe : ce sont des fonctions de parité déterminée. En effet, puisque l'opérateur parité Π commute avec \vec{L} et que le couple (l, m) spécifie complètement l'état propre de (\vec{L}^2, L_z) (il n'y a pas de dégénérescence), le vecteur $|lm\rangle$ est propre de Π ; comme les valeurs propres de Π sont $\varpi = \pm 1$, on a $\Pi|lm\rangle = \pm|lm\rangle$. Par ailleurs $[\Pi, L_\pm] = 0$, donc $|lm\rangle$ et $|l, m \pm 1\rangle$ ont la même valeur propre ϖ : tous les états $\{|lm\rangle\}_{-l \leq m \leq +l}$ ont donc la même parité, qui est celle de $Y_{ll} = C_l(\sin \theta)^l e^{il\phi}$. Dans l'inversion d'espace, le $\sin \theta$ est invariant et il vient seulement une phase $e^{il\pi} = (-1)^l$. Finalement :

$$\Pi Y_{lm}(\theta, \phi) \equiv Y_{lm}(\pi - \theta, \phi + \pi) = (-1)^l Y_{lm}(\theta, \phi) . \quad (3.68)$$

Il n'est pas non plus difficile d'établir le mode de transformation des Y_{lm} par la conjugaison complexe¹⁴ :

$$[Y_{lm}(\theta, \phi)]^* = (-1)^m Y_{l, -m}(\theta, \phi) , \quad (3.69)$$

qui se démontre en regardant de près les expressions des Y_{lm} engendrées les unes ($m \geq 0$) par $L_-^k Y_{l+l}$, les autres ($m \leq 0$) par $L_+^k Y_{l-l}$.

Pour usage ultérieur, donnons la relation précise ([22], p. 512-513) entre les harmoniques sphériques et des fonctions spéciales répertoriées, appelées fonctions associées de Legendre P_l^m :

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} (-1)^m e^{im\phi} P_l^m(\cos \theta) \quad (m \geq 0) . \quad (3.70)$$

P_l^m ($m \geq 0$) se construit à partir d'autres polynômes connus sous le nom de polynômes de Legendre, P_l :

$$P_l(X) = \frac{(-1)^l}{2^l l!} \frac{d^l}{dX^l} (1 - X^2)^l , \quad (3.71)$$

¹²On a alors $Y_{l, m=0}(0,0) = [(2l+1)/(4\pi)]^{1/2}$ et $Y_{l, m \neq 0}(0,0) = 0$, soit $Y_{lm}(0,0) = \delta_{m0} [(2l+1)/(4\pi)]^{1/2}$.

¹³On utilise :

$$\delta(f(x)) = \frac{1}{|f'(x_0)|} \delta(x - x_0) , \quad (3.64)$$

relation qui est vraie quand la fonction s'annule une seule fois, en x_0 – et se généralise immédiatement quand f a plusieurs zéros.

¹⁴La conjugaison complexe change m en $-m$, en accord avec le renversement du temps, qui inverse les moments cinétiques.

dont la fonction génératrice, définie comme :

$$F(X, t) = \sum_{l=0}^{+\infty} t^l P_l(X) , \quad (3.72)$$

est¹⁵ :

$$F(X, t) \equiv \frac{1}{\sqrt{1 - 2Xt + t^2}} . \quad (3.73)$$

Les polynômes P_l sont orthogonaux suivant :

$$\int_{-1}^{+1} P_l(X) P_{l'}(X) dX = \int_0^\pi P_l(\cos \theta) P_{l'}(\cos \theta) \sin \theta d\theta = \frac{2}{2l+1} \delta_{ll'} . \quad (3.74)$$

Les P_l^m sont alors définis comme¹⁶ :

$$P_l^m(X) = (1 - X^2)^{m/2} \frac{d^m}{dX^m} P_l(X) . \quad (3.75)$$

Les fonctions associées de même indice supérieur m sont orthogonales et satisfont ([17], eq. (9.62)) :

$$\int_{-1}^{+1} P_l^m(X) P_{l'}^m(X) dX = \frac{2}{2l+1} \frac{(l+m)!}{(l-m)!} \delta_{ll'} . \quad (3.76)$$

Enfin, les harmoniques sphériques satisfont une importante relation, dite théorème d'addition :

$$P_l(\cos \alpha) = \frac{4\pi}{2l+1} \sum_{m=-l}^{+l} Y_{lm}(\theta', \phi') Y_{lm}^*(\theta'', \phi'') = \frac{4\pi}{2l+1} \sum_{m=-l}^{+l} (-1)^m Y_{lm}(\theta', \phi') Y_{l-m}(\theta'', \phi'') , \quad (3.77)$$

où (θ', ϕ') et (θ'', ϕ'') spécifient deux directions de l'espace formant entre elles l'angle α . En particulier, en choisissant $\theta' = \theta, \phi' = \phi, \theta'' = \phi'' = 0$, on en déduit :

$$P_l(\cos \theta) = \frac{4\pi}{2l+1} \sum_{m=-l}^{+l} Y_{lm}(\theta, \phi) Y_{lm}^*(0, 0) , \quad (3.78)$$

mais comme $Y_{lm}(0, 0) = \delta_{m0} [(2l+1)/(4\pi)]^{1/2}$, il vient :

$$P_l(\cos \theta) = \frac{4\pi}{2l+1} Y_{l0}(\theta, \phi) , \quad (3.79)$$

où $Y_{l0}(\theta, \phi)$ est, en fait, une fonction indépendante de l'angle ϕ .

Les expressions des premières harmoniques sphériques sont :

$$Y_{00} = \frac{1}{\sqrt{4\pi}} , \quad Y_{1\pm 1} = \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi} , \quad Y_{10} = \sqrt{\frac{3}{4\pi}} \cos \theta , \quad (3.80)$$

$$Y_{20} = \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1) , \quad Y_{2\pm 1} = \mp \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\phi} , \quad Y_{2\pm 2} = \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi} . \quad (3.81)$$

¹⁵Plus généralement, la fonction $(1 - 2Xt + t^2)^{-\alpha}$ – où l'exposant positif α peut prendre diverses valeurs – définit plusieurs classes de polynômes remarquables.

¹⁶Noter que $P_l^{m=0}(X) = P_l(X)$.

À propos de l'équation de Legendre

L'équation différentielle (3.50) se résout par séparation des variables ; en effet, en posant :

$$Y_{lm}(\theta, \phi) = \Phi(\phi) \Theta(\theta) , \quad (3.82)$$

et en reportant, il vient :

$$-\Phi(\phi) \left(\frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} \right) \Theta(\theta) - \Theta(\theta) \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \Phi(\phi) = l(l+1) \Phi(\phi) \Theta(\theta) , \quad (3.83)$$

soit :

$$-\frac{1}{\Phi} \frac{\partial^2 \Phi}{\partial \phi^2} = \frac{\sin^2 \theta}{\Theta} \left(\frac{\partial^2 \Theta}{\partial \theta^2} + \cot \theta \frac{\partial \Theta}{\partial \theta} \right) + l(l+1) \sin^2 \theta . \quad (3.84)$$

Cette dernière équation montre clairement la séparation des variables : le premier membre ne dépend que de ϕ , le second terme ne dépend que de θ . Chacun d'entre eux doit donc être une seule et même constante, notée C . Par ailleurs, en dérivant une fois de plus (3.51), on obtient :

$$\frac{\partial^2}{\partial \phi^2} Y_{lm}(\theta, \phi) = -m^2 Y_{lm}(\theta, \phi) \iff \frac{\partial^2}{\partial \phi^2} \Phi(\phi) = -m^2 \Phi(\phi) . \quad (3.85)$$

La constante C est donc l'opposé du carré du nombre quantique magnétique m . D'où les deux équations :

$$\frac{\partial^2 \Phi}{\partial \phi^2} + m^2 \Phi = 0 . \quad (3.86)$$

$$\left(\frac{\partial^2 \Theta}{\partial \theta^2} + \cot \theta \frac{\partial \Theta}{\partial \theta} \right) - \frac{m^2}{\sin^2 \theta} \Theta + l(l+1) \Theta = 0 . \quad (3.87)$$

Le changement de variable :

$$\cos \theta = \xi \quad \Theta(\theta) = F(\xi) \quad (3.88)$$

transforme (3.87) en :

$$(1 - \xi^2) \frac{d^2 F}{d\xi^2} - 2\xi \frac{dF}{d\xi} + \left[l(l+1) - \frac{m^2}{1 - \xi^2} \right] F = 0 ; \quad (3.89)$$

pour $m = 0$, (3.89) est une équation dûment répertoriée, appelée l'équation de Legendre :

$$(1 - \xi^2) \frac{d^2 F}{d\xi^2} - 2\xi \frac{dF}{d\xi} + l(l+1) F = 0 \iff \frac{d}{d\xi} \left[(1 - \xi^2) \frac{dF}{d\xi} \right] + l(l+1) F = 0 . \quad (3.90)$$

On remarque que l'équation (3.90) est invariante si ξ est changé en $-\xi$, c'est-à-dire si θ est changé en $\pi - \theta$. On peut donc se borner à chercher des solutions qui sont soit paires soit impaires dans ce changement (correspondant à une symétrie-miroir par rapport au plan xOy).

L'équation de Legendre possède des solutions singulières en $\xi = \pm 1$, avec des singularités logarithmiques. En effet, prenons le voisinage de $\xi = 1$; on a alors approximativement :

$$2(1 - \xi) \frac{d^2 F}{d\xi^2} - 2 \frac{dF}{d\xi} + l(l+1) F \simeq 0 . \quad (3.91)$$

Reportons maintenant une fonction $f(\xi)$ qui se comporte comme $\ln(1 - \xi)$ dans ce voisinage :

$$f(\xi) = \ln(1 - \xi) \quad f'(\xi) = -\frac{1}{1 - \xi} \quad f''(\xi) = \frac{1}{(1 - \xi)^2} . \quad (3.92)$$

Il vient alors :

$$-2 \frac{1}{1 - \xi} + 2 \frac{1}{1 - \xi} + l(l+1) \ln(1 - \xi) \simeq 0 . \quad (3.93)$$

Les termes les plus divergents se compensent bien, le logarithme étant un infiniment grand d'ordre inférieur. Ainsi, pour $\xi \simeq 1$, $F(\xi) \simeq \ln(1 - \xi)$, $\Theta(\theta) \simeq \ln(1 - \cos \theta)$. La condition de normalisabilité de Θ introduit

forcément quelque part une intégrale du genre $\int_{-1}^{+1} d\xi [\ln(1-\xi)]^2$, qui est bel et bien sommable¹⁷ en $\xi = \pm 1$. Il faut donc chercher ailleurs la justification de l'affirmation “*such singular functions [...] are not acceptable eigenfunctions...*” ([17], p. 182). Une bonne raison est la suivante : l'opérateur donnant l'énergie cinétique de rotation contient des dérivées secondes en θ , qui introduit des intégrales du genre $\int_{-1}^{+1} d\xi [\ln(1-\xi)]/(1-\xi)^2$ qui, visiblement, sont divergentes. De telles solutions singulières ne sont donc pas acceptables physiquement.

Les solutions régulières, elles, peuvent être cherchées sous la forme d'une série entière :

$$F(\xi) = \sum_{k=0}^{+\infty} c_k \xi^k, \quad (3.94)$$

et la substitution dans (3.90) produit la relation de récurrence :

$$(k+1)(k+2)c_{k+2} = [k(k+1) - l(l+1)]c_k, \quad (3.95)$$

qui relie les coefficients de deux en deux : si on prend $c_0 \neq 0$, on obtient une fonction dont tous les termes sont pairs, c'est donc une fonction paire. À l'inverse, si on prend $c_1 \neq 0$, on obtient une fonction impaire. Dans les deux cas, la fonction est déterminée comme d'habitude à un coefficient multiplicatif global près.

Oublions un instant que l est un entier. Le rapport c_{k+2}/c_k est $\simeq k/(k+2)$ quand $k \gg 1$: il en résulte que, en $\xi = \pm 1$ (soit pour $\theta = 0$ ou $\theta = \pi$), la série se comporte comme la série harmonique, dont on sait qu'elle diverge. Pour éviter ce genre d'ennui, il faut que la série s'arrête à un certain rang k ; pour qu'il en soit ainsi, il faut et suffit précisément que l soit un entier : on retrouve une fois de plus la nécessaire quantification du moment cinétique en conséquence des conditions aux limites pour les fonctions propres.

3.4 Cas particuliers des moments cinétiques $j = \frac{1}{2}$, $j = 1$

Après les traitements formels des paragraphes précédents, il est bon de revenir à des cas un peu plus concrets, d'autant plus que les deux choix de j détaillés ci-dessous sont extrêmement fréquents en pratique et jouent de ce fait un rôle important dans les applications.

Commençons par traiter le cas $j = \frac{1}{2}$. Comme on le sait, il ne peut s'agir d'un moment cinétique orbital : c'est donc forcément un moment de spin et pour rappeler ceci, on désignera par \vec{S} (et non pas \vec{J}) le moment cinétique considéré.

Les états propres de (\vec{S}^2, S_z) sont notés $|s, m\rangle$ avec $s = \frac{1}{2}$, $m = \pm \frac{1}{2}$. L'espace des états de spin $\frac{1}{2}$ est donc de dimension 2. On a¹⁸ :

$$\vec{S}^2 \left| \frac{1}{2}, m \right\rangle = \frac{3}{4} \hbar^2 \left| \frac{1}{2}, m \right\rangle, \quad S_z \left| \frac{1}{2}, m \right\rangle = m \hbar \left| \frac{1}{2}, m \right\rangle \quad (m = \pm \frac{1}{2}). \quad (3.96)$$

Les opérateurs S_u ($u = x, y, z$) sont donc représentés par des matrices 2×2 , qu'il est facile d'écrire. En ce qui concerne \vec{S}^2 , c'est vite fait : tous les vecteurs propres correspondent à la même valeur propre $(3/4)\hbar^2$, d'où :

$$\vec{S}^2 = \frac{3}{4} \hbar^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (3.97)$$

\vec{S}^2 est réellement le prototype de l'opérateur scalaire (au sens où ceci a été défini au chapitre II) : sa matrice est proportionnelle à la matrice identité. Venons-en à S_z ; sur la base ordonnée $(|\frac{1}{2}, +\frac{1}{2}\rangle, |\frac{1}{2}, -\frac{1}{2}\rangle)$, il vient :

$$S_z = \hbar \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}. \quad (3.98)$$

¹⁷ $\int_0^a dx [\ln(x)]^2$, $a > 0$, existe comme on s'en convainc en effectuant des intégrations par parties.

¹⁸ $s(s+1)|_{s=1/2} = \frac{3}{4}$.

Pour avoir les matrices de S_x et S_y il suffit de passer par les opérateurs S_{\pm} . D'après les formules générales (3.37), on a :

$$S_- \left| \frac{1}{2}, \frac{1}{2} \right\rangle = \hbar \sqrt{\frac{1}{2}(\frac{1}{2} + 1) - \frac{1}{2}(\frac{1}{2} - 1)} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = \hbar \left| \frac{1}{2}, -\frac{1}{2} \right\rangle, \quad S_- \left| \frac{1}{2}, -\frac{1}{2} \right\rangle = 0. \quad (3.99)$$

Comme d'habitude, la matrice d'un opérateur contient en colonne le transformé du vecteur de la base ayant le même numéro d'ordre que la colonne considérée. D'où la matrice représentant S_- sur la base ordonnée comme ci-dessus :

$$S_- = \hbar \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}. \quad (3.100)$$

S_+ est l'adjoint de S_- ; la base est orthonormée, donc sa matrice est l'adjointe de celle de S_- :

$$S_+ = \hbar \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \quad (3.101)$$

Il suffit maintenant d'inverser $S_{\pm} = S_x \pm iS_y$ pour obtenir :

$$S_x = \frac{1}{2}(S_+ + S_-) = \frac{\hbar}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad S_y = \frac{1}{2i}(S_+ - S_-) = \frac{\hbar}{2} \begin{bmatrix} 0 & -i \\ +i & 0 \end{bmatrix}. \quad (3.102)$$

Pour se faire la main, il n'est pas inutile de vérifier que ces différentes matrices satisfont bien les relations de commutation caractéristiques d'un moment cinétique.

Les trois matrices 2×2 ainsi mises en évidence sont les célèbres matrices de Pauli, omniprésentes en Mécanique Quantique et qu'il vaut la peine d'écrire explicitement :

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ +i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (3.103)$$

Un spin \vec{S} , $S = 1/2$, peut donc s'écrire vectoriellement :

$$\vec{S} = \frac{\hbar}{2} \vec{\sigma}. \quad (3.104)$$

Les matrices de Pauli satisfont des relations de commutation qui ressemblent évidemment à celle d'un moment cinétique, à des facteurs près ; très précisément :

$$[\sigma_x, \sigma_y] = 2i\sigma_z, \quad [\sigma_y, \sigma_z] = 2i\sigma_x, \quad [\sigma_z, \sigma_x] = 2i\sigma_y. \quad (3.105)$$

En outre, il est facile de voir que les *anticommutateurs* des σ_u sont nuls¹⁹ :

$$\{\sigma_x, \sigma_y\} = 0, \quad \{\sigma_y, \sigma_z\} = 0, \quad \{\sigma_z, \sigma_x\} = 0, \quad (3.106)$$

de sorte que, par combinaison avec (3.105) :

$$\sigma_x \sigma_y = i\sigma_z, \quad \sigma_y \sigma_z = i\sigma_x, \quad \sigma_z \sigma_x = i\sigma_y. \quad (3.107)$$

On définit aussi les $\sigma_{\pm} = \sigma_x \pm i\sigma_y$, qui satisfont :

$$[\sigma_z, \sigma_{\pm}] = \pm 2i\sigma_{\pm}, \quad [\sigma_+, \sigma_-] = 4i\sigma_z. \quad (3.108)$$

Notons une égalité importante qu'il n'est pas difficile de démontrer²⁰. Soient \vec{A} et \vec{B} deux vecteurs scalaires²¹ vis-à-vis des opérateurs de spin :

$$(\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) = \vec{A} \cdot \vec{B} \mathbf{1} + i\vec{\sigma} \cdot (\vec{A} \times \vec{B}), \quad (3.109)$$

¹⁹ $\{\sigma_u, \sigma_v\} \equiv \sigma_u \sigma_v + \sigma_v \sigma_u$.

²⁰ Toutes les directions de l'espace jouent ici le même rôle (isotropie galiléenne !) : il suffit donc de démontrer (3.109) avec un choix particulier de l'un des vecteurs scalaires, par exemple $\vec{A} = A\vec{e}_x$. La distribution du commutateur et l'usage répété des relations (3.105)-(3.108) conduit aisément à (3.109).

²¹ au sens défini dans le ch. 2. Ainsi, les deux vecteurs \vec{A} et \vec{B} peuvent être des opérateurs ne dépendant que des grandeurs dynamiques spatiales (orbitales), auquel cas il est impératif de maintenir le même ordre entre eux dans les deux membres.

où $\mathbf{1}$ désigne l'identité.

Passons maintenant au moment cinétique $j = 1$, qui peut donc être soit orbital, soit de spin (il ne peut résulter de la combinaison des deux, puisque le moment orbital ne peut être demi-entier et que pour avoir $j = 1$ avec deux moments cinétiques il faut que les deux soient $1/2$). Toujours en excluant la considération d'autres degrés de liberté, la dimension de l'espace est égale à $(2j + 1)_{j=1} = 3$. Les opérateurs J_x , J_y et J_z sont donc représentés par des matrices 3×3 . La valeur propre de \vec{J}^2 est $[j(j + 1)]_{j=1}\hbar^2 = 2\hbar^2$ et le nombre magnétique m peut prendre les trois valeurs $-1, 0, +1$. Bien sûr :

$$\vec{J}^2 = \hbar^2 \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad (3.110)$$

qui est à nouveau le prototype de l'opérateur scalaire, cette fois dans l'espace à trois dimensions. En ordonnant la base suivant $|1, +1\rangle, |1, 0\rangle, |1, -1\rangle$, la matrice de J_z est :

$$J_z = \hbar \begin{bmatrix} +1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \quad (3.111)$$

Déterminons maintenant la matrice de J_- en appliquant J_- aux trois vecteurs de base :

$$J_-|1, +1\rangle = \hbar\sqrt{1(1+1) - 1(1-1)}|1, 0\rangle = \sqrt{2}|1, 0\rangle, \quad (3.112)$$

$$J_-|1, +0\rangle = \hbar\sqrt{1(1+1) - 0(1-1)}|1, 0\rangle = \sqrt{2}|1, -1\rangle, \quad (3.113)$$

et bien sûr $J_-|1, -1\rangle = 0$. D'où la matrice de J_- , puis celle de J_+ en prenant l'adjointe :

$$J_- = \hbar \begin{bmatrix} 0 & 0 & 0 \\ \sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{bmatrix}, \quad J_+ = \hbar \begin{bmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \\ 0 & 0 & 0 \end{bmatrix}. \quad (3.114)$$

Les deux matrices de $J_x = (1/2)(J_+ + J_-)$ et $J_y = (1/2i)(J_+ - J_-)$ sont donc :

$$J_x = \hbar \begin{bmatrix} 0 & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 0 \end{bmatrix}, \quad J_y = \hbar \begin{bmatrix} 0 & -i/\sqrt{2} & 0 \\ i/\sqrt{2} & 0 & -i/\sqrt{2} \\ 0 & i/\sqrt{2} & 0 \end{bmatrix}. \quad (3.115)$$

Ces résultats donnent les transformés des vecteurs propres par les différentes composantes du moment cinétique pour $j = 1$. Comme j est entier, il peut s'agir d'un moment orbital et on peut notamment les appliquer aux harmoniques sphériques $l = 1$; on a ainsi, par exemple :

$$J_x Y_{11}(\theta, \phi) = \frac{\hbar}{\sqrt{2}} Y_{10}, \quad J_x Y_{10}(\theta, \phi) = \frac{\hbar}{\sqrt{2}} [Y_{11}(\theta, \phi) + Y_{1-1}(\theta, \phi)], \quad J_x Y_{1-1}(\theta, \phi) = \frac{\hbar}{\sqrt{2}} Y_{10}, \quad (3.116)$$

$$J_y Y_{11}(\theta, \phi) = \frac{i\hbar}{\sqrt{2}} Y_{10}, \quad J_y Y_{10}(\theta, \phi) = \frac{i\hbar}{\sqrt{2}} [-Y_{11}(\theta, \phi) + Y_{1-1}(\theta, \phi)], \quad J_y Y_{1-1}(\theta, \phi) = -\frac{i\hbar}{\sqrt{2}} Y_{10}, \quad (3.117)$$

etc, relations que l'on peut aussi trouver – laborieusement – en utilisant les expressions (3.45).

3.5 Addition de deux moments cinétiques. Coefficients de Clebsch-Gordan

3.5.1 Comment apparaît la somme des moments cinétiques

L'addition (aussi appelée composition) des moments cinétiques est une notion très importante. Par exemple, pour une particule ayant un moment cinétique intrinsèque \vec{S} mais qui possède par ailleurs un mouvement orbital lui conférant un moment orbital \vec{L} , il est naturel de définir un moment cinétique total \vec{J} par la relation :

$$\vec{J} = \vec{L} + \vec{S}. \quad (3.118)$$

Cette relation est en un sens triviale mais, telle quelle, n'a pas de contenu physique et, au fond n'a rien d'évident : on verra au ch. 5 qu'elle s'impose comme une nécessité physique. Quoi qu'il en soit, une fois admise (3.118), la description quantique d'une particule avec spin exige la connaissance des états propres de ce moment total : il convient donc de savoir comment on les construit à partir des états propres de \vec{L} et de ceux de \vec{S} . Autre exemple : pour un système de N particules (supposées sans spin pour simplifier), possédant chacune un moment orbital (\vec{L}_n pour la $n^{\text{ème}}$ particule) il est possible de définir le moment cinétique total \vec{L} :

$$\vec{L} = \sum_{n=1}^N \vec{L}_n . \quad (3.119)$$

À nouveau, ceci n'a rien d'évident. Le contenu physique de (3.119) est finalement l'invariance euclidienne par rotation : l'opérateur de rotation pour plusieurs particules implique le produit des rotations individuelles $e^{\frac{\theta_n}{i\hbar} \vec{L}_n \vec{u}_n}$. Comme les opérateurs \vec{L}_n agissent sur des variables différentes, les commutateurs $[\vec{L}_n \vec{u}_n, \vec{L}_{n'} \vec{u}_{n'}]$ sont nuls si $n \neq n'$, de sorte que le produit des exponentielles est égal à l'exponentielle de la somme :

$$\prod_{n=1}^N e^{\frac{\theta_n}{i\hbar} \vec{L}_n \vec{u}_n} = e^{\frac{1}{i\hbar} \sum_{n=1}^N \theta_n \vec{L}_n \vec{u}_n} . \quad (3.120)$$

L'invariance n'est obtenue que si toutes les particules tournent de la même façon (même angle, même axe), ce qui alors fait naturellement apparaître la somme présente au second membre de (3.119). À nouveau, la question est alors la suivante : connaissant les valeurs des moments cinétiques individuels, quelles sont les valeurs du moment total ?

On imagine aisément que la procédure de composition est assez lourde quand il s'agit de combiner plusieurs moments cinétiques ; clairement, il suffit de savoir en additionner deux pour, de proche en proche, pouvoir en principe en combiner un nombre arbitraire. Dans toute la suite, on se limite à l'addition de *deux* moments cinétiques quelconques.

La toute première chose dont il faut se convaincre est que la somme de deux moments cinétiques est encore un moment cinétique. Avant d'établir ce résultat, il est utile d'affiner les notations et en particulier d'introduire à l'occasion la notion de produit (tensoriel) de deux espaces vectoriels. Pour fixer les idées, on raisonne dans le cas d'une particule possédant un moment orbital \vec{L} et un spin \vec{S} , mais ce qui suit vaut tout autant pour deux moments cinétiques quelconques \vec{J}_1 et \vec{J}_2 , quelle que soit leur origine physique.

Le moment orbital agit sur les degrés de liberté qui ont un équivalent classique (coordonnées au sens usuel), le plus souvent appelées variables d'espace. S'il n'était le spin, les états quantiques de cette particule seraient des fonctions d'onde ordinaires que l'on peut toujours représenter avec les états propres $\psi_n(\vec{r})$ du Hamiltonien H . En réalité, puisque la particule possède un spin, il convient de considérer tous les états formés en associant successivement un $\psi_n(\vec{r})$ donné à chacun des états de spin possibles. Pour un spin $S = \frac{1}{2}$, chaque $\psi_n(\vec{r})$ permet de construire deux états quantiques distincts, que l'on peut noter $\psi_{n+}(\vec{r})$ et $\psi_{n-}(\vec{r})$ en introduisant le signe de la composante S_z ; si l'on épuise tous les états les uns après les autres, on voit que le nombre total d'états quantiques "complets"²² est le produit du nombre d'états orbitaux par le nombre d'états de spin. L'espace vectoriel correspondant est le produit des espaces d'espace et de spin. Dans chacun de ces deux espaces, il existe un opérateur identité, $\mathbf{1}$, que l'on peut noter $\mathbf{1}_{\text{orb}}$ pour l'espace "orbital" et $\mathbf{1}_{\text{spin}}$ pour l'espace de spin.

Maintenant, quand on combine \vec{L} et \vec{S} pour fabriquer \vec{J} , l'écriture précise correcte est de fait :

$$\vec{J} = \vec{L} \otimes \mathbf{1}_{\text{spin}} + \mathbf{1}_{\text{orb}} \otimes \vec{S} , \quad (3.121)$$

où le symbole \otimes est utilisé pour distinguer l'opération en question ("produit" de deux opérateurs agissant dans des espaces différents) du produit ordinaire de deux opérateurs agissant dans un même espace vectoriel. De la même façon, si on combine deux moments orbitaux de deux particules distinctes, on écrira le moment total :

$$\vec{L} = \vec{L}_1 \otimes \mathbf{1}_2 + \mathbf{1}_1 \otimes \vec{L}_2 , \quad (3.122)$$

²²au sens : donnant une description complète de la particule, avec tous ses degrés de liberté.

où maintenant $\mathbf{1}_i$ désigne l'opérateur identité dans l'espace des états de la particule i ($i = 1, 2$) (l'ordre des opérateurs associés dans le produit tensoriel \otimes est visiblement sans importance). Ce luxe de précautions dans l'écriture est parfois nécessaire : quand les idées sont bien en place, quand aucune ambiguïté n'est à craindre, on s'en tiendra à des écritures plus simples.

Ceci étant, il est aisé de montrer que la somme de deux moments cinétiques est en effet un moment cinétique. Les deux moments cinétiques \vec{L} et \vec{S} agissant dans deux espaces différents, ils commutent entre eux :

$$[L_u, S_v] = 0 \quad \forall u, v, \quad (3.123)$$

puisque chaque opérateur ne peut modifier les degrés de liberté que l'autre affecte et réciproquement. Il résulte de (3.123) que²³ :

$$[J_u, J_v] = [L_u + S_u, L_v + S_v] = [L_u, L_v] + [S_u, S_v] = i\hbar\varepsilon_{uvw}L_w + i\hbar\varepsilon_{uvw}S_w \equiv i\hbar\varepsilon_{uvw}J_w \quad (3.125)$$

Les mêmes arguments valent pour les moments cinétiques orbitaux de deux particules différentes numérotées n et n' : agissant sur des variables d'espace différentes (\vec{r} et \vec{r}' , $n \neq n'$), ils commutent entre eux :

$$[L_{nu}, L_{n'v}] = 0 \quad \forall u, v, n \neq n' \quad (3.126)$$

et la somme d'un nombre quelconque d'entre eux satisfait $[L_u, L_v] = i\hbar\varepsilon_{uvw}L_w$. Au total, la somme des plusieurs moments cinétiques est bien un moment cinétique.

3.5.2 Somme de deux moments cinétiques

Abordons maintenant à proprement parler la composition de deux moments cinétiques \vec{J}_1 et \vec{J}_2 . Les idées étant en place, la notation précise :

$$\vec{J} = \vec{J}_1 \otimes \mathbf{1}_2 + \mathbf{1}_1 \otimes \vec{J}_2 \quad (3.127)$$

est désormais délaissée pour simplifier ; (3.127) sera noté comme suit :

$$\vec{J} = \vec{J}_1 + \vec{J}_2 . \quad (3.128)$$

Chacun des moments cinétiques a ses propres états propres, $|j_1 m_1\rangle$ et $|j_2 m_2\rangle$. Comme \vec{J}_1 et \vec{J}_2 commutent, on peut trouver des vecteurs propres communs à \vec{J}_1^2 , J_{1z} , \vec{J}_2^2 , J_{2z} , que l'on note $|j_1 j_2 m_1 m_2\rangle$ et qui sont le produit tensoriel des états de chaque moment cinétique :

$$|j_1 j_2 m_1 m_2\rangle = |j_1 m_1\rangle \otimes |j_2 m_2\rangle \quad (3.129)$$

ce que l'on écrit plus simplement, quand aucune ambiguïté n'est à craindre :

$$|j_1 j_2 m_1 m_2\rangle = |j_1 m_1\rangle |j_2 m_2\rangle . \quad (3.130)$$

j_1 et j_2 étant fixés, il existe $d(j_1, j_2) = (2j_1 + 1)(2j_2 + 1)$ vecteurs linéairement indépendants de ce type²⁴, constituant une base pour un sous-espace $\mathcal{E}_{j_1 j_2}$. Chacun d'entre eux est propre de $(\vec{J}_1^2, J_{1z}, \vec{J}_2^2, J_{2z})$.

À partir de maintenant, on raisonne exclusivement dans le sous-espace vectoriel $\mathcal{E}_{j_1 j_2}$ de dimension égale à $d(j_1, j_2)$. Il existe un autre ensemble d'observables qui commutent ; en effet, les \vec{J}_i^2 ($i = 1, 2$) commutent avec \vec{J}^2 :

$$[\vec{J}_1^2, \vec{J}^2] = [\vec{J}_1^2, \vec{J}_1^2 + \vec{J}_2^2 + 2\vec{J}_1 \cdot \vec{J}_2] = [\vec{J}_1^2, \vec{J}_2^2] + 2[\vec{J}_1^2, \vec{J}_1] \cdot \vec{J}_2 = 0 . \quad (3.131)$$

²³ ε_{uvw} est le tenseur antisymétrique qui permet d'écrire commodément les relations de commutation caractéristiques d'un moment cinétique :

$$[J_u, J_v] = i\hbar\varepsilon_{uvw} J_w . \quad (3.124)$$

$\varepsilon_{uvw} = +1$ si (uvw) se déduit de (xyz) par une permutation circulaire, $\varepsilon_{uvw} = -1$ dans le cas contraire et $\varepsilon_{uvw} = 0$ si deux indices (ou plus) sont égaux entre eux.

²⁴Chacun d'entre eux est propre d'un jeu d'observables qui commutent et donc est orthogonal à tous les autres.

et commutent évidemment entre eux, puisqu'ils agissent sur des variables différentes (en revanche²⁵, le commutateur $[J_{iz}, \vec{J}^2]$ n'est pas nul). Ainsi, \vec{J}_1^2 , \vec{J}_2^2 , \vec{J}^2 et J_z ont des vecteurs propres communs, que l'on note $|JMj_1j_2\rangle$. Les $\{|JMj_1j_2\rangle\}$ et les $\{|j_1j_2m_1m_2\rangle\}$ forment deux bases orthonormées de $\mathcal{E}_{j_1j_2}$, puisqu'elles sont constituées de vecteurs propres d'opérateurs hermitiques. j_1 et j_2 étant fixés – on reste bien dans $\mathcal{E}_{j_1j_2}$ –, il existe une transformation reliant ces deux jeux de vecteurs propres, que l'on peut écrire :

$$|JMj_1j_2\rangle = \sum_{m_1m_2} |j_1j_2m_1m_2\rangle \langle j_1j_2m_1m_2|JMj_1j_2\rangle . \quad (3.132)$$

Toute la question est donc de trouver les coefficients $\langle j_1j_2m_1m_2|JMj_1j_2\rangle$, qui s'appellent coefficients de Clebsch-Gordan (en abrégé : CG), ou encore coefficients de Wigner. Il est fréquent de les noter plus simplement²⁶, en sous-entendant j_1 et j_2 , qui apparaissent partout et sont les valeurs maximales de m_1 et m_2 . Les coefficients de Clebsch-Gordan seront aussi à l'occasion représentés par le symbole simplifié $\langle m_1m_2|JM\rangle$; ceci a toutefois l'inconvénient de faire disparaître j_1 et j_2 .

Les CG ont été calculés une fois pour toutes, et il existe des tables donnant leurs valeurs ; toutefois, dans les cas simples, il est facile et peu coûteux de les retrouver à chaque fois. Des indications seront données plus loin sur la méthode systématique qui permet de les obtenir. On peut d'ores et déjà noter les relations d'orthogonalité importantes :

$$\sum_{JM} \langle j_1j_2m_1m_2|JMj_1j_2\rangle \langle JMj_1j_2|j_1j_2m'_1m'_2\rangle = \delta_{m_1m'_1} \delta_{m_2m'_2} , \quad (3.133)$$

$$\sum_{m_1m_2} \langle JMj_1j_2|j_1j_2m_1m_2\rangle \langle j_1j_2m_1m_2|J'M'j_1j_2\rangle = \delta_{JJ'} \delta_{MM'} . \quad (3.134)$$

Ces relations résultent du fait que les CG sont les coefficients d'une transformation unitaire reliant entre elles deux bases orthonormées.

Valeurs possibles de J

Avant d'en venir à l'exposé du procédé de calcul général des CG, il convient de répondre à la question suivante : j_1 et j_2 étant fixés, quelles sont les valeurs possibles de M et J donnant les valeurs propres du couple (\vec{J}^2, J_z) associé au moment total ? La réponse à cette question constitue le théorème fondamental d'addition énoncé ci-dessous (voir (3.158)).

En ce qui concerne les valeurs possibles de M , donnant valeur propre $M\hbar$ de J_z , la réponse est évidente ; en effet, on a :

$$J_z |j_1j_2m_1m_2\rangle = J_{1z} |j_1j_2m_1m_2\rangle + J_{2z} |j_1j_2m_1m_2\rangle . \quad (3.135)$$

Par définition, chaque opérateur d'indice i agit exclusivement sur les variables de même indice figurant dans le vecteur produit tensoriel²⁷. Dans (3.135), on a à chaque fois $J_{iz} |j_1j_2m_1m_2\rangle = m_i\hbar |j_1j_2m_1m_2\rangle$, d'où :

$$J_z |j_1j_2m_1m_2\rangle = (m_1 + m_2)\hbar |j_1j_2m_1m_2\rangle . \quad (3.137)$$

Autrement dit, le nombre magnétique attaché à J_z , M , n'est autre que la somme :

$$M = m_1 + m_2 . \quad (3.138)$$

²⁵Cette non-commutation se vérifie trivialement. On peut aussi observer que J_{1z} est associé à une rotation qui ne change que \vec{J}_1 ; si on fait tourner \vec{J}_1 sans changer \vec{J}_2 , il est bien évident que la norme de la somme $\vec{J}_1 + \vec{J}_2$ n'est pas invariante.

²⁶C'est notamment la notation utilisée dans [20].

²⁷Plus généralement, un opérateur A_1 agissant seulement sur la première particule produit une combinaison linéaire où les nombres quantiques de la deuxième n'ont pas changé :

$$A_1 |j_1j_2m_1m_2\rangle = \sum_{j'_1m'_1} A_{1j'_1m'_1 ; j_1m_1} |j'_1j_2m'_1m_2\rangle , \quad (3.136)$$

d'où, si c'est nécessaire, la notation plus précise $A_1 \otimes \mathbf{1}_2$, pour un tel opérateur.

Inversement, un vecteur propre de J_z de valeur propre donnée $M\hbar$ est une combinaison linéaire des seuls $|j_1 j_2 m_1 m_2\rangle$ tels que $m_1 + m_2 = M$. Ceci permet déjà de voir que la valeur maximale de J est $j_1 + j_2$: en prenant $m_1 = j_1$ et $m_2 = j_2$, M est égal à $j_1 + j_2$; comme M doit varier d'une unité entre $\pm J$, bornes comprises, il doit bien exister (au moins une fois) la valeur $j_1 + j_2$ pour J . Rappelons que chaque m_i est entier ou demi-entier, selon que j_i est entier ou demi-entier ; le caractère entier ou demi-entier est donc fixé dès que les j_i sont fixés, ce qui est le cas dans l'analyse en cours.

Ceci étant, il est bien clair qu'une valeur donnée, M , est dégénérée : M étant fixé, il existe *a priori* plusieurs façons d'assurer que la somme $m_1 + m_2$ est de fait égale à M ; si on représente le couple (m_1, m_2) par un point dans le plan de coordonnées $x = m_1, y = m_2$, la droite $x + y = M$ contient (en général) plusieurs points de coordonnées entières ou demi-entières ; les j_i étant fixés, chaque m_i varie par pas d'une unité. Désignons par $d(M)$ la dégénérescence de la valeur M associée à J_z .

Pour l'instant, on ne sait rien de la dégénérescence $D(J)$ des différentes valeurs de J , que l'on ne connaît pas encore et qu'il s'agit précisément de trouver – on sait seulement que $0 \leq J \leq J_{\max} = j_1 + j_2$ et que, M variant au moins d'une unité, les seules valeurs envisageables sont $j_1 + j_2, j_1 + j_2 - 1, j_1 + j_2 - 2, \text{etc.}$ Dans ces notations, on peut alors dire, à J fixé, qu'il existe *a priori* $D(J)$ ensembles de $(2J + 1)$ vecteurs du type $|JM\rangle$ dont le M varie entre $-J$ et $+J$ par sauts d'une unité. Trouver l'ensemble des valeurs de J , c'est finalement trouver toutes les valeurs de la fonction $D(J)$ pour J entier ou demi-entier positif ou nul. Ainsi par exemple, trouver $D(J_\alpha) = 0$ signifie que la valeur J_α pour J n'existe pas avec les valeurs de (j_1, j_2) considérées. En définitive, $D(J)$ est la fonction indicatrice du domaine des valeurs possibles de J pour un couple (j_1, j_2) donné.

Pour savoir ce qu'il en est, donnons successivement deux arguments, l'un rapide, l'autre détaillé.

Argument rapide Il est indiscutable que la valeur maximum de M est $M_{\max} = j_1 + j_2$, d'où il résulte que c'est aussi la valeur maximum J_{\max} de J . L'idée à toujours avoir en tête est que chaque valeur possible de J arrive en bloc avec ses $2J + 1$ composantes $M = -J, -J + 1, \dots, J - 1, J$.

Il est facile de trouver le vecteur $|J = J_{\max} M = M_{\max} = J_{\max}\rangle$ puisqu'il n'y a qu'un seul *bra* ayant ces deux valeurs de J et M ; d'où l'égalité :

$$|J = j_1 + j_2, M = j_1 + j_2, j_1, j_2\rangle = |j_1, m_1 = j_1, j_2, m_2 = j_2\rangle . \quad (3.139)$$

Appliquons l'opérateur J_- au premier membre :

$$J_- |J = j_1 + j_2, M = j_1 + j_2, j_1, j_2\rangle = \hbar\sqrt{J(J+1) - M(M-1)} |J = j_1 + j_2, M - 1, j_1, j_2\rangle , \quad (3.140)$$

soit :

$$J_- |J = j_1 + j_2, M = j_1 + j_2, j_1, j_2\rangle = \hbar\sqrt{2J} |J = j_1 + j_2, M - 1, j_1, j_2\rangle , \quad (3.141)$$

Comme $J_- = J_{1-} + J_{2-}$, appliquons $J_{1-} + J_{2-}$ au second membre de (3.139) :

$$(J_{1-} + J_{2-}) |j_1, m_1 = j_1, j_2, m_2 = j_2\rangle = \hbar\sqrt{j_1(j_1+1) - m_1(m_1-1)} |j_1, m_1 - 1, j_2, m_2\rangle + \hbar\sqrt{j_2(j_2+1) - m_2(m_2-1)} |j_1, m_1, j_2, m_2 - 1\rangle , \quad (3.142)$$

soit :

$$(J_{1-} + J_{2-}) |j_1, m_1 = j_1, j_2, m_2 = j_2\rangle = \hbar\sqrt{2j_1} |j_1, m_1 - 1, j_2, m_2\rangle + \hbar\sqrt{2j_2} |j_1, m_1, j_2, m_2 - 1\rangle . \quad (3.143)$$

En rapprochant (3.141) et (3.143), il vient :

$$\sqrt{2J} |J = j_1 + j_2, M = j_1 + j_2 - 1, j_1, j_2\rangle = \sqrt{2j_1} |j_1, m_1 - 1, j_2, m_2\rangle + \sqrt{2j_2} |j_1, m_1, j_2, m_2 - 1\rangle . \quad (3.144)$$

Cette équation donne donc la composante $M = J - 1$ du vecteur ayant la valeur maximale pour J , qui apparaît comme la combinaison linéaire des deux vecteurs (connus) figurant au second membre de (3.144). Dans le

sous-espace de dimension 2, où tous les vecteurs ont le même $M = j_1 + j_2 - 1$, engendré par ces deux vecteurs, on peut fabriquer la combinaison orthogonale^{28, 29} :

$$\sqrt{2j_2}|j_1, m_1 - 1, j_2, m_2\rangle - \sqrt{2j_1}|j_1, m_1, j_2, m_2 - 1\rangle . \quad (3.145)$$

Ce vecteur correspond forcément à une autre valeur de J , mais son M vaut $J_{\max} - 1$: c'est donc la composante M maximum du vecteur ayant $J = j_1 + j_2 - 1$ (voir fig. 3.1 à gauche) ; d'où l'identification :

$$|J = j_1 + j_2 - 1, M = j_1 + j_2 - 1, j_1, j_2\rangle = \sqrt{2j_2}|j_1, m_1 - 1, j_2, m_2\rangle - \sqrt{2j_1}|j_1, m_1, j_2, m_2 - 1\rangle . \quad (3.146)$$

Il suffit de recommencer le même travail sur l'équation (3.144) : on applique J_- au premier membre, ce qui produit $|J = j_1 + j_2, M = j_1 + j_2 - 2, j_1, j_2\rangle$, à un facteur près. Appliquons $J_{1-} + J_{2-}$ au second membre de 3.144) : on trouve une combinaison linéaire de trois vecteurs seulement, soit $|j_1, m_1 - 2, j_2, m_2\rangle$, $|j_1, m_1 - 1, j_2, m_2 - 1\rangle$ et $|j_1, m_1, j_2, m_2 - 2\rangle$. Le sous-espace est donc de dimension égale à 3, tous les vecteurs correspondant à un M total égal à $j_1 + j_2 - 2$ (voir fig. 3.1 à droite) ; une direction de ce sous-espace est la composante $M = J - 2$ du J maximum, une deuxième dimension est occupée par la composante $M = J - 1$ du $J = J_{\max} - 1$. La dernière direction disponible est donc la composante (maximum) d'un état ayant $J = J_{\max} - 2$.

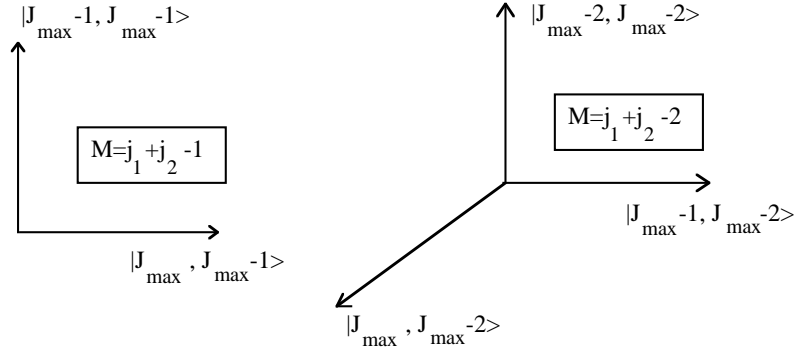


Figure 3.1: Interprétation géométrique des états construits par applications successives de J_- à partir de l'état $|J = j_1 + j_2, M = j_1 + j_2, j_1, t j_2\rangle$ et orthogonalisation.

De proche en proche, on montre ainsi que les valeurs possibles de J sont $j_1 + j_2, j_1 + j_2 - 1, j_1 + j_2 - 2, \dots$. Comme J est forcément positif ou nul, la plus petite valeur est $|j_1 - j_2|$. De plus, puisque pour chaque valeur J , il reste à chaque fois une seule direction disponible en dehors des directions déjà occupées par des composantes M appartenant à des multiplets J' déjà rencontrés, $J' > J$, on voit qu'il existe un seul³⁰ multiplet de $2J + 1$ composantes pour chacune des valeurs de J .

Argument détaillé Donnons maintenant l'argument détaillé, établissant la relation explicite entre $D(J)$ et $d(M)$. Comme $|M| \leq J$, le nombre d'états de valeur propre M est égal à la somme des nombres d'états ayant $J \geq |M|$. D'où :

$$d(M) = \sum_{J' \geq |M|} D(J') . \quad (3.147)$$

On peut dire aussi : pour chaque valeur de J , il existe une valeur de M égale à J . Cette relation est vraie pour toutes les valeurs de M et pour toutes les valeurs de J , sans aucune contrainte entre J et M , qui parcourent l'ensemble de leurs valeurs possibles indépendamment l'un de l'autre. Écrivons la relation (3.147) pour $M = J$ et pour la valeur consécutive $M = J + 1$:

$$d(J) = \sum_{J' \geq J} D(J') \quad d(J + 1) = \sum_{J' \geq J+1} D(J') , \quad (3.148)$$

²⁸ Deux vecteurs propres associés à deux valeurs propres distinctes d'un même opérateur hermitique (ici J^2) sont orthogonaux.

²⁹ Pour l'instant, on ne se soucie pas de normaliser à l'unité les différents bras qui apparaissent les uns après les autres.

³⁰ D'ailleurs, le calcul traduit par (3.159) montre bien que l'argument rapide ci-dessus n'oublie personne.

et retranchons membre à membre ; il vient :

$$d(J) - d(J + 1) = D(J) . \quad (3.149)$$

Ainsi, la connaissance de la fonction discrète $d(J)$ permet de trouver $D(J)$, que l'on est précisément en train de chercher ; $D(J)$ n'est autre que l'opposé de la dérivée discrète de $d(J)$.

Pour avoir $d(M)$, le plus commode est de raisonner géométriquement dans le plan, en représentant le couple (m_1, m_2) par un point dans le plan³¹. $d(M)$ est le nombre de points de coordonnées (m_1, m_2) situés sur la droite d'équation $x + y = M$ et confinés dans un rectangle défini par les valeurs j_1 et j_2 . Toutes les droites sont parallèles à la 2^{ème} bissectrice.

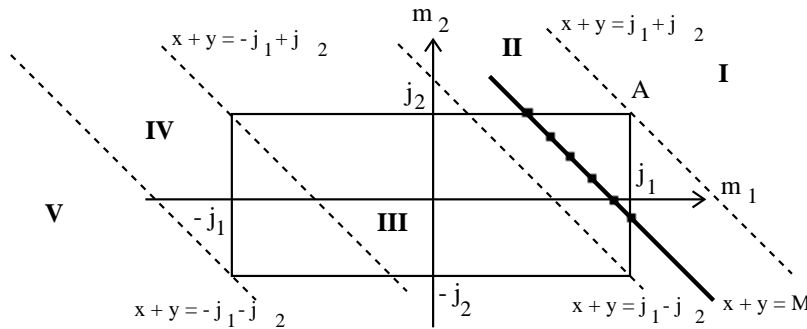


Figure 3.2: Graphique illustrant la dégénérescence $d(M)$ des valeurs de M ($j_1 > j_2$).

La figure (3.2) est tracée dans le cas $j_1 \geq j_2$. Les points “physiques” sont les points de coordonnées entières ou demi-entières situés à l’intérieur ou sur la périphérie du grand rectangle. Les différentes droites parallèles à la deuxième bissectrice délimitent cinq zones dans le plan. La question est de trouver le nombre de points $d(M)$ sur une droite quelconque de même direction et passant par des points de coordonnées entières ou demi-entières ; M lui-même varie par pas d’une unité puisque chaque m_i en fait autant : on doit donc imaginer un réseau de points sous-jacent (non tracé pour la lisibilité de la figure) dont les “sites” sont les points de coordonnées (m_1, m_2) . Considérons maintenant les différentes régions les unes après les autres.

- région I ($M > j_1 + j_2$) : par construction, cette région ne contient aucun point physique, puisque $j_1 + j_2$ est la valeur maximale de j et que l’on ne peut avoir $|M| > j$; il en va de même pour la région V.
- région II ($j_1 - j_2 \leq M \leq j_1 + j_2$) : une droite quelconque située dans cette bande contient $d(M)$ points. Visiblement, chaque fois que la droite est déplacée d’un cran vers la gauche (M décroît d’une unité), elle gagne un point de plus : $d(M)$ croît d’une unité quand M décroît d’une unité, $\Delta[d(M)] = -\Delta M$. $d(M)$ est donc une fonction *linéaire* de M de la forme :

$$d(M) = C - M \quad (3.150)$$

où C est une constante. Écrivons que la droite passant par le point A contient un seul point :

$$1 = d(M = j_1 + j_2) = C - j_1 - j_2 \iff C = j_1 + j_2 + 1 , \quad (3.151)$$

d’où :

$$d(M) = j_1 + j_2 + 1 - M . \quad (3.152)$$

- région III : dans toute cette région, le nombre de points situés sur une droite est visiblement constant ; il est égal au nombre de pas que l’on fait quand on part d’un point d’ordonnée $-j_2$ pour aller, le long de cette droite, en un point d’ordonnée $+j_2$ en sautant d’une unité vers le haut à chaque fois. Il y a évidemment $2j_2 + 1$ tels points.

³¹Pour bien comprendre la suite, il est vivement recommandé de faire une figure soignée sur du papier quadrillé, en choisissant des valeurs pour j_1 et j_2 .

- région IV : c'est l'analogue de ce qui se passe dans la région II, sauf que maintenant M est négatif, d'où la nécessité d'une valeur absolue :

$$d(M) = j_1 + j_2 + 1 - |M| . \quad (3.153)$$

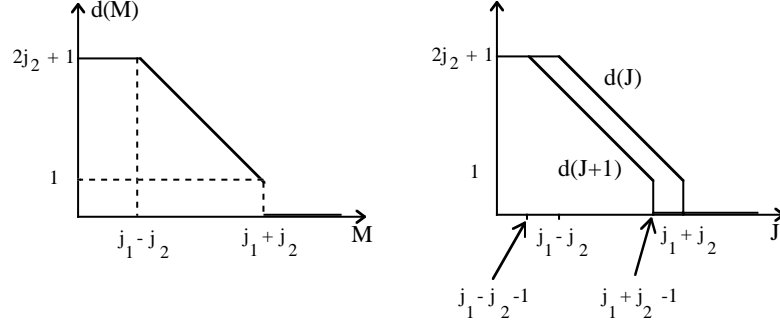


Figure 3.3: Variations de $d(M)$ (à gauche), et de $d(J)$ et $d(J + 1)$ (à droite) ($j_2 \leq j_1$).

Finalement, la fonction $d(M)$, pour $j_1 \geq j_2$ est définie comme suit :

$$d(M) = \begin{cases} 0 & \text{si } |M| > j_1 + j_2 \\ j_1 + j_2 + 1 - |M| & \text{si } j_1 - j_2 \leq |M| \leq j_1 + j_2 \\ 2j_2 + 1 & \text{si } 0 \leq |M| \leq j_1 - j_2 \end{cases} , \quad (3.154)$$

et est une fonction paire comme on pouvait s'y attendre. Ce résultat permet maintenant de tracer la fonction $d(J)$ (voit figure 3.3) ; la relation (3.149) montre alors que :

$$D(J) = \begin{cases} 0 & \text{si } J < j_1 - j_2 \\ 0 & \text{si } J > j_1 + j_2 \\ 1 & \text{si } j_1 - j_2 \leq J \leq j_1 + j_2 \end{cases} . \quad (3.155)$$

Evidemment, si $j_2 \geq j_1$, il suffit d'échanger j_1 et j_2 , et on a alors :

$$d(M) = \begin{cases} 0 & \text{si } |M| > j_1 + j_2 \\ j_2 + j_1 + 1 - |M| & \text{si } j_2 - j_1 \leq |M| \leq j_2 + j_1 \\ 2j_1 + 1 & \text{si } 0 \leq |M| \leq j_2 - j_1 \end{cases} . \quad (3.156)$$

ce qui ne change pas fondamentalement la conclusion pour $D(J) = d(J) - d(J + 1)$: il suffit de remplacer $j_2 - j_1$ par $|j_2 - j_1|$. Au total, dans tous les cas :

$$D(J) = \begin{cases} 0 & \text{si } J < j_1 - j_2 \\ 0 & \text{si } J > j_1 + j_2 \\ 1 & \text{si } |j_1 - j_2| \leq J \leq j_1 + j_2 \end{cases} . \quad (3.157)$$

D'où le théorème fondamental d'addition :

- dans le sous-espace $\{|j_1 j_2 m_1 m_2\rangle\}$ de dimension $d(j_1, j_2) = (2j_1 + 1)(2j_2 + 1)$, les valeurs possibles de J sont :

$$J = |j_1 - j_2|, |j_1 - j_2| + 1, |j_1 - j_2| + 2, \dots, j_1 + j_2 - 1, j_1 + j_2 \quad (3.158)$$

- pour chaque valeur de J , il existe un et un seul ensemble de $(2J + 1)$ vecteurs du type $|JM\rangle$, propres de (\vec{J}^2, J_z) .

La somme des dimensions des sous-espaces J fixé doit être égale à $d(j_1, j_2)$; de fait, il n'est pas difficile de démontrer que³² :

$$\sum_{k=0}^{2 \operatorname{Inf}[j_1, j_2]} [2(|j_1 - j_2| + k) + 1] = d(j_1, j_2) , \quad (3.159)$$

comme il se doit.

Outre les propriétés d'orthogonalité déjà écrites, les CG possèdent des propriétés importantes qu'il faut retenir :

1. Compte tenu des résultats précédents, pour que le CG $\langle j_1 j_2 m_1 m_2 | JM j_1 j_2 \rangle$ soit non-nul il faut que l'on ait simultanément :

$$|j_1 - j_2| \leq J \leq j_1 + j_2 , \quad m_1 + m_2 = M . \quad (3.160)$$

Ces conditions, très importantes, sont la base des règles de sélection classant les raies atomiques en raies permises et raies interdites³³.

2. Les CG sont des produits scalaires de vecteurs propres ; comme ces derniers sont définis à une phase près, il en va de même des CG. Par ailleurs, il existe, à J fixé, des relations de récurrence entre les CG, déduites de l'application répétée des opérateurs J_{\pm} à des états propres de moment cinétique judicieusement choisis. Il en résulte notamment que, J étant fixé, tous les CG peuvent se déduire de l'un d'entre eux ; ceci entraîne que l'arbitraire de phase n'existe que sur le CG qui sert de point de départ à la déduction des autres, les phases des autres se trouvant fixées automatiquement par la récurrence. La convention habituelle consiste à choisir ([17], pp. 392 - 393, eq. (16.72)) :

$$\langle j_1 j_2, m_1 = j_1, m_2 = J - j_1 | J J j_1 j_2 \rangle \text{ réel et positif} . \quad (3.161)$$

3. Avec cette convention, tous les CG sont *réels* (mais pas forcément positifs !). En effet, la récurrence résulte de l'application des formules générales (3.37), où tous les coefficients sont réels. Si le CG qui initie la procédure récursive est conventionnellement choisi réel, tous les autres le seront aussi. Il en résulte que, dans les relations d'orthogonalité (3.133) et (3.134), on peut échanger le bra et le ket d'un produit scalaire, si nécessaire.

Remarques

1. L'addition de *deux* moments cinétiques donne donc des valeurs équidistantes pour la somme, chacune de ces valeurs étant possible une seule fois. Cette dernière propriété disparaît dès que l'on additionne trois moments cinétiques. Par exemple, la somme de trois spins $\frac{1}{2}$ donne un état quadruplet ($S = \frac{3}{2}$) et *deux* états doublets ($S = \frac{1}{2}$). On retrouve bien toutes les dimensions : on part d'un espace d'états dont la dimension est $2^3 = 8$; le quadruplet a $g_{3/2} = 2(\frac{3}{2} + 1) = 4$ composantes, chaque doublet a $g_{1/2} = 2(\frac{1}{2} + 1) = 2$ composantes, et au total (un quadruplet et deux doublets), la dimension de l'espace des états propres du spin total est $g_{3/2} + 2g_{1/2} = 8$, égale à la dimension de l'espace produit tensoriel de trois spins $\frac{1}{2}$
2. Bien sûr, outre les nombres quantiques relatifs aux moments cinétiques, chaque état est éventuellement caractérisé par un ensemble α de valeurs propres d'observables qui, avec les moments cinétiques, constituent un ECOC. En pareil cas, la notation précise des états est ainsi $|\alpha_1 \alpha_2, j_1 j_2 m_1 m_2 \rangle \equiv |\alpha_1, j_1 m_1 \rangle \otimes |\alpha_2, j_2 m_2 \rangle$. En pareil cas, l'équation (3.132) s'écrit :

$$\langle \alpha JM j_1 j_2 \rangle = \sum_{m_1 m_2} |\alpha j_1 j_2 m_1 m_2 \rangle \langle j_1 j_2 m_1 m_2 | JM j_1 j_2 \rangle . \quad (3.162)$$

Les coefficients du développement de droite sont bien indépendants de α . En effet, les opérateurs des moments cinétiques sont représentés sur la base $|JM m_1 m_2 \rangle$ ou $|j_1 j_2 m_1 m_2 \rangle$ par des matrices indépendantes

³² $j_1 + j_2 - |j_1 - j_2| = 2 \operatorname{Inf}[j_1, j_2]$.

³³ Le plus souvent, sauf symétrie du genre "supersymétrie", la notion d'interdiction est relative à une approximation donnée de la description de l'interaction champ-matière (dipolaire électrique, etc).

des autres nombres quantiques α ; ceci traduit le caractère purement géométrique de la composition des moments cinétiques, qui ne dépend que de la valeur des moments cinétiques que l'on additionne, en aucune façon des valeurs propres des autres observables appartenant à l'ECOC. Il en va de même de leurs valeurs propres et de leurs vecteurs propres, d'où l'indépendance des coefficients du développement (3.162) par rapport à l'ensemble des autres nombres quantiques α .

Principes de la méthode de calcul des coefficients de Clebsch - Gordan

Donnons maintenant quelques indications sur la façon d'obtenir effectivement tous les CG dans les cas simples ("petites" valeurs de j_1 et j_2). On commence par déterminer les $2J + 1$ composantes associées à la valeur maximale de J , soit $J = j_1 + j_2$, en commençant par la composante $M = +J (= j_1 + j_2)$, que l'on peut trouver immédiatement. En effet, dans le sous-espace $\mathcal{E}_{\alpha j_1 j_2}$, il existe un seul vecteur ayant cette valeur de J et cette valeur de M ; il s'agit donc d'un vecteur appartenant à la fois à $\{|j_1 j_2 m_1 m_2\rangle\}$ et à $\{|JM j_1 j_2\rangle\}$. D'où l'identité :

$$|J = j_1 + j_2, M = j_1 + j_2, j_1, j_2\rangle = |j_1, j_2, m_1 = j_1, m_2 = j_2\rangle . \quad (3.163)$$

Appliquons maintenant J_- aux deux membres :

$$J_- |J = j_1 + j_2, M = j_1 + j_2, j_1, j_2\rangle = (J_{1-} + J_{2-}) |j_1, j_2, m_1 = j_1, m_2 = j_2\rangle . \quad (3.164)$$

Le premier membre est :

$$\hbar \sqrt{2(j_1 + j_2)} |J = j_1 + j_2, M = j_1 + j_2 - 1, j_1, j_2\rangle . \quad (3.165)$$

Le second membre est³⁴ :

$$\hbar \left[\sqrt{2j_1} |j_1 j_2, m_1 = j_1 - 1, m_2 = j_2\rangle + \sqrt{2j_2} |j_1, j_2, m_1 = j_1, m_2 = j_2 - 1\rangle \right] , \quad (3.166)$$

d'où :

$$\sqrt{2(j_1 + j_2)} |J = j_1 + j_2, M = j_1 + j_2 - 1, j_1, j_2\rangle = \sqrt{2j_1} |j_1 j_2, m_1 = j_1 - 1, m_2 = j_2\rangle + \sqrt{2j_2} |j_1, j_2, m_1 = j_1, m_2 = j_2 - 1\rangle . \quad (3.167)$$

Sur ce développement, on lit les CG correspondants, $\langle j_1 j_2 m_1 m_2 | JM j_1 j_2 \rangle$:

$$\langle j_1, j_2, j_1 - 1, j_2 | j_1 + j_2, j_1 + j_2 - 1, j_1, j_2 \rangle = \sqrt{\frac{j_1}{j_1 + j_2}} . \quad (3.168)$$

$$\langle j_1, j_2, j_1, j_2 - 1 | j_1 + j_2, j_1 + j_2 - 1, j_1, j_2 \rangle = \sqrt{\frac{j_2}{j_1 + j_2}} . \quad (3.169)$$

Une fois trouvé le vecteur $|J = j_1 + j_2, M = j_1 + j_2 - 1, j_1, j_2\rangle$, on peut obtenir la composante suivante $|J = j_1 + j_2, M = j_1 + j_2 - 2, j_1, j_2\rangle$ par une nouvelle application de J_- ; la lecture du développement fournit deux nouveaux CG, et ainsi de suite pour toutes les autres composantes du multiplet $J = j_1 + j_2$.

On passe maintenant à la valeur immédiatement inférieure de J , soit $J = j_1 + j_2 - 1$, et on cherche sa composante maximale $M = J = j_1 + j_2 - 1$. Il y a en réalité deux vecteurs (et deux seulement) ayant cette valeur de M , car il y a deux façons seulement de choisir m_1 et m_2 pour que leur somme soit égale à $j_1 + j_2 - 1$. L'un de ces vecteurs a déjà été trouvé : c'est la "deuxième" composante du multiplet $J = j_1 + j_2$, ayant précisément $M = j_1 + j_2 - 1$. Le deuxième est celui que l'on cherche maintenant : il se trouve par orthogonalité avec le précédent – à une phase près, fixée finalement par la convention énoncée plus haut. Il reste maintenant à effectuer une application répétée de J_- sur le vecteur qui vient d'être trouvé pour obtenir successivement toutes les $2(j_1 + j_2 - 1) + 1$ composantes du multiplet $j_1 + j_2 - 1$. Le développement de chaque composante ainsi trouvée sur les $\{|j_1 j_2 m_1 m_2\rangle\}$ fournit de nouveaux coefficients de Clebsch-Gordan. Puis on recommence avec le multiplet d'ordre immédiatement inférieur, $J = j_1 + j_2 - 2$, et ainsi de suite.

³⁴ $\sqrt{j(j+1) - j(j-1)} = \sqrt{2j}$.

Détaillons la méthode³⁵ schématisée à l'instant, qui permet de calculer en général les coefficients de Clebsch-Gordan. Les calculs suivants sont pénibles, mais ils montrent la puissance et le contenu des relations algébriques caractéristiques des moments cinétiques, qui a permis de constituer des tables donnant les CG.

Compte tenu du fait que tout vecteur propre de J_z de valeur propre donnée $\hbar M$ est une combinaison linéaire des seuls $|j_1 j_2 m_1 m_2\rangle$ tels que $m_1 + m_2 = M$, les CG sont nuls si cette égalité n'est pas vérifiée :

$$\langle j_1 j_2 m_1 m_2 | JM j_1 j_2 \rangle = 0 \quad \text{si } m_1 + m_2 \neq M . \quad (3.170)$$

Comme expliqué ci-dessus, J , j_1 et j_2 étant fixés, tous les CG peuvent se déduire de l'un d'entre eux en utilisant les relations caractéristiques faisant intervenir les opérateurs J_{\pm} . Appliquons d'abord J_- aux deux membres de la relation linéaire existant entre les deux types de vecteurs propres ; il vient :

$$J_- |JM j_1 j_2\rangle = \sum_{m'_1, m'_2} (J_{1-} + J_{2-}) |j_1 j_2 m'_1 m'_2\rangle \langle j_1 j_2 m'_1 m'_2 | JM j_1 j_2 \rangle . \quad (3.171)$$

Posons :

$$R_{\pm}(J, M) = \sqrt{J(J+1) - M(M \pm 1)} , \quad (3.172)$$

d'où résulte :

$$R_{\pm}(J, M \mp 1) = R_{\mp}(J, M) . \quad (3.173)$$

Alors (3.171) s'écrit :

$$\begin{aligned} R_-(J, M) |J, M-1, j_1, j_2\rangle = \\ \sum_{m'_1, m'_2} [R_-(j_1, m'_1) |j_1, j_2, m'_1-1, m'_2\rangle + R_-(j_2, m'_2) |j_1, j_2, m'_1, m'_2-1\rangle] \langle j_1, j_2, m'_1, m'_2 | JM j_1 j_2 \rangle , \end{aligned} \quad (3.174)$$

où la relation (3.173) a été utilisée. Multiplions maintenant à gauche par $\langle j_1 j_2 m_1 m_2 |$; il vient (en utilisant la notation simplifiée $\langle j_1 j_2 m_1 m_2 | JM j_1 j_2 \rangle = \langle m_1 m_2 | JM \rangle$) :

$$R_-(J, M) \langle m_1 m_2 | JM - 1 \rangle = R_+(j_1, m_1) \langle m_1 + 1 m_2 | JM \rangle + R_+(j_2, m_2) \langle m_1 m_2 + 1 | JM \rangle . \quad (3.175)$$

Faisons maintenant le choix :

$$M = J , \quad m_1 = j_1 . \quad (3.176)$$

Le premier terme au second membre de (3.175) est alors nul ($R_+(j_1, m_1 = j_1) = 0$) ; pour obtenir une relation non-triviale (du genre $0 = 0$!), il faut, en vertu de la règle générale : “ $M = m_1 + m_2$ sinon le CG est nul”, choisir $m_1 + m_2 + 1 = M$ (quand on regarde le CG du second membre) ou $m_1 + m_2 = M - 1$ (quand on regarde le CG du premier membre), soit :

$$m_2 = J - j_1 - 1 . \quad (3.177)$$

Il vient alors :

$$\sqrt{2J} \langle j_1, J - j_1 - 1 | J, J - 1 \rangle = \sqrt{j_2(j_2 + 1) - (J - j_1 - 1)(J - j_1)} \langle j_1 J - j_1 | J J \rangle . \quad (3.178)$$

Si donc on connaît $\langle j_1 J - j_1 | J J \rangle$, cette relation permet de calculer $\langle j_1 J - j_1 - 1 | J J - 1 \rangle$.

Recommençons maintenant avec J_+ ; on a :

$$J_+ |JM j_1 j_2\rangle = \sum_{m'_1, m'_2} (J_{1+} + J_{2+}) |j_1 j_2 m'_1 m'_2\rangle \langle j_1 j_2 m'_1 m'_2 | JM j_1 j_2 \rangle , \quad (3.179)$$

c'est-à-dire :

$$\begin{aligned} R_+(J, M) |JM + 1 j_1 j_2\rangle = \\ \sum_{m'_1, m'_2} [R_+(j_1, m'_1) |j_1 j_2 m'_1 + 1 m'_2\rangle + R_+(j_2, m'_2) |j_1 j_2 m'_1 m'_2 + 1\rangle] \langle j_1 j_2 m'_1 m'_2 | JM j_1 j_2 \rangle . \end{aligned} \quad (3.180)$$

³⁵L'exposé qui suit est dû à Merzbacher ([17], p. 390 et sq).

Multiplions maintenant à gauche par $\langle j_1 j_2 m_1 m_2 |$; il vient (en utilisant toujours la notation simplifiée) :

$$R_+(J, M) \langle m_1 m_2 | J M + 1 \rangle = R_-(j_1, m_1) \langle m_1 - 1 m_2 | J M \rangle + R_-(j_2, m_2) \langle m_1 m_2 - 1 | J M \rangle . \quad (3.181)$$

Faisons maintenant le choix :

$$m_1 = j_1 , \quad m_2 = J - j_1 . \quad (3.182)$$

Le 2^{ème} coefficient au second membre n'est pas nul si $m_1 + m_2 - 1 = M$, soit $M = J - 1$ et l'on obtient :

$$\begin{aligned} \sqrt{2J} \langle j_1 J - j_1 | J J \rangle &= \sqrt{2j_1} \langle j_1 - 1 J - j_1 | J J - 1 \rangle \\ &+ \sqrt{j_2(j_2 + 1) - (J - j_1)(J - j_1 - 1)(J - j_1)} \langle j_1 J - j_1 - 1 | J J - 1 \rangle ; \end{aligned} \quad (3.183)$$

$\langle j_1 J - j_1 - 1 | J J - 1 \rangle$ est obtenu de la première relation si l'on connaît $\langle j_1 J - j_1 | J J \rangle$; dès que ce dernier coefficient est connu, on peut alors trouver $\langle j_1 - 1 J - j_1 | J J - 1 \rangle$, et ainsi de suite. Ainsi, pour J, j_1 et j_2 fixés, tous les CG peuvent bien se déduire de l'un d'entre eux, choisi ici comme étant $\langle j_1 J - j_1 | J J \rangle \equiv \langle j_1 j_2 j_1 J - j_1 | J J \rangle \equiv \langle j_1 j_2 j_1 J - j_1 | J J j_1 j_2 \rangle$, conformément à l'usage. Ce coefficient est non nul si $J - j_1$ est compris entre $-j_2$ et $+j_2$:

$$-j_2 \leq J - j_1 \leq j_2 \quad \Longleftrightarrow \quad j_1 - j_2 \leq J \leq j_1 + j_2 . \quad (3.184)$$

Par les mêmes procédés, on peut bien évidemment exprimer tous les CG à partir de $\langle J - j_2 j_2 | J J \rangle \equiv \langle j_1 j_2 J - j_2 - 1 j_2 | J J j_1 j_2 \rangle$ et pour que la solution ne soit pas triviale, il faut :

$$-j_1 \leq J - j_2 \leq j_1 \quad \Longleftrightarrow \quad j_2 - j_1 \leq J \leq j_1 + j_2 . \quad (3.185)$$

Au total, j_1 et j_2 étant donnés, il n'y a de solution non-triviale que si l'inégalité triangulaire suivante est satisfaite :

$$|j_2 - j_1| \leq J \leq j_1 + j_2 , \quad (3.186)$$

ce que l'on savait déjà. Elle donne l'intervalle auquel appartient le nombre J associé à la longueur du moment cinétique total, j_1 et j_2 étant donnés. Dans cet intervalle, J ne varie pas continûment : tous les entiers M, m_1 et m_2 varient d'une unité dans leurs intervalles respectifs $\pm J, \pm j_1$ et $\pm j_2$. Il en résulte que J ne peut varier également que par unité et, finalement, on retrouve bien que les valeurs possibles pour le moment total sont :

$$J = |j_2 - j_1|, |j_2 - j_1| + 1, |j_2 - j_1| + 2, \dots, j_1 + j_2 - 1, j_1 + j_2 . \quad (3.187)$$

La méthode purement algébrique utilisée ci-dessus permet bien de retrouver le théorème fondamental d'addition donné plus haut, exprimé par (3.160).

3.6 Théorème de Wigner-Eckart. Règles de sélection

Les opérateurs scalaires et vectoriels ont été définis au chapitre II en relation avec leur mode de transformation par une rotation R ; les opérateurs scalaires sont invariants et les composantes V_u des opérateurs vectoriels se transforment comme les trois composantes cartésiennes d'un vecteur de la géométrie élémentaire dans \mathbb{R}^3 :

$$V'_u \equiv R V_u R^\dagger = \sum_v V_v R_{vu} , \quad (3.188)$$

où R_{vu} définit la matrice associée à la rotation considérée³⁶. Il s'agit ici d'une matrice 3×3 , tout comme les matrices des rotations géométriques.

³⁶Cette définition est celle qui est utilisée par la plupart des auteurs, et c'est celle qui va se généraliser pour les opérateurs tensoriels. En géométrie ordinaire, on aurait plutôt défini la matrice de rotation par :

$$V'_u \equiv R V_u R^\dagger = \sum_v R_{uv} V_v , \quad (3.189)$$

en rangeant comme d'habitude en colonne les composantes des vecteurs. Il est clair que la définition de la matrice associée à une rotation donnée est purement conventionnelle : on peut sans dommage faire un choix ou l'autre, l'essentiel est de s'y tenir.

D'un autre côté, on a vu que toute rotation géométrique $\mathcal{R}_{\vec{u}\theta}$ induit dans l'espace des états une transformation unitaire $R_{\vec{u}\theta}$ qui s'exprime à l'aide du moment cinétique :

$$R_{\vec{u}\theta} = e^{\frac{1}{i\hbar}\theta \vec{u} \cdot \vec{J}} . \quad (3.190)$$

Ceci entraîne, en particulier, qu'un vecteur $|jm\rangle$ se transforme suivant :

$$|jm\rangle' = e^{\frac{1}{i\hbar}\theta \vec{u} \cdot \vec{J}} |jm\rangle = \sum_{m'} |jm'\rangle \langle jm'| e^{\frac{1}{i\hbar}\theta \vec{u} \cdot \vec{J}} |jm\rangle . \quad (3.191)$$

Notons que la rotation ne peut changer que les nombres quantiques magnétiques : elle ne peut changer j , puisque les trois composantes du moment cinétique commutent avec le carré du module et que l'opérateur de rotation ne dépend que de ces trois composantes ; il en résulte que $|jm'\rangle$ n'a de composantes non nulles que sur des bras de même j et donc que la somme court seulement sur m' . Ce qui précède peut s'écrire :

$$|jm'\rangle = \sum_{m'} |jm'\rangle R_{m'm}^{(j)} , \quad R_{m'm}^{(j)} = \langle jm'| e^{\frac{1}{i\hbar}\theta \vec{u} \cdot \vec{J}} |jm\rangle , \quad (3.192)$$

et met à nouveau en évidence une matrice de rotation qui, elle, peut maintenant avoir une dimension quelconque (entière !), égale à $2j+1$. Avec cette notation, les matrices des rotations ordinaires (et celles qui donnent le mode de transformation des opérateurs vectoriels) peuvent se noter désormais $R_{m'm}^{(1)}$ ($3 = 2 \times 1 + 1$). En redéfinissant convenablement les composantes standard d'un opérateur vectoriel à partir de ses composantes cartésiennes, on peut toujours écrire la combinaison linéaire suivante :

$$V_m' = \sum_{m'} V_{m'} R_{m'm}^{(1)} , \quad R_{m'm}^{(1)} = \langle 1m'| e^{\frac{1}{i\hbar}\theta \vec{u} \cdot \vec{J}} |1m\rangle . \quad (3.193)$$

C'est ce mode de transformation qui est généralisé pour poser la définition d'un opérateur tensoriel irréductible³⁷ (OTI) d'ordre k : les $(2k+1)$ opérateurs $T_q^{(k)}$ (q variant par pas d'une unité de $-k$ à $+k$) sont les composantes standard d'un opérateur tensoriel irréductible d'ordre k , noté $\mathbf{T}^{(k)}$, s'ils se transforment par rotation suivant la loi :

$$[T_q^{(k)}]' \equiv R T_q^{(k)} R^\dagger = \sum_{q'} T_{q'}^{(k)} R_{q'q}^{(k)} . \quad (3.194)$$

Il est facile de montrer ([20], p. 488) que cette définition est équivalente à la satisfaction des relations de commutation suivantes :

$$[J_\pm, T_q^{(k)}]' = \hbar \sqrt{k(k+1) - q(q \pm 1)} T_{q \pm 1}^{(k)} , \quad [J_z, T_q^{(k)}]' = \hbar q T_q^{(k)} . \quad (3.195)$$

Les scalaires sont des OTI d'ordre 0 ; les opérateurs vectoriels sont des OTI d'ordre 1, dont les composantes standard s'expriment comme suit à l'aide des composantes cartésiennes :

$$V_{\pm 1}^{(1)} = \mp \frac{1}{\sqrt{2}} (V_x \pm iV_y) , \quad V_0^{(1)} = V_z . \quad (3.196)$$

On vérifie facilement que \vec{r} et \vec{p} sont des opérateurs vectoriels. Pour une particule sans spin, le moment cinétique orbital $\vec{L} = \vec{r} \times \vec{p}$ est également un opérateur vectoriel. Pour une particule avec spin, \vec{S} est un opérateur vectoriel vis-à-vis du moment cinétique *total*³⁸ $\vec{J} = \vec{L} + \vec{S}$. D'un autre côté, un champ externe (électrique, magnétique, ...) n'est pas un opérateur vectoriel, et pour cause : ce n'est pas un opérateur du tout ! En effet, un tel champ est supposé donné une fois pour toutes et n'appartient pas au système quantique ; toutes ses composantes commutent avec tous les opérateurs, ce sont des scalaires au sens défini plus haut. Un tel objet peut être appelé "vecteur scalaire", à condition de bien comprendre de quoi il s'agit. Bien évidemment, si on décide d'inclure le champ dans le système que l'on quantifie, la situation change complètement : alors les sources du champ et le système qui était seul quantifié auparavant forment un et un seul "super-système" quantifié. Toutes les grandeurs dynamiques deviennent alors des observables associées à des opérateurs et pour chacun de ceux-ci, la question de savoir s'ils sont, par exemple, des opérateurs vectoriels, n'est plus triviale.

³⁷Sur l'origine de la terminologie, voir [20], p. 487.

³⁸ \vec{S} commutant avec \vec{L} , les relations caractéristiques (3.195) ne sauraient être satisfaites avec le couple (\vec{L}, \vec{S}) .

Ces définitions étant posées, il est possible de démontrer l'important théorème de Wigner-Eckart³⁹, dont l'énoncé est le suivant :

Dans une représentation $|\alpha jm\rangle$, l'élément de matrice d'une composante standard d'OTI est proportionnel à un coefficient de Clebsch-Gordan⁴⁰ :

$$\langle \alpha jm | T_q^{(k)} | \alpha' j' m' \rangle = \langle \alpha j || \mathbf{T}_q^{(k)} || \alpha' j' \rangle \langle j' k m' q | j m \rangle , \quad (3.198)$$

où $\langle \alpha j || \mathbf{T}_q^{(k)} || \alpha' j' \rangle$ dépend de α, α', j, j' (et de l'OTI !) et s'appelle élément de matrice réduit.

Pour démontrer ce théorème, considérons les $(2k+1)(2j'+1)$ vecteurs $T_q^{(k)}|\alpha' j' m'\rangle$, qui ne sont d'ailleurs pas tous forcément linéairement indépendants, puis formons-en les combinaisons linéaires suivantes, fabriquées avec les CG :

$$|\psi_{j''m''}\rangle = \sum_{m', q'} T_{q'}^{(k)} |\alpha' j' m'\rangle \langle j' k m' q' | j'' m'' \rangle . \quad (3.199)$$

Comme les $T_q^{(k)}|\alpha' j' m'\rangle$ ne sont pas tous forcément linéairement indépendants, les $|\psi_{j''m''}\rangle$ peuvent être nuls. Multiplions membre à membre (3.199) par le CG $\langle j' k j'' m'' | j' k m'' q \rangle \equiv \langle j'' m'' | j' k m'' q \rangle$ et sommions sur j'', m'' :

$$\sum_{j'', m''} \langle j'' m'' | j' k m'' q \rangle |\psi_{j''m''}\rangle = \sum_{m', q'} T_{q'}^{(k)} |\alpha' j' m'\rangle \sum_{j'', m''} \langle j'' m'' | j' k m'' q \rangle \langle j' k m' q' | j'' m'' \rangle . \quad (3.200)$$

La dernière somme vaut $\delta_{m'm''} \delta_{qq'}$, compte tenu des relations d'orthogonalité des CG (3.133) et (3.134). D'où, réciproquement :

$$T_q^{(k)} |\alpha' j' m'\rangle = \sum_{j'', m''} |\psi_{j''m''}\rangle \langle j'' m'' | j' k m' q \rangle . \quad (3.201)$$

Il en résulte :

$$\langle \alpha jm | T_q^{(k)} | \alpha' j' m' \rangle = \sum_{j'', m''} \langle \alpha jm | \psi_{j''m''} \rangle \langle j'' m'' | j' k m' q \rangle . \quad (3.202)$$

Par ailleurs, en vertu de (3.195) :

$$\begin{aligned} J_+ T_q^{(k)} |\alpha' j' m'\rangle &= [J_+, T_q^{(k)}] |\alpha' j' m'\rangle + T_q^{(k)} J_+ |\alpha' j' m'\rangle = \\ &= \hbar \sqrt{k(k+1) - q(q+1)} T_{q+1}^{(k)} |\alpha' j' m'\rangle + \hbar \sqrt{j'(j'+1) - m'(m'+1)} T_q^{(k)} |\alpha' j' m'+1\rangle . \end{aligned} \quad (3.203)$$

D'où, repartant de (3.199) :

$$\begin{aligned} \hbar^{-1} J_+ |\psi_{j''m''}\rangle &= \sum_{m', q'} \langle j' k m' q' | j'' m'' \rangle [\sqrt{k(k+1) - q'(q'+1)} T_{q'+1}^{(k)} |\alpha' j' m'\rangle \\ &+ \sqrt{j'(j'+1) - m'(m'+1)} T_q^{(k)} |\alpha' j' m'+1\rangle] \\ &= \sum_{m', q'} T_{q'}^{(k)} |j' m'\rangle [\sqrt{k(k+1) - (q'-1)q'} \langle j' k m' q' - 1 | \alpha' j'' m'' \rangle \\ &+ \sqrt{j'(j'+1) - (m'-1)m'} \langle j' k m' - 1 q' | j'' m'' \rangle] . \end{aligned} \quad (3.204)$$

D'après (3.181) en remplaçant respectivement J par j'' et M par m'' :

$$\begin{aligned} \sqrt{j''(j''+1) - m''(m''+1)} \langle j_1 j_2 m_1 m_2 | j'' m'' + 1 \rangle &= \\ \sqrt{j_1(j_1+1) - m_1(m_1-1)} \langle j_1 j_2 m_1 - 1 m_2 | j'' m'' \rangle &+ \sqrt{j_2(j_2+1) - m_2(m_2-1)} \langle j_1 j_2 m_1 m_2 - 1 | j'' m'' \rangle . \end{aligned} \quad (3.205)$$

³⁹ j et m ont la signification habituelle ; α est un nombre quantique (ou un ensemble de nombres quantiques) nécessaire pour spécifier complètement l'état quantique, indépendamment des valeurs du moment cinétique. Ce peut être par exemple le nombre quantique n spécifiant la valeur de l'énergie pour le champ central Coulombien.

⁴⁰ Dans la suite, on utilise les notations simplifiées quand c'est possible :

$$\langle j_1 j_2 m_1 m_2 | JM j_1 j_2 \rangle \equiv \langle j_1 j_2 m_1 m_2 | JM \rangle = \langle m_1 m_2 | JM \rangle , \text{ ici } \langle j' k m' q | j m j' k \rangle \equiv \langle j' k m' q | j m \rangle = \langle m' q | j m \rangle . \quad (3.197)$$

La quantité entre crochets dans (3.204) est donc :

$$\sqrt{j''(j''+1) - m''(m''+1)} \langle j' km' q' | j'' m'' + 1 \rangle , \quad (3.206)$$

d'où :

$$\hbar^{-1} J_+ |\psi_{j'' m''}\rangle = \sum_{m', q'} T_q^{(k)} |\alpha' j' m'\rangle \sqrt{j''(j''+1) - m''(m''+1)} \langle j' km' q' | j'' m'' + 1 \rangle , \quad (3.207)$$

soit finalement :

$$\hbar^{-1} J_+ |\psi_{j'' m''}\rangle = \sqrt{j''(j''+1) - m''(m''+1)} |\psi_{j'' m''+1}\rangle . \quad (3.208)$$

Par des moyens analogues, il est facile d'établir les deux relations suivantes :

$$\hbar^{-1} J_- |\psi_{j'' m''}\rangle = \sqrt{j''(j''+1) - m''(m''-1)} |\psi_{j'' m''-1}\rangle \quad (3.209)$$

et :

$$\hbar^{-1} J_z |\psi_{j'' m''}\rangle = m'' |\psi_{j'' m''}\rangle . \quad (3.210)$$

Les trois relations (3.208), (3.209) et (3.210) montrent que, quand ils ne sont pas nuls, les $|\psi_{j'' m''}\rangle$ sont des vecteurs propres de (\vec{J}^2, J_z) . Il en résulte, dans tous les cas, que les produits scalaires $\langle \alpha j m | \psi_{j'' m''}\rangle$ sont nuls, sauf quand $|\psi_{j'' m''}\rangle$ n'est pas nul et si $j'' = j$, $m'' = m$. Avec l'expression de l'élément de matrice (3.202) on trouve alors :

$$\langle \alpha j m | T_q^{(k)} | \alpha' j' m' \rangle = \langle \alpha j m | \psi_{j m} \rangle \langle j m | j' k m' q \rangle . \quad (3.211)$$

Ce résultat montre que l'élément de matrice de l'opérateur tensoriel irréductible est proportionnel à un coefficient de Clebsch-Gordan⁴¹. La constante de proportionnalité, $\langle \alpha j m | \psi_{j m} \rangle$, est précisément :

$$\langle \alpha j m | \psi_{j m} \rangle = \sum_{m', q'} \langle \alpha j m | T_q^{(k)} | \alpha' j' m' \rangle \langle j' k m' q' | j m \rangle . \quad (3.212)$$

Elle est indépendante de m , comme tout produit scalaire, mais dépend de α , α' , j et j' – et bien sûr de l'OTI considéré ; c'est elle que l'on note⁴² $\langle \alpha j | \mathbf{T}^{(k)} | \alpha' j' \rangle$. Finalement :

$$\langle \alpha j m | T_q^{(k)} | \alpha' j' m' \rangle = \langle \alpha j | \mathbf{T}^{(k)} | \alpha' j' \rangle \langle j' k m' q | j m \rangle . \quad (3.213)$$

Donnons un exemple de l'importance et de l'utilité de ce théorème en calculant les éléments de matrice des opérateurs vectoriels. Écrivons d'abord les composantes standard du moment cinétique ; d'après (3.196) :

$$J_{+1}^{(1)} = -\frac{1}{\sqrt{2}} (J_x + i J_y) , \quad J_{-1}^{(1)} = +\frac{1}{\sqrt{2}} (J_x - i J_y) , \quad J_0^{(1)} = J_z . \quad (3.214)$$

Par le théorème de Wigner-Eckart, (3.213), on a :

$$\langle \alpha j j | J_z | \alpha' j j \rangle = \langle \alpha j | \vec{J} | \alpha' j \rangle \langle j 1 j 0 | j j \rangle . \quad (3.215)$$

Pour un système invariant par rotation, le premier membre est proportionnel à $\delta_{\alpha\alpha'}$ puisque α est le nombre quantique associé à l'énergie (ou un ensemble contenant celui qui est associé à l'énergie). Dans (3.215) le CG à droite vaut $\sqrt{j/(j+1)}$, d'où :

$$\langle \alpha j | \vec{J} | \alpha' j' \rangle = \hbar \sqrt{j(j+1)} \delta_{j j'} \delta_{\alpha\alpha'} . \quad (3.216)$$

Le moment magnétique d'un atome résulte tout à la fois du moment cinétique orbital et du moment cinétique intrinsèque de spin ; on peut donc l'écrire en combinaison linéaire des opérateurs vectoriels correspondants. Il se trouve que, outre le facteur gyromagnétique habituel, un facteur numérique⁴³ g_S doit être introduit (c'est l'expérience qui l'impose) ; on écrit comme suit le moment magnétique :

$$\vec{\mu} = \frac{e}{2m} (\vec{L} + g_S \vec{S}) . \quad (3.217)$$

⁴¹dont on rappelle au passage que l'usage est de les choisir tous réels, d'où $\langle J m | j k m q \rangle = \langle J k m q | j m \rangle$.

⁴²Messiah ([20]) introduit un facteur $(2j+1)^{-1/2}$, ce que ne font pas d'autres auteurs. Ceci est une pure affaire de définition : prendre garde à la définition utilisée dans un ouvrage donné.

⁴³ $g_S = 2.0023\dots$; en pratique, on prend le plus souvent $g_S = 2$.

\vec{L} et \vec{S} étant des opérateurs vectoriels, il est clair que $\vec{\mu}$ en est un aussi ($k = 1$). La moyenne de la valeur maximale de la projection μ_z (correspondant à $q = 0$) est conventionnellement appelée “moment magnétique de l’atome” dans l’état considéré ; d’après le théorème de Wigner-Eckart, elle vaut :

$$\mu_{\text{at}} = \langle \alpha j j | \mu_z | \alpha j j \rangle = \langle \alpha j | \vec{\mu} | \alpha' j \rangle \langle J j | j 1 j 0 \rangle , \quad \langle J j | j 1 j 0 \rangle = \sqrt{\frac{j}{j+1}} . \quad (3.218)$$

Chapitre 4

Potentiel central et atome d'hydrogène

*Ce chapitre est consacré à l'étude générale du champ central ;
après le traitement de quelques exemples,
dont certains généralisent à trois dimensions
des résultats obtenus dans \mathbb{R} ,
la théorie détaillée de l'atome d'hydrogène sera exposée.*

4.1 Définition du champ central et exemples

Un champ de force conservatif dérive par définition d'un potentiel V , fonction qui dépend du rayon-vecteur fixant le point de l'espace où on calcule la force ; la relation est bien connue :

$$\vec{F}(\vec{r}) = -\vec{\nabla}V(\vec{r}) . \quad (4.1)$$

Lorsque la force \vec{F} est colinéaire à \vec{r} , elle est dite centrale ; alors, par le théorème du moment cinétique, \vec{L} est une constante du mouvement :

$$\text{champ central} \iff \vec{F} \propto \vec{r} \iff \vec{L} \equiv \vec{r} \times \vec{p} = \text{C}^{\text{ste}} . \quad (4.2)$$

Dans ces conditions, seule la composante du gradient parallèle à \vec{r} , égale à $\partial/\partial r$, est non nulle. Il en résulte que l'énergie potentielle V ne dépend que de r , module de \vec{r} , et se note plus simplement $V(r)$. Pour un tel champ central, et pour une particule de masse m_p , le Hamiltonien s'écrit :

$$H = \frac{\vec{p}^2}{2m_p} + V(r) . \quad (4.3)$$

En réalité, le problème se présente rarement d'emblée de cette façon. Souvent, on part de deux particules ponctuelles, sans structure interne, de masses m_1 et m_2 en interaction par une force obéissant à la loi de l'action et de la réaction : la force exercée par la particule 1 sur la particule 2 est opposée (vectoriellement) à la force exercée par la particule 2 sur la particule 1. En l'absence de toute direction privilégiée, notamment pour deux particules constituant un système isolé, l'énergie potentielle dépend alors de la seule distance entre les deux particules ; en outre, le mouvement du centre de masse est une translation uniforme. On peut alors se ramener à un problème plus simple : une fois évacuée la translation uniforme du centre de masse, subsiste un problème à 3 degrés de liberté représentant le mouvement *relatif* des deux particules dans le repère de leur centre de masse, formellement identique à celui d'une particule unique de masse μ (masse réduite) égale à $m_1 m_2 / (m_1 + m_2)$ et d'énergie potentielle $V(r)$; c'est ce que l'on appelle la "réduction du problème à deux corps en interaction centrale".

L'exemple sans doute le plus connu est le potentiel Coulombien, d'ailleurs formellement identique au potentiel de gravitation¹, pour lequel $V(r)$ a la forme précise :

$$V(r) = \frac{q_1 q_2}{4\pi\epsilon_0 r} . \quad (4.4)$$

¹sauf que dans ce dernier cas les scalaires qui apparaissent (les masses) sont des quantités toujours essentiellement positives.

Dans le cas d'un atome hydrogénoïde de charge nucléaire² $Z|e|$, l'énergie potentielle d'interaction entre le noyau et l'électron est :

$$V(r) = -\frac{Ze^2}{4\pi\epsilon_0 r} \equiv -\frac{Ze'^2}{r} \quad (e'^2 = \frac{e^2}{4\pi\epsilon_0}) . \quad (4.5)$$

D'autres cas méritent l'intérêt : avec son énergie potentielle constante (en pratique nulle), la particule libre est un exemple de champ central, dont la trivialité n'est qu'apparente et fournit l'occasion d'une belle illustration des théorèmes de développement (on peut développer les ondes planes $e^{i\vec{k}\cdot\vec{r}}$ sur les états propres à variables séparées – qui sont également propres du moment cinétique). Un autre exemple intéressant est celui du puits fini sphérique défini par :

$$V(r) = \begin{cases} -V_0 & \text{si } r < a \\ 0 & \text{si } r > a \end{cases} . \quad (4.6)$$

Ces exemples seront abordés en temps utile, après la formation explicite du Hamiltonien exprimé en coordonnées sphériques – coordonnées adaptées à la symétrie du problème considéré. Bien évidemment, on rencontrera ici et là des applications concrètes de résultats énoncés dans les deux chapitres précédents, à propos de la symétrie en général et du rôle du moment cinétique – il s'agira alors du moment cinétique orbital, $\vec{L} = \vec{r} \times \vec{p}$.

4.1.1 Hamiltonien du problème central

Le Hamiltonien du problème central est :

$$H = \frac{\vec{p}^2}{2\mu} + V(r) \quad (4.7)$$

où μ est la masse réduite et $\vec{p} = -i\hbar\vec{\nabla} = -i\hbar(\partial/\partial x, \partial/\partial y, \partial/\partial z)$, conformément à la prescription standard de la représentation- q . Pour obtenir l'expression du terme cinétique en coordonnées sphériques, deux méthodes sont possibles :

- la première méthode consiste à effectuer le changement de variables dans les opérateurs différentiels représentant l'énergie cinétique T , pour passer des coordonnées cartésiennes (x, y, z) aux coordonnées sphériques (r, θ, ϕ) . Un calcul facile mais laborieux donne le résultat suivant :

$$\vec{p}^2 = -\frac{\hbar^2}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{\hbar^2}{r^2} \left(\frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right) . \quad (4.8)$$

On reconnaît dans la partie angulaire l'expression du carré du moment cinétique obtenue au ch. 3 ; d'où l'expression de l'énergie cinétique T :

$$T \equiv \frac{\vec{p}^2}{2\mu} = -\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{\vec{L}^2}{2\mu r^2} . \quad (4.9)$$

Le terme $\vec{L}^2/(2\mu r^2)$ 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

- la deuxième méthode part des expressions classiques ; on a :

$$\vec{L}^2 = (\vec{r} \times \vec{p})^2 = r^2 p^2 \sin^2 \Theta = r^2 p^2 (1 - \cos^2 \Theta) \equiv \vec{r}^2 \vec{p}^2 - (\vec{r} \cdot \vec{p})^2 , \quad (4.10)$$

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

$$\vec{p}^2 = \left(\frac{\vec{r} \cdot \vec{p}}{r} \right)^2 + \frac{\vec{L}^2}{r^2} . \quad (4.11)$$

²Dans tout le cours, e désigne la charge négative de l'électron : $e = -1.6 \times 10^{-19}$ C.

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 \vec{L}^2/r^2 car les deux opérateurs commutent entre eux. La seule difficulté réside dans le terme p_r^2 :

$$p_r^2 \equiv \left(\frac{\vec{r} \cdot \vec{p}}{r} \right)^2 \iff p_r = \frac{\vec{r} \cdot \vec{p}}{r}, \quad (4.12)$$

dont la symétrisation "naturelle" est :

$$p_r = \frac{1}{2} \left(\frac{\vec{r}}{r} \cdot \vec{p} + \vec{p} \cdot \frac{\vec{r}}{r} \right). \quad (4.13)$$

En vertu de (4.14), on peut écrire³ :

$$\vec{p} \cdot \frac{\vec{r}}{r} = [\vec{p}, \frac{\vec{r}}{r}] + \frac{\vec{r}}{r} \cdot \vec{p} = -i\hbar \left(\vec{\nabla} \cdot \frac{\vec{r}}{r} \right) + \frac{\vec{r}}{r} \cdot \vec{p}. \quad (4.15)$$

L'expression symétrisée (4.13) conduit alors à :

$$p_r = \frac{\vec{r}}{r} \cdot \vec{p} - \frac{i\hbar}{2} \left(\vec{\nabla} \cdot \frac{\vec{r}}{r} \right) = -i\hbar \left[\frac{\vec{r}}{r} \cdot \vec{\nabla} + \frac{1}{2} \left(\vec{\nabla} \cdot \frac{\vec{r}}{r} \right) \right]. \quad (4.16)$$

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

$$\left(\vec{\nabla} \cdot \frac{\vec{r}}{r} \right) = \frac{\partial}{\partial x} \frac{x}{r} + \frac{\partial}{\partial y} \frac{y}{r} + \frac{\partial}{\partial z} \frac{z}{r}. \quad (4.17)$$

Le terme en x est :

$$\frac{\partial}{\partial x} \frac{x}{r} = \frac{1}{r} - \frac{x^2}{r^3}. \quad (4.18)$$

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

$$\left(\vec{\nabla} \cdot \frac{\vec{r}}{r} \right) = \frac{3}{r} - \frac{x^2 + y^2 + z^2}{r^3} = \frac{2}{r}, \quad (4.19)$$

d'où, selon (4.16) :

$$p_r = -i\hbar \left(\frac{\partial}{\partial r} + \frac{1}{r} \right) \equiv -i\hbar \frac{1}{r} \frac{\partial}{\partial r} r. \quad (4.20)$$

On remarque au passage que p_r n'est pas égal à $-i\hbar(\partial/\partial r)$ en conséquence du fait que r n'est pas une coordonnée rectangulaire. Avec l'expression (4.20), 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 cartésiennes⁴.

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

$$\left(\frac{\partial}{\partial r} + \frac{1}{r} \right)^2 \bullet = \left(\frac{\partial}{\partial r} \frac{\partial}{\partial r} + \frac{\partial}{\partial r} \frac{1}{r} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \right) \bullet, \quad (4.21)$$

on trouve⁵ :

$$p_r^2 = -\hbar^2 \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} \right) \equiv -\hbar^2 \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right). \quad (4.23)$$

³Pour toute fonction vectorielle $\vec{f}(\vec{r})$ on a :

$$[\vec{p}, \vec{f}(\vec{r})] = -i\hbar \left(\vec{\nabla} \cdot \vec{f}(\vec{r}) \right) \mathbf{1}; \quad (4.14)$$

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.

⁴toute expression du genre $\varpi(r) \equiv -i\hbar[(\partial/\partial r) + \phi(r)]$ conduit d'ailleurs elle aussi à la relation de commutation $[r, \varpi(r)] = i\hbar$.

⁵D'une façon générale, dans \mathbb{R}^d , p_r^2 est donné par :

$$p_r^2 = -\hbar^2 \left(\frac{\partial^2}{\partial r^2} + \frac{d-1}{r} \frac{\partial}{\partial r} + \frac{(d-1)(d-3)}{4r^2} \right). \quad (4.22)$$

En ajoutant $\vec{L}^2/(2\mu r^2)$, l'expression (4.9) est retrouvée.

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

$$H = \frac{p_r^2}{2\mu} + \frac{\vec{L}^2}{2\mu r^2} + V(r) \equiv -\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{\vec{L}^2}{2\mu r^2} + V(r) \equiv T + V, \quad (4.24)$$

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 (4.24) 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 :

$$\langle \psi_1 | p_r \psi_2 \rangle = \langle p_r \psi_1 | \psi_2 \rangle \iff \int_0^{+\infty} r^2 dr \psi_1^*(r) [p_r \psi_2(r)] = \int_0^{+\infty} r^2 dr [p_r \psi_1(r)]^* \psi_2(r) \quad (4.25)$$

soit, explicitement compte tenu de (4.20) :

$$-i\hbar \int_0^{+\infty} r^2 dr \psi_1^*(r) \left[\left(\frac{\partial}{\partial r} + \frac{1}{r} \right) \psi_2(r) \right] = +i\hbar \int_0^{+\infty} r^2 dr \left[\left(\frac{\partial}{\partial r} + \frac{1}{r} \right) \psi_1^*(r) \right] \psi_2(r). \quad (4.26)$$

Le premier membre est :

$$-i\hbar \int_0^{+\infty} r^2 dr \psi_1^*(r) [\psi_2'(r) + r^{-1}\psi_2(r)]. \quad (4.27)$$

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

$$+i\hbar [r^2 \psi_1^*(r) \psi_2(r)]_0^{+\infty} - i\hbar \int_0^{+\infty} dr \psi_1^*(r) [r \psi_2(r) + r^2 \psi_2'(r)]. \quad (4.28)$$

En comparant (4.27) et (4.28), on voit que la condition d'hermiticité (4.26) s'écrit :

$$[r^2 \psi_1^*(r) \psi_2(r)]_0^{+\infty} = 0. \quad (4.29)$$

En particulier, pour tout élément diagonal, il faut :

$$[r^2 |\psi(r)|^2]_0^{+\infty} = 0. \quad (4.30)$$

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

$$\lim_{r \rightarrow 0} [r |\psi(r)|] = 0, \quad \lim_{r \rightarrow +\infty} [r |\psi(r)|] = 0. \quad (4.31)$$

Ces conditions suffisantes, notamment celle à l'infini, sont en général satisfaites pour les états liés⁶, 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 $1/r$ 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 (4.31).

⁶Pour 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 strict⁷ : 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 (4.20) l'équation propre pour p_r est :

$$-i\hbar \frac{1}{r} \frac{\partial}{\partial r} [rf(r)] = \lambda f(r) . \quad (4.32)$$

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

$$f(r) = \frac{C^{\text{ste}}}{r} e^{\frac{i}{\hbar} \lambda r} . \quad (4.33)$$

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 hermitique⁸. 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.

Le Hamiltonien H (4.24) possède visiblement la symétrie sphérique ; on est donc assuré d'avance que H et n'importe laquelle des composantes de \vec{L} commutent :

$$[H, \vec{L}] = 0 . \quad (4.34)$$

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 (4.34) 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 forme⁹ de H , tous les états propres de ce dernier opérateur peuvent s'obtenir comme des combinaisons linéaires¹⁰ de fonctions à variables séparées :

$$\psi(r, \theta, \phi) = R(r) Y_{lm}(\theta, \phi) , \quad (4.35)$$

où les Y_{lm} sont les harmoniques sphériques introduites antérieurement. En effet, reportons une telle forme dans $H\psi \equiv [T + V]\psi = E\psi$; il vient :

$$\left[-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{\vec{L}^2}{2\mu r^2} + V(r) \right] R(r) Y_{lm}(\theta, \phi) = E R(r) Y_{lm}(\theta, \phi) . \quad (4.36)$$

En faisant agir le carré du moment cinétique \vec{L}^2 sur ses états propres Y_{lm} , et après simplification¹¹, il reste l'équation dite radiale pour la seule fonction $R(r)$:

$$-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right) R(r) + \frac{\hbar^2 l(l+1)}{2\mu r^2} R(r) + V(r) R(r) = E R(r) . \quad (4.37)$$

Il est souvent avantageux d'introduire une autre fonction inconnue $u(r) = rR(r)$. On voit sans peine à partir de (4.37) que $u(r)$ satisfait l'équation suivante :

$$-\frac{\hbar^2}{2\mu} \frac{d^2 u}{dr^2} + \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) \right] u(r) \equiv -\frac{\hbar^2}{2\mu} \frac{d^2 u}{dr^2} + V_{\text{eff}}(r) u(r) = E u(r) . \quad (4.38)$$

⁷ voir [16], p. 293.

⁸ Cette propriété de p_r n'est pas si particulière que cela : p_x n'est pas non plus de ce point de vue 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 .

⁹ 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.

¹⁰ En raison de la symétrie de rotation autour notamment de Oz , tous les états propres du type (4.35), 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.

¹¹ L'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.

Cette équation, nettement plus simple que (4.37), 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 nul¹², 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 positive¹³, 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 :

$$\lim_{r \rightarrow 0} [rR(r)] = 0 \iff \lim_{r \rightarrow 0} [u(r)] = 0, \quad \lim_{r \rightarrow \infty} [rR(r)] = 0 \iff \lim_{r \rightarrow \text{infy}} [u(r)] = 0. \quad (4.39)$$

La sous-section 4.1.2 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 :

$$\int_0^{+\infty} r^2 dr R^2(r) = 1 \iff \int_0^{+\infty} dr u^2(r) = 1. \quad (4.40)$$

L'analyse de l'équation de conservation :

$$\frac{\partial \rho}{\partial t} + \text{div } \vec{j} = 0, \quad \rho = \Psi^* \Psi, \quad \vec{j} = \frac{\hbar}{2i\mu} [\Psi^* \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^*] \quad (4.41)$$

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 $\text{div } \vec{j} = 0$. 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 :

$$j_r = \frac{\hbar}{2i\mu} |Y_{lm}|^2 \left[R^* \frac{dR}{dr} - R \frac{dR^*}{dr} \right] \quad (4.42)$$

et il en résulte que :

$$R^* \frac{dR}{dr} = R \frac{dR^*}{dr} \iff \ln R = \ln R^* + C^{\text{ste}} \iff R^* \propto R. \quad (4.43)$$

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.

¹²Les états $l = 0, 1, 2, 3, \dots$ sont traditionnellement notés s, p, d, f, ...

¹³On 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.

La réalité de R peut se voir d'une autre façon, en invoquant l'absence de dégénérescence et la symétrie par renversement du temps¹⁴. La fonction $u(r)$ introduite ci-dessus satisfait de fait l'équation à une dimension sur \mathbb{R}_+ (4.38) ; 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 :

$$[W(u_1, u_2)]_a^b \equiv [u_1 u_2' - u_2 u_1']_a^b = (E_1 - E_2) \int_a^b dr u_1(r) u_2(r) . \quad (4.44)$$

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 :

$$[W(u_1, u_2)]_a^b = 0 , \quad (4.45)$$

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

$$W(u_1, u_2) = \text{C}^{\text{ste}} . \quad (4.46)$$

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, $\forall r$:

$$u_1(r) u_2'(r) - u_2(r) u_1'(r) = 0 \iff u_2(r) \propto u_1(r) . \quad (4.47)$$

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 ¹⁵. 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.

En ce qui concerne les états non-liés, ils ne sont pas non plus dégénérés (pour un l donné). Dans \mathbb{R} , tout mouvement non-lié est dégénéré deux fois, puisque l'inversion de la vitesse ne change pas l'énergie (encore une manifestation de l'invariance par renversement du temps). En revanche, l'introduction d'une barrière parfaitement réfléchissante brise la symétrie gauche-droite, et sur \mathbb{R}_+ le mouvement non-lié perd ainsi sa dégénérescence. Il en va de même ici. La fonction $u(r)$ doit s'annuler en $r = 0$, ce qui montre bien que même dans le cas libre et $l = 0$, on n'a pas la liberté de choisir une combinaison linéaire *arbitraire* des $e^{\pm ikr}$ (ce qui serait le cas en présence d'une dégénérescence égale à 2) : on doit prendre la seule et unique fonction nulle à l'origine soit $u_{l=0}(r) \propto \sin kr$, qui est bien essentiellement réelle.

4.1.2 Comportements de la fonction radiale

D'après la discussion précédente, la fonction d'onde du champ central peut être recherchée sous la forme ($R(r) = u(r)/r$) :

$$\psi(r, \theta, \phi) = \frac{u(r)}{r} Y_{lm}(\theta, \phi) , \quad (4.48)$$

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

$$-\frac{\hbar^2}{2\mu} \frac{d^2 u}{dr^2} + \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) \right] u(r) = E u(r) . \quad (4.49)$$

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).

¹⁴Cette 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 .

¹⁵ 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 (4.38) dépend de l par le terme centrifuge.

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 (4.40), ce qui élimine une grande partie des solutions mathématiques de l'équation différentielle (4.49). 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 (4.39)), 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-liés¹⁶, l'immense majorité des cas pratiques est couverte quand on fait les hypothèses additionnelles suivantes sur la fonction $V(r)$:

$$V(r) \rightarrow 0 \quad \text{si} \quad r \rightarrow +\infty, \quad (4.50)$$

$$V(r) \simeq A r^\alpha \quad \text{si} \quad r \rightarrow 0, \quad (4.51)$$

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 (4.51) est satisfaite. Partant de l'équation pour u , on l'intègre dans un voisinage à droite :

$$\int_0^{\delta r} dr \left[-\frac{\hbar^2}{2\mu} \frac{d^2 u}{dr^2} + \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) \right] u(r) \right] = E \int_0^{\delta r} dr u(r). \quad (4.52)$$

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

$$-\frac{\hbar^2}{2\mu} \left[\frac{du}{dr} \right]_0^{\delta r} + \int_0^{\delta r} dr \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} + A r^\alpha \right] u(r) = 0. \quad (4.53)$$

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 :

$$-\frac{\hbar^2}{2\mu} [\beta r^{\beta-1}]_0^{\delta r} + \int_0^{\delta r} dr \left[\frac{\hbar^2 l(l+1)}{2\mu} r^{\beta-2} + A r^{\alpha+\beta} \right] = 0. \quad (4.54)$$

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

$$-\frac{\hbar^2}{2\mu} \left[\left(\beta - \frac{l(l+1)}{\beta-1} \right) r^{\beta-1} \right]_0^{\delta r} + \left[\frac{A}{\alpha+\beta+1} r^{\alpha+\beta+1} \right]_0^{\delta r} = 0. \quad (4.55)$$

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 (4.55) fournit un terme logarithmique ; en effet, (4.54) donne alors :

$$\int_0^{\delta r} dr \left[\frac{\hbar^2}{2\mu} l(l+1) r^{-1} + A r^{\alpha+1} \right] = 0 \iff \left[\frac{\hbar^2}{2\mu} l(l+1) \ln r + \frac{A}{\alpha+2} r^{\alpha+2} \right]_0^{\delta r} = 0. \quad (4.56)$$

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 :

$$-\frac{\hbar^2}{2\mu} \beta(\beta-1) r^{\beta-2} + \frac{\hbar^2 l(l+1)}{2\mu} r^{\beta-2} + A r^{\alpha+\beta} = E r^\beta; \quad (4.57)$$

¹⁶L'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.

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

$$\beta(\beta - 1) = l(l + 1) \iff \beta = l + 1 . \quad (4.58)$$

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 $r = 0$ ($R(r) \propto r^0$) :

$$V(r) \propto r^\alpha \quad (r \sim 0, \alpha \geq -1) \implies R(r) \propto r^l \iff u(r) \propto r^{l+1} \quad (r \sim 0) . \quad (4.59)$$

Discutons maintenant brièvement le comportement de $R(r)$ à l'infini, en se cantonnant aux états liés ($E < 0$ compte tenu de (4.50)). 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 (4.38) avec les termes dominants. En revanche, ceci devient possible si u contient un facteur exponentiel ; il en résulte :

$$R(r) \propto r^\lambda e^{-kr} \quad (r \rightarrow +\infty, k = \sqrt{-2\mu E/\hbar^2}) \quad (4.60)$$

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 :

$$-\frac{\hbar^2}{2\mu} [u'(r_0 + 0) - u'(r_0 - 0)] + \int_{r_0-0}^{r_0+0} dr \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} + V(r) \right] u(r) = E \int_{r_0-0}^{r_0+0} dr u(r) . \quad (4.61)$$

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 $\lambda < 1/2$ et alors l'intégrale du second membre donne zéro. Dans l'intégrale au premier membre, le terme en $l(l+1)/r^2$ 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 $\lambda < 1/2$: son intégrale est nulle. Il reste finalement :

$$u'(r_0 + 0) - u'(r_0 - 0) = 0 ; \quad (4.62)$$

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.

4.1.3 La particule libre

Une particule libre possède une énergie potentielle constante, que l'on peut toujours prendre égale à zéro. Ce problème – dépourvu d'états liés pour des raisons physiques évidentes – peut évidemment se résoudre en coordonnées cartésiennes ; alors les états propres de H sont des ondes planes :

$$\psi_{\vec{k}}(\vec{r}) = C(\vec{k}) e^{i\vec{k} \cdot \vec{r}} \quad (4.63)$$

où \vec{k} est un vecteur de composantes réelles, faute de quoi $\psi_{\vec{k}}$ divergerait exponentiellement à l'infini. L'énergie est reliée à \vec{k} par la relation habituelle :

$$E(\vec{k}) = \frac{\hbar^2}{2\mu} k^2 \quad (4.64)$$

qui constitue la relation de dispersion. Chaque état est infiniment dégénéré, puisque l'énergie ne dépend que du module k de \vec{k} . Il n'y a pas de quantification de l'énergie, toutes les valeurs de k étant autorisées. Les états propres sont orthogonaux au sens où :

$$\int_{\mathbb{R}^3} d^3r \psi_{\vec{k}}^*(\vec{r}) \psi_{\vec{k}'}(\vec{r}) = 0 \quad \forall \vec{k}' \neq \vec{k} . \quad (4.65)$$

Souvent, la contante C est choisie égale¹⁷ à $(2\pi)^{-3/2}$, de sorte que :

$$\int_{\mathbb{R}^3} d^3r \psi_{\vec{k}}^*(\vec{r}) \psi_{\vec{k}'}(\vec{r}) = \delta(\vec{k} - \vec{k}') , \quad (4.66)$$

en vertu de la relation :

$$\int_{-\infty}^{+\infty} dt e^{i(\omega - \omega')t} = 2\pi\delta(\omega - \omega') . \quad (4.67)$$

Les ondes planes ne sont pas normalisables. Il est cependant toujours loisible de choisir dans un premier temps une grande boîte cubique de quantification, de côté L , permettant de manipuler uniquement des états discrets, moyennant le choix judicieux de conditions aux limites. En principe, la limite $L \rightarrow +\infty$ prise en fin de calcul restitue les mêmes résultats que le calcul direct¹⁸. La représentation en ondes planes correspond à un ensemble d'observables qui commutent (EOC) constitué des trois composantes de l'impulsion et de l'énergie.

Bien évidemment, rien n'interdit de résoudre le même problème en coordonnées sphériques : le potentiel nul est visiblement un cas particulier de potentiel à symétrie sphérique ; ce choix étant fait, on obtient une autre représentation de l'état quantique, correspondant *ipso facto* à un EOC constitué de $(\vec{L}^2, L_z$ et $H)$. Comme \vec{p} et \vec{L} ne commutent pas, ces deux représentations sont incompatibles, au sens des observables incompatibles. Dans le droit fil de ce qui a été dit au paragraphe précédent, on pose la factorisation (4.35) ; la seule fonction inconnue, $R(r)$, satisfait :

$$-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right) R(r) + \frac{\hbar^2 l(l+1)}{2\mu r^2} R(r) = E R(r) . \quad (4.68)$$

En posant :

$$k = \sqrt{2\mu E/\hbar^2} , \quad \rho = kr , \quad R(r) = v(\rho) , \quad (4.69)$$

on trouve l'équation pour la fonction v :

$$\frac{d^2 v}{d\rho^2} + \frac{2}{\rho} \frac{dv}{d\rho} + \left[1 - \frac{l(l+1)}{\rho^2} \right] v(\rho) = 0 . \quad (4.70)$$

Cette équation est bien connue ; ses solutions non-divergentes à l'origine (elles y sont en réalité nulles, sauf celle associée à $l=0$) sont par définition les fonctions de Bessel sphériques $j_l(r)$. Ces dernières peuvent être déduites des fonctions de Bessel ordinaires, J_ν , par la relation ([23], p. 437) :

$$j_l(\rho) = \sqrt{\frac{\pi}{2\rho}} J_{l+\frac{1}{2}}(\rho) . \quad (4.71)$$

Une autre définition de j_l , utile pour la suite, est :

$$j_l(\rho) = \frac{\rho^l}{2^{l+1} l!} \int_{-1}^{+1} ds e^{i\rho s} (1-s^2)^l . \quad (4.72)$$

En effectuant l intégrations par parties, on trouve :

$$j_l(\rho) = i^{-l} \frac{(-1)^l}{2^{l+1} l!} \int_{-1}^{+1} ds e^{i\rho s} \frac{d^l}{ds^l} (1-s^2)^l . \quad (4.73)$$

Si l'on se souvient alors de la définition des polynômes de Legendre obtenus lors de l'étude des harmoniques sphériques, on réalise le lien étroit entre le polynôme de degré l , P_l , et la fonction de Bessel sphérique j_l , lien exprimé par la relation :

$$j_l(\rho) = \frac{1}{2^l l!} \int_{-1}^{+1} ds e^{i\rho s} P_l(s) . \quad (4.74)$$

¹⁷auquel cas ψ est un nombre (sans dimension physique).

¹⁸Dans le cas contraire, le problème traité aurait, pour des raisons à élucider, une sensibilité à des conditions aux limites non pertinentes physiquement.

Les expressions des premières j_l sont :

$$j_0(\rho) = \frac{\sin \rho}{\rho} , \quad j_1(\rho) = \frac{\sin \rho}{\rho^2} - \frac{\cos \rho}{\rho} , \quad j_2(\rho) = \left(\frac{3}{\rho^3} - \frac{1}{\rho} \right) \sin \rho - \frac{3}{\rho^2} \cos \rho . \quad (4.75)$$

La solution de l'équation radiale pour la particule libre est donc de la forme :

$$R(r) = C_{kl} j_l(kr) , \quad (4.76)$$

où la constante C_{kl} sera déterminée par une convention de normalisation, par exemple¹⁹ :

$$C_{kl} C_{k'l'} \int_0^{+\infty} r^2 dr j_l(kr) j_{l'}(k'r) = \delta(k - k') . \quad (4.77)$$

Au contraire des ondes planes, le module de ces fonctions n'est pas constant et décroît même vers zéro en oscillant. Toutefois, les fonctions ne sont pas non plus normalisables car :

$$\int_0^{+\infty} r^2 dr |j_l(kr)|^2 = +\infty . \quad (4.78)$$

Par ailleurs, on vérifie facilement que $R(r)$ est bien en r^l quand $r \rightarrow 0$.

À ce stade, on dispose finalement de deux jeux de fonctions propres pour la particule libre d'énergie donnée E : les ondes planes ayant cette énergie, en nombre infini non-dénombrable (toutes les orientations possibles pour \vec{k}), et les fonctions $j_l(kr)Y_{lm}(\theta, \phi)$ qui forment au contraire un ensemble infini mais dénombrable. Il s'agit toutefois dans les deux cas d'un système complet de fonctions et on peut donc développer n'importe quelle fonction d'un jeu en série des fonctions de l'autre jeu. Évidemment, la première fonction étant choisie et paramétrée par k – qui représente l'énergie E –, elle se développe sur tous les états de l'autre jeu ayant la même énergie E ; c'est pourquoi on retrouve le même module k du vecteur \vec{k} dans les deux membres de la relation (4.79) ci-dessous. En tout cas, on doit pouvoir trouver des coefficients c_{lm} tels que :

$$e^{i\vec{k}\cdot\vec{r}} = \sum_{l=0}^{+\infty} \sum_{m=-l}^{+l} c_{lm} j_l(kr) Y_{lm}(\theta, \phi) . \quad (4.79)$$

Prenons en particulier le cas où le vecteur \vec{k} est dirigé le long de Oz ; alors le premier membre ne dépend que de θ , plus de ϕ , et seules les harmoniques sphériques $Y_{lm=0}$ apparaissent dans le développement ; on se souvient que l'harmonique sphérique $Y_{lm=0}$ est proportionnelle au polynôme de Legendre P_l , de sorte que l'on peut écrire :

$$e^{ikr \cos \theta} = \sum_{l=0}^{+\infty} c_l j_l(kr) P_l(\cos \theta) . \quad (4.80)$$

Compte tenu de l'orthogonalité des polynômes P_l :

$$\int_{-1}^{+1} dX P_l(X) P_{l'}(X) = \frac{2}{2l+1} \delta_{ll'} , \quad (4.81)$$

il est possible de trouver facilement les coefficients c_l . En effet, par l'orthogonalité, (4.81) donne :

$$c_l j_l(kr) = \frac{2l+1}{2} \int_{-1}^{+1} dX e^{ikrX} P_l(X) . \quad (4.82)$$

Prenant en compte (4.74), on déduit $c_l = (2l+1)!$ et finalement l'identité suivante, importante et particulièrement utile pour les problèmes de diffusion :

$$e^{ikz} = \sum_{l=0}^{+\infty} (2l+1)! j_l(kr) P_l(\cos \theta) \quad (z = r \cos \theta) . \quad (4.83)$$

¹⁹Il n'y a pas lieu d'attendre l'orthogonalité de j_l et de $j_{l'}$ ($l \neq l'$) ; en effet, ces deux fonctions sont propres de deux Hamiltoniens différents.

Évidemment, il est possible d'effectuer les développements inverses et de trouver les coefficients $d_{lm}(k)$ tels que :

$$j_l(kr)Y_{lm}(\theta, \phi) = \int_0^\pi \sin \Theta d\Theta \int_0^{2\pi} d\Phi d_{lm}(k) e^{i\vec{k}\cdot\vec{r}} . \quad (4.84)$$

où Θ et Φ sont les angles fixant la direction des vecteurs \vec{k} ayant tous le même module k (paramètre qui est, à nouveau, présent aux deux membres et représente l'énergie commune des états dégénérés que l'on développe les uns sur les autres). La base propre sphérique $j_l Y_{lm}$ permet également de représenter n'importe quel paquet d'ondes, en particulier un paquet gaussien.

4.1.4 Puits "carré" sphérique

Il s'agit ici de résoudre²⁰ l'équation radiale pour la fonction $R(r)$ avec l'énergie potentielle :

$$V(r) = \begin{cases} -V_0 & \text{si } r < a \\ 0 & \text{si } r > a \end{cases} \quad (V_0 > 0) . \quad (4.85)$$

V a une variation "carrée", mais, physiquement, le puits a évidemment une symétrie sphérique. L'équation aux valeurs propres radiale est :

$$\left[-\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d}{dr} \right) + \frac{\hbar^2 l(l+1)}{2\mu r^2} \right] R(r) = \begin{cases} (E + V_0) R(r) & \text{si } r < a \\ E R(r) & \text{si } r > a \end{cases} . \quad (4.86)$$

Bornons-nous aux états liés, pour lesquels $-V_0 < E < 0$. Dans la région $r < a$, on trouve la même équation que dans la sous-section 4.1.3, et il suffit de remplacer dans les résultats alors obtenus E par $E + V_0$, qui est positif. Dans cette région, la fonction R est donc :

$$r < a : \quad R(r) = C j_l(Kr) , \quad K^2 = \frac{2\mu(E + V_0)}{\hbar^2} > 0 . \quad (4.87)$$

Dans l'autre région, $r > a$, on retrouve à nouveau la même équation que précédemment, mais le facteur au second membre, l'énergie E , est maintenant négatif, ce qui signifie que, pour r grand, la fonction $u(r) = r R(r)$ satisfait :

$$-\frac{\hbar^2}{2\mu} u''(r) \simeq E u(r) \equiv -|E| u(r) . \quad (4.88)$$

u est donc essentiellement²¹ une combinaison d'exponentielles réelles ; comme toujours, seules les solutions *décroissantes* sont acceptables ; elles s'expriment à l'aide d'une fonction spéciale connue, appelée fonction de Hankel de première espèce, $h_l^{(1)}$ ([16], p. 302), qui a précisément ce comportement à l'infini :

$$r > a : \quad R_l(r) = C' h_l^{(1)}(ikr) , \quad k^2 = -\frac{2\mu E}{\hbar^2} > 0 . \quad (4.89)$$

Au total, on obtient des fonctions oscillantes (les j_l) à l'intérieur du puits, et des fonctions décroissant à l'infini à l'extérieur de celui-ci.

À l'instar du problème unidimensionnel, et comme montré plus haut (voir 4.62), il convient d'écrire les équations de continuité de R et de sa dérivée en $r = a$ (la condition de normalisabilité, évacuant d'emblée les solutions divergentes à l'infini, est déjà prise en compte), les relations qui en découlent conduisent finalement à une équation dont les zéros fournissent les seules valeurs possibles de k , donc de l'énergie. Une fois encore, les conditions physiques requises pour la fonction d'onde produisent la quantification spontanée de l'énergie des états liés.

Les conditions de raccordement produisent une équation dépendant de la valeur du moment cinétique orbital²², donnée par l'entier l . Pour les états s ($l = 0$), ces conditions s'expriment très simplement, puisqu'alors²³ :

$$R_{l=0}(r) = \begin{cases} C \frac{\sin Kr}{r} & \text{si } r < a \\ C' \frac{1}{r} e^{-kr} & \text{si } r > a \end{cases} . \quad (4.90)$$

²⁰[16], p. 304.

²¹*i. e.* à un préfacteur algébrique r^λ près.

²²voir [24], problème 63, p. 153.

²³On note que $R_0(r=0) = C^{\text{ste}} \neq 0$, conformément à ce qui a été dit plus haut pour les états s .

L'écriture explicite des conditions de continuité de R et de R' pour $l = 0$ conduit à :

$$ka = -\frac{Ka}{\tan Ka} \iff \sqrt{k_0^2 - K^2} a = -\frac{Ka}{\tan Ka} \quad (k_0^2 = 2\mu V_0/\hbar^2), \quad (4.91)$$

avec $K \leq k_0$ puisque l'énergie E est négative mais bornée inférieurement par $-V_0$. Cette équation est réminiscente de ce qui a été obtenu à une dimension d'espace ; ce n'est pas surprenant : pour $l = 0$, l'équation en u – fonction qui doit être continue et à dérivée continue – est identique à l'équation aux valeurs propres pour le même potentiel à $d = 1$ (complété avec la barrière en $r = 0$), puisque le terme centrifuge est nul dans le cas considéré ici ($l = 0$). Pour les états s , on obtient donc toujours, à trois dimensions, les mêmes résultats que pour une particule soumise au même potentiel mais confinée sur le demi-axe réel positif par une barrière infranchissable située à l'origine : u est donc $\propto \sin Kr$, et $R_0 \propto \frac{\sin Kr}{r}$.

Le puits carré, qui a toujours au moins un état lié sur tout \mathbb{R} , perd cette propriété si une telle barrière est présente : celle-ci a un effet répulsif²⁴, donc déstabilisant et il n'y a d'état lié que si le puits est assez profond et pas trop étroit. Il en va de même ici : il est facile de voir que l'équation ci-dessus n'a de solution que si $k_0 a > \pi/2$, c'est-à-dire :

$$V_0 a^2 > \frac{\pi^2 \hbar^2}{4 \cdot 2\mu} \iff V_0 > \frac{\pi^2 \hbar^2}{8\mu a^2} \iff V_0 a^2 > \frac{\pi^2 \hbar^2}{8\mu}. \quad (4.92)$$

Sans surprise, le nombre d'états liés N_b augmente quand V_0 augmente (avec a constant), très précisément d'une unité à chaque fois que k_0 franchit une valeur du genre $\frac{\pi}{2} + n\pi$ ($n \in \mathbb{N}$). Inversement, à V_0 constant, N_b décroît si la portée a diminue ; en dessous de la valeur-seuil défini par (4.92), il n'y a plus d'état lié.

Si la profondeur du puits V_0 est très grande devant toute autre énergie, et si son rayon a est beaucoup plus petit que toute autre longueur, il est tentant – comme on le fait sans difficulté à une dimension – de modéliser le puits par une fonction de Dirac. Une telle tentative se révèle problématique [25] car la réponse dépend du processus de limite. Si on pose :

$$V(r) = A \delta(\vec{r}) \quad , \quad (4.93)$$

la constante A a pour dimension [énergie] \times [longueur] ^{d} , où d est la dimension de l'espace physique ; compte tenu des seules échelles disponibles, il est licite de poser $A = V_0 a^d$, de sorte que²⁵ :

$$V(r) = V_0 a^d \delta(\vec{r}) \quad . \quad (4.95)$$

Formellement, la limite du puits δ s'obtient en faisant $a \rightarrow 0$ et $V_0 \rightarrow +\infty$; pour que la modélisation garde un sens, ces deux limites ne doivent pas indépendantes : il faut $V_0 a^d = C^{\text{ste}}$. En conséquence, la combinaison $V_0 a^2$ apparaissant dans (4.92) est proportionnelle à a^{2-d} ; l'analyse détaillée [26] montre que la régularisation est nécessaire dès que $d \geq 2$.

Un autre exemple de champ central intéressant est celui de l'oscillateur harmonique à trois dimensions – qui ne possède que des états liés puisque $V(r \rightarrow +\infty) \rightarrow +\infty$. Ce problème se sépare évidemment en trois oscillateurs indépendants le long de chaque direction cartésienne. Dans le cas où l'oscillateur est isotrope (même constante de raideur suivant les trois directions), la symétrie alors sphérique produit la séparation des variables sphériques²⁶, fournissant une nouvelle illustration de deux choix différents d'ensembles d'observables qui commutent : en cartésiennes, les éléments de l'ECOC sont les trois énergies de vibration le long des trois axes, en sphériques, l'ECOC est constitué par l'énergie totale, \vec{L}^2 et L_z . À nouveau, pour une énergie donnée, les deux types d'états propres se déduisent les uns des autres par des combinaisons linéaires exprimant la décomposition des vecteurs propres d'un ECOC sur les états propres de l'autre ECOC.

²⁴Un puits $V(x) = g\delta(x - x_0)$, $g < 0$, offre toujours un et un seul état lié à toute particule de masse m_p ; si on ajoute une barrière réfléchissante (par exemple en $x = 0$), l'état lié ne subsiste que si $|g| > \hbar^2/(2m_p x_0)$: si la barrière est trop proche du puits et/ou la particule est trop légère, il n'y a plus de piégeage possible.

²⁵L'expression (4.95) est la première qui vient à l'esprit. En fait, on pourrait tout autant poser :

$$V(r) = V_0 a^d f(\lambda) \delta(\vec{r}) \quad (4.94)$$

où f est une fonction sans dimension d'un paramètre adimensionné contenant a , par exemple $\lambda = Ka$. L'analyse précise du puits *fini* en dimension $d = 2$ permet de trouver la fonction f et de montrer que, dans ce cas, la modélisation régularisée introduit un puits δ "ramolli" (pour plus de détails, voir [26]).

²⁶Voir [4], complément B_{VII}.

Remarque

Les quelques exemples abordés montrent bien que 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 dimension²⁷ 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”²⁸, ce qui n'est pas possible dans \mathbb{R} pour des raisons purement topologiques.

4.2 Atome d'hydrogène : états propres liés

4.2.1 Résolution de l'équation radiale

Comme toujours, les conditions requises pour 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 qui engendre la quantification de la composante L_z du moment cinétique. 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 $u(r)$ est²⁹ :

$$-\frac{\hbar^2}{2\mu} \frac{d^2u}{dr^2} + \left[\frac{\hbar^2 l(l+1)}{2\mu r^2} - \frac{e'^2}{r} \right] u(r) = E u(r) . \quad (4.96)$$

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 :

$$-\frac{\hbar^2}{2\mu} \frac{d^2u}{dr^2} \simeq E u(r) . \quad (4.97)$$

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 4.1.2.

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 dimension³⁰. 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 pas acceptables. Quand on reporte dans (4.96) un comportement du type :

$$u(r) \simeq r^\lambda e^{-kr} \quad (r \rightarrow +\infty) , \quad (4.98)$$

²⁷sur *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.

²⁸Dans un état de l'atome d'hydrogène où L_z n'est pas nul, il y a une “boucle de courant” perpendiculaire à Oz .

²⁹Pour un ion hydrogénoïde de charge nucléaire Z , il suffit de remplacer partout e'^2 par Ze'^2 .

³⁰voir cours de Licence.

et ne retenant que les termes dominants³¹, on obtient :

$$-\frac{\hbar^2 k^2}{2\mu} r^\lambda e^{-kr} \simeq E r^\lambda e^{-kr} . \quad (4.99)$$

E étant négatif, la forme simplifiée (4.98) est donc solution si $k = \pm\sqrt{-2\mu E/\hbar^2}$; les seules solutions physiques se comportent donc comme suit à l'infini :

$$u(r) \simeq r^\lambda e^{-kr} , \quad k = \sqrt{-2\mu E/\hbar^2} > 0 \quad (r \rightarrow +\infty) , \quad (4.100)$$

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

Par ailleurs, on sait d'après (4.59) que le comportement de u à l'origine est :

$$u(r) \simeq r^{l+1} , \quad (r \simeq 0) . \quad (4.101)$$

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 maintenant déterminés en introduisant une nouvelle fonction $w(\rho)$:

$$u(r) = \rho^{l+1} e^{-\rho} w(\rho) , \quad \rho = kr . \quad (4.102)$$

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 (4.59), a été mis en exergue. La substitution dans l'équation complète (4.96) donne :

$$\rho \frac{d^2 w}{d\rho^2} + 2(l+1-\rho) \frac{dw}{d\rho} + (\lambda - 2l - 2)w = 0 \quad (4.103)$$

avec :

$$\lambda = \sqrt{\frac{2\mu e'^4}{-\hbar^2 E}} . \quad (4.104)$$

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

$$w(\rho) = \sum_{p=0}^{+\infty} c_p \rho^p . \quad (4.105)$$

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 (4.103) fournit la récurrence suivante³² :

$$(p+1)[p+2(l+1)]c_{p+1} + (\lambda - 2l - 2 - 2p)c_p = 0 , \quad p \in \mathbb{N} , \quad c_0 \neq 0 . \quad (4.106)$$

Ceci permet bien d'exprimer tous les coefficients en fonction du premier c_0 . La relation (4.106) 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 (4.105) 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 (4.102) finit par l'emporter et u , tendant vers zéro à l'infini, possède le bon comportement.

³¹Le 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 (4.96) sont négligeables.

³²On 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 (4.102) qui résulte des conditions physiques imposées aux solutions à retenir exclusivement.

La divergence exponentielle de la série (4.105) se met en évidence comme suit. À l et E fixés, pour p grand, la récurrence (4.106) s'écrit à peu près :

$$p^2 c_{p+1} - 2p c_p \simeq 0 \iff \frac{c_{p+1}}{c_p} \simeq \frac{2}{p} . \quad (4.107)$$

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

$$e^{2\rho} = \sum_{p=0}^{+\infty} \frac{1}{p!} (2\rho)^p \equiv \sum_{p=0}^{+\infty} \gamma_p \rho^p, \quad \frac{\gamma_{p+1}}{\gamma_p} = \frac{2}{p} . \quad (4.108)$$

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 (4.103) donne, après simplification par $\rho^{q-1} e^{\alpha\rho}$:

$$q(q-1) + 2q\alpha\rho + \alpha^2\rho^2 + 2(l+1-\rho)(q+\alpha\rho) + (\lambda-2l-2)\rho \sim 0 . \quad (4.109)$$

Seuls les termes les plus divergents doivent être considérés³³, soit le terme en ρ^2 ; son coefficient, égal à $(\alpha^2 - 2\alpha)$, doit être nul, ce qui fournit bien la seule valeur acceptable³⁴ $\alpha = 2$, confirmant l'analyse fondée sur la comparaison des séries.

Au total, suivant (4.102), on trouve que $u(r)$ se comporte comme :

$$u(r) \simeq \rho^{l+1} e^{-\rho} e^{+2\rho} = \rho^{l+1} e^{+\rho} \quad (4.110)$$

Si on garde la série (4.105) telle que, la fonction $u(r)$ diverge 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 :

$$u(r) \simeq \rho^{l+1+p_0} e^{-\rho} . \quad (4.111)$$

p_0 est visiblement un entier positif ou nul ; quelle que soit la valeur finie de p_0 , l'exponentielle dans (4.111) finit toujours par l'emporter. Les seules bonnes solutions sont donc celles pour lesquelles :

$$\lambda - 2l - 2 = 2p_0 , \quad p_0 \in \mathbb{N} . \quad (4.112)$$

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

$$\lambda = 2n , \quad n \in \mathbb{N}^* . \quad (4.113)$$

Compte tenu de la définition de λ , (4.104), ceci donne :

$$E \in \{E_n\} : \quad E_n = -\frac{\mu e'^4}{2n^2 \hbar^2} \quad (n \in \mathbb{N}^*) . \quad (4.114)$$

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 :

$$\left[\frac{\vec{p}^2}{2\mu} - \frac{e'^2}{r} \right] \psi_{nlm}(r, \theta, \phi) = E_n \psi_{nlm}(r, \theta, \phi) , \quad \psi_{nlm}(r, \theta, \phi) = R_{nl}(r) Y_{lm}(\theta, \phi) . \quad (4.115)$$

³³Les autres devraient être comparés avec des termes omis en posant $w \sim \rho^q e^{\alpha\rho}$.

³⁴L'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.

L'expression (4.114) de l'énergie est strictement identique à celle de la théorie de Bohr, dont on sait qu'elle reproduit³⁵ 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 surprenant³⁶.

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 caractéristiques³⁷, qui permettent de déterminer toutes les constantes d'intégration apparaissant dans la résolution de l'équation aux valeurs propres³⁸. 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 inférieurement³⁹ 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 caractéristique⁴⁰ des potentiels qui tendent vers zéro moins vite que $1/r^2$ ($r^2V(r) \rightarrow +\infty$ si $r \rightarrow +\infty$) ; pour un potentiel allant plus rapidement vers zéro que $1/r^2$ ($r^2V(r) \rightarrow 0$ 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, comme on l'a vu plus haut.
2. une partie continue (dense) s'étendant de 0 à $+\infty$, correspondant aux états non-liés (états stationnaires de diffusion).

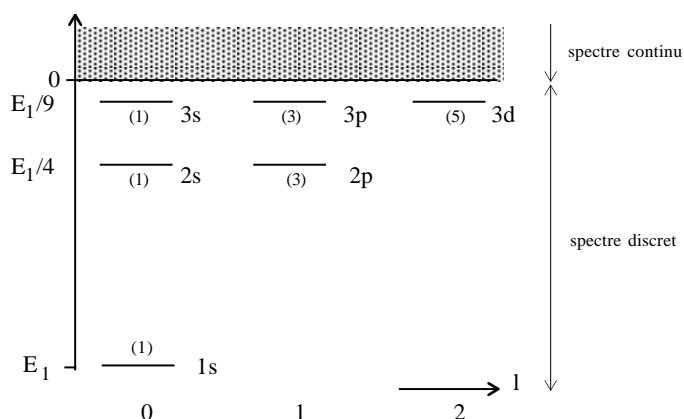


Figure 4.1: Spectre d'énergie de l'atome d'hydrogène et notation spectroscopique.

³⁵On se souvient que Bohr a construit sa théorie pour qu'il en soit ainsi.

³⁶Les 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 (4.114) : 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 $1/r$, qui donne lieu à la conservation dynamique du vecteur de Lenz - Runge (voir sous-section 4.2.2). Tout écart à la loi en $1/r$ – par exemple un terme en $1/r^2$ déstabilisant (parfois utilisé pour représenter un écrantage à longue portée) –, supprime cette propriété et l'énergie devient alors une fonction explicite de n et de l .

³⁷Par exemple le courant qu'elle émet.

³⁸Ainsi, 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, $j_s : (\hbar k/m)|A|^2 = j_s$.

³⁹Le 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 accélérée, rayonne et se précipite sur le noyau).

⁴⁰[16], p. 354

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) est⁴¹ :

$$a_0 = \frac{\hbar^2}{me'^2} \simeq 0.53 \text{ \AA} \quad (4.116)$$

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

$$E_n = -\frac{\mu}{m_e} \frac{e'^2}{2n^2 a_0} . \quad (4.117)$$

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

$$E_n = -\frac{\alpha^2}{2n^2} \mu c^2 \quad \alpha = \frac{e'^2}{\hbar c} = \frac{1}{137.007\dots} . \quad (4.118)$$

Comme l'entier p_0 , où la série (4.105) est tronquée, est positif ou nul, les seules valeurs possibles de l pour n fixé satisfont⁴² (voir (4.112) et (4.113)) :

$$2n - 2l - 2 \geq 0 \quad \Longleftrightarrow \quad l \leq n - 1 . \quad (4.119)$$

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 :

$$l = 0, 1, 2, \dots, n - 1 . \quad (4.120)$$

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 totale⁴³ est :

$$g_n = \sum_{l=0}^{n-1} (2l + 1) = 2 \frac{(n-1)n}{2} + n = n^2 . \quad (4.121)$$

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"⁴⁴. 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 :

$$l = 0 \leftrightarrow s \quad l = 1 \leftrightarrow p \quad l = 2 \leftrightarrow d \quad l = 3 \leftrightarrow f \quad l = 4 \leftrightarrow g \quad \dots \quad (4.122)$$

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.

⁴¹ m_e désigne la masse de l'électron, $m_e = 9 \times 10^{-31}$ kg.

⁴²Les états $l = 0$ correspondent aux orbites circulaires de la théorie de Bohr.

⁴³Compte non-tenu du spin ! La prise en compte du spin 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, ...).

⁴⁴La qualification "accidentelle" est source de contresens. Il existe en fait une symétrie subtile, comme expliqué plus bas (sous-section 4.2.2) ; cette symétrie, dans l'espace des impulsions, remarquée par Fock en 1935, a été étudiée notamment par McIntosh [27] et par Bander & Itzykson [28]. 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.

À propos des polynômes de Laguerre

Les fonctions radiales ainsi obtenues sont construites avec les *polynômes* satisfaisant (4.103) où λ est remplacé par sa valeur $\lambda = 2n$, compte tenu de (4.113) :

$$\rho \frac{d^2 w}{d\rho^2} + 2(l+1-\rho) \frac{dw}{d\rho} + (2n-2l-2)w = 0 . \quad (4.123)$$

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

$$z \frac{d^2 L}{dz^2} + (\beta - z) \frac{dL}{dz} - \alpha L = 0 \quad (4.124)$$

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

$$w(\rho) = L(z = 2\rho) \quad (4.125)$$

pour que l'équation (4.123) devienne :

$$z \frac{d^2 L}{dz^2} + [2(l+1) - z] \frac{dL}{dz} + (n-l-1)L = 0 , \quad (4.126)$$

qui, par comparaison avec (4.123), 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 définition⁴⁵ l'équation différentielle :

$$z \frac{d^2 L_p}{dz^2} + (1-z) \frac{dL_p}{dz} + p L_p = 0 . \quad (4.127)$$

On achève de définir les L_p en posant conventionnellement :

$$L_p(z=0) = p! . \quad (4.128)$$

Les L_p sont appelés polynômes de Laguerre. $L_p(z)$ est un polynôme de degré p , dont 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 :

$$z \frac{d^2 L^{(q)}}{dz^2} + (q+1-z) \frac{dL^{(q)}}{dz} + (p-q)L^{(q)} = 0 \quad (4.129)$$

qui est bien l'équation différentielle (4.126) obtenue plus haut avec l'identification $q+1 = 2l+1$, $p-q = n-l-1$. Ceci définit les polynômes associés de Laguerre, précisément définis comme :

$$L_p^q(z) = (-1)^q \frac{d^q}{dz^q} L_{p+q}^0(z) . \quad (4.130)$$

$L_p^q(z)$ 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 :

$$L_p(z) = \frac{p!}{2i\pi} \oint ds \frac{e^{-sz}}{s^{p+1}} (1+s)^p \quad (4.131)$$

⁴⁵Cette définition fixe les L_p à un facteur près, évidemment.

où l'intégrale est prise le long d'un petit contour entourant une fois l'origine du plan de s dans le sens positif⁴⁶. De l'expression intégrale (4.131), on déduit facilement :

$$L_p(z) = (-1)^p \frac{p!}{2i\pi} e^z \frac{d^p}{dz^p} \oint ds \frac{e^{-(s+1)z}}{s^{p+1}} = (-1)^p \frac{p!}{2i\pi} e^z \frac{d^p}{dz^p} \oint ds \frac{e^{-sz}}{s^{p+1}} e^{-z} \quad (4.133)$$

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

$$L_p(z) = e^z \frac{d^p}{dz^p} [z^p e^{-z}] , \quad (4.134)$$

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 :

$$L_p(z) = \sum_{r=0}^p \frac{(-1)^r p!^2}{(p-r)! r!^2} z^r . \quad (4.135)$$

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

$$F(t, z) = \sum_{p=0}^{+\infty} L_p(z) \frac{t^p}{p!} . \quad (4.136)$$

$F(t, z)$ est donc la fonction dont les coefficients de la série de Taylor autour de $t = 0$ sont⁴⁷ les $L_p(z)$:

$$L_p(z) = \left(\frac{\partial^p}{\partial t^p} F(t, z) \right)_{t=0} . \quad (4.138)$$

$F(t, z)$ s'obtient en utilisant l'expression (4.135) d'où :

$$F(t, z) = \sum_{r=0}^{+\infty} \frac{(-zt)^r}{r!} \sum_{p=0}^{+\infty} C_{p+r}^p t^p . \quad (4.139)$$

La sommation interne s'écrit :

$$\sum_{p=0}^{+\infty} C_{p+r}^p t^p = \sum_{p=0}^{+\infty} \frac{t^p (p+r)!}{p! r!} = \sum_{p=0}^{+\infty} (p+r)(p+r-1)\dots(r+2)(r+1) \frac{t^p}{p!} . \quad (4.140)$$

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

$$\frac{d^p}{dt^p} \frac{1}{(1-t)^{r+1}} = \frac{(r+1)(r+2)\dots(r+p)}{(1-t)^{r+1+p}} \implies (r+1)(r+2)\dots(r+p) = \left(\frac{d^p}{dt^p} \frac{1}{(1-t)^{r+1}} \right)_{t=0} , \quad (4.141)$$

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 (4.139) :

$$F(t, z) = \sum_{r=0}^{+\infty} \frac{(-zt)^r}{r!} \frac{1}{(1-t)^{r+1}} = \frac{1}{1-t} e^{-\frac{tz}{1-t}} . \quad (4.142)$$

⁴⁶L'expression (4.131) "sort" naturellement quand on résout l'équation (4.127) 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 :

$$L_p(z) = \oint_C ds e^{-sz} f(s) ; \quad (4.132)$$

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 [29], appendice a).

⁴⁷De ce fait, par la formule de Cauchy, on a :

$$L_p(z) = \frac{p!}{2i\pi} \oint \frac{F(\xi, z)}{\xi^{p+1}} d\xi , \quad (4.137)$$

où le contour est une petite boucle autour de l'origine. En injectant l'expression (4.142) de $F(t, z)$ et en posant $\frac{\xi}{1-\xi} = s$, on retombe bien sur l'expression intégrale (4.131) de $L_p(z)$.

La série de Taylor (4.136) 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 (4.135), on vérifie bien la valeur conventionnelle $L_p(0) = p!$:

$$F(t, z = 0) = \sum_{p=0}^{+\infty} L_p(0) \frac{t^p}{p!} = \frac{1}{1-t} = \sum_{p=0}^{+\infty} t^p \quad (|t| < 1) . \quad (4.143)$$

Le polynôme associé, L_p^q , est la dérivée $q^{\text{ème}}$ d'un polynôme de degré $p+q$, c'est donc encore un polynôme de degré p . En raison de la définition (4.130) des L_p^q , leur fonction génératrice est $F_q(t, z)$:

$$F_q(t, z) \equiv \sum_{p=0}^{+\infty} L_p^q(z) \frac{t^p}{(p+q)!} = \sum_{p=0}^{+\infty} (-1)^q \frac{d^q}{dz^q} L_{p+q}(z) \frac{t^p}{(p+q)!} = (-t)^{-q} \frac{d^q}{dz^q} \sum_{p=q}^{+\infty} L_p(z) \frac{t^p}{p!} . \quad (4.144)$$

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 :

$$F_q(t, z) = (-t)^{-q} \frac{d^q}{dz^q} \sum_{p=0}^{+\infty} L_p(z) \frac{t^p}{p!} = (-t)^{-q} \frac{\partial^q}{\partial z^q} F(t, z) = \frac{1}{(1-t)^{q+1}} e^{-\frac{tz}{1-t}} . \quad (4.145)$$

Faisant $z = 0$, on en déduit :

$$F_q(t, z = 0) = \sum_{p=0}^{+\infty} L_p^q(0) \frac{t^p}{(p+q)!} = \frac{1}{(1-t)^{q+1}} ; \quad (4.146)$$

reprenant alors le développement de Taylor de $(1-t)^{-(q+1)}$ (voir (4.141)), on voit que :

$$L_p^q(z = 0) = \frac{[p+q]!^2}{p!q!} . \quad (4.147)$$

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

$$R_{nl}(r) = C_{nl} r^l L_{n-l-1}^{2l+1}(2k_n r) e^{-k_n r} \quad (4.148)$$

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

$$\int_0^{+\infty} r^2 dr R_{nl}^2(r) = 1 , \quad (4.149)$$

la partie angulaire étant elle aussi normalisée à part. Les k_n sont les nombres d'onde déduits de la définition de k (voir (4.100)) et de l'expression de E_n donnée en (4.114) :

$$k_n = \frac{\mu}{m_e} \frac{1}{na_0} . \quad (4.150)$$

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 :

$$R_{nl}(r) = \left[(2k_n)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3} \right]^{1/2} (2k_n r)^l L_{n-l-1}^{2l+1}(2k_n r) e^{-k_n r} . \quad (4.151)$$

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 :

$$R_{nl=n-1}(r) = C_{nn-1} r^{n-1} e^{-k_n r} . \quad (4.152)$$

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 :

$$r_{\max} = \frac{n}{k_n} = \frac{m_e}{\mu} n^2 a_0 \simeq n^2 a_0 . \quad (4.153)$$

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 :

$$\psi_{nlm}(r, \theta, \phi) = R_{nl}(r) Y_{lm}(\theta, \phi) , \quad (4.154)$$

les Y_{lm} portent toute la dépendance angulaire ; les états $l = 0$ sont à symétrie sphérique, comme on le sait, les états $l = 1$ ont au contraire une symétrie axiale, en module (ils sont invariants 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 :

$$\vec{\nabla} = \left(\frac{\partial}{\partial r}, \frac{1}{r} \frac{\partial}{\partial \theta}, \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \right) . \quad (4.155)$$

Comme dans l'expression du courant on fait la différence entre un nombre complexe et son conjugué, seuls comptent les termes imaginaires purs – tous les autres se compensent deux à deux ; seule la dérivée 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 :

$$j_\phi = \frac{m\hbar}{\mu r \sin \theta} |\psi_{nlm}|^2 . \quad (4.156)$$

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 :

$$d\vec{K} = \mu \vec{r} \times \vec{j} d^3 r \quad (4.157)$$

et :

$$K_z = \mu \int_{\mathbb{R}^3} d^3 r j_\phi \sin \theta = \mu \int_{\mathbb{R}^3} d^3 r \frac{m\hbar}{\mu} |\psi_{nlm}|^2 = m\hbar . \quad (4.158)$$

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 $m = 0$ ($\langle L_z \rangle = 0$).

4.2.2 Symétrie dynamique du potentiel Coulombien

On a vu dans le chapitre 2 comment symétrie spatiale et dégénérescence étaient étroitement liées. 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 permettraient 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 $1/r$; 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 :

$$\vec{R} = \frac{\vec{p}}{\mu} \times \vec{L} - e'^2 \frac{\vec{r}}{r} . \quad (4.159)$$

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 :

$$\vec{R} = \frac{1}{2} \left(\frac{\vec{p}}{\mu} \times \vec{L} - \vec{L} \times \frac{\vec{p}}{\mu} \right) - e'^2 \frac{\vec{r}}{r} . \quad (4.160)$$

\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 :

$$[L_u, L_v] = i\hbar \epsilon_{uvw} L_w , \quad (4.161)$$

et on trouve :

$$[L_u, R_v] = i\hbar \epsilon_{uvw} R_w \quad [R_u, R_v] = -2 \frac{i\hbar}{\mu} \epsilon_{uvw} L_w H . \quad (4.162)$$

$$[L_u^2, R_v] = i\hbar \epsilon_{uvw} (L_u R_w + R_w L_u) \quad \vec{L} \cdot \vec{R} = \vec{R} \cdot \vec{L} = 0 . \quad (4.163)$$

En outre, bien sûr :

$$[H, R_v] = 0 . \quad (4.164)$$

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 (4.164), on a :

$$\langle n l m | [H, R_z] | n l' m \rangle = 0 . \quad (4.165)$$

Appelons provisoirement E_{nl} l'énergie d'un état propre de H ; (4.165) donne :

$$(E_{nl} - E_{n'l'}) \langle n l m | R_z | n l' m \rangle = 0 . \quad (4.166)$$

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 (4.166) que :

$$E_{nl} = E_{n'l'} , \quad (4.167)$$

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 connues⁴⁸. Dans ce but, commençons par redéfinir un autre vecteur \vec{R}' par :

$$\vec{R}' = \sqrt{\frac{-\mu}{2E}} \vec{R} , \quad (4.168)$$

puis renumérotons les différentes grandeurs dynamiques comme suit :

$$\vec{r} = (r_1, r_2, r_3) , \quad \vec{p} = (p_1, p_2, p_3) , \quad (4.169)$$

$$\vec{L} = (L_x, L_y, L_z) \equiv (L_{23}, L_{31}, L_{12}) . \quad (4.170)$$

Dans ces notations, on a les relations compactes :

$$L_{ij} = r_i p_j - r_j p_i , \quad [r_i, p_j] = i\hbar \delta_{ij} , \quad (4.171)$$

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 :

$$R'_x = L_{14} , \quad R'_y = L_{24} , \quad R'_z = L_{34} . \quad (4.172)$$

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

$$[L_{ij}, L_{kl}] = i\hbar (L_{li} \delta_{jk} + L_{kj} \delta_{il} + L_{ik} \delta_{jl} + L_{jl} \delta_{ik}) , \quad (4.173)$$

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 .

⁴⁸Voir [18], p. 234.

Chapitre 5

Le spin

Le but de ce chapitre est d'introduire le spin d'une particule et les conséquences de son existence sur le magnétisme atomique. Il s'agira également d'une illustration simple mais importante de la théorie du moment cinétique développée antérieurement.

5.1 Insuffisances de la description par une seule fonction d'onde

Jusqu'à présent, toute particule a été considérée comme ponctuelle 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 $\Psi(x, y, z ; t)$.

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-même¹. 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 absurdités². Il reste cependant que le spin d'une particule est son moment cinétique dans le référentiel où elle est au repos ([30], p. II. 4). L'hypothèse du spin de l'électron a été formulée par Uhlenbeck et Goudsmit en 1925 pour rendre compte 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 bosons et les fermions ; c'est cette distinction qui permet, par le jeu de statistiques bien différenciées, 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

¹On parle du spin de la Terre pour évoquer sa révolution diurne.

²Par exemple, on trouve qu'un point situé à la périphérie de l'électron a une vitesse linéaire très supérieure à c !!!

sur l'écran est incompatible avec un moment cinétique orbital : 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'observation 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 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 mouvement³, 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} et il ressort finalement dans le Hamiltonien un terme du genre :

$$V_{\text{magn}} = a(r)\vec{L} \cdot \vec{\mu} . \quad (5.1)$$

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 $\vec{\mathcal{B}}$ ($\rho(\vec{r}') = \delta(\vec{r}')$ puisque le noyau est pris ponctuel et situé à l'origine) :

$$\vec{\mathcal{B}}(\vec{r}) = \frac{\mu_0}{4\pi} \int_{\mathbb{R}^3} d^3r' \frac{\rho(\vec{r}')\vec{v}_N \times (\vec{r} - \vec{r}')}{\|\vec{r} - \vec{r}'\|^3} = \frac{\mu_0}{4\pi} \frac{Z|e|\vec{v}_N \times \vec{r}}{r^3} . \quad (5.2)$$

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

$$\vec{\mathcal{B}}(\vec{r}) = -\frac{\mu_0}{4\pi} \frac{Z|e|\vec{v} \times \vec{r}}{r^3} \equiv \frac{\mu_0}{4\pi} \frac{1}{m} \frac{Z|e|}{r^3} \vec{L} , \quad (5.3)$$

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 $U(r) = Z|e|/(4\pi\epsilon_0 r)$ le potentiel électrostatique du noyau, (5.3) s'écrit⁴ :

$$\vec{\mathcal{B}}(\vec{r}) = -\frac{1}{mc^2} \frac{1}{r} \frac{dU}{dr} \vec{L} . \quad (5.4)$$

L'électron se couple à ce champ magnétique par l'intermédiaire de son moment magnétique supposé ; le terme d'interaction est⁵ :

$$-\vec{\mu} \cdot \vec{\mathcal{B}} = +\frac{1}{mc^2} \frac{1}{r} \frac{dU}{dr} \vec{L} \cdot \vec{\mu} \equiv a(r)\vec{L} \cdot \vec{\mu} \equiv V_{\text{magn}} . \quad (5.6)$$

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 (5.6) et en admettant que $\vec{\mu}$ est lié à un moment cinétique par un facteur gyromagnétique du genre $|e|/m$ (à des facteurs près)⁶, on trouve, en ordre de grandeur :

$$V_{\text{magn}} \sim \frac{1}{mc^2} \frac{1}{a_0^3} \frac{Z|e|}{4\pi\epsilon_0} \hbar \frac{|e|\hbar}{2m} \sim \left(\frac{\hbar}{mca_0} \right) E_{n=1} \quad (5.7)$$

³Tout comme une charge électrique q en mouvement dans un champ magnétique $\vec{\mathcal{B}}$ ressent une force $\vec{F} = q\vec{v} \times \vec{\mathcal{B}}$.

⁴ $\epsilon_0\mu_0c^2 = 1$.

⁵En raison d'un effet subtil appelé précession de Thomas, l'expression correcte de $a(r)$ contient un facteur $\frac{1}{2}$ additionnel :

$$a(r) = +\frac{1}{2mc^2} \frac{1}{r} \frac{dU}{dr} . \quad (5.5)$$

⁶L'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 $E_{n=1} \sim e'^2/a_0$ 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, $a_0 = \hbar^2/(me'^2)$, il vient :

$$V_{\text{magn}} \sim \alpha^2 E_{n=1} \simeq 5 \times 10^{-5} E_{n=1} \sim 7 \times 10^{-4} \text{ eV} \quad \left(\alpha = \frac{1}{137.007\dots} \right). \quad (5.8)$$

α est la constante dite de structure fine : c'est la constante fondamentale de l'Électrodynamique quantique. Les déplacements en énergie varient certainement d'un état à l'autre, mais ils restent petits vis-à-vis des énergies purement électrostatiques (voir aussi (5.18)). 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 fine⁷. 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 hasard⁸.

L'existence du moment magnétique propre de l'électron étant admise sous la pression des faits expérimentaux, un argument⁹ 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 :

$$V_{\text{magn}} = A(r) \vec{L} \cdot \vec{S}. \quad (5.9)$$

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

$$\vec{J} = \vec{L} + \vec{S} \quad (5.10)$$

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 (5.10) est une constante du mouvement quand le Hamiltonien est complété par un terme du genre V_{magn} . On a :

$$\vec{L} \cdot \vec{S} = \frac{1}{2} \left[(\vec{L} + \vec{S})^2 - \vec{L}^2 - \vec{S}^2 \right] = \frac{1}{2} \left[\vec{J}^2 - \vec{L}^2 - \vec{S}^2 \right] \quad (5.11)$$

d'où :

$$[\vec{J}, \vec{L} \cdot \vec{S}] = \frac{1}{2} [\vec{J}, \vec{J}^2 - \vec{L}^2 - \vec{S}^2] = -\frac{1}{2} [\vec{J}, \vec{L}^2 + \vec{S}^2]. \quad (5.12)$$

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

$$[\vec{J}, \vec{L} \cdot \vec{S}] = -\frac{1}{2} [\vec{L}, \vec{L}^2] - \frac{1}{2} [\vec{S}, \vec{S}^2] = 0 \quad (5.13)$$

\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 facile de vérifier que le commutateur $[\vec{V}, \vec{J}]$, avec $\vec{V} = \vec{L}$ ou \vec{S} ,

⁷La 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.

⁸Comme on le verra plus loin, section 5.6, le spin est d'origine relativiste.

⁹voir [17] p. 255.

est bien égal à ce qui a été énoncé au ch. 3 pour élire un opérateur à trois composantes en tant qu'opérateur vectoriel¹⁰.

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 $|e|/T$. Si R est le rayon de l'orbite, cette période vaut $2\pi R/v$ et la surface est πR^2 . Le module du moment magnétique vaut alors¹¹ :

$$\mu = \frac{|e|}{(2\pi R/v)} \pi R^2 = \frac{|e|}{2m} mvR \quad . \quad (5.14)$$

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

$$\vec{\mu} = \frac{e}{2m} \vec{L} \quad . \quad (5.15)$$

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

$$\vec{\mu} = \frac{q}{2M} \vec{L} \equiv \gamma \vec{L} \quad , \quad (5.16)$$

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 :

$$\vec{\mu} = \frac{e\hbar}{2m} \frac{\vec{L}}{\hbar} \quad , \quad (5.17)$$

écriture où apparaît le magnéton¹² de Bohr, μ_B :

$$\mu_B = \frac{e\hbar}{2m} = -0.9274 \times 10^{-23} \text{ Joule/Tesla} = -5.789 \times 10^{-5} \text{ eV/Tesla} \quad . \quad (5.18)$$

Alors (5.17) devient :

$$\vec{\mu} = \mu_B \hbar^{-1} \vec{L} \quad . \quad (5.19)$$

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

$$\mu_N = \frac{|e|\hbar}{2M_p} = 5.051 \times 10^{-27} \text{ Joule/Tesla} = 3.153 \times 10^{-8} \text{ eV/Tesla} \quad . \quad (5.20)$$

Tout ceci repose sur des notions classiques et se réfère par conséquent à un moment magnétique issu d'un moment cinétique orbital. Etant 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 :

$$\vec{\mu}_S = g_e \frac{e}{2m} \vec{S} = g_e \mu_B \hbar^{-1} \vec{S} \equiv \gamma_e \vec{S} \quad . \quad (5.21)$$

¹⁰On 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}).

¹¹Comme 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 (5.14), voir cours de Licence.

¹²On adopte partout la convention par laquelle μ_B est *négatif* ; il semble que ce soit l'usage le plus répandu ([17], [4], *etc.*).

Dans (5.21), g_e est donc un nombre pur ($g_e = 2.002319314\dots$) pris égal à 2 sauf cas exceptionnel¹³ ; 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 donc dirigés en sens contraires. Au total, γ_e est une caractéristique intrinsèque de l'électron. Pour le proton¹⁴, de façon analogue :

$$\vec{\mu}_S = g_p \frac{|e|}{2M_p} \vec{S} = g_p \mu_N \hbar^{-1} \vec{S} \equiv \gamma_p \vec{S} \quad (5.22)$$

avec $g_p = 5.59$. Pour le neutron¹⁵, $g_n = -3.82630$, 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*¹⁶) peut sembler mystérieuse au vu de la relation classique (5.16) – 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)¹⁷.

Le spin d'une particule est toujours demi-entier ou entier, c'est un fait d'expérience. Le caractère entier ou demi-entier du spin définit deux grandes classes de particules, les bosons (spin entier) et les fermions (spin demi-entier), obéissant à des statistiques très différentes (d'où l'existence d'une relation appelée *théorème spin-statistique*¹⁸). Par ailleurs, S peut prendre des valeurs relativement élevées : on connaît des noyaux qui, dans leur état fondamental, ont un spin égal à $\frac{1}{2}$. Toutefois, il semble difficile de concevoir des spins arbitrairement grands à l'instar des moments cinétiques orbitaux¹⁹ ; le spin ne se prête pas à la limite quasi-classique, ce qui n'est finalement pas très surprenant : de nature essentiellement quantique, il s'évanouit dans cette limite²⁰.

Revenons au cas de l'électron. Les deux valeurs possibles révélées par une mesure de S_z sont donc $\pm\hbar/2$, associées aux deux valeurs possibles d'un nombre quantique $M_S = \pm\frac{1}{2}$, lui-même associé à un nombre $S = \frac{1}{2}$ donnant le module d'un moment cinétique ($\|\vec{S}\| = \sqrt{S(S+1)}\hbar = \sqrt{3}\hbar/2$). Une description complète de l'état de l'électron contient nécessairement une fonction d'onde donnant comme d'habitude la densité de probabilité de présence, mais prenant également en compte le degré de liberté de spin, d'où la notation $\Psi(x, y, z, M_S; t)$. Si les coordonnées d'espace prennent des valeurs réelles continues, en revanche la variable de spin est essentiellement discrète et prend ici exclusivement les deux valeurs ci-dessus. En maintenant l'interprétation usuelle, la quantité $|\Psi(x, y, z, M_S; t)|^2 d^3r$ est la probabilité de présence autour du point choisi avec la valeur $M_S\hbar$ pour le spin. La condition de normalisation des probabilités introduit comme toujours une sommation, qui porte non seulement

¹³La théorie de Dirac donne $g_e = 2$ (voir section (5.6)). L'écart fini $g_e - 2$ résulte de corrections électrodynamiques ; son calcul est l'un des enjeux de l'Electrodynamique Quantique, dont le succès à cet égard est proprement fabuleux : les valeurs théoriques et expérimentales coïncident sur une douzaine de chiffres significatifs... On a [31] $g_{th} = 2.002319304402 \pm 6 \times 10^{-11}$, cependant que la valeur expérimentale connue est $g_{exp} = 2.002319304376 \pm 8 \times 10^{-12}$...

¹⁴également 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 $S(S+1)\hbar^2$.

¹⁵également 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.

¹⁶La charge du neutron est expérimentalement inférieure à 2×10^{-22} charge élémentaire [32].

¹⁷Cette 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.

¹⁸À propos du statut de cette relation (règle empirique, théorème, ...), voir [4], XIV. C. 1.

¹⁹Dans la limite des très grands nombres quantiques pour le moment orbital, la fonction de Brillouin rencontrée dans l'étude du paramagnétisme se confond pratiquement avec la fonction classique de Langevin.

²⁰Cela n'aurait pas de sens de prendre une limite du genre $M_S \rightarrow +\infty$, $\hbar \rightarrow 0$, $M_S\hbar = C^{ste}$ – comme on peut le faire pour un moment orbital –, qui transforme les fonctions de Brillouin en fonctions de Langevin.

sur les degrés “orbitaux” (sommation continue, c’est-à-dire intégration) et mais également sur les degrés de spin (sommation discrète) :

$$\int_{\mathbb{R}^3} d^3r |\Psi(\vec{r}, +\frac{1}{2})|^2 + \int_{\mathbb{R}^3} d^3r |\Psi(\vec{r}, -\frac{1}{2})|^2 = 1 \quad (5.23)$$

exprimant notamment le fait que l’on épuise toutes les possibilités du spin en sommant sur les deux valeurs possibles. L’équation (5.23) implique que si l’on fait une mesure simultanée de position *et* de spin – plus précisément de S_z , on peut trouver l’électron près de \vec{r} avec le spin $+\hbar/2$ *ou* avec le spin $-\hbar/2$. En termes plus concis :

$$\sum_{M_S = \pm \frac{1}{2}} \int_{\mathbb{R}^3} d^3r |\Psi(\vec{r}, M_S)|^2 = 1 . \quad (5.24)$$

En tout état de cause, l’électron n’a plus une mais deux fonctions d’onde, une pour chaque valeur²¹ de S_z . De la même façon, pour une particule de spin $S = 1$, M_S prend les trois valeurs $0, \pm 1$ et on écrit :

$$\sum_{M_S = 0, \pm 1} \int_{\mathbb{R}^3} d^3r |\Psi(\vec{r}, M_S)|^2 = 1 . \quad (5.25)$$

Il y a maintenant trois fonctions d’onde, et ainsi de suite : une particule de spin S aura $2S + 1$ fonctions d’onde ou, mieux, une fonction d’onde à $2S + 1$ composantes.

En effet, la notation précédente n’est pas forcément la meilleure. S’agissant d’une variable prenant des valeurs discrètes, il est tout aussi légitime (et même plus naturel) de mettre M_S en indice et de poser $\Psi_{M_S}(\vec{r}; t)$. Enfin, il est commode d’utiliser une notation matricielle, rangeant en colonne les différentes fonctions correspondant aux valeurs possibles de la variable discrète M_S . Ainsi, pour l’électron, on admettra désormais que toute l’information au sens de la Mécanique Quantique est contenue dans un vecteur-colonne à deux lignes appelé spineur²² et noté :

$$\begin{bmatrix} \Psi_{+\frac{1}{2}}(\vec{r}; t) \\ \Psi_{-\frac{1}{2}}(\vec{r}; t) \end{bmatrix} , \quad (5.26)$$

ou, encore plus simplement :

$$\begin{bmatrix} \Psi_+(\vec{r}; t) \\ \Psi_-(\vec{r}; t) \end{bmatrix} . \quad (5.27)$$

L’état d’un spin $S = \frac{1}{2}$ est donc complètement représenté non par une mais par deux fonctions d’onde. Pour une particule de spin $S = 1$, on posera de même :

$$\begin{bmatrix} \Psi_{+1}(\vec{r}; t) \\ \Psi_0(\vec{r}; t) \\ \Psi_{-1}(\vec{r}; t) \end{bmatrix} , \quad (5.28)$$

ou, quand aucune ambiguïté n’est à craindre :

$$\begin{bmatrix} \Psi_+(\vec{r}; t) \\ \Psi_0(\vec{r}; t) \\ \Psi_-(\vec{r}; t) \end{bmatrix} . \quad (5.29)$$

Il y a maintenant trois fonctions d’onde – pour un spin S , l’état est un spineur à $2S + 1$ composantes. Seules les particules de spin nul restent décrites complètement par une seule fonction d’onde.

La notation matricielle ci-dessus se transcrit immédiatement en notation de Dirac ; pour une particule de spin $S = \frac{1}{2}$, le vecteur d’état complet (*i. e.* incluant la variable de spin) s’écrit :

$$|\Psi(\vec{r}; t)\rangle = \Psi_{+\frac{1}{2}}(\vec{r}; t) |+\frac{1}{2}\rangle + \Psi_{-\frac{1}{2}}(\vec{r}; t) |-\frac{1}{2}\rangle . \quad (5.30)$$

²¹ou deux combinaisons linéaires indépendantes des deux états propres de S_z .

²²C’est ainsi que l’on appelle un objet se comportant d’une façon bien déterminée par rotation – voir plus loin.

$|\pm \frac{1}{2}\rangle$ désignent les états propres de S_z (et de \vec{S}^2) et forment une base de l'espace des états de spin ; les $\Psi_{\pm \frac{1}{2}}(\vec{r}; t)$ sont les coefficients du développement de l'état global. On peut d'ailleurs monter d'un cran dans l'usage de la notation de Dirac en identifiant :

$$\langle \vec{r}, M_S | \Psi(t) \rangle \equiv \Psi_{M_S}(\vec{r}; t) . \quad (5.31)$$

$|\vec{r}, M_S\rangle$ désigne alors un état produit tensoriel espace \otimes spin. Ceci étant, les relations habituelles se généralisent immédiatement ; par exemple, la fermeture s'écrit :

$$\sum_{M_S=-S}^{+S} \int_{\mathbb{R}^3} d^3r |\vec{r}, M_S\rangle \langle \vec{r}, M_S| = \mathbf{1} . \quad (5.32)$$

Si $\{\psi_n(\vec{r})\}$ désigne une base complète pour les variables d'espace, par exemple la base propre d'une observable, on peut noter Ψ_{n, M_S} le spineur (vecteur-colonne) ayant des zéros partout sauf dans la $(S + 1 - M_S)$ ème ligne :

$$\Psi_{n, M_S}(\vec{r}) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \psi_n(\vec{r}) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{--- ligne } S + 1 - M_S \quad , \quad (5.33)$$

et son adjoint par le vecteur-ligne obtenu par transposition et conjugaison complexe. La relation de fermeture prend alors la forme :

$$\sum_n \Psi_{n, M_S}^*(\vec{r}) \Psi_{n, M'_S}(\vec{r}') = \delta(\vec{r} - \vec{r}') \delta_{M_S M'_S} . \quad (5.34)$$

Le produit scalaire de deux spineurs Ψ et Φ est :

$$\langle \Psi | \Phi \rangle = \sum_{M_S=-S}^{+S} \int_{\mathbb{R}^3} d^3r \Psi_{M_S}^*(\vec{r}) \Phi_{M_S}(\vec{r}) . \quad (5.35)$$

Ceci peut aussi s'écrire sous forme matricielle ; par exemple, pour un spin $S = \frac{1}{2}$, (5.35) s'écrit aussi :

$$\langle \Psi | \Phi \rangle = \int_{\mathbb{R}^3} d^3r \begin{bmatrix} \Psi_+^*(\vec{r}; t) & \Psi_-^*(\vec{r}; t) \end{bmatrix} \begin{bmatrix} \Phi_+(\vec{r}; t) \\ \Phi_-(\vec{r}; t) \end{bmatrix} . \quad (5.36)$$

Il convient de signaler une autre notation, commode dans la pratique et couramment utilisée. Pour fixer les idées, on considère à nouveau un spin $S = \frac{1}{2}$. Soit les deux fonctions définies comme suit :

$$\alpha(M_S) = \begin{cases} 1 & \text{si } M_S = +\frac{1}{2} \\ 0 & \text{si } M_S = -\frac{1}{2} \end{cases} , \quad \beta(M_S) = \begin{cases} 0 & \text{si } M_S = +\frac{1}{2} \\ 1 & \text{si } M_S = -\frac{1}{2} \end{cases} . \quad (5.37)$$

Les valeurs de α et de β reflètent les composantes des vecteurs-colonne représentant les deux états $|\pm\rangle$; en fait, en introduisant les deux kets $|\alpha\rangle$ et $|\beta\rangle$, on peut écrire l'identification :

$$\alpha(M_S) \equiv \langle M_S | \alpha \rangle , \quad \beta(M_S) \equiv \langle M_S | \beta \rangle ; \quad (5.38)$$

l'état complet d'un électron peut alors se noter :

$$\Psi(\vec{r}, M_S; t) = \Psi_{+\frac{1}{2}}(\vec{r}; t) \alpha(M_S) + \Psi_{-\frac{1}{2}}(\vec{r}; t) \beta(M_S) . \quad (5.39)$$

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 :

$$\sum_{M_S = \pm \frac{1}{2}} \alpha(M_S) \beta(M_S) = 0 . \quad (5.40)$$

Parmi toutes les fonctions du type (5.39), il en est où les variables sont séparées. En effet, imaginons que les $\Psi_{\pm \frac{1}{2}}$, soient à un facteur près une seule et même fonction :

$$\Psi_{\pm \frac{1}{2}}(\vec{r}; t) = c_{\pm \frac{1}{2}} \psi(\vec{r}; t) . \quad (5.41)$$

Alors, d'après (5.39) :

$$\Psi(\vec{r}, M_S; t) = \psi(\vec{r}; t) [c_{+\frac{1}{2}} \alpha(M_S) + c_{-\frac{1}{2}} \beta(M_S)] \equiv \psi(\vec{r}; t) \otimes \chi(M_S) \quad (5.42)$$

et l'état total est bien un simple produit tensoriel espace \otimes spin (pour la simplicité d'écriture, on omet le plus souvent le symbole \otimes). La norme se calcule suivant :

$$\langle \Psi | \Psi \rangle = \int_{\mathbb{R}^3} d^3r |\psi(\vec{r}; t)|^2 [|c_{+\frac{1}{2}}|^2 + |c_{-\frac{1}{2}}|^2] ; \quad (5.43)$$

l'usage consiste naturellement à normaliser chaque facteur à l'unité :

$$\int_{\mathbb{R}^3} d^3r |\psi(\vec{r}; t)|^2 = 1 , \quad |c_{+\frac{1}{2}}|^2 + |c_{-\frac{1}{2}}|^2 = 1 . \quad (5.44)$$

Il n'est pas difficile de se convaincre que l'on peut toujours utiliser comme base de l'espace \mathcal{E} des états du genre (5.42) où les degrés de liberté d'espace et de spin sont séparés. Les états *propres* ne seront pas toujours de cette forme, en particulier quand il existe un couplage entre spin et espace²³, mais on peut toujours les représenter en combinaison linéaire d'états factorisés. Dans la suite, l'hypothèse d'une base séparée pourra toujours être faite pour la commodité et l'espace des états E pourra être considéré comme le produit tensoriel $E^{(\text{orb})} \otimes E^{(\text{spin})}$.

Bien sûr, toutes les observables n'impliquent pas le spin, loin s'en faut. Par exemple, une coordonnée cartésienne est, en représentation- q , une simple multiplication, qui n'affecte pas le degré de spin. Pour un spin $\frac{1}{2}$, en ordonnant les lignes et les colonnes toujours de la même façon, l'observable x (opérateur agissant dans l'espace des états E) est ainsi représentée par la matrice 2×2 :

$$x = \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} , \quad (5.45)$$

ce qui signifie que le vecteur résultant de l'action de x sur le spineur de rang 2 $|\psi\rangle$ est :

$$x|\Psi\rangle \longleftrightarrow \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} \begin{bmatrix} \Psi_+(\vec{r}) \\ \Psi_-(\vec{r}) \end{bmatrix} = \begin{bmatrix} x \Psi_+(\vec{r}) \\ x \Psi_-(\vec{r}) \end{bmatrix} . \quad (5.46)$$

De tels opérateurs, qui ne touchent pas au spin, sont appelés opérateurs orbitaux. Le moment conjugué est de ceux-ci ; sa composante p_x sera ainsi représentée par la matrice :

$$p_x = -i\hbar \begin{bmatrix} \partial/\partial x & 0 \\ 0 & \partial/\partial x \end{bmatrix} , \quad (5.47)$$

avec la signification :

$$p_x|\Psi\rangle \longleftrightarrow -i\hbar \begin{bmatrix} (\partial/\partial x) & 0 \\ 0 & (\partial/\partial x) \end{bmatrix} \begin{bmatrix} \Psi_+(\vec{r}) \\ \Psi_-(\vec{r}) \end{bmatrix} = -i\hbar \begin{bmatrix} (\partial/\partial x) \Psi_+(\vec{r}) \\ (\partial/\partial x) \Psi_-(\vec{r}) \end{bmatrix} . \quad (5.48)$$

²³D'ailleurs, même en l'absence d'interaction spin-orbite, les états acceptables physiquement pour un ensemble de N particules identiques ne sont presque jamais de cette forme. On verra par la suite comment, pour un ensemble de fermions, le postulat d'antisymétrisation entraîne que, pour $N > 2$, seul l'état “ferro” – tous les spins alignés – se présente sous la forme séparée, quelle que soit la composante M_S ($-N/2 \leq M_S \leq +N/2$). Au contraire, pour le cas $N = 2$, tous les états physiques, dans l'approximation électrostatique, sont séparés.

De par leur définition, les opérateurs orbitaux sont représentés dans l'espace des états de spin $|\pm \frac{1}{2}\rangle$ par des matrices diagonales scalaires. En désignant par $\mathbf{1}_S$ l'identité dans cet espace, un opérateur orbital est toujours de la forme $f(\vec{r}, \vec{p}) \otimes \mathbf{1}_S$, ou plus simplement $f(\vec{r}, \vec{p}) \mathbf{1}_S$; par exemple :

$$p_x = -i\hbar \frac{\partial}{\partial x} \mathbf{1}_S . \quad (5.49)$$

Le plus souvent, on omet aussi $\mathbf{1}_S$, pour la simplicité de l'écriture.

Les opérateurs qui au contraire affectent seulement le spin sont appelés opérateurs de spin et peuvent tous, pour un spin $\frac{1}{2}$, s'exprimer à l'aide des trois matrices de Pauli définies comme suit :

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} , \quad \sigma_y = \begin{bmatrix} 0 & -i \\ +i & 0 \end{bmatrix} , \quad \sigma_z = \begin{bmatrix} +1 & 0 \\ 0 & -1 \end{bmatrix} . \quad (5.50)$$

Ces matrices satisfont des relations remarquables qu'il faut connaître :

$$[\sigma_x, \sigma_y] = 2i\sigma_z \quad (+ \text{permutations circulaires}) . \quad (5.51)$$

$$\sigma_x \sigma_y + \sigma_y \sigma_x \equiv \{\sigma_x, \sigma_y\} = 0 \quad (+ \text{permutations circulaires}) . \quad (5.52)$$

$$\sigma_x \sigma_y = i\sigma_z \quad (+ \text{permutations circulaires}) . \quad (5.53)$$

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \mathbf{1} . \quad (5.54)$$

Une autre relation remarquable, souvent utile dans les applications, est la suivante :

$$(\vec{\sigma} \cdot \vec{A})(\vec{\sigma} \cdot \vec{B}) = (\vec{A} \cdot \vec{B}) \mathbf{1} + i\vec{\sigma} \cdot (\vec{A} \times \vec{B}) . \quad (5.55)$$

Dans (5.55), \vec{A} et \vec{B} désignent deux vecteurs *scalaires* vis-à-vis du spin. Le moment cinétique de spin pour $S = \frac{1}{2}$ s'exprime comme :

$$\vec{S} = \frac{\hbar}{2} \vec{\sigma} . \quad (5.56)$$

Il existe enfin des opérateurs mixtes, modifiant simultanément les degrés orbitaux et le spin. Soit par exemple le produit $L_z S_x$ agissant sur $\Psi_{\pm \frac{1}{2}}(\vec{r})$; en coordonnées sphériques, on peut écrire²⁴ :

$$L_z S_x \Psi_{\pm \frac{1}{2}}(\vec{r}) = -i\hbar \frac{\partial}{\partial \phi} \frac{\hbar}{2} \sigma_x \Psi_{\pm \frac{1}{2}}(\vec{r}) = -\frac{i}{2} \hbar^2 \frac{\partial}{\partial \phi} \Psi_{\mp \frac{1}{2}}(\vec{r}) . \quad (5.57)$$

Il en résulte que la matrice représentant cet opérateur est :

$$L_z S_x = -\frac{i}{2} \hbar^2 \begin{bmatrix} 0 & \frac{\partial}{\partial \phi} \\ \frac{\partial}{\partial \phi} & 0 \end{bmatrix} . \quad (5.58)$$

L'hermiticité de cette matrice ne saute pas aux yeux²⁵, bien que l'opérateur au premier membre est clairement hermitique puisque L_z et S_x sont hermitiques et commutent entre eux.

²⁴Noter l'association des signes : $\pm \leftrightarrow \mp$.

²⁵La démonstration de l'hermiticité de (5.58) est laissée à titre d'exercice. Il suffit d'expliciter la condition d'hermiticité :

$$(L_z S_x |\Phi\rangle, |\Psi\rangle) = (|\Phi\rangle, L_z S_x |\Psi\rangle) . \quad (5.59)$$

en manipulant explicitement les deux composantes Φ_{\pm} et Ψ_{\pm} des deux spineurs, et en faisant des intégrations par parties sur l'angle ϕ , toutes les fonctions de ϕ étant supposées périodiques et continues.

5.2 Lagrangien d'une particule dans un champ électromagnétique

Soit une particule de masse m et de charge q soumise à un champ électromagnétique $(\vec{\mathcal{E}}, \vec{\mathcal{B}})$. L'équation fondamentale de la dynamique prend la forme :

$$m \frac{d\vec{v}}{dt} = q(\vec{\mathcal{E}} + \vec{v} \times \vec{\mathcal{B}}) . \quad (5.60)$$

La partie magnétique de la force ne peut dériver d'une énergie potentielle au sens usuel : habituellement l'énergie potentielle ne dépend que de la coordonnée d'espace, alors que la force de Lorentz fait intervenir la vitesse de la particule. Remarquons en outre que la force magnétique ne donne lieu à aucun travail puisqu'elle est perpendiculaire au déplacement.

Comme on le sait, il est commode de définir un potentiel vecteur \vec{A} et un potentiel scalaire U permettant de calculer les champs par :

$$\vec{\mathcal{B}} = \vec{\nabla} \times \vec{A} , \quad \vec{\mathcal{E}} = -\vec{\nabla}U - \frac{\partial \vec{A}}{\partial t} . \quad (5.61)$$

Il n'est pas difficile de montrer que l'équation fondamentale de la dynamique (5.60) est reproduite par les équations de Lagrange habituelles :

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{u}} - \frac{\partial L}{\partial u} = 0 \quad (u = x, y, z) \quad (5.62)$$

appliquées au Lagrangien suivant :

$$L = \frac{1}{2} m \vec{v}^2 - qU + q\vec{A} \cdot \vec{v} . \quad (5.63)$$

Le terme potentiel de ce Lagrangien présente la particularité de dépendre de la vitesse. Pour cette raison, le moment conjugué (impulsion) \vec{p} , toujours défini par :

$$p_u = \frac{\partial L}{\partial \dot{u}} \quad (u = x, y, z) \iff \vec{p} = \frac{\partial L}{\partial \vec{v}} \equiv \vec{\nabla}_{\vec{v}} L \quad (5.64)$$

ne coïncide pas avec la quantité de mouvement ; compte tenu de (5.63) et (5.64), on trouve :

$$\vec{p} = m\vec{v} + q\vec{A} \neq m\vec{v} . \quad (5.65)$$

Le Hamiltonien se forme comme d'habitude :

$$H = \vec{p} \cdot \vec{v} - L = \frac{1}{2m} (\vec{p} - q\vec{A})^2 + qU . \quad (5.66)$$

Le premier terme représente bien l'énergie cinétique habituelle $m\vec{v}^2/2$. L'expression (5.66) vaut pour une particule soumise exclusivement au champ $(\vec{\mathcal{E}}, \vec{\mathcal{B}})$; bien sûr, si la particule est en outre soumise à un champ de force conservatif au sens usuel et donnant lieu à l'énergie potentielle V , le Hamiltonien complet est :

$$H = \frac{1}{2m} (\vec{p} - q\vec{A})^2 + qU + V . \quad (5.67)$$

Tout ceci est classique (ou formellement quantique). En particulier, les équations canoniques de Hamilton sont :

$$\frac{d\vec{r}}{dt} = \vec{\nabla}_{\vec{p}} H , \quad \frac{d\vec{p}}{dt} = -\vec{\nabla}_{\vec{r}} H \quad (5.68)$$

soit :

$$\frac{d\vec{r}}{dt} = \frac{1}{m} (\vec{p} - q\vec{A}) , \quad \frac{dp_u}{dt} = \frac{q}{m} [(\vec{p} - q\vec{A}) \cdot \vec{\nabla}] A_u - q \frac{\partial U}{\partial u} - \frac{\partial V}{\partial u} . \quad (5.69)$$

Ces équations reproduisent l'équation fondamentale de la dynamique où seules les forces apparaissent. Il est évident physiquement que la position et la vitesse de la particule sont des grandeurs physiques indépendantes de la jauge choisie : on peut les mesurer directement à l'aide d'expériences simples et cela n'aurait aucun sens de trouver que l'une ou l'autre dépend de la jauge choisie. Il en résulte que, en Mécanique Classique, l'impulsion \vec{p} dépend de la jauge, tout comme le moment cinétique $\vec{L} = \vec{r} \times \vec{p}$. Comme on l'a vu au chapitre 2, la situation semble inversée en Mécanique Quantique : l'opérateur $\vec{p} = -i\hbar \vec{\nabla}$ est invariant de jauge et, par voie de conséquence, l'opérateur vitesse $\vec{v} = (\vec{p} - q\vec{A})/m$ ne l'est pas – mais tout revient dans l'ordre au niveau des valeurs moyennes et c'est ceci qui compte.

5.3 Magnétisme atomique : l'atome d'hydrogène

Il s'agit de mettre en évidence sur l'exemple le plus simple – l'atome d'hydrogène –, les sources fondamentales du magnétisme. En particulier, on va voir que dans son l'état fondamental – mais la conclusion est la même pour tous les états sphériques de type S (*i. e.* $L = 0$) –, c'est le spin à lui seul qui explique le paramagnétisme observable de l'hydrogène. Plus généralement, le degré de liberté de spin est indispensable pour expliquer et décrire le magnétisme d'un atome complexe, notamment si ce dernier ne possède qu'un électron célibataire dans une couche s . Si le spin n'existait pas, un tel atome serait dépourvu de moment magnétique permanent ; mettant à part la nature inattendue du spectre des impacts observés dans l'expérience de Stern et Gerlach (nombre discret de taches alors que classiquement, postulant que le moment magnétique *existe*, on prévoit une répartition continue), aucune déviation ne devrait de toute façon se produire.

Pour l'atome d'hydrogène, mettant de côté pour l'instant la structure fine liée à V_{magn} , le Hamiltonien en l'absence de champ magnétique est :

$$H_0 = \frac{\vec{p}^2}{2m} + V(r) , \quad V(r) = -\frac{e'^2}{r} ; \quad (5.70)$$

e désigne toujours la charge de l'électron. En l'absence d'une composante électrique ($\vec{\mathcal{E}} \equiv 0$) et pour un champ magnétique statique, le potentiel scalaire U peut être pris nul (voir (5.61)) ; \vec{A} désignant toujours le potentiel-vecteur, le Hamiltonien classique est d'après (5.66) :

$$H_{\text{classique}} = \frac{1}{2m} (\vec{p} - e\vec{A})^2 + V(r) . \quad (5.71)$$

Le Hamiltonien quantique en représentation- q correspondant s'obtient en remplaçant \vec{p} par $-i\hbar\vec{\nabla}$. En prenant garde à l'ordre des opérateurs, il vient ainsi :

$$H = \frac{\vec{p}^2}{2m} - \frac{e}{2m} (\vec{p} \cdot \vec{A} + \vec{A} \cdot \vec{p}) + \frac{e^2 \vec{A}^2}{2m} + V(r) \equiv H_0 - \frac{e}{2m} (\vec{p} \cdot \vec{A} + \vec{A} \cdot \vec{p}) + \frac{e^2 \vec{A}^2}{2m} . \quad (5.72)$$

En prenant l'axe Oz le long du champ magnétique, un choix possible de potentiel-vecteur²⁶ est le suivant :

$$\vec{A} = \frac{1}{2} \vec{\mathcal{B}} \times \vec{r} \quad \implies \quad \vec{\mathcal{B}} = (0, 0, \mathcal{B}) . \quad (5.73)$$

Ce potentiel-vecteur est visiblement à divergence nulle ; comme pour toute fonction vectorielle $\vec{f}(\vec{r})$, on a :

$$[\vec{p}, \vec{f}] = -i\hbar \nabla \cdot \vec{f} , \quad (5.74)$$

il résulte du choix de la jauge (5.73) que l'opérateur \vec{p} commute ici avec \vec{A} . En outre, le report de l'expression (5.73) dans (5.72), suivi d'une permutation circulaire du produit mixte, fait apparaître le moment cinétique orbital \vec{L} ; on obtient :

$$H = H_0 - \frac{e}{2m} \vec{L} \cdot \vec{\mathcal{B}} + \frac{e^2 \vec{\mathcal{B}}^2}{8m} r^2 \sin^2 \theta , \quad (5.75)$$

θ étant l'angle en \vec{r} et $\vec{\mathcal{B}}$. Le terme linéaire en champ a la forme classique $-\vec{\mu}_{\text{orb}} \cdot \vec{\mathcal{B}}$ et a tout ce qu'il faut pour donner du paramagnétisme – pourvu que le moment cinétique orbital n'ait pas une moyenne nulle. Le terme quadratique²⁷ en champ, quant à lui, est de nature diamagnétique : il donne visiblement une variation positive d'énergie et est donc déstabilisant. En fait, dans les conditions usuelles, il est très petit devant le terme linéaire ; l'ordre de grandeur du rapport de ces termes est :

$$\left\langle \frac{|e|\mathcal{B}}{4\hbar} r^2 \sin^2 \theta \right\rangle \sim \frac{|e|\mathcal{B}}{4\hbar} a_0^2 \sim 10^{-7} \mathcal{B}_{\text{Tesla}} . \quad (5.76)$$

²⁶Cette jauge est parfois appelée "jauge cylindrique".

²⁷Ce caractère à lui-seul montre que l'interaction ainsi représentée se produit entre le champ et un moment *induit* par le champ.

Compte tenu de ceci, on néglige ce terme dans toute la suite²⁸.

L'expression (5.75), obtenue à partir du Hamiltonien classique, ignore par construction un aspect essentiel : le magnétisme lié au spin. L'interaction correspondante doit être ajoutée "à la main" ; elle a l'expression usuelle $-\vec{\mu}_{\text{spin}} \cdot \vec{\mathcal{B}}$, $\vec{\mu}_{\text{spin}}$ étant relié au spin \vec{S} suivant (5.21). Cette adjonction étant faite, l'expression (5.75) (débarassée du terme diamagnétique) devient²⁹ :

$$H = H_0 + H_{\text{Zeeman}} , \quad (5.77)$$

où H_{Zeeman} est³⁰ :

$$H_{\text{Zeeman}} = -\frac{e}{2m} (\vec{L} + g_e \vec{S}) \cdot \vec{\mathcal{B}} \equiv -\mu_B \hbar^{-1} (\vec{L} + g_e \vec{S}) \cdot \vec{\mathcal{B}} \quad (g_e \simeq 2) . \quad (5.78)$$

Ce terme correctif est petit devant H_0 , dans les conditions usuelles. En effet, en ordre de grandeur, il vaut (voir (5.18)) :

$$\langle H_{\text{Zeeman}} \rangle \sim |\mu_B| \mathcal{B} \sim 10^{-5} \text{eV } \mathcal{B}_{\text{Tesla}} . \quad (5.79)$$

Comme les différences d'énergie pour les premiers niveaux atomiques sont de l'ordre de l'eV ou plus, il est clair que H_{Zeeman} sera, pour les champs ordinaires, une petite correction. D'un autre côté, l'estimation faite en (5.8) montre que $\langle H_{\text{Zeeman}} \rangle$ est *a priori* comparable avec l'interaction spin-orbite, responsable de la structure fine. Pour éviter toute complication ici hors de propos – tout en gardant la cohérence –, on supposera le champ assez fort pour que le terme Zeeman domine largement les variations produisant cette structure fine³¹, ce qui revient finalement à oublier cette dernière.

L'ensemble d'opérateurs $(H, \vec{L}^2, L_z, \vec{S}^2, S_z)$ constitue visiblement un ensemble d'observables qui commutent. Les états propres de H_0 , notés $|n, l, m_l, m_s\rangle$, forment une base sur laquelle on peut donc d'emblée écrire :

$$\langle n, l, m_l, m_s | H_{\text{Zeeman}} | n, l', m_{l'}, m_{s'} \rangle \propto \delta_{ll'} \delta_{m_l m_{l'}} \delta_{m_s m_{s'}} . \quad (5.80)$$

D'un autre côté, rien ne permet d'affirmer que les éléments de matrice de H_{Zeeman} entre deux états ne diffèrent que par le nombre quantique n sont nuls – d'ailleurs, il ne le sont pas en général. Il en résulte que la résolution exacte de l'équation aux vecteurs propres de H_{Zeeman} exige la considération simultanée de tous les états $\{|n, l, m_l, m_s\rangle\}_n, l, m_l$ et m_s pouvant être fixés. En pratique, compte tenu de la petitesse de H_{Zeeman} par rapport aux différences d'énergie des niveaux de H_0 , on résout cette question par la technique de perturbation. Ici, la simple question est de mettre en évidence la valeur non-nulle du moment magnétique atomique. On se doute qu'elle est donnée, à une très bonne approximation³² par les valeurs moyennes du facteur de $-\vec{\mathcal{B}}$ dans H_{Zeeman} , calculées sur les états propres de H_0 . Ces valeurs moyennes sont en fait le moment magnétique de l'atome dans la limite du champ nul et représentent donc sa valeur intrinsèque, en-dehors de toute contribution induite. Ainsi, la valeur moyenne du moment magnétique atomique est :

$$\langle \vec{\mu}_{\text{at}} \rangle_{n, l, m_l, m_s} = \langle n, l, m_l, m_s | \frac{\mu_B}{\hbar} (\vec{L} + g_e \vec{S}) | n, l, m_l, m_s \rangle . \quad (5.81)$$

Dans un état propre de (\vec{S}^2, S_z) , les valeurs moyennes de S_x et de S_y sont nulles, d'où :

$$\langle \vec{\mu}_{\text{at}} \rangle_{n, l, m_l, m_s} = \mu_B (m_l + g_e m_s) (0, 0, 1) . \quad (5.82)$$

²⁸Évidemment, si le terme linéaire – qui donne la contribution paramagnétique – est nul de fait (exemple : atome dans l'état singulet de spin (si nombre pair d'électrons) et dans un état S ($L = 0$) pour le moment cinétique orbital), alors le terme diamagnétique à lui seul représente tout l'effet du champ magnétique. On sait bien que le diamagnétisme est toujours présent – au contraire du paramagnétisme – et qu'il produit des effets très faibles.

²⁹La limite de l'équation relativiste de Dirac produit tout naturellement le terme ici ajouté à la main et donne le Hamiltonien dit de Pauli.

³⁰C'est ce terme qui est responsable de l'effet Zeeman (normal ou anormal), d'où la notation employée.

³¹Dans le cas opposé (champ très faible en présence de structure fine), la bonne base est celle du moment cinétique total \vec{J} , soit $|n J M_J L S\rangle$. Chaque niveau de structure fine éclate en plusieurs niveaux très proches, l'écart étant toujours proportionnel à \mathcal{B} . En tant que combinaison linéaire d'opérateurs vectoriels, $\vec{\mu}_{\text{at}}$ est un opérateur vectoriel, dont le théorème de Wigner - Eckart permet de calculer la valeur moyenne dans un sous-espace n, J donnés.

³²Plus précisément à des corrections quadratiques en \mathcal{B} près, venant de la correction au premier ordre des états propres ; ces corrections quadratiques représentent une part de l'effet inductif global.

Ceci montre que quand Oz est choisi comme axe de quantification, seule la composante μ_z est non-nulle en moyenne et vaut :

$$\langle \mu_{\text{at}, z} \rangle_{n, l, m_l, m_s} = \mu_B (m_l + g_e m_s) \equiv \frac{e\hbar}{2m} (m_l + g_e m_s) . \quad (5.83)$$

Une telle moyenne est *toujours* différente de zéro : même dans un état S ($L = 0$) – le fondamental par exemple –, pour lequel m_l est forcément nul, il reste le nombre m_s qui prend lui les deux seules valeurs $m_s = \pm \frac{1}{2}$: quel que soit son état, l'atome d'hydrogène possède donc toujours – grâce au spin de l'électron – un moment magnétique non-nul ; sa composante suivant Oz est bornée inférieurement en module par $(g_e/2) |e\hbar/(2m)|$:

$$|\langle \mu_{\text{at}, z} \rangle| > \frac{1}{2} g_e \frac{|e\hbar|}{2m} \simeq \frac{|e\hbar|}{2m} \equiv |\mu_B| . \quad (5.84)$$

Le magnéton de Bohr, $|\mu_B|$, ressort bien comme l'unité fondamentale de moment magnétique atomique.

5.4 Rotation d'un spin

On sait que la rotation d'un angle θ autour de l'axe défini par le vecteur unitaire \vec{n} est représentée dans l'espace des états \mathcal{E} par un opérateur linéaire unitaire associé au moment cinétique, noté ici simplement R :

$$R = e^{\frac{1}{i\hbar} \theta \vec{J} \cdot \vec{n}} , \quad (5.85)$$

où \vec{J} est le moment cinétique. Pour un atome, quand on tient compte du spin, \vec{J} est la somme donnée en (5.10) : c'est bien vis-à-vis de ce moment total que \vec{r} , \vec{p} , \vec{L} et \vec{S} satisfont les relations caractéristiques requises pour les opérateurs vectoriels.

Autant le mode de transformation par rotation d'un moment cinétique orbital est clair, puisque ce dernier résulte d'un *mouvement* dont on peut se faire une représentation imagée et s'exprime en $\vec{r} \times \vec{p}$, autant celui d'un moment cinétique de spin est loin d'être évident³³. Pour les mêmes raisons, la transformation des états de spin n'est pas intuitive. Afin de déterminer ces modes de transformation, on s'appuie sur le fait que $\vec{\mu}_{\text{at}}$ est un opérateur vectoriel ; de ce fait, les valeurs moyennes de ses composantes se comportent par rotation comme celles d'un vecteur de la géométrie ordinaire. Alors, l'opérateur de rotation dans l'espace du spin se construit avec \vec{S} exactement comme la rotation dans \mathbb{R}^3 se construit avec \vec{L} .

En vertu de ceci, de (5.85), de (5.10) et de la relation de commutation $[\vec{L}, \vec{S}] = 0$, on a :

$$R = e^{\frac{1}{i\hbar} \theta (\vec{L} + \vec{S}) \cdot \vec{n}} = e^{\frac{1}{i\hbar} \theta \vec{L} \cdot \vec{n}} e^{\frac{1}{i\hbar} \theta \vec{S} \cdot \vec{n}} \equiv R^{(\text{orb})} \otimes R^{(\text{spin})} . \quad (5.86)$$

Par la rotation R , l'état $|\Psi\rangle = |\psi\rangle \otimes |\chi\rangle$ devient $|\Psi'\rangle$:

$$|\Psi'\rangle = R |\Psi\rangle = R^{(\text{orb})} |\psi\rangle \otimes R^{(\text{spin})} |\chi\rangle \quad (5.87)$$

montrant que l'état de spin est bel et bien affecté par la rotation. Comme la transformation de la partie orbitale a déjà été trouvée au chapitre 3, on examine dans la suite la seule transformation dans l'espace de spin.

Pour simplifier, on fixe la valeur du spin à $S = \frac{1}{2}$ – c'est aussi le cas le plus important en pratique. Dans ces conditions, on a précisément (voir (5.56)) :

$$R^{(\text{spin})} = e^{\frac{1}{i\hbar} \theta \frac{\hbar}{2} \vec{n} \cdot \vec{\sigma}} . \quad (5.88)$$

En développant l'exponentielle en série et en utilisant $\sigma_u^2 = \mathbf{1}_S$, on voit que :

$$R^{(\text{spin})} = \cos \frac{\theta}{2} \mathbf{1}_S - i \vec{n} \cdot \vec{\sigma} \sin \frac{\theta}{2} . \quad (5.89)$$

³³Concernant le comportement d'un spin par réflexion, voir [17], chapitre 12, section 8.

Il est facile d'écrire la matrice de rang 2 représentant cet opérateur sur la base propre de S_z ; en utilisant les expressions des matrices de Pauli (5.50), on obtient :

$$R^{(\text{spin})} = \begin{bmatrix} \cos \frac{\theta}{2} - in_z \sin \frac{\theta}{2} & -i(n_x - in_y) \sin \frac{\theta}{2} \\ -i(n_x + in_y) \sin \frac{\theta}{2} & \cos \frac{\theta}{2} + in_z \sin \frac{\theta}{2} \end{bmatrix} \equiv R^{(1/2)}(\theta, \vec{n}) . \quad (5.90)$$

La présence de l'angle moitié $\theta/2$ est visiblement liée à la valeur demi-entière du spin et a une conséquence un peu surprenante. En effet, si on prend $\theta = 2\pi$ (un tour complet), on s'attendrait à ce que la rotation soit la transformation identité. Il n'en est rien ; les formules précédentes donnent dans ce cas :

$$R^{(1/2)}(\theta = 2\pi, \vec{n}) = \cos \pi \mathbf{1}_S = -\mathbf{1}_S . \quad (5.91)$$

Ainsi, quand, par exemple, l'appareil de mesure fait un tour complet, le spineur de rang 2 change de signe et l'état total $|\Psi\rangle = |\psi\rangle \otimes |\chi\rangle$ en fait autant. Ceci n'est pas gênant : les deux états $\pm |\psi\rangle \otimes |\chi\rangle$ ne diffèrent que par une phase globale et représentent en fait le même état physique. La description de l'état physique est donc bien inchangée par la rotation d'un tour complet *tant qu'un seul état est impliqué* (voir remarque 3 ci-dessous) – heureusement.

Remarques

1. Cette bizarrerie peut être énoncée sous forme plus générale ([33], § 7-6). En effet, pour un moment cinétique \vec{J} quelconque, d'états propres $|j m\rangle$, on a :

$$R^{(j)}(\theta = 2\pi, \vec{n})|j m\rangle = (-1)^{2j} |j m\rangle \quad \forall j, m . \quad (5.92)$$

Pour démontrer cette relation, commençons par prendre une rotation autour de Oz ; alors, on a :

$$R^{(j)}(\theta, Oz)|j m\rangle = e^{\frac{i}{\hbar}\theta J_z}|j m\rangle = e^{-im\theta} |j m\rangle \implies R^{(j)}(\theta = 2\pi, \vec{n})|j m\rangle = e^{-2im\pi} |j m\rangle , \quad (5.93)$$

puisque le vecteur est propre de J_z . Toutes les valeurs de m ne diffèrent entre elles que par un entier relatif, l'égalité suivante est aussi vraie :

$$e^{-2im\pi} = e^{-2ij\pi} = (-1)^{2j} \quad \forall m \in \{-j, -j+1, \dots, j\} . \quad (5.94)$$

Pour un autre axe de rotation, le résultat de $R^{(j)}(2\pi, \vec{n})$ sur $|j m\rangle$ est une combinaison linéaire des seuls $(2j+1)$ vecteurs j fixé, m variable, $\{|j m\rangle\}_{-j \leq m \leq +j}$, puisque $\vec{n} \cdot \vec{J}$ commute avec \vec{J}^2 quel que soit \vec{n} ; mais la même phase $(-1)^{2j}$ apparaît pour chacun d'entre eux, et se met donc en facteur – ce qui achève d'établir (5.92). En conséquence, si j est entier, la rotation de 2π est bien l'identité ; dans le cas contraire :

$$R^{(j)}(2\pi, \vec{n})|j m\rangle = -|j m\rangle \quad (\text{pour tout } j \text{ demi-entier}) ; \quad (5.95)$$

ceci généralise (5.91) pour j demi-entier quelconque.

Notons que si le vecteur d'état peut éventuellement être affecté d'un signe lors d'une rotation de 2π , il n'en va pas de même pour une observable transformée ; $A' = RAR^\dagger$ étant une expression quadratique en R , tout signe finit par disparaître. Pour toute observable, on a donc :

$$[A, R^{(j)}(2\pi, \vec{n})] = 0 \iff AR^{(j)}(2\pi, \vec{n}) = R^{(j)}(2\pi, \vec{n})A . \quad (5.96)$$

2. Si l'on considère *tous* les vecteurs de type $|j m\rangle$, $\forall j, m$, on voit que l'espace vectoriel qu'ils engendrent peut être structuré en deux sous-espaces³⁴, l'un contenant tous les vecteurs ayant j entier, l'autre tous les vecteurs ayant j demi-entier³⁵. Pour la simplicité, on note $|\mathfrak{a}\rangle$ un vecteur quelconque du premier sous-espace (alors noté $\mathcal{E}_{\mathfrak{a}}$), et $|\mathfrak{b}\rangle$ un vecteur quelconque du second sous-espace noté $\mathcal{E}_{\mathfrak{b}}$; de la sorte, on a :

$$R^{(j)}(2\pi, \vec{n})|\mathfrak{a}\rangle = +|\mathfrak{a}\rangle , \quad R^{(j)}(2\pi, \vec{n})|\mathfrak{b}\rangle = -|\mathfrak{b}\rangle . \quad (5.97)$$

³⁴On dit parfois *secteurs*.

³⁵Ces deux sous-espaces sont mutuellement orthogonaux : en tant que vecteurs propres d'observables associés à des valeurs propres distinctes, tout vecteur $|j m\rangle$ est orthogonal à tout vecteur $|j' m'\rangle$ si $j \neq j'$ (et/ou) $m \neq m'$.

Les deux sous-espaces \mathcal{E}_a et \mathcal{E}_b s'ignorent complètement au sens où aucune observable ne permet de passer de l'un à l'autre. En effet, d'après (5.96), toute observable A est insensible à une rotation d'un tour complet. Prenons l'élément de matrice de (5.96) entre deux vecteurs $|a\rangle$ et $|b\rangle$; il vient :

$$\langle b|AR^{(j)}(2\pi, \vec{n})|a\rangle = \langle b|R^{(j)}(2\pi, \vec{n})A|a\rangle \iff \langle b|A|a\rangle = -\langle b|A|a\rangle , \quad (5.98)$$

d'où résulte :

$$\langle b|A|a\rangle = 0 \quad \forall A . \quad (5.99)$$

Ainsi, aucune observable ne peut transmuter un état $|a\rangle$ en un état $|b\rangle$ et inversement. Cette impossibilité est un exemple de règle de *supersélection*³⁶. Il est ainsi impossible de transformer un système contenant un nombre impair de fermions en un système contenant un nombre pair de fermions ; un fermion ne peut se transformer en boson, et inversement. Schématiquement, la fission du neutron en un proton et un électron (rayonnement β^-) exige la participation d'au moins un autre partenaire dans les produits finals (en l'occurrence un antineutrino).

Une telle règle a pour vertu de rendre inessentielle une phase non-globale qui, dans un autre contexte, ne peut ni ne doit être ignorée³⁷. En effet, soit un vecteur du genre :

$$|\Psi\rangle = |a\rangle + e^{i\alpha}|b\rangle . \quad (5.100)$$

Compte tenu de (5.99), la valeur moyenne d'une observable dans l'état $|\Psi\rangle$ est :

$$\langle \Psi|A|\Psi\rangle = \langle a|A|a\rangle + e^{-i\alpha}e^{i\alpha}\langle b|A|b\rangle = \langle a|A|a\rangle + \langle b|A|b\rangle . \quad (5.101)$$

Cette valeur moyenne est indépendante de la phase α – ce qui montre que cette dernière ne peut recevoir *ici* le moindre sens physique et qu'elle peut être complètement ignorée.

3. À l'inverse, pour *un* j demi-entier *donné*, la phase $i\pi$ donnant le signe $-$ dans (5.95) est visible dans des expériences d'interférométrie, comme l'ont montré Overhauser *et al.* en 1975 [34] (diffraction puis recombinaison de faisceaux de neutrons).
4. L'ensemble des matrices $R^{(1/2)}(\theta, \vec{n})$ définies par (5.90) peut être doté d'une structure de groupe avec la multiplication des matrices usuelle ; ce groupe est noté³⁸ $SU(2)$. Les $R^{(1/2)}$ sont unitaires ; leur déterminant est de surcroît³⁹ égal à $+1$:

$$\text{Det}R^{(1/2)}(\theta, \vec{n}) = \cos^2 \frac{\theta}{2} + n_z^2 \sin^2 \frac{\theta}{2} - i^2(n_x^2 + n_y^2) \sin^2 \frac{\theta}{2} = 1 . \quad (5.103)$$

$SU(2)$ est appelé groupe des matrices unitaires unimodulaires de rang 2.

5.5 Retour sur le renversement du temps

On a vu au chapitre 2 que, pour une particule décrite par une seule fonction d'onde (particule sans spin), l'opération de renversement du temps, alors désignée par K , est :

$$\Psi_{\text{renv}}(\vec{r}, t) = K\Psi(\vec{r}, t) = \Psi^*(\vec{r}, -t) . \quad (5.104)$$

³⁶Un autre exemple, très familier, d'une telle règle est fourni par la conservation de l'énergie : une transition qui viole la conservation de l'énergie est *strictement* interdite.

³⁷On sait bien que, en général, une phase dans un terme d'une combinaison linéaire est essentielle par le rôle qu'elle joue dans les interférences quantiques.

³⁸L'abréviation est pour *Special Unitary*.

³⁹D'une façon générale, le déterminant d'une matrice unitaire est un complexe de module 1 :

$$UU^\dagger = \mathbf{1} \iff \text{Dét}(UU^\dagger) = 1 \iff |\text{Dét } U| |\text{Dét } U^\dagger| = 1 \iff |\text{Dét } U| = |\text{Dét } U^\dagger| = 1 , \quad (5.102)$$

puisque $|\text{Dét } U| = (\text{Dét } U^\dagger)^*$.

Pour toute fonction indépendante du temps, l'opération de renversement du temps (en l'absence de degré de liberté de spin) est strictement identique à la conjugaison complexe. Dès lors, si on désigne par K_0 l'opération de conjugaison complexe définie comme :

$$K_0\psi(\vec{r}) = \psi^*(\vec{r}) \quad \forall\psi \quad (5.105)$$

alors, on a :

$$K = K_0 \quad (\text{particule sans spin}) \quad (5.106)$$

et, de toute évidence⁴⁰ :

$$K^2 \equiv K_0^2 = \mathbf{1} . \quad (5.107)$$

H lui-même est invariant par renversement du temps quand seules des puissances paires du moment conjugué y apparaissent et ceci assure que H , en représentation- q , est représenté par un opérateur différentiel réel⁴¹. La transformation des observables fondamentales s'écrit⁴² :

$$\vec{r}' = K \vec{r} K^\dagger = \vec{r} , \quad \vec{p}' = K \vec{p} K^\dagger = -\vec{p} . \quad (5.108)$$

Ainsi, par renversement du temps, les relations de commutation fondamentales – qui contiennent le scalaire imaginaire pur i – changent de signe, rappelant au passage que K est un opérateur antiunitaire⁴³. Lorsque la représentation est réelle, la coordonnée est représentée par une matrice réelle, le moment conjugué est représenté par une matrice imaginaire pure : en pareil cas, K coïncide strictement avec l'opération de conjugaison complexe effectuée cette fois sur les matrices ; cette équivalence opératoire n'est pas vraie en général⁴⁴.

De la transformation de \vec{r} et \vec{p} (5.108), il résulte immédiatement que le moment cinétique orbital est lui aussi changé en son opposé :

$$K \vec{L} K^\dagger = K (\vec{r} \times \vec{p}) K^\dagger = (K \vec{r} K^\dagger) \times (K \vec{p} K^\dagger) = \vec{r} \times (-\vec{p}) \equiv -\vec{L} ; \quad (5.110)$$

physiquement, ceci signifie simplement que la particule “renversée” tourne dans le sens opposé. En revanche, la transformation du moment cinétique de spin est loin d'être évidente, puisque l'on n'en possède pas une définition “fondamentale”, en termes de grandeurs dynamiques. Cependant, comme il s'agit précisément d'un moment cinétique, on *postule* :

$$K \vec{S} K^\dagger = -\vec{S} . \quad (5.111)$$

Ceci assure, notamment, que le moment cinétique total \vec{J} est changé en son opposé par renversement du temps⁴⁵ :

$$K \vec{J} K^\dagger = -\vec{J} \iff K J_z K^\dagger = -J_z , \quad K J_\pm K^\dagger = -J_\mp . \quad (5.112)$$

Tout ceci est raisonnable : on imagine mal que le moment cinétique total se transforme autrement, gardant en tête l'argument physique qui a conduit de l'évidence expérimentale d'un moment magnétique propre de l'électron – impliquant nécessairement un moment cinétique intrinsèque – à la nécessaire conservation du moment angulaire total d'un système isolé. En outre, K commute visiblement avec toutes les transformations spatiales (rotations, translations, etc.) : toucher au temps ne saurait affecter les coordonnées spatiales ; compte tenu du caractère particulier de K , il est utile de vérifier la cohérence de ce qui précède. \vec{J} étant le générateur des rotations, on a – précisément parce que K est antiunitaire :

$$K e^{\frac{i}{\hbar}\theta \vec{n} \cdot \vec{J}} K^\dagger = e^{-\frac{i}{\hbar}\theta \vec{n} \cdot (-\vec{J})} K K^\dagger = e^{\frac{i}{\hbar}\theta \vec{n} \cdot \vec{J}} \iff [K, e^{\frac{i}{\hbar}\theta \vec{n} \cdot \vec{J}}] = 0 . \quad (5.113)$$

⁴⁰(5.107) entraîne aussi que $K_0^\dagger = K_0$.

⁴¹C'est aussi pourquoi un Hamiltonien hermitique ne peut engendrer un mouvement irréversible, comme celui résultant d'une dissipation d'énergie d'un petit système vers un grand système.

⁴²En représentation- p on a aussi $K \vec{r} K^\dagger = K i \hbar \frac{\partial}{\partial \vec{p}} K^\dagger = -i \hbar \frac{\partial}{\partial (-\vec{p})} K K^\dagger = i \hbar \frac{\partial}{\partial \vec{p}} = \vec{r}$.

⁴³La deuxième relation dans (5.108) montre que l'anticommutateur de K et du moment conjugué est nul :

$$K \vec{p} K^\dagger = -\vec{p} \iff K \vec{p} + \vec{p} K \equiv \{K, \vec{p}\} = 0 . \quad (5.109)$$

⁴⁴Tout dépend de la représentation choisie. Par exemple, considérons les trois matrices 3×3 d'un moment angulaire $J = 1$, écrites au ch. 3. J_x et J_z ont des matrices réelles, cependant que celle de J_y est imaginaire pure. Il est bien clair que conjuguer ces matrices ne change pas \vec{J} en $-\vec{J}$. La même remarque vaut dans le cas $S = 1/2$, représenté par les trois matrices de Pauli.

⁴⁵Se souvenir que si Ω est antiunitaire, \forall scalaire λ , $\Omega \lambda = \lambda^* \Omega$, d'où l'équation le plus à droite dans (5.112).

La même vérification peut être faite avec \vec{p} , générateur des translations.

Il s'agit maintenant de construire l'opérateur de renversement du temps pour une particule pouvant avoir un spin, c'est-à-dire pour une particule ayant un moment cinétique \vec{J} avec j quelconque, entier ou demi-entier⁴⁶. Dans la suite, on se met dans une représentation dite standard, dont la base est l'ensemble des vecteurs $|\alpha j m\rangle$ qui sont propres d'un ECOC⁴⁷ contenant d'une part un ensemble d'observables invariantes par renversement du temps, dont les nombres quantiques sont collectivement dénotés α , d'autre part le couple (\vec{J}^2, J_z) . K change le signe de chaque composante de \vec{J} , mais laisse \vec{J}^2 invariant, donc $[K, \vec{J}^2] = 0$; en prenant l'élément de matrice entre deux états ne différant que par j , on a :

$$\langle \alpha j m | K \vec{J}^2 | \alpha j' m \rangle = \langle \alpha j m | \vec{J}^2 K | \alpha j' m \rangle \iff [j(j+1) - j'(j'+1)] \langle \alpha j m | K | \alpha j' m \rangle = 0 . \quad (5.114)$$

K n'a donc pas d'élément de matrice non-nul entre deux vecteurs de j différents. On peut donc fixer une fois pour toute la valeur de j ; le même argument vaut pour les observables faisant apparaître les nombres α : on peut donc tout autant fixer α . Ce faisant, on se place dans le sous-espace de dimension $2j+1$ engendré par les vecteurs ayant tous les mêmes j et α et différant seulement par leur m . Pour simplifier les notations, on pose désormais :

$$|\alpha j m\rangle \equiv |m\rangle . \quad (5.115)$$

L'opérateur K sera complètement défini quand on aura trouvé le résultat de l'action de K sur ces vecteurs. Physiquement, renverser le temps renverse les moments cinétiques – c'est bien ce qu'exprime (5.112) ; on s'attend donc à ce que $K|m\rangle \propto |-m\rangle$, ce qui assure déjà que la valeur propre de J_z change bien de signe. En fait, un peu de réflexion montre que ceci reste vrai avec un facteur de phase, pour l'instant indéterminé, dépendant *a priori* de m . On essaie donc :

$$K|m\rangle = e^{i\lambda(m)} |-m\rangle , \quad (5.116)$$

et l'objectif est maintenant de trouver la fonction $\lambda(m)$. (5.116) définit *ipso facto* K^\dagger . En effet :

$$(K^\dagger|m\rangle, |m'\rangle) = (|m\rangle, K|m'\rangle)^* = (|m\rangle, e^{i\lambda(m')} |-m'\rangle)^* = e^{-i\lambda(m')} (|m\rangle, |-m'\rangle)^* = e^{-i\lambda(m')} \delta_{m-m'} . \quad (5.117)$$

En d'autres termes, prenant les complexes conjugués :

$$(|m'\rangle, K^\dagger|m\rangle) = e^{i\lambda(-m)} \delta_{m-m'} \iff K^\dagger|m\rangle = e^{i\lambda(-m)} |-m\rangle . \quad (5.118)$$

Explicitons maintenant l'équation le premier membre de (5.112) :

$$K J_z K^\dagger |m\rangle = K J_z e^{i\lambda(-m)} |-m\rangle = -m\hbar K e^{i\lambda(-m)} |-m\rangle . \quad (5.119)$$

Comme K est antilinéaire :

$$K J_z K^\dagger |m\rangle = -m\hbar e^{-i\lambda(-m)} K |-m\rangle = -m\hbar e^{-i\lambda(-m)} e^{i\lambda(-m)} |+m\rangle = -m\hbar |m\rangle \equiv -J_z |m\rangle . \quad (5.120)$$

Comme attendu, J_z est transformé en $-J_z$. Maintenant :

$$\begin{aligned} K J_+ K^\dagger |m\rangle &= K J_+ e^{i\lambda(-m)} |-m\rangle = \hbar e^{-i\lambda(-m)} \sqrt{j(j+1) - (-m)(-m+1)} K |-m+1\rangle \\ &= \hbar e^{-i\lambda(-m)} \sqrt{j(j+1) - m(m-1)} e^{i\lambda(-m+1)} |m-1\rangle . \end{aligned} \quad (5.121)$$

Conformément à (5.112), le point terminal doit être $-J_- |m\rangle$, ce qui impose quel que soit m :

$$-\lambda(-m) + \lambda(-m+1) = \pi \quad (2\pi) \iff \lambda(m+1) - \lambda(m) = \pi \quad (2\pi) . \quad (5.122)$$

La dernière condition $K J_- K^\dagger = -J_+$ fournit la même équation, qui dit que d'une valeur de m à la suivante – que j soit entier ou demi-entier –, la phase λ augmente de π ; l'origine des phases étant inessentielle, on peut donc prendre :

$$\lambda(m) = m\pi \quad (2\pi) , \quad (5.123)$$

⁴⁶Pour une autre présentation de la suite, voir [20], XV-18, p. 571.

⁴⁷Le Hamiltonien n'est pas forcément dans l'ECOC ; l'argument qui suit vaut donc tout autant que le système ait ou n'ait pas la symétrie sphérique.

ce qui achève de définir l'opérateur de renversement du temps pour un moment cinétique quelconque :

$$K|\alpha j m\rangle = e^{im\pi} |\alpha j - m\rangle . \quad (5.124)$$

Comme la base est propre de J_z , ceci s'énonce encore :

$$K|\alpha j m\rangle = e^{-\frac{1}{i\hbar}\pi J_z} |\alpha j - m\rangle . \quad (5.125)$$

Formellement, et avec ce choix d'origine pour la phase, l'action de K sur $|\alpha j m\rangle$ est la même que celle d'une rotation de π (2π) autour de l'axe Oz effectuée sur $|\alpha j - m\rangle$. Ces résultats sont vrais $\forall j$ et en particulier pour j entier auquel cas il peut s'agir d'un moment cinétique orbital. Alors, la représentation- q des $|\alpha j m\rangle$ fait apparaître les Y_{lm} et (5.124) se transcrit comme suit :

$$KY_{lm} = e^{im\pi} Y_{l-m} = (-1)^m Y_{l-m} . \quad (5.126)$$

Le second membre n'est autre que⁴⁸ $(Y_{lm})^*$: dans ce cas précis, l'opérateur K se réduit à la conjugaison complexe.

Ceci étant acquis, que vaut le carré de K ? Par (5.124), on a :

$$K^2|m\rangle = K K|m\rangle = K(e^{im\pi} |-m\rangle) = e^{-im\pi} K|-m\rangle = e^{-im\pi} e^{-mi\pi} |m\rangle = e^{-2im\pi} |m\rangle , \quad (5.127)$$

ce que l'on peut aussi voir avec l'expression (5.125), compte tenu de la commutation (5.113). Comme, dans tous les cas, m ne diffère de j que par un entier, on a :

$$e^{2im\pi} = e^{2ij\pi} \quad \forall j, m = -j, -j+1, \dots, +j . \quad (5.128)$$

En définitive, sur cette base standard, le carré de K a la même action que :

$$K^2 = e^{2ij\pi} \mathbf{1} \quad (5.129)$$

qui donne finalement l'important résultat⁴⁹ :

$$K^2 = \begin{cases} +\mathbf{1} & \text{si spin entier} \\ -\mathbf{1} & \text{si spin demi-entier} \end{cases} . \quad (5.130)$$

Comme le caractère entier ou demi-entier du spin d'un système composite est lié biunivoquement au nombre de particules de spin demi-entier, $N_{\frac{1}{2}}$, il vient finalement :

$$K^2 = (-1)^{N_{\frac{1}{2}}} \mathbf{1} . \quad (5.131)$$

Ce résultat permet de démontrer l'existence de la dégénérescence de Kramers pour les systèmes à spin demi-entier. En effet, dans le cas où $K^2 = -\mathbf{1}$, on a :

$$K = -K^{-1} = -K^\dagger , \quad (5.132)$$

d'où, en partant de la définition de l'adjoint d'un opérateur antiunitaire :

$$(K^\dagger|\Psi\rangle, |\Psi\rangle) = (|\Psi\rangle, K|\Psi\rangle)^* = (K|\Psi\rangle, |\Psi\rangle) = -(K^\dagger|\Psi\rangle, |\Psi\rangle) . \quad (5.133)$$

Les termes extrêmes de cette chaîne montrent que le produit scalaire $(K^\dagger|\Psi\rangle, |\Psi\rangle)$ est nul :

$$(K^\dagger|\Psi\rangle, |\Psi\rangle) = 0 \iff (|\Psi\rangle, K|\Psi\rangle) = 0 . \quad (5.134)$$

Si donc H est invariant par renversement du temps, $|\Psi\rangle$ et $K|\Psi\rangle$ d'une part ont la même énergie et d'autre part sont orthogonaux : il existe donc en ce cas une dégénérescence forcément paire, c'est la dégénérescence de Kramers. Un champ magnétique brise l'invariance par renversement du temps et lève donc la dégénérescence de Kramers.

⁴⁸voir ch 3.

⁴⁹Comme un moment demi-entier ne peut provenir que du spin, on peut remplacer *spin* par *moment cinétique* dans l'affirmation traduite par (5.130), et inversement.

Remarque

Il est possible de d'écrire K sous forme d'un produit d'opérateurs⁵⁰, comme on l'a fait partiellement pour écrire (5.125). Il suffit pour cela d'introduire l'opérateur qui transforme $|\alpha j m\rangle$ en $|\alpha j - m\rangle$. Sur la base standard, cet opérateur est représenté par une matrice qui a des 1 partout sur l'antidiagonale⁵¹ et des zéros partout ailleurs⁵². Visiblement, le carré de cet opérateur est l'identité $\mathbf{1}$: il est donc assez naturel de le noter⁵³ conventionnellement $\sqrt{\mathbf{1}}$. Ainsi, par définition :

$$\sqrt{\mathbf{1}}|\alpha j m\rangle = |\alpha j - m\rangle . \quad (5.135)$$

Dès lors, l'opérateur K s'écrit :

$$K = e^{\frac{i}{\hbar}\pi J_z} \sqrt{\mathbf{1}} . \quad (5.136)$$

Dans le cas particulier $S = \frac{1}{2}$, $\sqrt{\mathbf{1}} = \sigma_x$ et on a simplement :

$$K = e^{\frac{i}{\hbar}\pi S_z} \sigma_x = e^{-\frac{i\pi}{2}\sigma_z} \sigma_x = -i\sigma_z\sigma_x = \sigma_y , \quad (5.137)$$

soit explicitement :

$$K|\frac{1}{2}\rangle = i|-\frac{1}{2}\rangle , \quad K|-\frac{1}{2}\rangle = -i|\frac{1}{2}\rangle \quad (5.138)$$

en conformité avec (5.124).

5.6 Origine relativiste du spin

Il s'agit de présenter – sous forme élémentaire – les arguments mettant en évidence la nature relativiste du degré de liberté interne que constitue le spin. Afin de rester au niveau le plus simple et pour aller droit au but, 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 (5.5).

5.6.1 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 :

$$i\hbar \frac{\partial}{\partial t} \bullet = \frac{1}{2m} (-i\hbar \vec{\nabla})^2 \bullet \quad (5.139)$$

s'obtient à partir de l'expression classique :

$$H = \frac{p^2}{2m} , \quad (5.140)$$

en y faisant la substitution :

$$H \longrightarrow i\hbar \frac{\partial}{\partial t} , \quad \vec{p} \longrightarrow -i\hbar \vec{\nabla} . \quad (5.141)$$

⁵⁰Une telle factorisation apparaît aussi dans [20], mais ce n'est pas la même (4×3 est aussi égal à 2×6 !).

⁵¹*i. e.* la diagonale SW - NE.

⁵²En dimension 2, cet opérateur n'est autre que σ_x .

⁵³Dans un espace vectoriel, il existe une infinité de racines de l'unité. Par exemple, en dimension 2, on connaît déjà les 8 matrices : $\mathbf{1}$, les trois matrices de Pauli et leurs opposés. Ceci étant, toutes les matrices s'en déduisant par une transformation (unitaire par exemple) sont encore des racines de l'unité.

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

$$H^2 = p^2 c^2 + m^2 c^4 ; \quad (5.142)$$

En y faisant la substitution (5.141), on obtient :

$$-\hbar^2 \frac{\partial^2}{\partial t^2} = -\hbar^2 c^2 \Delta + m^2 c^4 \iff \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta + \frac{m^2 c^2}{\hbar^2} \right) \Psi = 0 . \quad (5.143)$$

$\frac{\hbar}{mc}$, 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 (5.143) 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 continuité⁵⁴, 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 :

$$m\ddot{x} = -kx . \quad (5.144)$$

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

$$\dot{x} = \frac{p}{m} , \quad \dot{p} = -kx , \quad (5.145)$$

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

$$\frac{d}{dt} \begin{bmatrix} x \\ p \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{m} \\ -k & 0 \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix} \quad (5.146)$$

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, (5.143) s'écrit aussi :

$$\left(1 + \frac{i\hbar}{mc^2} \frac{\partial}{\partial t} \right) \left(1 - \frac{i\hbar}{mc^2} \frac{\partial}{\partial t} \right) \Psi = \frac{\hbar^2}{mc^2} \Delta \Psi , \quad (5.147)$$

ce qui suggère d'introduire les deux combinaisons :

$$\Psi_a = \Psi + \frac{i\hbar}{mc^2} \frac{\partial \Psi}{\partial t} , \quad \Psi_b = \Psi - \frac{i\hbar}{mc^2} \frac{\partial \Psi}{\partial t} , \quad (5.148)$$

d'où résulte :

$$\Psi = \frac{1}{2} (\Psi_a + \Psi_b) , \quad \frac{\partial \Psi}{\partial t} = \frac{mc^2}{2i\hbar} (\Psi_a - \Psi_b) . \quad (5.149)$$

Dès lors, (5.147) peut s'écrire de deux façons :

$$\left(1 - \frac{i\hbar}{mc^2} \frac{\partial}{\partial t} \right) \Psi_a = \frac{\hbar^2}{2mc^2} \Delta (\Psi_a + \Psi_b) , \quad \left(1 + \frac{i\hbar}{mc^2} \frac{\partial}{\partial t} \right) \Psi_b = \frac{\hbar^2}{2mc^2} \Delta (\Psi_a + \Psi_b) \quad (5.150)$$

⁵⁴Bien 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 :

$$i\hbar \frac{\partial}{\partial t} \begin{bmatrix} \Psi_a \\ \Psi_b \end{bmatrix} = \begin{bmatrix} mc^2 - \frac{\hbar^2}{2m} \Delta & -\frac{\hbar^2}{2m} \Delta \\ +\frac{\hbar^2}{2m} \Delta & -mc^2 + \frac{\hbar^2}{2m} \Delta \end{bmatrix} \begin{bmatrix} \Psi_a \\ \Psi_b \end{bmatrix} . \quad (5.151)$$

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 (5.143) 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 (5.143). 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 :

$$\vec{j} = \frac{\hbar}{2im} (\Psi^* \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^*) . \quad (5.152)$$

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 :

$$\rho = \frac{i\hbar}{2mc^2} \left(\Psi^* \frac{\partial\Psi}{\partial t} - \Psi \frac{\partial\Psi^*}{\partial t} \right) , \quad (5.153)$$

alors l'équation de Klein - Gordon (5.143) prend la la forme souhaitée d'une équation de conservation :

$$\frac{\partial\rho}{\partial t} + \text{div } \vec{j} = 0 . \quad (5.154)$$

Le seul ennui est que la densité ρ définie en (5.153)... 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 ([20], 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 ([35], 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 (5.151), 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 ([35], 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 (5.151), 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 , $\mu = 1, 2, 3, 4$:

$$\Psi = \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \Psi_3 \\ \Psi_4 \end{bmatrix} . \quad (5.155)$$

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

$$\rho(\vec{r}, t) = \sum_{\mu=1}^4 |\Psi_\mu(\vec{r}, t)|^2 . \quad (5.156)$$

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 :

$$i\hbar \frac{\partial}{\partial t} \Psi(\vec{r}, t) = (c \vec{\alpha} \cdot \vec{p} + \beta mc^2) \Psi(\vec{r}, t) \equiv H_D \Psi(\vec{r}, t) . \quad (5.157)$$

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 (5.157) est⁵⁵ :

$$i\hbar \frac{\partial}{\partial t} \Psi(\vec{r}, t) = [c \vec{\alpha} \cdot \vec{p} + V(\vec{r}) \mathbf{1} + \beta mc^2] \Psi(\vec{r}, t) . \quad (5.158)$$

L'équation (5.157) s'écrit aussi :

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \vec{\alpha} \cdot \vec{\nabla} + \frac{imc}{\hbar} \beta \right) \Psi(\vec{r}, t) = 0 , \quad (5.159)$$

soit, sous forme plus explicite :

$$\left[\frac{1}{c} \frac{\partial}{\partial t} + \left(\sum_{u=x,y,z} \alpha_u \frac{\partial}{\partial u} \right) + \frac{imc}{\hbar} \beta \right] \Psi(\vec{r}, t) = 0 . \quad (5.160)$$

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 (5.142), qui est vraie pour *toute* particule et qui produit l'équation de Klein - Gordon (5.143) par simple substitution des opérateurs aux grandeurs classiques⁵⁶. Une façon de maintenir le contact est de faire des manipulations sur l'opérateur entre crochets dans (5.160) pour lui donner l'allure de l'opérateur apparaissant dans (5.143). 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 (5.160) 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 (5.142), 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 (5.160), un peu de réflexion montre qu'il faut introduire l'opérateur de la forme $A - B$, soit :

$$\frac{1}{c} \frac{\partial}{\partial t} - \left(\sum_{v=x,y,z} \alpha_v \frac{\partial}{\partial v} \right) - \frac{imc}{\hbar} \beta \quad (5.161)$$

⁵⁵Dans la suite, $\mathbf{1}$ note tantôt la matrice identité 4×4 , tantôt la matrice identité 2×2 ; le contexte permet de lever toute ambiguïté.

⁵⁶Pour 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 (5.160). On obtient ainsi :

$$\left[\frac{1}{c} \frac{\partial}{\partial t} - \left(\sum_v \alpha_v \frac{\partial}{\partial v} \right) - \frac{imc}{\hbar} \beta \right] \left[\frac{1}{c} \frac{\partial}{\partial t} + \left(\sum_u \alpha_u \frac{\partial}{\partial u} \right) + \frac{imc}{\hbar} \beta \right] \Psi(\vec{r}, t) = 0 . \quad (5.162)$$

Il suffit maintenant d'expliciter les termes au premier membre, et de jouer avec les indices muets de sommation pour obtenir⁵⁷ :

$$\left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{1}{2} \sum_{u,v} (\alpha_u \alpha_v + \alpha_v \alpha_u) \frac{\partial^2}{\partial u \partial v} - \frac{imc}{\hbar} \sum_u (\alpha_u \beta + \beta \alpha_u) \frac{\partial}{\partial u} + \left(\frac{mc}{\hbar} \right)^2 \beta^2 \right] \Psi(\vec{r}, t) = 0 . \quad (5.163)$$

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

$$\alpha_u \alpha_v + \alpha_v \alpha_u = 2 \delta_{uv} \iff \alpha_u^2 = \mathbf{1} \quad (u = x, y, z) \quad \text{et} \quad \alpha_u \alpha_v + \alpha_v \alpha_u = 0 \quad (u \neq v) \quad (5.164)$$

et :

$$\beta^2 = \mathbf{1} , \quad \alpha_u \beta + \beta \alpha_u = 0 \quad (u = x, y, z) . \quad (5.165)$$

Les α_u et β sont bien des matrices : les relations ci-dessus ne sauraient être satisfaites par des nombres ordinaires⁵⁸.

Les relations (5.164) et (5.165) étant admises, il n'est pas difficile de montrer que les matrices 4×4 cherchées peuvent s'écrire par blocs 2×2 comme suit⁵⁹ :

$$\vec{\alpha} = \begin{bmatrix} \mathbf{0} & \vec{\sigma} \\ \vec{\sigma} & \mathbf{0} \end{bmatrix} , \quad \beta = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} \quad (5.166)$$

où les $\vec{\sigma}$ sont les matrices de Pauli (5.50) 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 (5.166) est appelé "représentation-standard", étant entendu qu'il existe une infinité de couples $(\vec{\alpha}, \beta)$ vérifiant (5.164) et (5.165), déductibles de (5.166) par une transformation unitaire quelconque. En tout cas, dans cette représentation, H_D s'écrit :

$$H_D = \begin{bmatrix} mc^2 \mathbf{1} & c \vec{\sigma} \cdot \vec{p} \\ c \vec{\sigma} \cdot \vec{p} & -mc^2 \mathbf{1} \end{bmatrix} . \quad (5.167)$$

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 (5.143).

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 :

$$\vec{\Sigma} = \begin{bmatrix} \vec{\sigma} & \mathbf{0} \\ \mathbf{0} & \vec{\sigma} \end{bmatrix} \quad (5.168)$$

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 :

$$\vec{J} = \vec{L} + \frac{\hbar}{2} \vec{\Sigma} , \quad (5.169)$$

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 (5.158), soit :

$$H_D = [c \vec{\alpha} \cdot \vec{p} + V(\vec{r}) \mathbf{1} + \beta mc^2] . \quad (5.170)$$

⁵⁷ *A priori* les matrices α_u ne commutent pas entre elles, alors que $\frac{\partial^2}{\partial u \partial v} = \frac{\partial^2}{\partial v \partial u}$.

⁵⁸ 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.

⁵⁹ $\mathbf{0}$ 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 :

$$[\vec{J}, H_D] = [\vec{J}, c\vec{\alpha}\cdot\vec{p} + \beta mc^2] . \quad (5.171)$$

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 :

$$\vec{L} = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} (\vec{r} \times \vec{p}) , \quad (5.172)$$

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

$$[\vec{L}, H_D] = [(\vec{r} \times \vec{p}) \mathbf{1}, c\vec{\alpha}\cdot\vec{p} + \beta mc^2] = [(\vec{r} \times \vec{p}) \mathbf{1}, c\vec{\alpha}\cdot\vec{p}] . \quad (5.173)$$

Par exemple, on a :

$$[L_z, H_D] = c[xp_y - yp_x, \sum_u \alpha_u p_u] = i\hbar c(\alpha_x p_y - \alpha_y p_x) \equiv i\hbar c(\vec{\alpha} \times \vec{p})_z , \quad (5.174)$$

d'où :

$$[\vec{L}, H_D] = i\hbar c(\vec{\alpha} \times \vec{p}) . \quad (5.175)$$

Par ailleurs⁶⁰ :

$$[\vec{\Sigma}, H_D] = \left[\begin{bmatrix} \vec{\sigma} & \mathbf{0} \\ \mathbf{0} & \vec{\sigma} \end{bmatrix}, \begin{bmatrix} mc^2 \mathbf{1} & c\vec{\sigma}\cdot\vec{p} \\ c\vec{\sigma}\cdot\vec{p} & -mc^2 \mathbf{1} \end{bmatrix} \right] = \begin{bmatrix} \mathbf{0} & c[\vec{\sigma}, \vec{\sigma}\cdot\vec{p}] \\ c[\vec{\sigma}, \vec{\sigma}\cdot\vec{p}] & \mathbf{0} \end{bmatrix} . \quad (5.176)$$

On a, par exemple :

$$[\sigma_x, \vec{\sigma}\cdot\vec{p}] = [\sigma_x, \sum_u \sigma_u p_u] = 2i(p_y \sigma_z - p_z \sigma_y) \equiv 2i(\vec{p} \times \vec{\sigma})_x , \quad (5.177)$$

d'où :

$$[\vec{\Sigma}, H_D] = 2ic \begin{bmatrix} \mathbf{0} & \vec{p} \times \vec{\sigma} \\ \vec{p} \times \vec{\sigma} & \mathbf{0} \end{bmatrix} \equiv 2ic \vec{p} \times \vec{\alpha} . \quad (5.178)$$

Prenant en compte (5.175) et (5.178), le commutateur de \vec{J} avec le Hamiltonien de Dirac est :

$$[\vec{J}, H_D] \equiv [\vec{L} + \frac{\hbar}{2} \vec{\Sigma}, H_D] = i\hbar c(\vec{\alpha} \times \vec{p} + \vec{p} \times \vec{\alpha}) = 0 . \quad (5.179)$$

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 :

$$\vec{S} = \frac{\hbar}{2} \vec{\Sigma} = \frac{\hbar}{2} \begin{bmatrix} \vec{\sigma} & \mathbf{0} \\ \mathbf{0} & \vec{\sigma} \end{bmatrix} , \quad \vec{J} = \vec{L} + \vec{S} . \quad (5.180)$$

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 ([20], p. 807, et voir la Remarque plus loin) fournit le Hamiltonien :

$$H_{c \rightarrow \infty} = \frac{1}{2m} (\vec{p} - e\vec{A})^2 - \frac{e\hbar}{2m} \vec{\sigma}\cdot\vec{B} + eU . \quad (5.181)$$

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 :

$$\vec{\mu} = 2 \frac{e\hbar}{2m} \vec{S} \iff g_e = 2 . \quad (5.182)$$

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 (5.181) 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 (5.78) – à $g_e - 2$ près, évidemment.

⁶⁰Penser à utiliser la multiplication par bloc des matrices.

5.6.2 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 :

$$\Psi(\vec{r}, t) = e^{\frac{i}{\hbar} Et} \psi(\vec{r}) , \quad (5.183)$$

(5.157) devient :

$$E \psi(\vec{r}) = H_D \psi(\vec{r}) \quad (5.184)$$

et constitue l'équation aux valeurs et fonctions propres de Dirac pour un électron libre. En l'absence de champ deforces, 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} :

$$\psi(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} C(\vec{k}) . \quad (5.185)$$

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 (5.185), $\psi(\vec{r})$ et $C(\vec{k})$ sont des matrices-colonne à quatre éléments ; sous forme explicite :

$$\psi_\mu(\vec{r}) = e^{i\vec{k}\cdot\vec{r}} C_\mu(\vec{k}) \quad (\mu = 1, 2, 3, 4) . \quad (5.186)$$

Bien évidemment, C_μ est indépendant de \vec{r} . En reportant dans (5.184) et en utilisant :

$$\vec{p} e^{i\vec{k}\cdot\vec{r}} = -i\hbar \vec{\nabla} e^{i\vec{k}\cdot\vec{r}} = \hbar \vec{k} e^{i\vec{k}\cdot\vec{r}} \quad (5.187)$$

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

$$(c \hbar \vec{k} \cdot \vec{\alpha} + \beta mc^2) C(\vec{k}) = E C(\vec{k}) , \quad (5.188)$$

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

$$\begin{bmatrix} mc^2 \mathbf{1} & c \hbar \vec{k} \cdot \vec{\sigma} \\ c \hbar \vec{k} \cdot \vec{\sigma} & -mc^2 \mathbf{1} \end{bmatrix} \begin{bmatrix} C_1(\vec{k}) \\ C_2(\vec{k}) \\ C_3(\vec{k}) \\ C_4(\vec{k}) \end{bmatrix} = E \begin{bmatrix} C_1(\vec{k}) \\ C_2(\vec{k}) \\ C_3(\vec{k}) \\ C_4(\vec{k}) \end{bmatrix} \iff H_D C(\vec{k}) = E C(\vec{k}) . \quad (5.189)$$

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

$$H_D^2 C(\vec{k}) = E^2 C(\vec{k}) ; \quad (5.190)$$

en calculant (toujours par blocs) le carré de la matrice figurant dans (5.189), on trouve facilement la forme matricielle⁶¹ de (5.190) :

$$\begin{bmatrix} (mc^2)^2 \mathbf{1} + \hbar^2 c^2 (\vec{k} \cdot \vec{\sigma})(\vec{k} \cdot \vec{\sigma}) & \mathbf{0} \\ \mathbf{0} & (mc^2)^2 \mathbf{1} + \hbar^2 c^2 (\vec{k} \cdot \vec{\sigma})(\vec{k} \cdot \vec{\sigma}) \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} = E^2 \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} . \quad (5.191)$$

Compte tenu de (5.55), $(\vec{k} \cdot \vec{\sigma})(\vec{k} \cdot \vec{\sigma})$ est égal à $\vec{k}^2 \mathbf{1}$. Il en résulte que le carré de l'énergie est :

$$E^2 = (mc^2)^2 + \hbar^2 \vec{k}^2 c^2 . \quad (5.192)$$

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 :

$$E^2 = (mc^2)^2 + (pc)^2 , \quad (5.193)$$

⁶¹Dans la suite, la dépendance en \vec{k} des C_μ est sous-entendue pour alléger les notations.

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

$$E = \pm \sqrt{(mc^2)^2 + (pc)^2} \equiv \pm E_p , \quad (5.194)$$

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 séparées⁶² l'une de l'autre par un *gap* égal à $2mc^2$; l'interprétation de ce fait sera donnée plus loin.

Il s'avère commode de paramétrer comme suit :

$$p = mc \sinh \phi , \quad E = \pm mc^2 \cosh \phi , \quad (5.195)$$

et alors la limite non relativiste $c \rightarrow +\infty$ 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 ECOC ; 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, (5.189) prend la forme explicite plus simple :

$$\begin{bmatrix} mc^2\mathbf{1} & pc\sigma_z \\ pc\sigma_z & -mc^2\mathbf{1} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} = E \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} . \quad (5.196)$$

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

$$\begin{cases} (mc^2 - E)C_1 + pcC_3 = 0 \\ pcC_1 - (mc^2 + E)C_3 = 0 \end{cases} , \quad \begin{cases} (mc^2 - E)C_2 - pcC_4 = 0 \\ -pcC_2 - (mc^2 + E)C_4 = 0 \end{cases} . \quad (5.197)$$

Ces deux systèmes ont la *même* équation caractéristique en E , dont les racines sont données par les expressions (5.194) ; cette identité n'est autre que la dégénérescence annoncée plus haut. Les valeurs propres étant déterminées, il reste à trouver les vecteurs (spineurs) propres. De toute évidence, $\vec{\Sigma} \cdot (\vec{p}/p) = \Sigma_z$ commute avec H_D , puisque (voir 5.176) :

$$[\Sigma_z, H_D] = \left[\begin{bmatrix} \sigma_z & \mathbf{0} \\ \mathbf{0} & \sigma_z \end{bmatrix}, \begin{bmatrix} mc^2\mathbf{1} & cp\sigma_z \\ cp\sigma_z & -mc^2\mathbf{1} \end{bmatrix} \right] = \begin{bmatrix} \mathbf{0} & c[\sigma_z, p\sigma_z] \\ c[\sigma_z, p\sigma_z] & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} . \quad (5.198)$$

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 :

$$\begin{bmatrix} \sigma_z & \mathbf{0} \\ \mathbf{0} & \sigma_z \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} = \sigma \begin{bmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{bmatrix} \quad (\sigma = \pm 1) . \quad (5.199)$$

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 (5.197), 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 (5.194)), les spineurs propres seront donc étiquetés, à impulsion $\hbar\vec{k}$ donnée, par le couple de nombres quantiques (ε, σ) et notés $\Psi_{\varepsilon, \sigma}$.

⁶²Pour une particule de masse nulle (le photon par exemple), on a simplement $E = \pm \hbar kc$ et le *gap* est nul.

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

$$\Psi_{+,+1} = \frac{1}{\sqrt{\cosh \phi}} \begin{bmatrix} \cosh \frac{\phi}{2} \\ 0 \\ \sinh \frac{\phi}{2} \\ 0 \end{bmatrix}, \quad \Psi_{+,-1} = \frac{1}{\sqrt{\cosh \phi}} \begin{bmatrix} 0 \\ \cosh \frac{\phi}{2} \\ 0 \\ -\sinh \frac{\phi}{2} \end{bmatrix}. \quad (5.200)$$

De même, pour les énergies négatives, on trouve :

$$\Psi_{-,+1} = \frac{1}{\sqrt{\cosh \phi}} \begin{bmatrix} -\sinh \frac{\phi}{2} \\ 0 \\ \cosh \frac{\phi}{2} \\ 0 \end{bmatrix}, \quad \Psi_{-,-1} = \frac{1}{\sqrt{\cosh \phi}} \begin{bmatrix} 0 \\ \sinh \frac{\phi}{2} \\ 0 \\ \cosh \frac{\phi}{2} \end{bmatrix}. \quad (5.201)$$

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 Σ avec le spin (au facteur $\hbar/2$ 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 :

$$\Psi_{+,+1} \simeq \begin{bmatrix} 1 \\ 0 \\ \frac{p}{2mc} \\ 0 \end{bmatrix}, \quad \Psi_{+,-1} \simeq \begin{bmatrix} 0 \\ 1 \\ 0 \\ \frac{-p}{2mc} \end{bmatrix} \quad (p \ll mc). \quad (5.202)$$

Dans cette limite, chaque état propre de (H_D, Σ_z) a une composante d'ordre zéro en 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 deux⁶³ :

$$\Psi_{+,+1} \rightarrow \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \Psi_{+,-1} \rightarrow \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (c \rightarrow +\infty). \quad (5.203)$$

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 :

$$\Psi_{-,+1} \simeq \begin{bmatrix} \frac{-p}{2mc} \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \Psi_{-,-1} \simeq \begin{bmatrix} 0 \\ \frac{p}{2mc} \\ 0 \\ 1 \end{bmatrix} \quad (p \ll mc). \quad (5.204)$$

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

$$\Psi_{-,+1} \rightarrow \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \Psi_{-,-1} \rightarrow \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (c = +\infty). \quad (5.205)$$

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

⁶³On 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.

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.

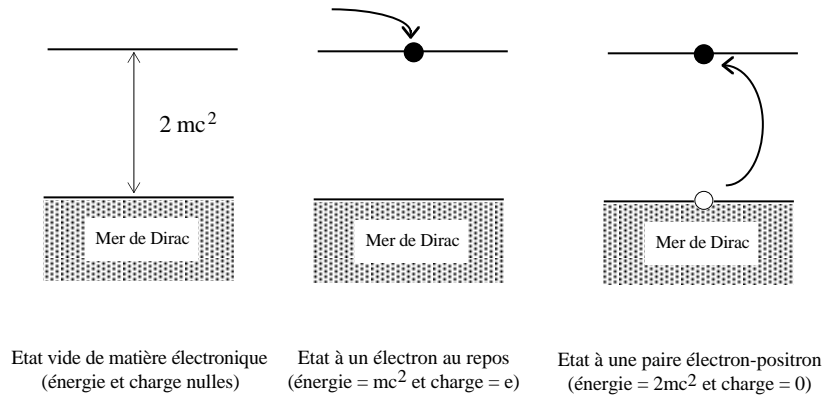


Figure 5.1: Illustration schématique de l'interprétation de Dirac.

Dirac postule l'existence d'un nombre (infini !) de particules occupant tous les états d'énergie négative⁶⁴ ; 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 zéro⁶⁵, 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 $mc^2 \leq \frac{mc^2}{\sqrt{1-(v/c)^2}}$ (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. 5.1).

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 positive⁶⁶ : ceci coûte donc une énergie au moins égale à $2mc^2$. Après cette opération, on se retrouve 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 : Hamiltonien de Pauli

Explicitons l'autre justification du fait que l'opérateur $\vec{\Sigma}$ (5.168) apporte le degré de spin, suivant la

⁶⁴Ceci 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).

⁶⁵Seules les *différences d'énergie* ont un sens physique.

⁶⁶Cette 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.

relation précise :

$$\vec{S} = \frac{\hbar}{2} \vec{\Sigma} , \quad (5.206)$$

en examinant en détail la limite faiblement relativiste en présence d'un champ électromagnétique $\vec{\mathcal{E}}, \vec{\mathcal{B}}$ dérivant du potentiel (\vec{A}, ϕ) . On repart du Hamiltonien de Dirac en y faisant la substitution habituelle :

$$E \longrightarrow E - e\phi , \quad \vec{p} \longrightarrow \vec{p} - e\vec{A} . \quad (5.207)$$

L'équation de Dirac (5.157) prend alors la forme :

$$[E - e\phi - c\vec{\alpha}(\vec{p} - e\vec{A}) - \beta mc^2] \Psi = 0 , \quad (5.208)$$

où E et \vec{p} sont les opérateurs :

$$E = i\hbar \frac{\partial}{\partial t} , \quad \vec{p} = -i\hbar \vec{\nabla} . \quad (5.209)$$

Multiplions maintenant (5.208) à gauche par $[E - e\phi + c\vec{\alpha}(\vec{p} - e\vec{A}) - \beta mc^2]$. Il vient :

$$\left[(E - e\phi)^2 - c^2 [\vec{\alpha}(\vec{p} - e\vec{A})]^2 - m^2 c^4 \right] \Psi = \left[c(E - e\phi)\vec{\alpha}(\vec{p} - e\vec{A}) - c\vec{\alpha}(\vec{p} - e\vec{A})(E - e\phi) \right] \Psi . \quad (5.210)$$

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

$$(\vec{V} \cdot \vec{\alpha})(\vec{W} \cdot \vec{\alpha}) = \vec{V} \cdot \vec{W} + i\vec{\Sigma} \cdot (\vec{V} \times \vec{W}) . \quad (5.211)$$

De la sorte, et prenant en compte la relation $\vec{\mathcal{B}} = \vec{\nabla} \times \vec{A}$, on voit que le premier membre de (5.210) est égal à :

$$\left[(E - e\phi)^2 - c^2 (\vec{p} - e\vec{A})^2 - m^2 c^4 + e\hbar c^2 \vec{\Sigma} \cdot \vec{\mathcal{B}} \right] \Psi . \quad (5.212)$$

Par ailleurs, des manipulations simples sur le second membre faisant usage de (5.209) et également de $\vec{\mathcal{E}} = -(\partial\vec{A})/(\partial t) - \vec{\nabla}\phi$ le transforment en $+ie\hbar c \vec{\alpha} \cdot \vec{\mathcal{E}} \Psi$. Finalement, l'équation de Dirac s'écrit :

$$\left[(E - e\phi)^2 - m^2 c^4 - c^2 (\vec{p} - e\vec{A})^2 + e\hbar c^2 \vec{\Sigma} \cdot \vec{\mathcal{B}} - ie\hbar c \vec{\alpha} \cdot \vec{\mathcal{E}} \right] \Psi = 0 . \quad (5.213)$$

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

$$E = E_{\text{fr}} + mc^2 , \quad (5.214)$$

avec l'hypothèse supplémentaire $E_{\text{fr}}, e\phi \ll mc^2$. Les deux premiers termes de (5.213) se simplifient en :

$$(E - e\phi)^2 - m^2 c^4 \simeq 2mc^2 (E_{\text{fr}} - e\phi) , \quad (5.215)$$

ce qui permet de réécrire (5.213) comme suit :

$$E_{\text{fr}} \Psi = \left[\frac{1}{2m} (\vec{p} - e\vec{A})^2 + e\phi - \frac{e\hbar}{2m} \vec{\Sigma} \cdot \vec{\mathcal{B}} + i \frac{e\hbar}{2mc} \vec{\alpha} \cdot \vec{\mathcal{E}} \right] \Psi . \quad (5.216)$$

Ψ est toujours un vecteur à 4 composantes. Comme il s'agit maintenant de la limite faiblement relativiste et compte tenu du développement en 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 v/c par rapport au premières), l'équation (5.216) devient à cet ordre :

$$E_{\text{fr}} \Psi_{\text{g}} = \left[\frac{1}{2m} (\vec{p} - e\vec{A})^2 + e\phi - \frac{e\hbar}{2m} \vec{\sigma} \cdot \vec{\mathcal{B}} \right] \Psi_{\text{g}} , \quad (5.217)$$

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 :

$$\vec{S} = \frac{\hbar}{2} \vec{\sigma} , \quad (5.218)$$

justifiant l'interprétation de $\vec{\Sigma}$. A nouveau, par comparaison avec (5.21), 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 (5.217) s'appelle le Hamiltonien de Pauli :

$$H_{\text{Pauli}} = \frac{1}{2m} (\vec{p} - e\vec{A})^2 + e\phi - \frac{e}{m} \vec{S} \cdot \vec{\mathcal{B}} , \quad (5.219)$$

compte tenu de (5.218). 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} , (5.5). 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 :

$$E \Psi = [c \vec{\alpha} \cdot \vec{p} + \beta mc^2 + V(\vec{r})] \Psi . \quad (5.220)$$

En désignant par Ψ_g et Ψ_p les couples de grandes et petites composantes, et en introduisant à nouveau E_{fr} comme en (5.214), l'équation (5.220) s'explique en deux équations couplées :

$$E_{\text{fr}} \Psi_g = c \vec{\sigma} \cdot \vec{p} \Psi_p + (\beta mc^2 + V) \Psi_g , \quad (5.221)$$

$$E_{\text{fr}} \Psi_p = c \vec{\sigma} \cdot \vec{p} \Psi_g + (-\beta mc^2 + V) \Psi_p . \quad (5.222)$$

Maintenant, par élimination de Ψ_p , on obtient :

$$(E_{\text{fr}} + 2mc^2) \Psi_g = \frac{1}{2m} \vec{\sigma} \cdot \vec{p} \frac{1}{1 + \frac{E_{\text{fr}} - V}{2mc^2}} \vec{\sigma} \cdot \vec{p} \Psi_g . \quad (5.223)$$

Le développement de la fraction au second membre et quelques manipulations algébriques⁶⁷ conduisent à :

$$E_{\text{fr}} \Psi_g = \left(\frac{\vec{p}^2}{2m} - \frac{p^4}{8m^3 c^2} + V - \frac{\hbar^2}{4m^2 c^2} \frac{dV}{dr} \frac{\partial}{\partial r} + \frac{1}{2m^2 c^2} \frac{1}{r} \frac{dV}{dr} \vec{L} \cdot \vec{S} \right) \Psi_g , \quad (5.225)$$

où $\vec{L} = \vec{r} \times \vec{p}$ et $\vec{S} = (\hbar/2)\vec{\sigma}$. Les deuxième et quatrième termes sont des corrections relativistes aux énergies cinétique et potentielle (celui-ci sans équivalent classique, alors que le terme en p^4 est une simple "correction de masse"). Le dernier terme donne, comme souhaité, l'expression convenable de V_{magn} .

⁶⁷Utiliser notamment :

$$(\vec{\sigma} \cdot \vec{\nabla} V) (\vec{\sigma} \cdot \vec{p}) = (\vec{\nabla} V) \vec{p} + i \vec{\sigma} \cdot [\vec{\nabla} V \times \vec{p}] , \quad (5.224)$$

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

Chapitre 6

Méthodes d'approximation pour les états propres

Le but de ce chapitre est de présenter deux méthodes d'approximation d'usage courant pour les états propres liés d'un Hamiltonien que l'on ne sait pas traiter exactement : la méthode variationnelle et la théorie des perturbations stationnaires.

6.1 Méthode variationnelle

6.1.1 Formulation variationnelle de l'équation aux valeurs et vecteurs propres

L'équation aux valeurs et vecteurs propres :

$$H \psi(\vec{r}) = E \psi(\vec{r}) , \quad (6.1)$$

apparaît après la séparation espace - temps dans l'équation de Schrödinger, $\Psi(\vec{r}, t) = e^{\frac{i}{\hbar}Et} \psi(\vec{r})$, et peut être considérée comme l'équation d'Euler-Lagrange d'un problème variationnel¹. En effet, soit à trouver la fonction $\psi(\vec{r})$ satisfaisant la condition de normalisation :

$$\int_{\mathbb{R}^3} d^3r |\psi(\vec{r})|^2 = 1 \quad (6.2)$$

et rendant stationnaire l'expression suivante :

$$\int_{\mathbb{R}^3} d^3r \left[-\frac{\hbar^2}{2m} \psi^*(\vec{r}) \Delta \psi(\vec{r}) + \psi^*(\vec{r}) V(\vec{r}) \psi(\vec{r}) \right] . \quad (6.3)$$

Ceci n'est autre que la valeur moyenne du Hamiltonien H :

$$H = \frac{\vec{p}^2}{2m} + V(\vec{r}) \quad (6.4)$$

dans l'état normalisé ψ , que l'on peut noter $\langle H \rangle$ ou, plus conformément à l'usage dans un contexte variationnel, $E[\psi]$; les crochets rappellent qu'il s'agit d'une *fonctionnelle* de la fonction ψ : E est un *nombre* (réel) attaché à une famille de *fonctions*.

¹Historiquement, c'est d'ailleurs sous la forme variationnelle que Schrödinger a d'abord introduit son équation [1] dans le premier [36] de ses quatre articles fondateurs de la Mécanique Quantique.

Avec les conditions habituelles (annulation de ψ aux bornes de l'intervalle, ce qui est toujours réalisé pour les états liés), une intégration par parties transforme l'intégrale (6.3) en² :

$$E[\psi] = \int_{\mathbb{R}^3} d^3r \left[+\frac{\hbar^2}{2m} \vec{\nabla}\psi^*(\vec{r}) \cdot \vec{\nabla}\psi(\vec{r}) + \psi^*(\vec{r})V(\vec{r})\psi(\vec{r}) \right] . \quad (6.6)$$

On va montrer que la fonctionnelle $E[\psi]$ est extrémale pour les états propres du Hamiltonien ; pour la simplicité de l'écriture, la démonstration est faite à une dimension d'espace.

Soit donc à rendre extrémale la fonctionnelle :

$$E[\psi] = \int_{-\infty}^{+\infty} dx \left[+\frac{\hbar^2}{2m} \psi'^*(x) \psi'(x) + \psi^*(x)V(x)\psi(x) \right] \quad (6.7)$$

avec la condition :

$$\int_{-\infty}^{+\infty} dx \psi^*(x)\psi(x) = 1 . \quad (6.8)$$

Comme on le sait, la variation d'une fonctionnelle avec contrainte(s) se fait commodément en introduisant un (des) multiplicateur(s) de Lagrange – qui est (sont) autant d'inconnue(s) supplémentaire(s) à déterminer – mais qui permettent un calcul plus souple en général et surtout plus symétrique. Il y a ici une seule contrainte – c'est l'équation (6.8) – donc un seul multiplicateur, noté λ . La fonctionnelle à rendre extrémale est alors :

$$F[\psi] = \int_{-\infty}^{+\infty} dx \left[+\frac{\hbar^2}{2m} \psi'^*(x) \psi'(x) + \psi^*(x)V(x)\psi(x) - \lambda\psi^*(x)\psi(x) \right] \equiv \int_{-\infty}^{+\infty} dx \Phi(\psi, \psi', \psi^*, \psi'^*) . \quad (6.9)$$

Oublions provisoirement que les fonctions ψ et ψ^* sont complexes conjuguées l'une de l'autre – ce qui semble rendre leurs variations non-indépendantes. En considérant ψ et ψ^* comme indépendantes, on peut écrire deux équations d'Euler-Lagrange³ pour $F[\psi] \equiv \int dx \Phi$:

$$\frac{d}{dx} \frac{\partial \Phi}{\partial \psi'^*} - \frac{\partial \Phi}{\partial \psi^*} = 0 , \quad \frac{d}{dx} \frac{\partial \Phi}{\partial \psi'} - \frac{\partial \Phi}{\partial \psi} = 0 , \quad (6.10)$$

qui s'explicitent comme suit :

$$\frac{d}{dx} \frac{\hbar^2}{2m} \psi' - V\psi + \lambda\psi = 0 , \quad \frac{d}{dx} \frac{\hbar^2}{2m} \psi'^* - V\psi^* + \lambda\psi^* = 0 , \quad (6.11)$$

soit :

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V\psi = \lambda\psi , \quad -\frac{\hbar^2}{2m} \frac{d^2\psi^*}{dx^2} + V\psi^* = \lambda\psi^* . \quad (6.12)$$

Ces deux équations montrent que le multiplicateur λ n'est autre que l'énergie, qui est réelle. De surcroît, elles se déduisent l'une de l'autre par conjugaison puisque V est réel⁴ ; tout est donc cohérent et il apparaît après coup qu'il était parfaitement licite de considérer les variations de ψ et de ψ^* comme indépendantes⁵. Finalement, la condition d'extremum pour $F[\psi]$ est :

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V\psi = \lambda\psi . \quad (6.13)$$

Ceci n'est autre que l'équation aux valeurs propres (6.1).

²L'expression (6.6) montre que la valeur moyenne de l'énergie cinétique est :

$$E_{\text{cin}}[\psi] = \frac{\hbar^2}{2m} \int_{\mathbb{R}^3} d^3r |\vec{\nabla}\psi(\vec{r})|^2 \quad (6.5)$$

et il saute aux yeux que cette quantité est positive, comme il se doit.

³Les équations (6.10) sont structurellement identiques aux équations de Lagrange, où ψ est remplacé par une coordonnée q et x est le temps ; la fonctionnelle minimisée est alors l'action S et Φ est le Lagrangien L .

⁴ V réel assure que l'équation de conservation de la probabilité est satisfaite, que H est hermitique, que l'évolution est unitaire, etc.

⁵Ceci tient au fait que la variation d'une fonction à valeurs complexes implique la variation de deux fonctions (la partie réelle et la partie imaginaire) ; l'annulation de la variation de la fonctionnelle conduit donc de fait à deux équations réelles.

Remarque

Le calcul a été fait dans le cas où H est de la forme (6.4). Tout ceci se généralise aux cas plus complexes, par exemple lorsqu'un champ magnétique est présent. En outre, la généralisation au cas tridimensionnel est un pur jeu d'écriture.

L'équivalence :

$$\delta E[\psi] = 0 \iff H\psi = E\psi . \quad (6.14)$$

est un résultat majeur, qui peut aussi s'obtenir de façon plus formelle ([20], p. 654), en utilisant la notation de Dirac. Soit la fonctionnelle définie sur l'espace des états :

$$E[|\psi\rangle] = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} . \quad (6.15)$$

On va montrer que tout $|\psi\rangle$ rendant E stationnaire est vecteur propre du spectre discret de H – et réciproquement –, étant entendu que tout état discret est normalisable, assurant que le dénominateur dans (6.15) existe et est fini ; la valeur propre associée n'est autre que la valeur de la fonctionnelle $E[|\psi\rangle]$ pour l'état propre en question. Afin que la définition de E (6.15) ait un sens, seuls des vecteurs de norme *finie* sont considérés dans la suite.

Quand $|\psi\rangle$ varie de $|\delta\psi\rangle$, $E[|\psi\rangle]$ varie de δE :

$$\delta E = E[|\psi\rangle + |\delta\psi\rangle] - E[|\psi\rangle] ; \quad (6.16)$$

L'expression de δE s'obtient comme quand on calcule explicitement la variation d'une fonction ordinaire $f(x)$ quand son argument varie de dx . En calculant sur l'expression de la dérivée d'une fraction, on voit que la variation de $E[|\psi\rangle]$ s'écrit :

$$\delta E[|\psi\rangle] = \frac{\langle\delta\psi|H|\psi\rangle + \langle\psi|H|\delta\psi\rangle}{\langle\psi|\psi\rangle} - \langle\psi|H|\psi\rangle \frac{\langle\delta\psi|\psi\rangle + \langle\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle^2} + \mathcal{O}(\delta\psi^2) , \quad (6.17)$$

soit :

$$\begin{aligned} \langle\psi|\psi\rangle \delta E[|\psi\rangle] &= \langle\delta\psi|H|\psi\rangle + \langle\psi|H|\delta\psi\rangle - E[|\psi\rangle] (\langle\delta\psi|\psi\rangle + \langle\psi|\delta\psi\rangle) + \mathcal{O}(\delta\psi^2) \\ &= \langle\delta\psi|(H - E)|\psi\rangle + \langle\psi|(H - E)|\delta\psi\rangle + \mathcal{O}(\delta\psi^2) . \end{aligned} \quad (6.18)$$

Tant que $\langle\psi|\psi\rangle$ est non nul et fini, la condition d'extremum, $\delta E = 0$, s'écrit en annulant le facteur de l'incrément linéaire en $\delta\psi$:

$$\langle\delta\psi|(H - E)|\psi\rangle + \langle\psi|(H - E)|\delta\psi\rangle = 0 . \quad (6.19)$$

À nouveau se pose la question de l'indépendance des deux variations $\langle\delta\psi|$ et $|\delta\psi\rangle$; l'argument antérieur vaut toujours : puisque la théorie s'appuie sur le corps des complexes, il est licite de les considérer comme indépendantes⁶. De fait, remplaçons $|\delta\psi\rangle$ par $i|\delta\psi\rangle$; par l'antilinearité du produit scalaire, la condition (6.19) devient :

$$-i \langle\delta\psi|(H - E)|\psi\rangle + i \langle\psi|(H - E)|\delta\psi\rangle = 0 . \quad (6.21)$$

La combinaison des deux dernières équations montre bien que la condition d'extremum, quelles que soient les deux variations $|\delta\psi\rangle$ et $\langle\delta\psi|$, s'exprime par les deux équations :

$$(H - E)|\psi\rangle = 0 , \quad \langle\psi|(H - E) = 0 \quad (6.22)$$

mais comme H est hermitique, l'une d'entre elles suffit.

⁶Pour se convaincre autrement ([4], p. 1138) de l'indépendance de fait des deux variations, on peut aussi poser $|\delta\psi\rangle = \epsilon|\phi\rangle$ où $|\phi\rangle$ est donné et où ϵ est un infiniment petit réel ; alors $\langle\delta\psi| = \epsilon\langle\phi|$ et la condition devient :

$$2 \langle\delta\phi|(H - E)|\psi\rangle = 0 . \quad (6.20)$$

Réciproquement, soit $|\psi\rangle$ un état propre de H associé à l'énergie E . Par définition on a $H|\psi\rangle = E|\psi\rangle$, d'où $(H - E)|\psi\rangle = 0$.

En résumé, tant qu'il s'agit d'états normalisables, il y a équivalence entre la propriété d'être un état propre et celle de rendre $E[|\psi\rangle]$ extrémale. Cette affirmation est parfois appelée "Théorème de Ritz" [4].

Ce théorème est très utilement complété par le corollaire suivant ; pour un système donné, quel que soit l'état $|\phi\rangle$, on a :

$$E[\phi] \geq E_{\text{fond}} \equiv E_1 . \quad (6.23)$$

Ceci signifie que, pour tout Hamiltonien, sa valeur moyenne avec n'importe quelle fonction $|\phi\rangle$ normalisable est toujours plus grande (ou égale) que l'énergie de l'état fondamental exact. Pour établir ce résultat, calculons la différence $E[\phi] - E_{\text{fond}}$ sur la base propre du Hamiltonien du système considéré⁷ :

$$\begin{aligned} E[\phi] - E_{\text{fond}} &= \frac{1}{\langle\phi|\phi\rangle} [\langle\phi|H|\phi\rangle - E_1 \langle\phi|\phi\rangle] = \frac{1}{\langle\phi|\phi\rangle} \langle\phi| \left[\sum_n (E_n |E_n\rangle \langle E_n|) - E_1 \right] |\phi\rangle \\ &= \frac{1}{\langle\phi|\phi\rangle} \langle\phi| \left[\sum_n (E_n - E_1) |E_n\rangle \langle E_n| \right] |\phi\rangle . \end{aligned} \quad (6.24)$$

Le préfacteur est positif (c'est l'inverse du carré d'une norme), ainsi que chaque terme de la série⁸, d'où l'inégalité annoncée. On peut aussi raisonner comme suit ; soit le développement :

$$|\phi\rangle = \sum_n c_n |E_n\rangle . \quad (6.25)$$

Alors :

$$E[\phi] = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2} \geq \frac{\sum_n |c_n|^2 E_1}{\sum_n |c_n|^2} \equiv E_1 . \quad (6.26)$$

L'inégalité devient une égalité dans le seul cas où tous les c_n sont nuls, sauf c_1 qui vaut 1, auquel cas $|\phi\rangle$ n'est autre que l'état exact $|E_1\rangle$.

6.1.2 Calcul variationnel des états discrets (états liés)

L'usage de la méthode variationnelle et la résolution directe de l'équation aux valeurs propres sont donc deux démarches totalement équivalentes. Il en résulte que leur achèvement exact est d'une égale difficulté et, d'une façon générale, quand on ne sait pas trouver les solutions exactes d'une façon, on ne sait pas plus les obtenir par l'autre.

En pratique, la méthode variationnelle est cependant utile pour obtenir des valeurs propres et des fonctions propres approchées. En effet, en vertu du corollaire (6.23), on peut adopter la méthodologie suivante. Une famille de fonctions étant donnée (ce n'est pas forcément un espace vectoriel – elles peuvent dépendre non-linéairement d'un ou plusieurs paramètres), on peut espérer obtenir une approximation convenable de l'état fondamental en choisissant dans cette famille la fonction $|\psi_{\text{opt}}\rangle$ qui donne sa plus *petite* valeur à la fonctionnelle $E[|\psi\rangle]$. Par définition, de toutes les fonctions de cette famille, $|\psi_{\text{opt}}\rangle$ sera celle qui donne l'énergie la plus proche de E_{fond} . Ce critère vaut ce qu'il vaut – voir plus loin.

À titre d'illustration, prenons un exemple, académique certes puisqu'il traite un cas où l'on connaît la solution exacte, mais qui a ses vertus pédagogiques puisqu'il permet précisément de faire une comparaison systématique entre les fonctions d'onde exacte et approchée. Soit à trouver par ce type de variation la fonction d'onde radiale fondamentale de l'atome d'hydrogène. On pose :

$$R(r) = a_0^{-3/2} f(\rho) \quad (\rho = r/a_0) \quad (6.27)$$

⁷La démonstration suppose pour simplifier que tous les états discrets forment une base complète.

⁸L'opérateur entre crochets dans (6.24) est un opérateur défini positif ; sa valeur moyenne dans n'importe quel état est un nombre positif.

et il s'agit de trouver la meilleure fonction $f(\rho)$ dans une famille à préciser par la suite. La fonctionnelle E est :

$$E[R] = \frac{1}{\int_0^{+\infty} dr r^2 R^* R} \int_0^{+\infty} dr r^2 R^* \left[\frac{-\hbar^2}{2m} \left(R'' + \frac{2}{r} R' \right) - \frac{e'^2}{r} \right] , \quad (6.28)$$

et s'écrit à l'aide de f comme suit :

$$E[f] = E_1 \frac{\int_0^{+\infty} d\rho f^* [\rho^2 f'' + 2\rho f' + 2\rho f]}{\int_0^{+\infty} d\rho \rho^2 f^* f} \quad (6.29)$$

où E_1 est l'énergie du fondamental de l'hydrogène :

$$E_1 = -\frac{me'^4}{2\hbar^2} = -\frac{e'^2}{2a_0} \simeq -13.6 \text{ eV} , \quad a_0 = \frac{\hbar^2}{me'^2} \quad (6.30)$$

Choisissons maintenant des familles de fonctions, étant entendu que le choix est illimité ; pour l'exemple, considérons⁹ :

$$f_a(\rho) = e^{-a\rho} , \quad f_b(\rho) = \frac{1}{\rho^2 + b^2} \quad (6.31)$$

où a et b ont le sens de longueurs adimensionnées (longueurs en unités a_0), constituant pour chaque fonction le paramètre variationnel. Quand ce paramètre varie, la fonction correspondante décrit une famille de fonctions qui peut toujours être dotée d'une structure d'espace vectoriel, mais dont la famille ne saurait visiblement constituer ici une base.

$E[f_a]$ et $E[f_b]$ seront donc ici de simples fonctions d'un paramètre variationnel, a ou b . Le calcul est facile et on trouve :

$$E[f_a] = (2a - a^2) E_1 \equiv \varepsilon_a(a) , \quad E[f_b] = \frac{8b - \pi}{2\pi b^2} E_1 \equiv \varepsilon_b(b) . \quad (6.32)$$

La minimisation donne :

$$a_{\min} = 1, \quad \varepsilon_a(a_{\min}) = E_1 , \quad b_{\min} = \frac{\pi}{4}, \quad \varepsilon_b(b_{\min}) = \frac{8}{\pi^2} E_1 \simeq 0.81 E_1 . \quad (6.33)$$

Comme $E_1 < 0$, on a bien $\varepsilon_b(b_{\min}) > E_1$.

En ce qui concerne f_a , on voit que la fonction optimisée $f_{a, \text{opt}}$ donne la valeur exacte de l'énergie, c'est-à-dire que cette fonction *est* la fonction exacte. De fait, la famille choisie contenait la fonction exacte $e^{-\rho}$ et la méthode des variations est allée la "pêcher" ; le but de l'exercice était de montrer ce phénomène.

En revanche, la fonction $f_{b, \text{opt}}$, si elle fournit une valeur à peu près convenable de l'énergie (avec évidemment $E_{\text{opt}} > E_1$ puisque $E_1 < 0$), est loin d'être satisfaisante et représente très mal la vraie fonction : notamment, elle décroît (à grande distance) très lentement, comme r^{-2} , alors que la décroissance exacte est exponentielle. Ceci a des conséquences assez catastrophiques : la valeur moyenne exacte de r , distance au noyau, vaut $1.5a_0$ (voir chapitre 4), alors que toutes les fonctions f_b , (l'optimisée comme les autres), donnent ... $\langle r \rangle = +\infty$ (à l'infini, l'intégrand donnant $\langle r \rangle$ calculé avec f_b varie comme $r^2 r^{-4} r = r^{-1}$: la divergence de $\langle r \rangle$ est logarithmique).

Il est clair sur ce dernier exemple que la valeur de l'énergie ne saurait définir un critère sûr de qualité pour la fonction d'onde ainsi trouvée. Ceci n'est pas surprenant : la valeur de l'énergie variationnelle résulte d'une intégration où des erreurs grossières de comportement de la fonction d'onde peuvent avoir le bon goût de se compenser, même si ce n'est pas toujours le cas, donnant finalement pour l'énergie une valeur acceptable en dépit de la médiocrité locale de la fonction.

Une façon plus sûre d'apprécier la qualité, bonne ou mauvaise, d'une fonction variationnelle repose sur l'évaluation du produit scalaire S :

$$S = \int_{\mathbb{R}^3} d^3r \psi_{\text{opt}}^* \psi_1 . \quad (6.34)$$

⁹Ces fonctions ne sont pas normalisées.

Visiblement, plus $|S|$ est proche de 1, meilleure est la fonction ψ_{opt} . Le malheur est qu'un tel calcul n'est possible ... que si l'on connaît la fonction exacte ψ_1 ! Cependant, le calcul de S dans quelques cas d'école permet de réaliser qu'il n'y a pas de corrélation directe entre l'erreur sur l'énergie et l'erreur sur la fonction d'onde telle que S permet de la définir.

La méthode variationnelle est une méthode puissante et relativement efficace, mais il est le plus souvent difficile d'évaluer la qualité des résultats qu'elle produit. On doit se laisser guider, lors du choix de la famille d'essai, par des considérations fondées sur des arguments généraux ou sur des propriétés dont on sait que la fonction exacte inconnue doit les posséder : nombre de nœuds, symétrie, comportement à l'origine et comportement asymptotique, *etc.* ; d'une façon générale, des considérations physiques sont toujours fort utiles.

Compte tenu de l'inégalité (6.23), la méthode variationnelle convient surtout à la description de l'état fondamental ; en effet, on ne dispose *a priori* d'aucune relation d'ordre analogue pour les états excités et la minimisation peut faire "plonger" l'énergie, loin de l'énergie d'un état excité donné, que l'on veut décrire.

D'un autre côté, dans le cas exceptionnel où l'on connaît le fondamental *exact* $|E_1\rangle$, on peut tout redémontrer en se plaçant d'emblée dans l'espace orthogonal à ce vecteur ; en particulier :

$$E[|\phi\rangle] \geq E_{1^{\text{er}} \text{ état excité}} \quad \forall |\phi\rangle \perp |E_1\rangle, \quad (|E_1\rangle = \text{état fondamental exact}) . \quad (6.35)$$

La méthode variationnelle joue également un grand rôle pour la compréhension du comportement des assemblées de particules en interaction (Problème à N corps). L'application la plus fameuse est probablement celle qui conduit aux équations de Hartree - Fock (voir par exemple [20] p. 662, [17] p. 535), qui joue un rôle de premier plan en Physique (nucléaire, atomique, moléculaire et du solide). Cette méthode donne une description en terme de champ moyen – chaque particule se mouvant dans le champ moyen des autres –, ce champ étant déterminé de façon auto-cohérente (*self-consistent field*) par une méthode itérative.

6.2 Théorie des perturbations stationnaires

La théorie des perturbations stationnaires est une autre méthode de résolution approchée de l'équation aux valeurs et vecteurs propres :

$$H|\psi\rangle = E|\psi\rangle . \quad (6.36)$$

Cette approche peut être envisagée à chaque fois que le Hamiltonien peut être décomposé en deux morceaux comme suit :

$$H = H_0 + gV \quad (6.37)$$

où H_0 est un Hamiltonien dont les solutions exactes sont connues, appelé Hamiltonien non-perturbé. V est un opérateur hermitique (ici indépendant du temps) homogène à une énergie ; g est un nombre réel sans dimension introduit pour la commodité et constituant également une mesure du terme additionnel gV , appelé perturbation¹⁰. Au total, pour que la méthode fonctionne bien, il faut que gV soit "petit", dans un sens à préciser autant que faire se peut¹¹. Par exemple, H_0 est le Hamiltonien d'un atome isolé et gV décrit l'interaction de celui-ci avec un champ extérieur, ici supposé statique, de petite intensité ; en pareil cas, g contiendra typiquement le module du champ extérieur – plus précisément, g sera le rapport entre l'énergie typique de couplage avec le champ et une énergie relative au système non-perturbé. Pour un Hamiltonien de la forme $H = H_0 + V$, il suffit de faire $g = 1$ dans toutes les formules ci-dessous.

La théorie des perturbations joue un rôle majeur en Physique Quantique : compte tenu de la relative complexité du formalisme, la classe de problèmes que l'on sait résoudre exactement est très réduite. Bien sûr, une telle situation n'est cependant pas propre au domaine quantique : on ne sait pas, par exemple, résoudre

¹⁰ g est aussi là pour compter facilement les puissances. À la fin, on fera formellement $g = 1$ pour se recaler sur l'écriture usuelle : il faudra alors comprendre que se tient caché dans V un petit paramètre multiplicatif qui autorise (sans le justifier complètement) l'usage d'une théorie de perturbation.

¹¹La "petitesse" est visiblement une condition *nécessaire* ; formellement, elle n'est pas en général suffisante – voir les remarques ultérieures à ce sujet.

exactement le problème des trois corps en interaction gravitationnelle – pour lequel ont d’ailleurs été développées des méthodes de type perturbatif. En outre, les méthodes de perturbation s’inscrivent très naturellement dans la démarche des physiciens : après avoir “dégrossi” un problème donné en retenant les phénomènes dominants (représentés par H_0), on affine la description en incluant de “petits” effets qui, malgré leur petitesse, peuvent jouer un rôle important, parfois spectaculaire (éclatement des raies spectrales, réduction de la durée de vie d’un niveau excité par contamination, *etc.*). D’une façon générale, la Théorie des perturbations joue un rôle de premier plan dans le problème à N -corps, pour lequel il a d’ailleurs fallu mettre au point des méthodes plus raffinées pour rendre compte d’effets collectifs¹².

L’idée fondamentale sur laquelle repose la méthode des perturbations est qu’il existe un développement en puissances *entières* de la constante de coupure g pour les modes propres de H . Dans toute la suite, on suppose donc que tout état propre $|\psi\rangle$ de H et la valeur propre E associée admettent des développements du genre :

$$|\psi\rangle = |\psi^{(0)}\rangle + g|\phi_1\rangle + g^2|\phi_2\rangle + \dots, \quad (6.38)$$

$$E = E^{(0)} + g\varepsilon_1 + g^2\varepsilon_2 + \dots. \quad (6.39)$$

Les $\{|\phi_n\rangle\}$ et les $\{\varepsilon_n\}$, $n \geq 1$ constituent les corrections successives à l’ordre n pour le vecteur et l’énergie propres. Une telle hypothèse est loin d’être évidente : l’obligation pour tout mode propre de H de se confondre avec ceux de H_0 pour $g = 0$ peut être satisfaite de bien d’autres manières. Par exemple, l’énergie E pourrait avoir un développement du genre :

$$E = E^{(0)}(1 + g \ln g + \dots), \quad (6.40)$$

qui se réduit bien à $E^{(0)}$ pour $g = 0$ mais qui n’est pas pour autant développable en série entière au voisinage de $g = 0$: si on essaie formellement un développement de Taylor, on voit que chaque coefficient du développement est infini... Autre exemple ; soit :

$$E = E^{(0)}(1 + e^{-1/g} + \dots). \quad (6.41)$$

Une telle expression a un développement en série entière ... dont tous les coefficients sont cette fois nuls, traduisant le fait que $e^{-1/g}$ tend vers zéro plus vite que toute puissance finie de g . De tels développements sont loin d’être rares et surviennent, par définition, pour les perturbations dites singulières, qui exigent la mise en œuvre d’une méthode particulière, éventuellement spécifique au cas traité ; il sera clair que la méthode exposée ci-dessous est totalement inopérante en pareille situation. Dans le cas contraire où les développements en puissances (6.38) et (6.39) existent, la perturbation est dite régulière ; seules les perturbations de ce dernier type sont considérées dans la suite.

En outre, même quand un développement en puissances entières existe ou semble exister – au sens où les premiers coefficients sont finis – il conviendrait d’examiner la convergence de la série qui s’amorce¹³. On découvrirait alors qu’il s’agit souvent en fait d’un développement *asymptotique*, fort utile, mais qu’il convient de manier avec précaution¹⁴. Toutes ces subtilités seront écartées dans la suite, au nom du pragmatisme, mais il

¹²Essentiellement, il s’agit de méthodes qui mettent le doigt sur les termes les plus importants et qui, moyennant des resommations à tous les ordres (à l’aide de la résolvante par exemple), dépassent et de loin la technique élémentaire exposée en détail dans ce chapitre.

¹³Cette question est, dans la quasi-totalité des cas, totalement académique. Examiner la convergence suppose que l’on connaît l’expression générale du terme d’ordre n , ce qui n’est vérifié en pratique que pour des cas presque triviaux.

¹⁴Un exemple aide à comprendre ce point. Soit la fonction $I(g)$ définie, pour $g > 0$, par :

$$I(g) = \int_0^{+\infty} dx \frac{e^{-x/g}}{1+x}. \quad (6.42)$$

De toute évidence, $\lim_{g \rightarrow 0^+} I(g) = 0$. L’exponentielle agit comme une fonction de coupure et l’intégrand est quasiment nul dès que x est à peine supérieur à g ; donc quand $0 < g \ll 1$, seul le voisinage de l’origine contribue notablement à l’intégrale. On peut alors tenter une opération illicite, consistant à développer $(1+x)^{-1}$ en série entière, bien que ce développement ne converge que pour $|x| < 1$ et alors que l’intervalle d’intégration va jusqu’à $+\infty$. En poussant l’audace jusqu’à échanger sommation et intégration, on obtient :

$$I(g) = \sum_{n=0}^{+\infty} (-1)^n \int_0^{+\infty} dx x^n e^{-x/g} = \sum_{n=0}^{+\infty} (-1)^n n! g^{n+1}. \quad (6.43)$$

La série de droite est visiblement divergente (mais vaut bien zéro en $g = 0$), son rayon de convergence est nul. Toutefois, pour $0 < g \ll 1$, le développement (6.43) judicieusement tronqué à ses N premiers termes (N est d’une certaine façon fonction de la valeur de g !) donne une approximation remarquable de la valeur prise par la fonction $I(g)$.

L’absence d’un développement en série entière pour $I(g)$ est d’ailleurs évident quand on observe que $I(0) = 0$ et que si on fait

convient de les avoir présentes à l'esprit. Quoi qu'il en soit, le physicien est supposé avoir assez de flair pour se convaincre d'avance de la validité de la méthode qu'il choisit au cas par cas (voir à titre d'exemple la remarque p. 184 ci-dessous concernant l'effet Stark).

Dans le but de sérier les difficultés, il est préférable de traiter séparément le cas de la perturbation d'un niveau non-dégénéré et celui où existe une dégénérescence. De surcroît, afin de s'en tenir à une théorie élémentaire, on suppose le spectre de H_0 entièrement discret ; il en résulte que les $\{|\psi_n^{(0)}\rangle\}$ forment une base complète¹⁵.

6.2.1 Cas d'un niveau non-dégénéré

On suppose donc connues toutes les solutions propres de H_0 :

$$H_0|\psi_m^{(0)}\rangle = E_m^{(0)}|\psi_m^{(0)}\rangle . \quad (6.44)$$

Les $|\psi_n^{(0)}\rangle$ sont supposés normalisés à l'unité, et on se propose de calculer l'effet de gV sur un état propre donné, choisi une fois pour toutes, soit $|\psi_n^{(0)}\rangle$; par hypothèse, l'énergie $E_n^{(0)}$ est ici non-dégénérée. Il s'agit alors de résoudre :

$$(H_0 + gV)|\psi_n\rangle = E_n|\psi_n\rangle \quad (6.45)$$

où E_n et $|\psi_n\rangle$ notent les énergie et vecteurs propres qui tendent vers $E_n^{(0)}$ et $|\psi_n^{(0)}\rangle$ lorsque g tend vers zéro. On admet donc en outre l'existence des développements :

$$|\psi_n\rangle = |\psi_n^{(0)}\rangle + g|\phi_{n1}\rangle + g^2|\phi_{n2}\rangle + \dots \equiv |\psi_n^{(0)}\rangle + |\phi_n\rangle , \quad (6.46)$$

$$E_n = E_n^{(0)} + g\varepsilon_{n1} + g^2\varepsilon_{n2} + \dots \equiv E_n^{(0)} + \varepsilon_n . \quad (6.47)$$

En reportant ces développements dans (6.45), et en identifiant les termes de même puissance de g , on obtient la suite d'équations :

$$H_0|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(0)}\rangle , \quad (6.48)$$

$$H_0|\phi_{n1}\rangle + V|\psi_n^{(0)}\rangle = E_n^{(0)}|\phi_{n1}\rangle + \varepsilon_{n1}|\psi_n^{(0)}\rangle , \quad (6.49)$$

$$H_0|\phi_{n2}\rangle + V|\phi_{n1}\rangle = E_n^{(0)}|\phi_{n2}\rangle + \varepsilon_{n1}|\phi_{n1}\rangle + \varepsilon_{n2}|\psi_n^{(0)}\rangle , \quad (6.50)$$

et ainsi de suite. Ces équations se réécrivent¹⁶ :

$$(H_0 - E_n^{(0)})|\psi_n^{(0)}\rangle = 0 , \quad (6.51)$$

$$(H_0 - E_n^{(0)})|\phi_{n1}\rangle = (\varepsilon_{n1} - V)|\psi_n^{(0)}\rangle , \quad (6.52)$$

$$(H_0 - E_n^{(0)})|\phi_{n2}\rangle = (\varepsilon_{n1} - V)|\phi_{n1}\rangle + \varepsilon_{n2}|\psi_n^{(0)}\rangle , \dots \quad (6.53)$$

La première équation n'apprend rien. Toutes les équations suivantes ont la particularité d'avoir au premier membre la correction d'ordre $k + 1$ à l'état propre, exprimée en termes des corrections d'ordre $\leq k$, présentes au second membre. Par rapport à l'inconnue $|\phi_{nk+1}\rangle$, chaque équation est *inhomogène*. Cette série d'équations permet, en principe, de calculer de proche en proche toutes les corrections les unes après les autres ; ainsi, après avoir trouvé ε_{n1} , (6.52) permettra d'obtenir la 1^{ère} correction à l'état propre $|\phi_{n1}\rangle$, laquelle permettra, par l'équation suivante de trouver la correction d'énergie ε_{n2} , et ainsi de suite.

brutalement $g = 0-$ dans (6.42), on trouve $I(-0) = +\infty \dots$: $g = 0$ est donc un point singulier de la fonction I – que l'on connaît d'ailleurs en tant que fonction spéciale bien répertoriée (I s'exprime à l'aide de $\text{Ei}(x)$, fonction appelée exponentielle intégrale). Cette fonction a un comportement logarithmique près de $g = 0$ et on peut la prolonger dans le demi-plan complexe $\Re g < 0$, coupé du demi-axe réel négatif (ce prolongement est possible : il suffit d'incliner continûment la demi-droite d'intégration en rapport avec l'argument de g de sorte qu'à l'infini l'argument du produit gx soit positif ; l'origine est un point de singularité en ce sens qu'il s'agit d'un point de branchement.

¹⁵Pour tout dire, cette restriction est de pure forme. En pratique, cela n'empêchera nullement d'utiliser la théorie des perturbations dans des cas où l'on sait pertinemment que cette hypothèse n'est pas vérifiée. Par exemple, on appliquera cette théorie pour donner une description de l'effet Stark sur l'hydrogène... dont le spectre n'est pourtant pas entièrement discret. Ceci peut se justifier en présence d'un champ faible. Comme toujours, les subtilités mathématiques pèsent peu devant des arguments physiques fondés.

¹⁶Pour simplifier les écritures, on omet l'opérateur identité $\mathbf{1}$ en facteur des scalaires : $H_0 - E_n^{(0)} \equiv H_0 - E_n^{(0)}\mathbf{1}$, $\varepsilon_{n1} - V \equiv \varepsilon_{n1}\mathbf{1} - V$, etc.

À ce stade, une remarque permet de simplifier un peu l'algèbre. Supposons le problème résolu ; on va notamment trouver les $|\phi_{nk}\rangle$ exprimés sur la base propre $\{|\psi_m^{(0)}\rangle\}_m$ et ayant en particulier *a priori* une composante non nulle sur l'état $|\psi_n^{(0)}\rangle$ dont on cherche les corrections ; cette dernière peut être exhibée en écrivant les choses comme suit :

$$|\phi_{nk}\rangle = c_{nnk}|\psi_n^{(0)}\rangle + \sum_{m \neq n} c_{nmk}|\psi_m^{(0)}\rangle, \dots \quad (6.54)$$

En reportant une telle expression dans le développement (6.38) donnant le vecteur propre perturbé et en rassemblant tous les termes qui sont une composante le long de $|\psi_n^{(0)}\rangle$, on trouve :

$$|\psi_n\rangle = (1 + gc_{nn1} + g^2c_{nn2} + \dots)|\psi_n^{(0)}\rangle + \text{termes en } |\phi_{nk \geq 1}\rangle. \quad (6.55)$$

Compte tenu du regroupement qui vient d'être effectué, tous les termes en $|\phi_{nk \geq 1}\rangle$ ne contiennent plus $|\psi_n^{(0)}\rangle$ et sont donc tous orthogonaux à $|\psi_n^{(0)}\rangle$. Par ailleurs, comme toute équation aux valeurs et vecteurs propres, l'équation que l'on cherche à résoudre (approximativement) est une équation homogène ; toutes ses solutions sont donc au mieux définies à un facteur global près, qui doit être déterminé en temps utile sur d'autres considérations (par exemple, normalisation à l'unité). Ceci permet de *convenir* que la composante de $|\psi_n\rangle$ sur l'état non-perturbé $|\psi_n^{(0)}\rangle$ est égale à l'unité : il suffit en effet de remultiplier $|\psi_n\rangle$ par $(1 + gc_{nn1} + \dots)^{-1/2}$ pour qu'il en soit bien ainsi. Cette convention étant définitivement adoptée, il en résulte immédiatement que la correction globale de $|\psi_n\rangle$ est orthogonale à $|\psi_n^{(0)}\rangle$, ce que l'on peut résumer par :

$$|\psi_n\rangle = |\psi_n^{(0)}\rangle + \text{corrections orthogonales à } |\psi_n^{(0)}\rangle. \quad (6.56)$$

En particulier, il en résulte :

$$|\phi_{n1}\rangle \perp |\psi_n^{(0)}\rangle. \quad (6.57)$$

Examinons maintenant l'équation (6.52). Son inversion doit être conduite avec précaution ; en effet, la relation $(H_0 - E_n^{(0)})|\psi_n^{(0)}\rangle = 0$ montre que l'opérateur $(H_0 - E_n^{(0)})$ est un opérateur singulier, qui n'a pas d'inverse dans *tout* l'espace des états. Compte tenu de la propriété d'orthogonalité (6.57), une méthode de partition utilisant les projecteurs s'avère très utile. Posons :

$$P_n = |\psi_n^{(0)}\rangle \langle \psi_n^{(0)}|, \quad Q_n = \mathbf{1} - P_n \equiv \sum_{m \neq n} |\psi_m^{(0)}\rangle \langle \psi_m^{(0)}|. \quad (6.58)$$

Comme les $\{|\psi_m^{(0)}\rangle\}$ sont orthogonaux et normalisés à l'unité, P_n est un projecteur :

$$P_n^2 = P_n, \quad P_n^\dagger = P_n \quad (6.59)$$

et il en va évidemment de même pour Q_n ; en outre $P_n Q_n = Q_n P_n = 0$. Par ailleurs :

$$P_n H_0 = |\psi_n^{(0)}\rangle \langle \psi_n^{(0)}| H_0 = |\psi_n^{(0)}\rangle \langle \psi_n^{(0)}| E_n^{(0)} = E_n^{(0)} P_n; \quad (6.60)$$

et, par l'orthogonalité (6.57) :

$$P_n |\phi_{n1}\rangle = 0. \quad (6.61)$$

En multipliant maintenant à gauche par P_n l'équation (6.52), il vient :

$$P_n (H_0 - E_n^{(0)}) |\phi_{n1}\rangle = P_n (\varepsilon_{n1} - V) |\psi_n^{(0)}\rangle. \quad (6.62)$$

Le premier membre est nul en vertu de (6.60) et de (6.61) ; il en résulte :

$$P_n (\varepsilon_{n1} - V) |\psi_n^{(0)}\rangle = 0. \quad (6.63)$$

En multipliant par $\langle \psi_n^{(0)}|$ à gauche, il vient :

$$\langle \psi_n^{(0)}| (\varepsilon_{n1} - V) |\psi_n^{(0)}\rangle = 0 \iff \varepsilon_{n1} = \langle \psi_n^{(0)}| V |\psi_n^{(0)}\rangle. \quad (6.64)$$

Ce résultat s'énonce :

pour un niveau non-dégénéré, la première correction ε_{n1} à l'énergie s'obtient en prenant la valeur moyenne de la perturbation sur l'état non-perturbé

Cette affirmation semble après coup presque évidente. En faisant formellement $g = 1$, l'expression de l'énergie perturbée pour l'état $|\psi_n\rangle$ est donc :

$$H = H_0 + V, \quad E_n = E_n^{(0)} + \langle \psi_n^{(0)} | V | \psi_n^{(0)} \rangle + \text{termes d'ordre supérieur} . \quad (6.65)$$

Bien sûr, il peut arriver que la moyenne de V dans l'état non perturbé $|\psi_n^{(0)}\rangle$ soit nulle (pour des raisons de symétrie, par exemple) et alors il n'y a pas de correction au premier ordre ; en pareil cas, le *shift* en énergie est au moins du second ordre. En tout état de cause, la première correction formelle à l'énergie ne dépend que de l'état à l'ordre zéro.

Une fois obtenue la première correction pour l'énergie, on peut trouver la première correction à l'état propre. Revenant à (6.52) et en multipliant cette fois à gauche par Q_n , on obtient :

$$Q_n(H_0 - E_n^{(0)})|\phi_{n1}\rangle = Q_n(\varepsilon_{n1} - V)|\psi_n^{(0)}\rangle . \quad (6.66)$$

Maintenant on peut insérer $\mathbf{1} = P_n + Q_n$ au premier membre, juste devant $|\phi_{n1}\rangle$. Comme H_0 est diagonal, $Q_n H_0 P_n = 0$ et il reste :

$$Q_n(H_0 - E_n^{(0)})Q_n|\phi_{n1}\rangle = Q_n(\varepsilon_{n1} - V)|\psi_n^{(0)}\rangle . \quad (6.67)$$

La même opération peut être faite au second membre, juste devant $|\psi_n^{(0)}\rangle$; $Q_n|\psi_n^{(0)}\rangle$ est nul, d'où :

$$Q_n(H_0 - E_n^{(0)})Q_n|\phi_{n1}\rangle = Q_n(\varepsilon_{n1} - V)P_n|\psi_n^{(0)}\rangle = -Q_n V|\psi_n^{(0)}\rangle \quad (6.68)$$

(on peut omettre P_n devant $|\psi_n^{(0)}\rangle$ puisque $P_n|\psi_n^{(0)}\rangle = |\psi_n^{(0)}\rangle$). $Q_n(H_0 - E_n^{(0)})Q_n$ est la projection de $H_0 - E_n^{(0)}$ dans le sous-espace orthogonal à $|\psi_n^{(0)}\rangle$. Ceci peut maintenant s'inverser : puisque l'énergie $E_n^{(0)}$ est non-dégénérée, il n'y a aucune valeur propre égale à $E_n^{(0)}$ dans le sous-espace sur lequel Q_n projette, de sorte que l'opérateur au premier membre de (6.68) n'a aucune valeur propre nulle et possède donc un inverse. On en déduit :

$$|\phi_{n1}\rangle = -[Q_n(H_0 - E_n^{(0)})Q_n]^{-1} V|\psi_n^{(0)}\rangle . \quad (6.69)$$

Par définition, on a :

$$[Q_n(H_0 - E_n^{(0)})Q_n]^{-1} = \sum_{m \neq n} |\psi_m^{(0)}\rangle \frac{1}{E_m^{(0)} - E_n^{(0)}} \langle \psi_m^{(0)} | , \quad (6.70)$$

d'où :

$$|\phi_{n1}\rangle = \sum_{m \neq n} \frac{\langle \psi_m^{(0)} | V | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} |\psi_m^{(0)}\rangle . \quad (6.71)$$

En faisant maintenant $g = 1$, on voit que l'état de $H_0 + V$ perturbé au premier ordre s'exprime comme :

$$|\psi_n\rangle = |\psi_n^{(0)}\rangle + \sum_{m \neq n} \frac{\langle \psi_m^{(0)} | V | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} |\psi_m^{(0)}\rangle + \text{corrections d'ordre supérieur} . \quad (6.72)$$

Cette expression fournit une indication de la validité de la méthode. En pratique, celle-ci est particulièrement utile si les corrections restent petites, signifiant que les séries convergent rapidement. L'expression (6.72) montre que ce sera le cas si les éléments de matrice de V restent petits quand les différences d'énergie augmentent¹⁷. D'une façon tout à fait générale, une perturbation ne produit d'effet notable que pour les couples d'états tels que :

$$|\langle \psi_m^{(0)} | V | \psi_n^{(0)} \rangle| \sim |E_n^{(0)} - E_m^{(0)}| ; \quad (6.73)$$

cette condition signifie que seuls les états séparés d'une énergie comparable (ou inférieure) à la perturbation sont sensiblement affectés par cette dernière. Elle constitue un critère permettant de tronquer *a priori* l'espace

¹⁷L'expression de la correction d'énergie au second ordre, (6.76), confirme cette affirmation.

des états de façon raisonnable quand il s'agit de simplifier un problème complexe. En outre, on voit clairement pourquoi il est nécessaire d'envisager un traitement spécial pour le cas où le niveau non-perturbé présente une dégénérescence : alors, les dénominateurs d'énergie peuvent s'annuler et il convient de tout reprendre.

Une fois obtenues toutes les corrections relatives au premier ordre en g , on peut continuer et passer à l'ordre suivant. Il faut maintenant résoudre (6.53) où les inconnues sont ε_{n2} et $|\phi_{n2}\rangle$. ε_{n2} s'obtient immédiatement en multipliant à gauche par P_n . À nouveau le premier membre est nul ; après simplification du second membre, il reste :

$$0 = -P_n V |\phi_{n1}\rangle + \varepsilon_{n2} |\psi_n^{(0)}\rangle \quad (6.74)$$

d'où, après multiplication à gauche par $\langle \psi_n^{(0)} |$:

$$\varepsilon_{n2} = \langle \psi_n^{(0)} | V | \phi_{n1} \rangle . \quad (6.75)$$

La correction du second ordre à l'énergie s'exprime à l'aide de la fonction perturbée au premier ordre. En reportant l'expression (6.71), on en déduit :

$$\varepsilon_{n2} = \sum_{m \neq n} \frac{\langle \psi_n^{(0)} | V | \psi_m^{(0)} \rangle \langle \psi_m^{(0)} | V | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} = \sum_{m \neq n} \frac{|\langle \psi_n^{(0)} | V | \psi_m^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}} . \quad (6.76)$$

Il est utile de remarquer que cette correction du second ordre s'écrit de façon compacte à l'aide du projecteur Q_n :

$$\varepsilon_{n2} = \langle \psi_n^{(0)} | V [Q_n (E_n^{(0)} - H_0) Q_n]^{-1} V | \psi_n^{(0)} \rangle . \quad (6.77)$$

Cette correction peut être positive ou négative, puisque le dénominateur dans (6.76) n'est pas d'un signe donné. Toutefois, s'il s'agit de la correction pour l'état fondamental ($n = 1$, conventionnellement), tous les dénominateurs sont négatifs. On en déduit que pour l'état fondamental la correction du second ordre à l'énergie est toujours *négative* :

$$\varepsilon_{n=1,2} = - \sum_{m \neq n} \frac{|\langle \psi_1^{(0)} | V | \psi_m^{(0)} \rangle|^2}{E_m^{(0)} - E_1^{(0)}} < 0 . \quad (6.78)$$

Il en résulte que, quand la correction au premier ordre est nulle, le déplacement d'énergie du fondamental est toujours vers le bas, comme si tous les niveaux excités "pesaient" sur le fondamental pour l'abaisser. Il en ira ainsi par exemple pour les corrections radiatives¹⁸ de l'état fondamental d'un atome couplé au rayonnement (dans une description où le champ est lui aussi quantifié). En ce qui concerne la correction du second ordre pour $|\psi_n\rangle$, $|\phi_{n2}\rangle$, elle est facile à écrire mais de peu d'intérêt en pratique.

En résumé : le développement de l'énergie perturbée pour $H = H_0 + V$, quand il existe, commence comme suit :

$$E_n = E_n^{(0)} + \langle \psi_n^{(0)} | P_n V P_n | \psi_n^{(0)} \rangle + \langle \psi_n^{(0)} | V [Q_n (E_n^{(0)} - H_0) Q_n]^{-1} V | \psi_n^{(0)} \rangle + \dots , \quad (6.79)$$

ou, sous forme explicite :

$$E_n = E_n^{(0)} + \langle \psi_n^{(0)} | V | \psi_n^{(0)} \rangle + \sum_{m \neq n} \frac{|\langle \psi_n^{(0)} | V | \psi_m^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}} + \dots . \quad (6.80)$$

Remarque

On a supposé que le niveau auquel on s'intéresse (identifié par n) est non-dégénéré, mais rien n'a été dit à ce sujet pour les autres niveaux, qui peuvent bien sûr présenter de la dégénérescence sans pour autant altérer la validité des résultats ci-dessus. Lorsqu'une telle dégénérescence existe, soit $\{|\psi_m^{(0)\alpha}\rangle\}$ l'ensemble des états correspondants, tous associés à la même valeur de l'énergie, $E_m^{(0)}$, l'indice α allant de 1 à g_m , ordre de dégénérescence du niveau d'énergie $E_m^{(0)}$. Dans les formules ci-dessus, l'indice m est une étiquette pour

¹⁸Le déplacement de l'énergie d'un niveau atomique sous l'effet de l'émission-absorption de photons (virtuels) s'appelle le déplacement de Lamb (*Lamb shift*) ; le *Lamb shift* de l'état fondamental est toujours négatif.

tous les autres états en-dehors de n ; en présence de dégénérescence, les mêmes dénominateurs d'énergie vont se retrouver ici et là dans la sommation et on peut les regrouper, ce qui revient à remplacer la somme sur m par une somme sur (p, α) où p repère les énergies toutes différentes les unes des autres et où α distingue les états dégénérés entre eux. Ainsi, la correction du second ordre obtenue ci-dessus s'écrit tout autant :

$$\varepsilon_{n2} = \sum_{p \neq n} \frac{\sum_{\alpha=1}^{g_p} |\langle \psi_{p\alpha}^{(0)} | V | \psi_n^{(0)} \rangle|^2}{E_n^{(0)} - E_p^{(0)}} \equiv \sum_{p \neq n} \sum_{\alpha=1}^{g_p} \frac{|\langle \psi_{p\alpha}^{(0)} | V | \psi_n^{(0)} \rangle|^2}{E_p^{(0)} - E_m^{(0)}} . \quad (6.81)$$

6.2.2 Cas d'un niveau dégénéré

Dans le cas où le niveau d'intérêt est dégénéré, il faut utiliser une notation plus précise, pour distinguer entre eux les états propres associés à une même valeur de l'énergie ; pour la suite, on adopte la notation suivante :

$$H_0 |\psi_{n\alpha}^{(0)}\rangle = E_n^{(0)} |\psi_{n\alpha}^{(0)}\rangle \quad (\alpha = 1, 2, \dots, g_n) . \quad (6.82)$$

L'ordre de dégénérescence, g_n , est supposé fini quelque soit n ; les $\{|\psi_{n\alpha}^{(0)}\rangle\}$ engendrent le sous-espace propre \mathcal{E}_n associé à $E_n^{(0)}$, de dimension g_n . Dans ces conditions, n'importe quelle combinaison linéaire de ces g_n vecteurs satisfait aussi l'équation aux vecteurs propres (6.82) :

$$H_0 \sum_{\alpha=1}^{g_n} c_\alpha |\psi_{n\alpha}^{(0)}\rangle = \sum_{\alpha=1}^{g_n} c_\alpha H_0 |\psi_{n\alpha}^{(0)}\rangle = \sum_{\alpha=1}^{g_n} c_\alpha E_n^{(0)} |\psi_{n\alpha}^{(0)}\rangle = E_n^{(0)} \sum_{\alpha=1}^{g_n} c_\alpha |\psi_{n\alpha}^{(0)}\rangle \quad \forall c_\alpha \in \mathbb{C} . \quad (6.83)$$

Bien évidemment, si l'on part de g fini et que l'on prend la limite $g \rightarrow 0$, il y a g_n vecteurs perturbés qui vont se réduire, par continuité, à *certaines* combinaisons linéaires bien déterminées des $\{|\psi_{n\alpha}^{(0)}\rangle\}$, mais, justement, on ne connaît pas ces combinaisons linéaires particulières, notées $|\phi_{n0\alpha}\rangle$. En présence de dégénérescence, on ne connaît pas d'avance les états propres à l'ordre zéro – au sens : états vers lesquels tendent les états propres de H quand gV tend vers zéro. Ces états sont *a priori* différents des $\{|\psi_{n\alpha}^{(0)}\rangle\}$: on sait seulement que tous ces vecteurs appartiennent au même sous-espace vectoriel \mathcal{E}_n , par définition. En l'absence de dégénérescence, une telle question ne se pose pas : il n'y a qu'un seul vecteur perturbé $|\psi_n\rangle$ et il tend forcément vers $|\psi_n^{(0)}\rangle$

Pour illustrer ceci, considérons un espace de dimension 2, engendré par deux vecteurs $|\psi_{n1}^{(0)}\rangle, |\psi_{n2}^{(0)}\rangle$, tous deux propres d'un Hamiltonien H_0 avec la même valeur propre $E_n^{(0)}$. La matrice de H_0 est :

$$H_0 = \begin{bmatrix} E_n^{(0)} & 0 \\ 0 & E_n^{(0)} \end{bmatrix} . \quad (6.84)$$

Soit maintenant une perturbation gV dont la matrice, sur la même base, est :

$$V = \begin{bmatrix} 0 & gv \\ gv & 0 \end{bmatrix} . \quad (6.85)$$

Toute combinaison linéaire quelconque des deux vecteurs $|\psi_{n\alpha=1,2}^{(0)}\rangle$ est encore vecteur propre de H_0 avec toujours la même valeur propre $E_n^{(0)}$. D'un autre côté, les valeurs propres (ici exacts !) de $H = H_0 + gV$ s'obtiennent immédiatement :

$$E_{n\pm} = E_n^{(0)} \pm gv , \quad (6.86)$$

et les vecteurs propres normalisés correspondants sont :

$$|\phi_{n\pm}\rangle = \frac{1}{\sqrt{2}} (\pm |\psi_{n1}^{(0)}\rangle + |\psi_{n2}^{(0)}\rangle) . \quad (6.87)$$

Maintenant, quand $g \rightarrow 0$, $E_{n\pm} \rightarrow E_n^{(0)}$, mais les deux vecteurs $|\phi_{n\pm}\rangle$ (qui ici ne dépendent¹⁹ même pas de g) restent ce qu'ils sont : les combinaisons linéaires particulières qui sont les limites des états propres de H ,

¹⁹Ceci vient du fait que l'exemple est trop simple, et que les solutions exactes correspondent à une *symétrie* donnée, brisée dès que $g \neq 0$, aussi petit soit-il.

calculés à un ordre supérieur, quand la perturbation tend vers zéro. Ils n'ont aucune raison de coïncider avec les $|\psi_{n\alpha=1,2}^{(0)}\rangle$ préalablement définis, et ce sont pourtant bien des états propres à l'ordre zéro.

Revenant au cas général, on voit qu'il existe donc exactement g_n vecteurs perturbés, $|\psi_{n\alpha}\rangle$, qui vont tendre vers des combinaisons linéaires $|\phi_{n0\alpha}\rangle$ à déterminer ; pour chacun de ces $|\psi_{n\alpha}\rangle$, on suppose toujours qu'il est possible d'écrire un développement en puissances entières de g . À chacun de ces vecteurs correspondra une énergie perturbée E_n ; les différentes E_n seront ou ne seront pas toutes distinctes. Si elles sont toutes différentes, on dit que la perturbation a complètement levé toute la dégénérescence ; dans le cas contraire, la levée de dégénérescence est seulement partielle. Quoi qu'il en soit, on a maintenant, repérant toujours par n le niveau auquel on s'intéresse particulièrement :

$$E_{n\alpha} = E_n^{(0)} + g \varepsilon_{n1\alpha} + g^2 \varepsilon_{n2\alpha} + \dots \quad (6.88)$$

et :

$$|\psi_{n\alpha}\rangle = |\phi_{n0\alpha}\rangle + g |\phi_{n1\alpha}\rangle + g^2 |\phi_{n2\alpha}\rangle + \dots \quad (6.89)$$

Comme précédemment, les corrections $|\phi_{nk\alpha}\rangle$, $k \geq 1$ peuvent être choisies orthogonales à tous les $|\psi_{n\alpha}^{(0)}\rangle$ (donc également aux $|\phi_{n0\alpha}\rangle$), c'est-à-dire appartenant au complément orthogonal de \mathcal{E}_n . En reportant ces développements dans l'équation aux vecteurs propres et en identifiant les termes de même puissance en g , on trouve successivement :

$$H_0 |\phi_{n0\alpha}\rangle = E_n^{(0)} |\phi_{n0\alpha}\rangle, \quad (6.90)$$

$$H_0 |\phi_{n1\alpha}\rangle + V |\phi_{n0\alpha}\rangle = E_n^{(0)} |\phi_{n1\alpha}\rangle + \varepsilon_{n1\alpha} |\phi_{n0\alpha}\rangle, \quad (6.91)$$

$$H_0 |\phi_{n2\alpha}\rangle + V |\phi_{n1\alpha}\rangle = E_n^{(0)} |\phi_{n2\alpha}\rangle + \varepsilon_{n1\alpha} |\phi_{n1\alpha}\rangle + \varepsilon_{n2\alpha} |\phi_{n0\alpha}\rangle, \quad (6.92)$$

et ainsi de suite. À nouveau, la première équation n'apprend rien. La seconde, issue du terme en g , s'écrit :

$$(H_0 - E_n^{(0)}) |\phi_{n1\alpha}\rangle = (\varepsilon_{n1\alpha} - V) |\phi_{n0\alpha}\rangle. \quad (6.93)$$

Introduisons maintenant le projecteur P_n sur le sous-espace propre associé à $E_n^{(0)}$:

$$P_n = \sum_{\alpha=1}^{g_n} |\psi_{n\alpha}^{(0)}\rangle \langle \psi_{n\alpha}^{(0)}|. \quad (6.94)$$

Comme les $|\psi_{n\alpha}^{(0)}\rangle$ et les $|\phi_{n0\alpha}\rangle$ sont dans le même sous-espace \mathcal{E}_n , on a tout autant :

$$P_n = \sum_{\alpha=1}^{g_n} |\phi_{n0\alpha}\rangle \langle \phi_{n0\alpha}|. \quad (6.95)$$

Le projecteur dans l'espace orthogonal complémentaire est :

$$Q_n = \mathbf{1} - P_n = \sum_{m \neq n} \sum_{\beta=1}^{g_m} |\psi_{m\beta}^{(0)}\rangle \langle \psi_{m\beta}^{(0)}| \equiv \sum_{m \neq n} P_m. \quad (6.96)$$

P_n n'est autre que l'identité dans le sous-espace dégénéré \mathcal{E}_n . En multipliant (6.93) à gauche par P_n :

$$P_n (H_0 - E_n^{(0)}) |\phi_{n1\alpha}\rangle = P_n (\varepsilon_{n1\alpha} - V) |\phi_{n0\alpha}\rangle. \quad (6.97)$$

Le premier membre est nul, puisque $P_n H_0 - E_n^{(0)} P_n = 0$ (et d'ailleurs les $|\phi_{n1\alpha}\rangle$ sont dans le complément orthogonal) ; d'où :

$$P_n (\varepsilon_{n1\alpha} - V) |\phi_{n0\alpha}\rangle = 0. \quad (6.98)$$

En injectant $\mathbf{1} = P_n + Q_n$ juste devant $|\psi_{n0\alpha}^{(0)}\rangle$ et en tenant compte de $Q_n |\phi_{n0\alpha}\rangle = 0$, il vient :

$$P_n (V - \varepsilon_{n1\alpha}) P_n |\phi_{n0\alpha}\rangle = 0, \quad (6.99)$$

soit :

$$(P_n V P_n) |\phi_{n0\alpha}\rangle = \varepsilon_{n1\alpha} |\phi_{n0\alpha}\rangle \quad (6.100)$$

Cette équation établit de fait un résultat important : elle montre que $|\phi_{n0\alpha}\rangle$ est vecteur propre de $P_n V P_n$ avec la valeur propre $\varepsilon_{n1\alpha}$; les corrections au premier ordre $\varepsilon_{n1\alpha}$ ne sont donc rien d'autre que les valeurs propres de l'opérateur $P_n V P_n$, et les bons états à l'ordre zéro²⁰ – dont on rappelle qu'ils étaient eux aussi inconnus – sont les vecteurs propres associés. L'équation (6.100) définit *ipso facto* la méthode de perturbation au premier ordre pour un niveau dégénéré : on écrit la matrice de V dans le seul sous-espace de \mathcal{E}_n et on la diagonalise. Notons que le résultat (6.64) n'est qu'un cas particulier quand $g_n = 1$ (pas de dégénérescence).

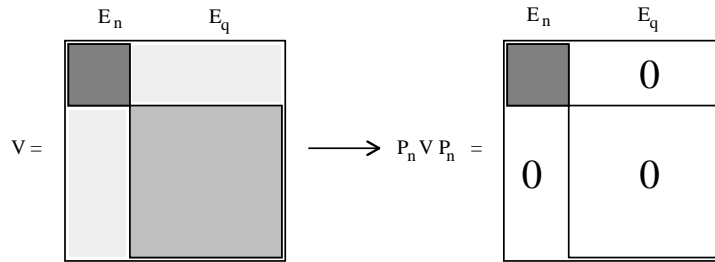


Figure 6.1: Illustration de la projection de l'opérateur de perturbation dans le sous-espace dégénéré \mathcal{E}_n .

L'opérateur $P_n V P_n$ est la *projection* de l'opérateur de perturbation V dans le sous-espace propre \mathcal{E}_n . En effet, les éléments de matrice de $P_n V P_n$ sont tels que :

$$\langle \psi_{n\alpha}^{(0)} | P_n V P_n | \psi_{n\alpha'}^{(0)} \rangle = \langle \psi_{n\alpha}^{(0)} | V | \psi_{n\alpha'}^{(0)} \rangle \quad \forall \alpha, \alpha' , \quad \langle \psi_{n\alpha}^{(0)} | V | \psi_{m\beta}^{(0)} \rangle = 0 \quad \forall n \neq m, \quad \forall \alpha, \beta . \quad (6.101)$$

Autrement dit, ils coïncident avec ceux de V dans le sous-espace \mathcal{E}_n et sont tous nuls partout ailleurs. Ceci peut s'illustrer par un dessin où sont schématiquement représentés deux sous-espaces dégénérés distincts (voir fig. 6.1).

En considérant successivement tous les sous-espaces dégénérés, on voit que la résolution du problème donnant les états à l'ordre zéro et les valeurs propres à l'ordre 1 implique non pas l'opérateur V présent au départ mais la somme (directe) de ses projections dans chaque sous-espace \mathcal{E}_n ; algébriquement parlant : on remplace V par un opérateur approché, V_{ap} , tel que :

$$V_{\text{ap}} = \sum_n P_n V P_n \neq V = \sum_n P_n V P_n + \sum_{n \neq m} (P_n V P_m + P_m V P_n) , \quad (6.102)$$

où tous les blocs non-diagonaux connectant deux sous-espaces dégénérés distincts dans V ont été annulés identiquement dans V_{ap} .

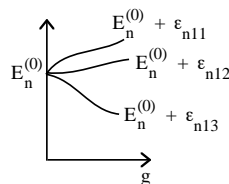


Figure 6.2: Illustration de la levée de dégénérescence par une perturbation.

En résumé, la méthode de résolution “minimale” pour un niveau dégénéré consiste à diagonaliser la matrice de la perturbation V , considérée seulement dans le sous-espace dégénéré et en ignorant tout le reste²¹. Les valeurs propres ainsi obtenues sont les premières corrections à l'énergie. L'opérateur V étant hermitique, toutes les corrections $\varepsilon_{n1\alpha}$ sont réelles ; les énergies correspondantes $E_n^{(0)} + g \varepsilon_{n1\alpha}$ sont donc les approximations

²⁰ au sens : limite des états perturbés lorsque la perturbation tend vers zéro.

²¹ On imagine d'avance le rôle de premier plan joué en pareille circonstance par le théorème de Wigner - Eckart.

des E_n . La perturbation a donc, en général, la vertu de scinder en plusieurs niveaux distincts les g_n états dégénérés en l'absence de V (voir fig. 6.2) ; la levée de dégénérescence est complète si l'on obtient exactement g_n énergies distinctes).

Pour aller plus loin dans les corrections, il convient de faire le tri : au premier ordre pour l'énergie, certains états sont encore dégénérés, d'autres ne le sont plus, quand la levée est partielle. Les derniers peuvent être alors traités par le schéma simple relatif à un niveau non-dégénéré ; les autres au contraire relèvent du traitement qui vient d'être exposé. Il est clair que la discussion générale des corrections d'ordre supérieur est sans intérêt mais que celles-ci sont calculables au coup par coup, parfois laborieusement.

Remarque

La distinction entre niveaux dégénérés et niveaux non-dégénérés n'est pas toujours aussi nette. On se trouve parfois dans une situation de quasi-dégénérescence, lorsque les éléments de matrice de V couplant deux états non-dégénérés sont du même ordre de grandeur que la différence d'énergie de ces deux états (voir la brève discussion autour de (6.73)). La sûreté exige alors que l'on applique la technique élaborée pour le cas dégénéré. En outre, il arrive que les niveaux dépendent continûment d'un paramètre (le module d'un champ extérieur, par exemple) et que, ce paramètre variant, ils en viennent à se croiser (croisement de niveaux). En pareil cas, au *voisinage* du point de croisement (et pas seulement au point de croisement !) il convient de traiter le problème dans un cadre unique – indépendant de la valeur du paramètre – et c'est à nouveau le cadre de la méthode de la situation dégénérée qu'il faut choisir : diagonaliser la matrice de V réduite au sous-espace des états pouvant se croiser. Selon les cas (et notamment la symétrie²² de V), le point de croisement peut disparaître – on parle alors de *croisement évité* – ou au contraire persister.

6.2.3 Développement systématique à l'aide de la résolvante

La résolvante $G(z)$ est définie comme l'opérateur²³ :

$$G(z) = \frac{1}{z\mathbf{1} - H} . \quad (6.103)$$

Dans toute la suite, pour simplifier l'écriture, on omettra $\mathbf{1}$ en facteur de tous les scalaires. z est un nombre complexe quelconque. G est relié simplement à la transformée de Laplace de l'opérateur d'évolution avancé $U_+(t)$:

$$U_+(t) = \begin{cases} e^{\frac{1}{i\hbar}Ht} & \text{si } t > 0 \\ 0 & \text{si } t < 0 \end{cases} , \quad U_+(t=0) = \mathbf{1} . \quad (6.104)$$

En effet, soit $K(z)$ le transformé de Laplace de l'opérateur U_+ défini par (6.104) :

$$K(z) \equiv \mathcal{L}(U_+) = \int_0^{+\infty} dt e^{-zt} e^{\frac{1}{i\hbar}Ht} . \quad (6.105)$$

L'intégrale peut se calculer comme avec des scalaires et on obtient :

$$K(z) = \frac{1}{z - (H/i\hbar)} \equiv i\hbar G(i\hbar z) , \quad (6.106)$$

soit, à l'envers :

$$G(z) = (i\hbar)^{-1} K(-i\hbar^{-1}z) . \quad (6.107)$$

Explicitement, on a la relation intégrale inverse :

$$U_+(t) = \int_{C_+} \frac{dz}{2i\pi} e^{\frac{1}{i\hbar}zt} G(z) , \quad (6.108)$$

²²Si l'élément de matrice de V est nul par symétrie, le point de croisement demeure.

²³Tout ce qui suit n'a évidemment de sens que pour un Hamiltonien H indépendant du temps.

où C_+ est un contour situé juste au-dessus de l'axe réel, parcouru dans le sens des parties réelles décroissantes, de $+\infty$ à $-\infty$.

L'intérêt de la résolvante est sa forme algébrique simple qui se prête remarquablement aux développements comme on va le voir. En outre, les propriétés analytiques de G reflètent la nature du spectre de H et des arguments généraux utilisant l'arsenal de la Théorie des fonctions analytiques permettent souvent d'obtenir des résultats qualitatifs de première importance sans effectuer de lourds calculs. En effet, en supposant que le spectre est composé d'un ensemble discret $\{E_n\}$ et d'une portion continue allant de E_b à l'infini (par exemple)²⁴, $G(z)$ prend la forme explicite :

$$G(z) = \sum_n \frac{P_n}{z - E_n} + \int_{E_b}^{+\infty} dE \frac{P(E)}{z - E} \quad (6.109)$$

où les $P_n = \sum_\alpha |E_{n\alpha}\rangle\langle E_{n\alpha}|$ et $P(E) = \int d\alpha |E, \alpha\rangle\langle E, \alpha|$ sont à nouveau des projecteurs, précisément ce sont les projecteurs sur les états propres de H – la trace de P_n est la dimension g_n du sous-espace dégénéré associé à l'énergie E_n . Ainsi, $G(z)$ a un pôle simple pour chaque valeur discrète du spectre, et une coupure sur l'intervalle continu : les propriétés analytiques de $G(z)$ sont donc le reflet fidèle du spectre du Hamiltonien. Notons que l'on peut facilement inverser la relation (6.109) grâce au théorème des résidus ; par exemple :

$$P_n = \frac{1}{2i\pi} \oint_{C_n} dz G(z) \quad (6.110)$$

où C_n est un contour fermé quelconque, parcouru dans le sens positif, contenant exclusivement le pôle E_n . Il est clair que connaître les modes propres de H ou ceux de $G(z)$ est équivalent, à des difficultés techniques près²⁵.

Dans le cas où $H = H_0 + gV$, on peut définir deux résolvantes $G_0 = (z - H_0)^{-1}$ et $G(z) = (z - H)^{-1}$. Ceci étant posé, il est alors facile d'établir une équation fondamentale, dite équation de Dyson, reliant G et G_0 . Écrivons :

$$z - H_0 = z - H_0 - gV + gV = z - H + gV, \quad (6.111)$$

soit :

$$G_0^{-1} = G^{-1} + gV. \quad (6.112)$$

En multipliant à gauche par G_0 , on a :

$$\mathbf{1} = G_0 G^{-1} + gG_0 V, \quad (6.113)$$

puis à droite par G :

$$G(z) = G_0(z) + gG_0(z) V G(z). \quad (6.114)$$

Ceci est l'équation de Dyson ; c'est une équation exacte implicite, où l'inconnue G figure dans les deux membres. L'équation (6.114) s'itère aisément :

$$G = G_0 + gG_0 V [G_0 + gG_0 V G]. \quad (6.115)$$

En itérant à l'infini, on en déduit le développement exact :

$$G = G_0 + gG_0 V G_0 + g^2 G_0 V G_0 V G_0 + \dots \equiv G_0 + G_0 \sum_{k=1}^{+\infty} g^k (V G_0)^k. \quad (6.116)$$

On se doute que ce développement, tronqué à un certain ordre, est plus ou moins équivalent à une théorie de perturbation tronquée au même ordre²⁶. De façon plus explicite, la méthode consiste à partir de (6.110) – où C_n est un contour entourant maintenant la valeur propre non perturbée $E_n^{(0)}$ et²⁷ celles qui tendent vers elle lorsque $g \rightarrow 0$ – et à essayer de calculer les différents termes :

$$P_n = \frac{1}{2i\pi} \oint_{C_n} dz [G_0 + gG_0 V G_0 + g^2 G_0 V G_0 V G_0 + \dots]. \quad (6.117)$$

²⁴Le plus souvent, on a de surcroît : $\max_n \{E_n\} < E_b$.

²⁵ $G(z)$ est d'un emploi très commode ; il reste que le retour à H , via $U(t)$ et ses fréquences d'oscillation, nécessite une transformation de Laplace inverse qui n'est pas toujours aisée.

²⁶Voir [20], p. 611.

²⁷faute de quoi chaque terme de la série dans (6.117) aurait un résidu nul.

On en déduit alors l'expression des projecteurs, P_n par exemple, sous la forme d'un développement en puissances de g :

$$P_n = P_n^{(0)} + \sum_{k=1}^{+\infty} \frac{g^k}{2i\pi} \oint_{C_n} dz G_0(z) [VG_0(z)]^k . \quad (6.118)$$

À l'intérieur de C_n , chaque terme de la série a pour seule singularité un pôle d'ordre $k+1$ en $E_n^{(0)}$. Puis, pour un niveau non-dégénéré, les approximations de l'énergie s'obtiennent en partant de :

$$E_n = \text{Tr}(HP_n) \quad (6.119)$$

puisque $HP_n = E_n P_n$ et que $\text{Tr}P_n = 1$. Formellement :

$$E_n = \text{Tr} \left((H_0 + gV) \left[P_n^{(0)} + \sum_{k=1}^{+\infty} \frac{g^k}{2i\pi} \oint_{C_n} dz G_0(z) [VG_0(z)]^k \right] \right) \equiv E_n^{(0)} + \sum_{k=1}^{+\infty} g^k \varepsilon_{nk} \quad (6.120)$$

Pour terminer, montrons sur un exemple simple un autre intérêt de G . Soit une perturbation V définie comme²⁸ :

$$V = v |\psi\rangle\langle\psi| \quad (6.121)$$

où $|\psi\rangle$ est un vecteur donné de l'espace des états²⁹. Soit à calculer l'élément diagonal $\langle\phi|G|\phi\rangle$ où $|\phi\rangle$ est *a priori* quelconque. L'équation itérée (6.116) permet d'écrire :

$$\langle\phi|G|\phi\rangle = \langle\phi|G_0|\phi\rangle + v \langle\phi|G_0|\psi\rangle\langle\psi|G_0|\phi\rangle + v^2 \langle\phi|G_0|\psi\rangle\langle\psi|G_0|\psi\rangle\langle\psi|G_0|\phi\rangle + \dots . \quad (6.122)$$

Posons :

$$f_0(z) = \langle\phi|G_0|\phi\rangle , \quad h(z) = \langle\phi|G_0|\psi\rangle\langle\psi|G_0|\phi\rangle , \quad k(z) = \langle\psi|G_0|\psi\rangle . \quad (6.123)$$

Alors :

$$\langle\phi|G(z)|\phi\rangle = f_0(z) + v h(z)h^*(z) + v^2 h(z)k(z)h^*(z) + v^3 h(z)[k(z)]^2 h^*(z) + \dots \quad (6.124)$$

où h^* doit être compris comme la fonction complexe conjuguée de h lorsque z est réel. On voit apparaître une série géométrique qui se resomme immédiatement³⁰ :

$$\langle\phi|G(z)|\phi\rangle = f_0(z) + v h(z) [1 + vk(z) + v^2[k(z)]^2 + \dots] h^*(z) = f_0(z) + v \frac{h(z)h^*(z)}{1 - vk(z)} \quad (6.125)$$

et donne une expression compacte et exacte de l'élément de matrice de G , resommée à tous les ordres de la théorie de perturbation. Le même problème, posé en théorie de perturbation standard aurait fourni des expressions tronquées à un ordre donné et n'aurait sûrement pas fourni le dénominateur présent dans (6.125), qui donne à $\langle\phi|G(z)|\phi\rangle$ ses propriétés analytiques. La puissance de la résolvante peut se mesurer à l'aune de cette dernière remarque. Plus généralement, la résolvante se prête bien à l'imagination de méthodes approchées nettement plus séduisantes que la théorie des perturbations standard, l'objectif étant au coup par coup de parvenir à fabriquer une résolvante (approchée) G_{ap} ayant la forme algébrique d'une résolvante, soit $G_{\text{ap}} = \frac{1}{z - H_{\text{ap}}}$, H_{ap} jouant le rôle d'un Hamiltonien effectif ; généralement parlant, une telle forme est obtenue par resommation à l'infini d'une parties des termes de la série de perturbation.

6.2.4 Exemples d'application : effet Stark et effet Zeeman pour l'hydrogène

Il s'agit ici de traiter deux applications classiques de la théorie des perturbations, donnant une description élémentaire de deux effets importants : l'effet Stark (éclatement des sous-niveaux atomiques par application d'un champ électrique) et l'effet Zeeman (phénomène analogue provoqué par un champ magnétique).

²⁸Cette perturbation peut par exemple représenter une impureté localisée en phase solide.

²⁹Ce peut être un vecteur propre de H_0 , mais alors le problème est sans intérêt.

³⁰On suppose bien sûr que $|vk(z)| < 1$.

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 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 :

$$H_0 = \frac{\vec{p}^2}{2m} - \frac{e'^2}{r} . \quad (6.126)$$

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 :

$$E_n^{(0)} = -\frac{\alpha^2}{2n^2} mc^2 = -\frac{e'^2}{2n^2 a_0} \simeq -\frac{13.6}{2n^2} \text{ eV} . \quad (6.127)$$

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

$$V = -e\mathcal{E}z \quad (6.128)$$

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 (intuitives) de validité de la méthode de perturbation. On doit avoir :

$$\frac{|\langle n, l, m | V | n', l', m' \rangle|}{|E_n^{(0)} - E_{n'}^{(0)}|} \ll 1 \iff |\langle n, l, m | e\mathcal{E}z | n', l', m' \rangle| \ll |E_n^{(0)} - E_{n'}^{(0)}| . \quad (6.129)$$

L'ordre de grandeur du premier membre est $|e\mathcal{E}a_0|$; le second est de l'ordre de e'^2/a_0 . La condition (6.129) se réécrit comme suit :

$$\mathcal{E} \ll \frac{|e|}{4\pi\epsilon_0 a_0^2} \equiv \mathcal{E}_{\text{atomique}} . \quad (6.130)$$

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 contenant l'axe Oz ; dans cette opération, L_z change de signe³¹. Les états propres communs³² 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.

Une dernière remarque concernant la symétrie : V est changé en son opposé par parité³³ :

$$V' \equiv \Pi V \Pi^\dagger = -V . \quad (6.131)$$

Par ailleurs :

$$\Pi |n, l, m\rangle = (-1)^l |n, l, m\rangle . \quad (6.132)$$

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 :

$$\langle n, l, m | z | n', l', m' \rangle = 0 \quad \text{si } l \text{ et } l' \text{ de même parité} . \quad (6.133)$$

³¹tout comme la composante L_x située dans ce plan, cependant que la composante perpendiculaire au plan, L_y , est inaltérée.

³²De toute évidence, \vec{L}^2 est encore une constante du mouvement.

³³Le champ extérieur n'appartient pas au système quantique et se factorise hors de tous les éléments de matrice.

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 est³⁴, d'après (6.76) ou (6.78) :

$$\varepsilon_{\text{fond 2}} = - \sum_{n=2}^{+\infty} \sum_{l,m} \frac{|\langle 1, 0, 0| - e\mathcal{E}z |n, l, m\rangle|^2}{E_n^{(0)} - E_1^{(0)}} < 0, \quad (6.134)$$

où les $E_n^{(0)}$ sont données par (6.127). Grâce à une astuce très ingénieuse, il est possible de sommer cette série ([18], p. 266) ; on trouve :

$$\varepsilon_{\text{fond 2}} = -\frac{1}{2} \alpha \mathcal{E}^2, \quad \alpha = 18\pi\epsilon_0 a_0^3. \quad (6.135)$$

L'écriture $-(\alpha/2)\mathcal{E}^2$ est choisie pour rappeler qu'il s'agit d'un effet de polarisabilité³⁵, la correction d'énergie quadratique en champ devant pouvoir s'écrire $-\int_0^{\mathcal{E}} d\mathcal{E}' \alpha \mathcal{E}'$.

Considérons maintenant le premier groupe d'états excités $n = 2$, définissant un sous-espace dégénéré de dimension égale à 4. Il convient donc de mettre en œuvre la technique exposée dans la sous-section 6.2.2. 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 (voir (6.102) et la fig. 6.1), la méthode de perturbation néglige ce type de couplage et projette 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 :

$$\langle 2, 0, 0|V|2, 1, 0\rangle = 3e\mathcal{E}a_0 \equiv -v < 0, \quad (6.136)$$

le signe venant de la charge de l'électron e . Il en résulte que les valeurs propres de l'opérateur projeté sont $\varepsilon_{21} = 0, \pm v$; les vecteurs propres s'en déduisent :

$$|+v\rangle = \frac{1}{\sqrt{2}} (|2, 0, 0\rangle + |2, 1, 0\rangle), \quad |-v\rangle = \frac{1}{\sqrt{2}} (-|2, 0, 0\rangle + |2, 1, 0\rangle). \quad (6.137)$$

$|\pm v\rangle$ a pour énergie $\pm v$ (l'état $|+v\rangle$ est au-dessus de l'état $|-v\rangle$). Ces deux états correspondent à la levée de dégénérescence, et sont les bons états propres dès que le champ est présent, aussi petit soit-il (c'est à nouveau un exemple de brisure de symétrie). 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 (6.137) s'écrivent :

$$|+v\rangle = \frac{1}{\sqrt{2}} (|2S\rangle + |2P_0\rangle), \quad |-v\rangle = \frac{1}{\sqrt{2}} (-|2S\rangle + |2P_0\rangle). \quad (6.138)$$

Ainsi, le champ électrique mélange deux états de symétrie différente vis-à-vis du moment cinétique orbital³⁶, en conséquence de la dégénérescence "accidentelle" du champ coulombien et de la brisure de la symétrie sphérique (\vec{L}^2 n'est plus une constante du mouvement). Ceci a des conséquences importantes, en particulier pour la durée de vie de certains états excités.

L'état $|2S\rangle$ a une durée de vie extraordinairement longue, de l'ordre de la seconde³⁷, 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 Lyman³⁸, appelée L_α , correspond précisément à la transition $|2P\rangle \rightarrow |1S\rangle$. 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

³⁴Ceci 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.

³⁵ α désigne ponctuellement la polarisabilité (notation traditionnelle), pas la constante de structure fine !

³⁶Un tel mélange d'états s'appelle traditionnellement *hybridation*, terme que l'on retrouve également en Chimie à propos des orbitales atomiques ou moléculaires.

³⁷Ceci 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*.

³⁸ tout entière dans l'U.V.

l'état $2S$ et qu'un champ électrique est ajouté, volontairement ou non, la métastabilité est anéantie et l'atome peut rayonner presque normalement grâce à sa composante P . 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 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} :

$$V_{\text{tot}} = -\frac{e'^2}{r} - e\mathcal{E}z = -\frac{e'^2}{r} + |e|\mathcal{E}z . \quad (6.139)$$

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 :

$$\lim_{z \rightarrow +\infty} V_{\text{tot}} = +\infty , \quad \lim_{z \rightarrow -\infty} V_{\text{tot}} = -\infty \quad (6.140)$$

avec toujours $\lim_{|z| \rightarrow 0} V_{\text{tot}} = -\infty$. Un potentiel ayant de tels comportements n'a pas d'état liés ; ceci traduit la possibilité pour l'électron de passer par effet-tunnel de n'importe quel état localisé à un état dissocié (ionisé). 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'épaisseur³⁹ : 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 problème⁴⁰, on se doute que le calcul ci-dessus est validé par le fait que des temps de l'ordre de la seconde ou plus sont des temps "infinis" vis-à-vis des échelles de temps pertinentes lors d'une expérience de spectroscopie atomique.

Effet Zeeman

À nouveau, on choisit l'axe Oz le long du champ, ici magnétique et noté \vec{B} . On conçoit qu'une description cohérente exige de prendre en compte tous les effets magnétiques simultanément et c'est pourquoi le Hamiltonien H_0 doit ici contenir le couplage spin-orbite, responsable de la structure fine. Ceci étant, H_0 s'écrit :

$$H_0 = \frac{\vec{p}^2}{2m} - \frac{e'^2}{r} + a(r) \vec{L} \cdot \vec{S} . \quad (6.141)$$

Il n'est pas ici nécessaire de préciser la fonction $a(r)$. En présence du champ \vec{B} , et en négligeant le terme diamagnétique, ce Hamiltonien doit être complété par V :

$$V = -\frac{e}{2m} (L_z + g_e S_z) B \quad (g_e = 2.0023 \dots \simeq 2) . \quad (6.142)$$

³⁹d'où l'extraordinaire sensibilité du microscope à effet-tunnel.

⁴⁰Si 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 !

De la sorte, le Hamiltonien en présence du champ magnétique est :

$$H = \frac{\vec{p}^2}{2m} - \frac{e'\prime^2}{r} + a(r) \vec{L} \cdot \vec{S} - \frac{e}{2m} (L_z + g_e S_z) B ; \quad (6.143)$$

comme le terme de structure fine est explicitement introduit, on n'est plus tenu⁴¹ de considérer seulement la situation de champ fort.

Il ne fait aucun doute que si le champ est très faible, le couplage V 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, dans la terminologie orthodoxe, porte le nom d'effet Zeeman⁴². Ce choix étant fait, les états propres à l'ordre zéro sont ceux de H_0 et se notent $|n, j, m_j, l, s\rangle$; ils sont 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 V dans le sous-espace auquel on s'intéresse, oublier le reste et diagonaliser. Pour trouver la matrice de V dans ce sous-espace, deux méthodes peuvent être suivies :

- V 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 V 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 $\{|n, j, m_j, l, s\rangle\}$, $\vec{L} + g_e \vec{S}$ est un opérateur vectoriel proportionnel à \vec{J} , avec un facteur de proportionnalité bien déterminé :

$$\vec{L} + g_e \vec{S} = \frac{\langle (\vec{L} + g_e \vec{S}) \cdot \vec{J} \rangle}{\hbar^2 j(j+1)} \vec{J} \equiv g_{lsj} \vec{J} . \quad (6.144)$$

g_{lsj} désigne un nombre pur (appelé *facteur de Landé*), qui s'exprime à l'aide des nombres quantiques l, s et j . En écrivant $\vec{L} \cdot \vec{S} = \frac{1}{2}(\vec{J}^2 - \vec{L}^2 - \vec{S}^2)$, il vient :

$$(\vec{L} + g_e \vec{S})(\vec{L} + \vec{S}) = \frac{1}{2} \left[(g_e + 1) \vec{J}^2 - (g_e - 1) \vec{L}^2 + (g_e - 1) \vec{S}^2 \right] . \quad (6.145)$$

Si maintenant on choisit $g_e = 2$, g_{lsj} est un rationnel simple :

$$g_{lsj} \simeq 1 + \frac{j(j+1) - l(l+1) + s(s+1)}{2j(j+1)} \quad (6.146)$$

Dans la limite des très grands nombres quantiques, s est négligeable devant l et g_{lsj} est presque égal à 1. De (6.144) et (6.142), il résulte immédiatement que la matrice du V projeté, V_{ap} , est déjà diagonale. Les corrections au premier ordre pour l'énergie se lisent sur la diagonale et sont donc :

$$\varepsilon_{nj1} = -g_{lsj} \frac{eB}{2m} m_j \hbar \equiv -g_{lsj} m_j \mu_B B . \quad (6.147)$$

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 B du champ appliqué. Pour l'atome d'hydrogène avec son unique électron de spin $1/2$, 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 nul.

⁴¹ comme c'était le cas dans la présentation du ch. 5.

⁴² L'autre cas où V_{magn} est petit devant le couplage avec \vec{B} donne lieu à l'effet dit Paschen - Back. Il est bien clair que l'on passe continûment d'un cas à l'autre en augmentant graduellement le champ extérieur.

Chapitre 7

Théorie des perturbations dépendant du temps

Le but de ce chapitre est d'exposer les bases de la théorie des perturbations dépendant du temps et de présenter ses premières applications à quelques cas simples mais importants.

7.1 Présentation générale

L'étude expérimentale d'un système, noté (Sy) dans la suite, consiste presque toujours à faire interagir ce système (que l'on peut appeler "cible") avec une sonde extérieure dont les caractéristiques physiques sont contrôlées. Très souvent, la sonde est un faisceau de particules d'énergie connue (excitation "monochromatique") ; si ces dernières sont douées d'un caractère vectoriel intrinsèque (un spin, par exemple), le faisceau peut, si besoin est, être préparé dans un état de polarisation bien déterminée (lumière polarisée pour sélectionner certaines transitions atomiques, neutrons polarisés pour mettre en évidence un ordre magnétique en phase condensée, *etc.*). Le faisceau diffusé¹, en raison de son interaction avec la cible, présente des caractéristiques différentes dont l'analyse permet précisément de remonter aux propriétés intrinsèques de la cible ; en quelque sorte, le faisceau diffusé constitue une image au sens large de la cible.

Dans la situation la plus courante, le faisceau est constitué soit de particules matérielles, soit de photons et l'expérience consiste à mesurer des sections efficaces. La dépendance de ces dernières vis-à-vis de l'énergie (ou de la fréquence) et de l'angle de diffusion fournit de précieux renseignements sur l'interaction entre les projectiles et la cible et permet de ce fait de mieux connaître la dynamique interne de celle-ci – en définitive d'affiner la description théorique (fondamentale ou phénoménologique) des interactions élémentaires, ou des interactions effectives lorsque les projectiles "moyennent" d'une certaine façon une partie des degrés de liberté de la cible.

Traditionnellement, on parle de collisions quand les projectiles sont des particules matérielles, et de spectroscopie (nucléaire, atomique, moléculaire, du solide, *etc.*) quand ce sont des photons qui sont utilisés. Cette terminologie est évidemment un peu arbitraire : on sait bien toutes les particules ont un comportement complexe qui ne permet pas une distinction instinctive aussi nette, héritée des conceptions anciennes. Suivant les circonstances définies par l'expérience, une même particule (l'électron par exemple) se comportera comme

¹En français, le vocable diffusion a un double sens :

- diffusion au sens d'éparpillement, c'est le sens employé premièrement en Théorie des Collisions
- diffusion au sens de l'équation de la diffusion déduite de la loi de Fick ($\partial P/\partial t = D\Delta P$)

En anglais, on dispose de deux substantifs, respectivement *scattering* et *diffusion*.

une onde ou comme une petite bille – ce sont des cas extrêmes – mais plus souvent les deux aspects ondulatoire et corpusculaire se superposent, ou plutôt s’imbriquent l’un dans l’autre.

Les applications de la spectroscopie sont innombrables et en font un outil d’usage universel pour l’étude de la matière, qu’il s’agisse de recherche fondamentale ou appliquée. Les mesures utilisant la spectroscopie atomique sont d’une précision diabolique : c’est grâce à certaines qu’il a été possible de vérifier expérimentalement avec une précision rarement atteinte les prévisions de l’Electrodynamique Quantique, notamment l’écart $g - 2$ du facteur anormal de l’électron. Dans un tout autre champ d’application, les méthodes d’imagerie – telles qu’on les utilise aujourd’hui en Médecine par exemple – sont fondées sur l’aptitude de la matière à réagir sélectivement et de façon spécifique à une petite perturbation dûment contrôlée.

D’une façon ou d’une autre, le problème théorique se présente schématiquement comme suit. Le système (Sy) , caractérisé par son Hamiltonien H_0 , est à un certain instant dans l’un de ses états propres $|\psi_i\rangle$, constituant un état stationnaire². À partir d’un instant t_0 , (Sy) est soumis à une perturbation *extérieure* non quantifiée ; l’interaction entre (Sy) et cette perturbation est représentée par un opérateur V où apparaissent conjointement une (ou plusieurs) observables de (Sy) et les grandeurs physiques caractéristiques de la perturbation externe. Pour $t \geq t_0$, l’état du système (Sy) obéit ainsi à l’équation de Schrödinger :

$$i\hbar \frac{\partial}{\partial t} |\Psi(t)\rangle = [H_0 + V(t)] |\Psi(t)\rangle . \quad (7.1)$$

Comme, d’une façon générale, l’état initial $|\psi_i\rangle$ n’est pas propre de $H(t)$ ³, cet état, qui serait stationnaire en l’absence de perturbation, va se mettre à évoluer sous l’effet de $V(t)$. Il est toujours loisible de garder comme base de l’espace des états de (Sy) , \mathcal{E} , les états propres de H_0 , les $\{|E_k\rangle\}_k$; dès lors, l’effet de la perturbation est d’induire des transitions entre l’état initial $|\psi_i\rangle$ et les autres états propres. On peut se représenter l’effet de la perturbation comme une promenade forcée du système dans son espace d’états \mathcal{E} .

La question typique que l’on se pose alors est la suivante : $|\psi_f\rangle$ désignant l’un quelconque des états propres $\{|E_k\rangle\}_k$, quelle est la probabilité $P_{i \rightarrow f}(t)$ pour que le système se retrouve à l’instant t dans l’état $|\psi_f\rangle$ sous l’effet de la perturbation ? Cette question a une réponse et une seule : l’équation (7.1) est du premier ordre en temps et on s’est donné un état initial, $|\psi_i\rangle$. En supposant effectuée l’intégration en temps, donc ayant déterminé le vecteur d’état à l’instant t , $|\Psi(t)\rangle$, la probabilité cherchée s’exprime, en vertu des postulats, sous la forme du module au carré d’une certaine amplitude de probabilité :

$$P_{i \rightarrow f}(t) = |A_{i \rightarrow f}(t)|^2 \equiv |\langle \psi_f | \Psi(t) \rangle|^2 \quad (|\Psi(t=t_0)\rangle = |\psi_i\rangle) . \quad (7.2)$$

Si $U(t, t_0)$ désigne l’opérateur d’évolution associé à $H(t)$ entre les instants t_0 et t , l’amplitude de probabilité $A_{i \rightarrow f}$ est donnée par :

$$A_{i \rightarrow f}(t) = \langle \psi_f | U(t, t_0) | \psi_i \rangle . \quad (7.3)$$

Les amplitudes étant trouvées, on peut alors en déduire notamment la variation d’énergie (moyenne) du système résultant du passage d’une perturbation provisoire ; en désignant par $A_{i \rightarrow k}(t)$ l’amplitude de probabilité de transition entre l’état initial d’énergie E_i et l’état propre $|E_k\rangle$ d’énergie E_k , la variation d’énergie est :

$$\Delta E \equiv E_{\text{finale}} - E_i = \left(\sum_k E_k |A_{i \rightarrow k}(t)|^2 \right) - E_i = \sum_k (E_k - E_i) P_{i \rightarrow k} . \quad (7.4)$$

La dernière égalité résulte du fait que la somme des probabilités de transitions de l’état initial aux différents états finals doit être égale à 1 :

$$\sum_k P_{i \rightarrow k} = 1 \quad (7.5)$$

puisque le système doit bien se trouver “quelque part”.

Le problème est ainsi bien posé, mais n’est malheureusement pas soluble exactement, mis à part quelques cas simples – qui demeurent cependant très instructifs. Comme exemples de problèmes exactement solubles,

²Cette hypothèse n’est pas nécessaire ; elle est admise pour fixer les idées et aussi parce qu’elle correspond à la situation la plus courante.

³Si c’était le cas, le problème serait sans grand intérêt.

on peut citer : un oscillateur harmonique chargé soumis à un champ électrique variable de façon quelconque dans le temps, un spin 1/2 soumis à une perturbation harmonique⁴, deux spins $S_i = 1/2$ en interaction par un couplage du genre $f(t)\vec{S}_1 \cdot \vec{S}_2$ où $f(t)$ est une fonction réelle quelconque.

Fort heureusement, et c'est presque une nécessité méthodologique, la perturbation qui permet de sonder le système d'intérêt doit rester "petite" en un sens à préciser. En effet, ce que l'on cherche à observer c'est le système lui-même et il convient donc de le "déranger" le moins possible. Quand on en a décidé ainsi, il y a forcément dans le terme de couplage $V(t)$ un petit paramètre sans dimension, g , et on peut espérer pouvoir construire un développement de l'amplitude $A_{i \rightarrow f}(t)$ en série entière de g ; si $g \ll 1$, le ou les tout premiers termes seront en général suffisants. Ceci est l'esprit de toutes les méthodes de perturbation, classiques ou quantiques d'ailleurs⁵.

La décomposition présupposée du Hamiltonien total $H(t) = H_0 + V(t)$ est toujours le résultat d'une approximation. Fondamentalement, on suppose qu'il est possible de séparer et d'identifier clairement d'une part le système (Sy), décrit par H_0 , et la perturbation extérieure, représentée par $V(t)$. Il y a bel et bien une dissymétrie entre l'objet d'observation et l'outil utilisé pour cette observation. De fait, la perturbation est considérée comme un champ extérieur donné qui "bouscule" le système mais n'en subit aucune contre-réaction ; on se doute physiquement que ceci sera licite si le champ extérieur est un "grand" système, constituant notamment un réservoir infini d'énergie, totalement désincarné physiquement et supposé complètement décrit par un (petit) nombre de grandeurs physiques données une fois pour toutes. En d'autres termes, seul le système (Sy) fait l'objet d'une description quantique, le champ étant décrit classiquement ; c'est pourquoi une telle approche est appelée semi-classique.

Construire une justification théorique de ce type d'approche – tout en restant à un niveau élémentaire⁶ – c'est énoncer les conditions permettant de se contenter de l'approche semi-classique et, *a contrario*, préciser les circonstances où une description complètement quantique est requise. Dans ce but, on se met d'emblée dans le cadre le plus rigoureux, en quantifiant à la fois le système d'intérêt et la perturbation avec sa source, génériquement appelée champ dans la suite ; le Hamiltonien du supersystème [$(Sy) + \text{champ}$] – qui constitue alors un système isolé – peut toujours être mis sous la forme :

$$H = H_0 + H_{\text{champ}} + H_{\text{int}} . \quad (7.6)$$

H , décrivant un système isolé, est alors indépendant du temps et est fonction des degrés de liberté de (Sy), notés collectivement $\{x\}$, et de ceux du champ, $\{X\}$. H_0 ne dépend que de $\{x\}$, H_{champ} ne dépend que de $\{X\}$; quant à H_{int} , s'agissant de l'opérateur décrivant l'interaction entre les deux sous-systèmes, c'est bien sûr une fonction des deux jeux de variables $\{x\}$ et $\{X\}$.

Le système (Sy) a son propre espace d'états, \mathcal{E}_{Sy} , celui du champ est $\mathcal{E}_{\text{champ}}$; le système total a pour espace d'états le produit tensoriel $\mathcal{E}_{Sy} \otimes \mathcal{E}_{\text{champ}}$. H_{int} couple entre eux ces états produits tensoriels – qui ne sont propres, donc stationnaires, qu'en l'absence précisément de H_{int} , quand chacun des partenaires évolue sans se soucier de l'autre, faute d'interaction.

Pour résoudre la dynamique du système total, un état initial étant donné, on peut écrire les équations de Heisenberg pour les variables $\{x\}$ et $\{X\}$; on obtient alors, formellement :

$$i\hbar \frac{d}{dt} x_H(t) = [x, H_0]_H + [x, H_{\text{int}}]_H , \quad i\hbar \frac{d}{dt} X_H(t) = [X, H_{\text{champ}}]_H + [X, H_{\text{int}}]_H . \quad (7.7)$$

Le second membre de chaque équation contient d'une part l'évolution libre (premier terme), d'autre part un terme de couplage impliquant les opérateurs des deux partenaires, (Sy) et champ ; par exemple, l'équation pour x est plus précisément de la forme :

$$i\hbar \frac{d}{dt} x_H(t) = [x, H_0]_H + W(\{x_H\}, \{X_H\}) \quad (7.8)$$

⁴à condition de délaissier les terme antirésonnants.

⁵Bien sûr, il existe des cas où la perturbation est délibérément choisie intense, par exemple pour observer les fragments d'un système qui n'aura pas résisté et se sera décomposé ; cette situation ne relève en aucune façon d'une théorie de perturbation, au sens précisé ci-dessus.

⁶La justification ci-dessous est élémentaire au sens où elle ignore les cas où H_0 contient déjà en fait certaines modifications prenant la perturbation en compte ("renormalisation" de la masse, de la charge, ...).

où W est un opérateur agissant dans l'espace produit tensoriel \mathcal{E} du supersystème. La description semi-classique sera justifiable si, pour une raison à préciser, les variables $\{X_H\}$ peuvent être remplacées par leurs valeurs moyennes. Ceci implique que les *fluctuations* quantiques du champ peuvent être négligées ; alors, on obtient :

$$i\hbar \frac{d}{dt} x_H(t) = [x, H_0]_H + W(\{x_H\}, \{\langle X_H \rangle\}) \equiv [x, H_0]_H + F(\{x_H\}; t) \quad (7.9)$$

où maintenant la représentation de Heisenberg, définie par référence au Hamiltonien total H^7 , peut en réalité être élaguée de tous les termes propres au champ. Cependant, la théorie reste ici purement mécanique⁸, et il sera toujours possible de réécrire le terme d'interaction simplifié F sous la forme d'un commutateur :

$$F(\{x_H\}; t) \equiv [x, V(t)]_H . \quad (7.10)$$

Dans ces conditions, l'équation (7.7) pour x devient :

$$i\hbar \frac{d}{dt} x_H(t) = [x, H_0 + V(t)]_H \quad (7.11)$$

et il apparaît naturel de définir un Hamiltonien dépendant du temps, $H(t) = H_0 + V(t)$, décrivant l'interaction du système en interaction avec le champ extérieur décrit classiquement.

Donnons quelques exemples. Soit un système quantique bien identifié (atome, noyau, molécule), couplé à un champ électromagnétique⁹ piloté de l'extérieur. Dans la situation la plus courante, l'atome est initialement dans son état fondamental et, sous l'effet du champ, peut "grimper" dans ses états excités ; c'est l'image la plus simple expliquant l'existence de raies spectrales d'absorption, caractérisées par les fréquences déduites de la loi de Bohr :

$$\nu_{nm} = \frac{1}{h} (E_n - E_m) \quad (7.12)$$

et par des intensités plus ou moins grandes que l'on essaie de déduire des caractéristiques quantiques de l'atome.

Si on se souvient que le champ électromagnétique est constitué de photons, le procédé semi-classique peut sembler très cavalier. Une justification un peu simpliste s'énonce ainsi : avec les sources ordinaires, il y a en général un nombre énorme de photons présents¹⁰ et les fluctuations quantiques sont souvent négligeables¹¹. Inversement, il est essentiel de quantifier le champ pour rendre compte d'un phénomène aussi important que l'émission spontanée ; une fois l'atome passé à l'état excité par l'action d'une perturbation transitoire, il n'y reste pas : à l'instar d'un noyau radioactif, il retombe à l'état fondamental au bout d'un temps aléatoire, dont la moyenne (statistique) constitue par définition la durée de vie de l'état excité. La description de ce phénomène échappe totalement à l'approche semi-classique – et pour cause : il se produit en l'absence de champ "extérieur", une situation où le nombre de photons "présents" est *en moyenne* nul. C'est évidemment une situation où les fluctuations vont jouer un rôle éminent. De fait, l'émission spontanée résulte de ce que l'on appelle les "fluctuations du vide", preuve que le champ électromagnétique est bel et bien un objet relevant d'une description quantique ; les modules des champs électrique et magnétique peuvent être nuls, mais la moyenne de leur carré ne peut évidemment pas l'être, d'où l'existence de fluctuations au sens statistique. Le photon peut

⁷Si tel n'était pas le cas, il y aurait encore des opérateurs du champ dans les équations simplifiées. En réalité, à partir du moment où les opérateurs $\{X_H\}$ ont été remplacés par leurs valeurs moyennes, toutes les grandeurs relatives au champ sont des scalaires ; dans la transformation de Heisenberg définie par H , tous les termes propres au champ disparaissent de fait en faisant commuter les scalaires. Une telle représentation n'est pas la "représentation-interaction" définie plus bas, puisque outre H_0 , il subsiste le terme $H_{\text{int}} + H_{\text{champ}}$ moyenné sur tous les degrés de liberté X du champ.

⁸par opposition à une théorie dissipative donnant lieu à une évolution irréversible ; une telle évolution résulte de termes non-hamiltoniens donnant au second membre de l'équation de Heisenberg des termes d'évolution n'ayant pas la forme d'un commutateur.

⁹Un tel énoncé est déjà, en soi, approximatif : l'atome étant constitué de particules chargées, on ne peut en toute rigueur séparer "matière" et "champ", les deux sont imbriqués pour, précisément, assurer la stabilité de la matière telle qu'elle est perçue ordinairement.

¹⁰Une lampe de 100 W émettant dans le jaune donne, à 1 m de distance, un courant de photons de l'ordre de 2×10^{15} par seconde et par cm^2 .

¹¹Cette dernière affirmation est usuelle dans un contexte purement classique : quand un grand nombre de variables aléatoires entre en jeu, les fluctuations régressent et, à la limite, les processus deviennent certains. On peut aussi invoquer ici le fait que lorsque champ contient beaucoup de photons, il se trouve de fait dans la limite des grands nombres quantiques.

Il ne faut toutefois pas en déduire pour autant que l'intervention d'un grand nombre de particules justifie toujours d'oublier les effets quantiques ! Il existe des transitions de phase (exemple : la condensation de Bose) dont l'existence repose sur la symétrie des fonctions d'onde de bosons ; cette symétrie fondamentale fonctionne par tout ou rien et ne peut être ignorée.

“s'échapper” de l'atome excité grâce à une telle fluctuation, un peu à la manière d'une particule classique qui, sous l'effet d'une fluctuation thermique, peut franchir une barrière d'énergie. D'ailleurs, l'argument vaut tout autant si le photon “émis” ne quitte pas le système, ce dernier le réabsorbant : ce jeu permanent d'émission et d'absorption virtuelles est responsable des corrections radiatives des niveaux atomiques (*Lamb shift*), de l'interaction de van der Waals entre deux objets neutres (atomes ou molécules) à grande distance, *etc.*

Un second exemple, fourni par l'excitation Coulombienne, montre d'ailleurs que le nombre de particules en jeu n'est pas forcément la variable à considérer quand il s'agit de “déquantifier” une partie des degrés de liberté. Ce phénomène est la transition vers un état excité d'un atome ou d'une molécule, sous l'effet de l'interaction de Coulomb, lors du passage à proximité¹² d'une particule de charge Q . La description la plus simple consiste à écrire l'interaction entre les particules chargées q_i de l'atome avec le champ électrostatique dont la source n'est autre que la particule chargée passant au voisinage ; on écrit ainsi :

$$V(t) = \sum_i \frac{q_i Q}{4\pi\epsilon_0 \|\vec{R}(t) - \vec{r}_i\|} \quad (7.13)$$

où \vec{R} est la coordonnée du projectile, considéré comme un point matériel classique, $\{\vec{r}_i\}$ celles des électrons atomiques. Cette description se doit d'être justifiée ; notamment il faut que, pour la cible, la structure physique du projectile soit invisible. Pour qu'il en soit ainsi, plusieurs conditions doivent être satisfaites. Par exemple, si la particule chargée est elle-même un électron, il y a visiblement un problème vis-à-vis de l'indiscernabilité entre cet électron et ceux de l'atome – sauf bien sûr si celui qui constitue le projectile passe effectivement très loin, auquel cas il peut de fait être distingué de ceux de la cible (non-recouvrement des fonctions d'onde). D'un autre côté, on peut se poser la question de l'effet inverse : l'excitation Coulombienne peut tout autant s'effectuer “dans l'autre sens”, la cible excitant le projectile supposé muni d'une structure interne¹³. Ceci ne peut se produire si les niveaux internes du projectile sont très séparés en énergie, caractérisés par une échelle très grande devant les énergies internes de la cible et devant la perturbation. Au total, la description ci-dessus doit bien fonctionner pour une particule assez massive (pour avoir une petite longueur d'onde associée) et très robuste (par exemple, une particule α traversant un gaz).

Remarque

Dans le cas d'une description semi-classique où $H(t) = H_0 + V(t)$, il est loin d'être évident que $H(t)$ représente l'énergie du système. En effet, il s'agit toujours, comme on l'a vu, d'une réduction d'un problème indépendant du temps (où, pour de bonnes raisons, un autre système a été désincarné physiquement). Sur le plan des principes, il existe donc plus fondamentalement une interaction notée H_{int} dans l'analyse ci-dessus. À partir de ce fait, il n'est plus possible à proprement parler de parler de l'énergie de l'une des parties du système, sauf à convenir que H_{int} doit être coupé en deux moitiés égales dans un jugement de Salomon que rien ne justifie à ce niveau.

D'un autre côté, dans la formulation semi-classique, il est parfaitement possible de définir sans aucune ambiguïté la variation d'énergie d'un système donné sous l'effet du passage d'une perturbation de durée finie. Si t_{avant} et $t_{\text{après}}$ désignent deux instants où la perturbation n'est pas présente et encadrant son passage, on a :

$$E(t_{\text{avant}}) - E(t_{\text{après}}) \equiv E_f - E_i = \Delta E \quad (7.14)$$

où la variation d'énergie ΔE est donnée par l'expression (7.4). En toutes circonstances, comme d'habitude, ce sont les *différences* d'énergie qui comptent.

7.2 Calcul systématique de l'opérateur d'évolution et approximations successives

Soit un système quantique de Hamiltonien H_0 , d'états propres $\{|E_k\rangle\}_k$, soumis à partir d'un instant pris comme origine des temps à une perturbation extérieure se couplant au système par l'opérateur $V(t)$. À $t \geq 0$, le

¹²Ce phénomène d'excitation des atomes est le principal mécanisme par lequel une particule chargée de haute énergie (~ 1 MeV) perd peu à peu son énergie en traversant la matière ordinaire.

¹³Cette question ne se pose évidemment pas si le projectile est élémentaire au sens strict, comme l'électron.

Hamiltonien décrivant le système perturbé est :

$$H(t) = H_0 + V(t) . \quad (7.15)$$

L'état à l'instant t , $|\Psi(t)\rangle$, peut se déduire de l'opérateur d'évolution $U(t)$, pourvu que l'on connaisse une condition initiale, spécifiée par exemple selon :

$$|\Psi(t = 0_-)\rangle = |\psi_i\rangle . \quad (7.16)$$

Alors :

$$|\Psi(t)\rangle = U(t)|\psi_i\rangle . \quad (7.17)$$

Résoudre le problème de l'évolution en présence de la perturbation, c'est donc trouver l'opérateur $U(t)$, qui obéit à l'équation :

$$i\hbar \frac{\partial}{\partial t} U(t) = [H_0 + V(t)] U(t) \quad (7.18)$$

avec la condition initiale $U(t = 0) = \mathbf{1}$. C'est cette équation que, dans l'immense majorité des cas, on résout faute de mieux par approximations successives. Dans ce but, il est naturel et commode de mettre en évidence l'évolution libre¹⁴ en posant :

$$U(t) = U_0(t) U_1(t) , \quad U_0(t) = e^{\frac{i}{\hbar} H_0 t} . \quad (7.19)$$

En reportant cette forme dans (7.18), on obtient :

$$i\hbar \left[\frac{\partial U_0}{\partial t} U_1 + U_0 \frac{\partial U_1}{\partial t} \right] = [H_0 + V(t)] U_0 U_1 \quad (7.20)$$

et comme $i\hbar \partial U_0 / \partial t = H_0 U_0$, il reste :

$$i\hbar U_0(t) \frac{\partial}{\partial t} U_1(t) = V(t) U_0(t) U_1(t) , \quad (7.21)$$

d'où il résulte que U_1 satisfait l'équation :

$$i\hbar \frac{\partial}{\partial t} U_1(t) = U_0^\dagger(t) V(t) U_0(t) U_1(t) \equiv V_I(t) U_1(t) , \quad ; \quad (7.22)$$

$V_I(t)$ est appelé "représentation-interaction" de $V(t)$:

$$V_I(t) = U_0^\dagger(t) V(t) U_0(t) . \quad (7.23)$$

Par la définition même de U_1 , la condition initiale complétant (7.22) est évidemment :

$$U_1(t = 0) = \mathbf{1} . \quad (7.24)$$

L'intégration formelle de (7.22) membre à membre, compte tenu de la condition initiale (7.24) pour U_1 , fournit :

$$U_1(t) = \mathbf{1} + \frac{1}{i\hbar} \int_0^t dt_1 V_I(t_1) U_1(t_1) . \quad (7.25)$$

On peut maintenant itérer et écrire :

$$\begin{aligned} U_1(t) &= \mathbf{1} + \frac{1}{i\hbar} \int_0^t dt_1 V_I(t_1) \left[\mathbf{1} + \frac{1}{i\hbar} \int_0^{t_1} dt_2 V_I(t_2) U_1(t_2) \right] \\ &= \mathbf{1} + \frac{1}{i\hbar} \int_0^t dt_1 V_I(t_1) + \frac{1}{(i\hbar)^2} \int_0^t dt_1 \int_0^{t_1} dt_2 V_I(t_1) V_I(t_2) U_1(t_2) . \end{aligned} \quad (7.26)$$

¹⁴c'est-à-dire en l'absence de perturbation.

En remplaçant à nouveau $U_1(t_2)$ par son expression intégrale (7.25) et en continuant, on met finalement en évidence le développement infini :

$$U_1(t) = \mathbf{1} + \sum_{n=1}^{+\infty} \frac{1}{(i\hbar)^n} \int_0^t dt_1 \int_0^{t_1} dt_2 \dots \int_0^{t_{n-1}} dt_n V_I(t_1)V_I(t_2)\dots V_I(t_n) \equiv \mathbf{1} + \sum_{n=1}^{+\infty} U_1^{(n)}(t) , \quad (7.27)$$

avec :

$$U_1^{(n)}(t) = (i\hbar)^{-n} \int_0^t dt_1 \int_0^{t_1} dt_2 \dots \int_0^{t_{n-1}} dt_n V_I(t_1)V_I(t_2)\dots V_I(t_n) . \quad (7.28)$$

Il est clair que chacun des $U_1^{(n)}$ représente un ordre donné de la série de perturbation : si V contient un petit paramètre g en facteur qui mesure la “petitesse” de la perturbation, $U_1^{(n)}$ est d'ordre g^n . Il est par ailleurs essentiel de se souvenir que les $V_I(t_k)$ sont des *opérateurs* et que leur ordre dans le produit ne doit pas être modifié. C'est la raison pour laquelle, en dépit de permutations de variables muettes toujours possibles¹⁵, l'opérateur d'évolution total associé à $H(t)$, $U(t)$, n'est pas égal à $\exp[(i\hbar)^{-1} \int_0^t dt' H(t')]$.

L'amplitude de probabilité de la transition $i \rightarrow f$, $A_{i \rightarrow f}(t)$, (7.3), admet ainsi le développement :

$$A_{i \rightarrow f}(t) = \langle \psi_f | U_0(t) \left[\mathbf{1} + \sum_{n=1}^{+\infty} U_1^{(n)} \right] | \psi_i \rangle = e^{\frac{1}{i\hbar} E_f t} \left[\delta_{if} + \sum_{n=1}^{+\infty} \langle \psi_f | U_1^{(n)} | \psi_i \rangle \right] \equiv e^{\frac{1}{i\hbar} E_f t} a_{i \rightarrow f}(t) , \quad (7.31)$$

avec :

$$a_{i \rightarrow f}(t) = \sum_{n=0}^{+\infty} a_{i \rightarrow f}^{(n)}(t) = \delta_{if} + \sum_{n=1}^{+\infty} \langle \psi_f | U_1^{(n)} | \psi_i \rangle . \quad (7.32)$$

On note que $|A_{i \rightarrow f}(t)| = |a_{i \rightarrow f}(t)|$. Selon (7.26), on a :

- à l'ordre zéro :

$$a_{i \rightarrow f}^{(0)} = \delta_{if} . \quad (7.33)$$

Comme il se doit, l'amplitude de transition à l'ordre zéro, à partir d'un état propre de H_0 , est nulle en l'absence de perturbation

- au premier ordre :

$$a_{i \rightarrow f}^{(1)} = (i\hbar)^{-1} \int_0^t dt_1 \langle \psi_f | V_I(t_1) | \psi_i \rangle . \quad (7.34)$$

Dans le cas où $|\psi_i\rangle$ et $|\psi_f\rangle$ sont des états propres¹⁶ de H_0 , l'action de U_0 contenu dans V_I peut être explicitée. En introduisant les pulsations :

$$\omega_i = \hbar^{-1} E_i \quad \omega_f = \hbar^{-1} E_f \quad \omega_k = \hbar^{-1} E_k \quad \omega_{fi} = \omega_f - \omega_i \quad \omega_{kl} = \omega_k - \omega_l , \quad (7.35)$$

on trouve :

$$a_{i \rightarrow f}^{(1)} = (i\hbar)^{-1} \int_0^t dt_1 e^{i\omega_f t_1} \langle \psi_f | V(t_1) | \psi_i \rangle e^{-i\omega_i t_1} , \quad (7.36)$$

soit :

$$a_{i \rightarrow f}^{(1)} = (i\hbar)^{-1} \int_0^t dt_1 e^{i\omega_{fi} t_1} \langle \psi_f | V(t_1) | \psi_i \rangle . \quad (7.37)$$

¹⁵On introduit parfois, lors de développements formels, l'opérateur chronologique de Dyson défini comme suit :

$$\mathcal{T} [V_I(t_{i_1})V_I(t_{i_2})\dots V_I(t_{i_n})] = V_I(t_1)V_I(t_2)\dots V_I(t_n) \quad t_1 > t_2 > \dots > t_n . \quad (7.29)$$

Il s'agit d'un opérateur “courtois” qui, agissant sur un produit où les temps sont dans un ordre quelconque, rétablit la chronologie – et les préséances – en faisant d'abord agir les opérateurs les plus anciens. Alors, il est licite d'écrire l'expression intégrée :

$$U_1(t) = \mathcal{T} \exp[(i\hbar)^{-1} \int_0^t dt' V_I(t')] , \quad (7.30)$$

la présence de \mathcal{T} devant l'exponentielle étant essentielle.

¹⁶C'est, de loin, le cas le plus fréquent.

• au second ordre :

$$\begin{aligned}
 a_{i \rightarrow f}^{(2)} &= (i\hbar)^{-2} \int_0^t dt_1 \int_0^{t_1} dt_2 \langle \psi_f | V_I(t_1) V_I(t_2) | \psi_i \rangle \\
 &= (i\hbar)^{-2} \int_0^t dt_1 \int_0^{t_1} dt_2 \langle \psi_f | U_0^\dagger(t_1) V(t_1) U_0(t_1 - t_2) V(t_2) U_0(t_2) | \psi_i \rangle .
 \end{aligned}
 \tag{7.38}$$

En utilisant la propriété de groupe de U_0 , il vient :

$$a_{i \rightarrow f}^{(2)} = (i\hbar)^{-2} \int_0^t dt_1 \int_0^{t_1} dt_2 \langle \psi_f | U_0^\dagger(t_1) V(t_1) U_0(t_1 - t_2) V(t_2) U_0(t_2) | \psi_i \rangle .
 \tag{7.39}$$

En introduisant maintenant la décomposition de l'unité $\sum_k |E_k\rangle\langle E_k|$ juste à la droite de $V(t_1)$ (ou juste à gauche de $V(t_2)$), ou en écrivant explicitement :

$$U_0(t_1 - t_2) = \sum_k |E_k\rangle e^{\frac{i}{\hbar} E_k(t_1 - t_2)} \langle E_k| ,
 \tag{7.40}$$

on trouve :

$$a_{i \rightarrow f}^{(2)} = (i\hbar)^{-2} \sum_k \int_0^t dt_1 \int_0^{t_1} dt_2 e^{i\omega_{fk}t_1} e^{i\omega_{ki}t_2} \langle \psi_f | V(t_1) | E_k \rangle \langle E_k | V(t_2) | \psi_i \rangle ,
 \tag{7.41}$$

et ainsi de suite. Sur un plan mnémotechnique, on remarque que l'ordre des indices des pulsations est le reflet exact de celui des éléments de matrice de la perturbation.

En résumé, les résultats pour les premiers ordres sont :

$$A_{i \rightarrow f}^{(0)} = e^{\frac{i}{\hbar} E_f t} \delta_{if} ,
 \tag{7.42}$$

$$A_{i \rightarrow f}^{(1)} = \frac{1}{i\hbar} e^{\frac{i}{\hbar} E_f t} \int_0^t dt_1 e^{i\omega_{fi}t_1} \langle \psi_f | V(t_1) | \psi_i \rangle ,
 \tag{7.43}$$

$$A_{i \rightarrow f}^{(2)} = -\frac{1}{(i\hbar)^2} e^{\frac{i}{\hbar} E_f t} \sum_k \int_0^t dt_1 \int_0^{t_1} dt_2 e^{i\omega_{fk}t_1} e^{i\omega_{ki}t_2} \langle \psi_f | V(t_1) | E_k \rangle \langle E_k | V(t_2) | \psi_i \rangle .
 \tag{7.44}$$

Il faut se souvenir que tous ces résultats supposent que les vecteurs $|\psi_i\rangle$ et $|\psi_f\rangle$ sont des états propres du Hamiltonien non-perturbé H_0 , et définit le cadre naturel de raisonnement dans les situations physiques usuelles.

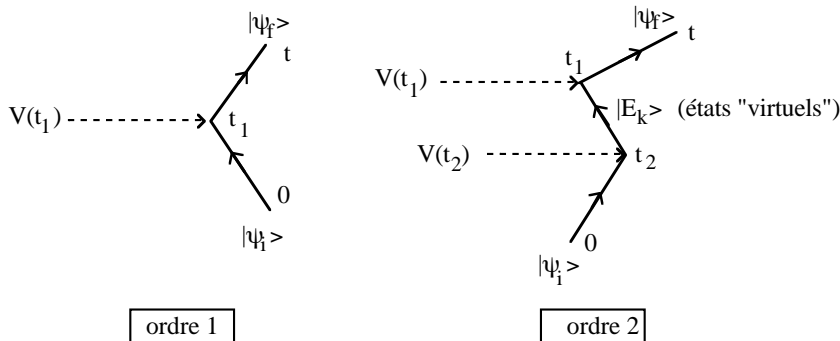


Figure 7.1: Illustration du mode de construction des amplitudes $A_{i \rightarrow f}^{(1)}$ et $A_{i \rightarrow f}^{(2)}$.

Les contributions aux ordres successifs peuvent être représentées par un diagramme comme montré sur la fig. 7.1. Chaque interaction est signalée par le sommet d'une ligne brisée. Pour le premier ordre, il suffit de

sommer l'instant intermédiaire t_1 entre les deux temps fixés 0 et t . Au contraire, dès le second ordre, outre les intégrations sur les temps intermédiaires t_1 et t_2 , apparaissent des sommations sur tous les états du système. Tous ces états, dits "virtuels", sont à l'œuvre entre les deux instants variables t_1 et t_2 (avec la contrainte $t_2 < t_1$). Leur "virtualité" tient au fait qu'ils ne sont pas observables dans le processus dont ils contribuent à l'amplitude – quoiqu'ils soient états propres du système libre. Leur apparition dans le formalisme fait suspecter une violation de la conservation de l'énergie, mais bien sûr il n'en est rien. De telles représentations graphiques de la série de perturbation, et leurs généralisations, portent génériquement le nom de diagrammes de Feynman.

Les résultats précédents permettent d'écrire l'expression de la probabilité de transition $P_{i \rightarrow f}(t)$, $i \neq f$ entre deux états propres de H_0 . S'en tenant au premier ordre, on a :

$$P_{i \rightarrow f}^{(1)} = \hbar^{-2} \left| \int_0^t dt_1 e^{i\omega_{fi}t_1} \langle \psi_f | V(t_1) | \psi_i \rangle \right|^2 . \quad (7.45)$$

À cet ordre, on a $P_{i \rightarrow f}^{(1)} = P_{f \rightarrow i}^{(1)}$, une symétrie qui ne tient plus aux ordres supérieurs¹⁷. Toujours à cet ordre, la probabilité $P_{i \rightarrow f}^{(1)}$ est finalement donnée par le module carré de la transformée de Fourier de la perturbation vue par le système depuis le début jusqu'à l'instant t , transformée qui est prise à la valeur de la pulsation de Bohr associée aux deux états propres considérés.

Il n'est pas très aisé de discuter en général la validité de l'approximation du premier ordre. Il est clair de toute façon que s'agissant en fait du deuxième terme d'un développement en série¹⁸, et puisque la perturbation est réputée *petite*, les probabilités de transition doivent également être très petites, soit $P_{i \rightarrow f} \ll 1$ à tout ordre, et notamment $P_{i \rightarrow f}^{(1)} \ll 1$. En particulier, il faut que l'intervalle de temps T pendant lequel la perturbation est appliquée ne soit pas trop long ; intuitivement, on attend une condition du genre :

$$|\langle \psi_f | \bar{V} | \psi_i \rangle| T \ll \hbar , \quad (7.46)$$

où \bar{V} est l'ordre de grandeur de la perturbation (ou de sa moyenne). Ceci se résume schématiquement en disant : la perturbation doit être assez faible et/ou de courte durée.

7.3 Amplitudes de transition entre états discrets : exemples

Ce paragraphe est consacré à l'application des résultats précédents à deux cas concrets très importants ; on se bornera à ne retenir que le premier ordre et on ne considérera que des transitions entre états bien séparés en énergie. Le cas où les états finals possibles sont au contraire très serrés fera l'objet d'un paragraphe particulier où sera obtenu le résultat dit "règle d'or de Fermi".

7.3.1 Perturbation constante

Cette situation est celle où la perturbation est nulle jusqu'à un certain instant (pris comme origine) et est constante après. $V(t)$ contient en facteur la fonction échelon-unité $\Theta(t)$ et peut s'écrire $V(t) = \Theta(t)V$, V étant un opérateur *indépendant* du temps. Alors, l'intégration en temps est immédiate ; en partant de¹⁹ :

$$|A_{i \rightarrow f}^{(1)}| = \hbar^{-1} \left| \int_0^t dt_1 e^{i\omega_{fi}t_1} \langle \psi_f | V(t_1) | \psi_i \rangle \right| , \quad (7.47)$$

et en posant $\langle \psi_f | V | \psi_i \rangle = V_{fi}$, on trouve :

$$|A_{i \rightarrow f}^{(1)}| = \hbar^{-1} |V_{fi}| \left| \int_0^t dt_1 e^{i\omega_{fi}t_1} \right| = \hbar^{-1} |V_{fi}| \left| \frac{e^{i\omega_{fi}t} - 1}{i\omega_{fi}} \right| , \quad (7.48)$$

¹⁷D'une façon générale, la série de perturbation tronquée à un ordre donné ne fournit pas un opérateur d'évolution unitaire. Cependant, on sait construire des développements unitaires à tout ordre, soit pour $U(t)$ en s'y prenant autrement (on fait un développement en puissances dans l'*argument* de l'exponentielle, voir [37] – l'idée est simple, sa mise en œuvre l'est moins), soit en passant par la résolvante à laquelle on impose d'emblée des propriétés d'analyticité assurant l'unitarité.

¹⁸Le premier terme est le terme d'ordre zéro, égal à δ_{if} .

¹⁹voir (7.43).

d'où la probabilité, au premier ordre :

$$P_{i \rightarrow f}^{(1)}(t) = \hbar^{-2} |V_{fi}|^2 f(t, \omega_{fi}) \quad , \quad f(T, \omega) = \frac{4}{\omega^2} \sin^2 \frac{\omega T}{2} \quad . \quad (7.49)$$

La fonction $f(T, \omega)$ est homogène au carré d'un temps et assure que la probabilité est un nombre pur, comme il se doit. Elle a l'allure indiquée sur la figure 7.2, présentant un maximum – très aigu si $\omega T \gg 1$ – pour $\omega = 0$ où elle vaut T^2 , et des maxima secondaires pour les valeurs de ω :

$$\omega = \text{entier} \times \frac{2\pi}{T} \quad . \quad (7.50)$$

Son enveloppe décroît assez vite, comme ω^{-2} . En outre, il est facile de montrer par résidus que :

$$\int_{-\infty}^{+\infty} d\omega f(T, \omega) = 2\pi T \quad ; \quad (7.51)$$

en définitive, f joue le rôle d'un filtre. Le cas échéant et à condition de bien préciser les échelles de fréquences pertinentes, on pourra faire la substitution :

$$f(T, \omega) \rightarrow 2\pi T \delta(\omega) \quad . \quad (7.52)$$

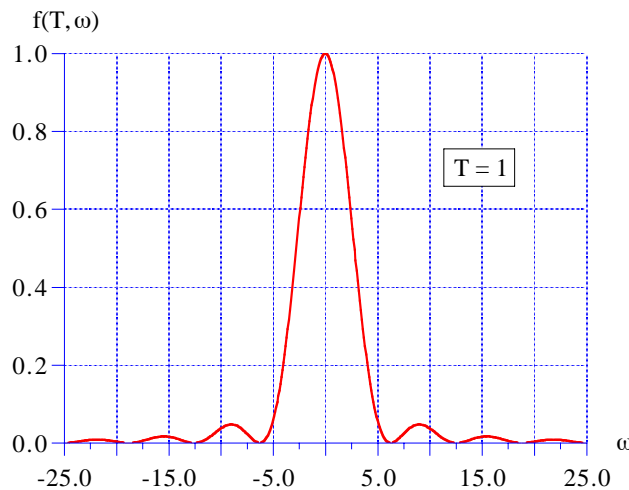


Figure 7.2: Allure typique de la fonction $f(t, \omega_{fi})$ définie en (7.49).

Supposons que la perturbation dure pendant un temps T donné ; une fois qu'elle a cessé, les probabilités de transitions entre états propres de H_0 deviennent constantes²⁰ et on a :

$$P_{i \rightarrow f}(t \geq T) = P_{i \rightarrow f}(T) \simeq P_{i \rightarrow f}^{(1)}(T) = \hbar^{-2} |V_{fi}|^2 f(T, \omega_{fi}) \quad . \quad (7.53)$$

Ainsi, sous l'effet d'une perturbation constante²¹, la probabilité de transition est maximale pour deux états de même énergie ($\omega_{fi} = 0 \iff E_f = E_i$) : pour un champ statique, les transitions se produisent majoritairement entre états dégénérés, ou quasi-dégénérés. Plus généralement, les transitions s'effectuent principalement à l'intérieur d'une bande d'énergie $\delta E \sim \hbar/T$. Pour les états d'énergie voisine, tels que $|\omega_{fi}|T \ll 1$, la probabilité au premier ordre vaut à peu près :

$$P_{i \rightarrow f}^{(1)}(T) \simeq \hbar^{-2} |V_{fi}|^2 \frac{4}{\omega_{fi}^2} \left(\frac{\omega_{fi} T}{2} \right)^2 = \hbar^{-2} |V_{fi}|^2 T^2 \quad , \quad (7.54)$$

²⁰ puisqu'alors le système évolue sous l'action de H_0 seul.

²¹ à ceci près qu'elle est allumée à partir d'un certain instant pris comme origine.

et varie donc en temps comme le *carré* de la durée de la perturbation²² ; cette variation est typique des perturbations de très courte durée²³. L'expression (7.54) est la valeur exacte de la fonction f à son maximum ; s'agissant d'une probabilité, ce résultat n'a de sens que dans la mesure où :

$$\frac{|V_{\text{fi}}|}{\hbar} T \ll 1 . \quad (7.57)$$

Il s'agit aussi d'une condition *nécessaire* pour que les termes suivants de la série de perturbation puissent être négligés ; cette condition confirme l'affirmation exprimée qualitativement par (7.46).

Enfin, on note que la largeur de la bande d'états concernés par la transition à partir de $|\psi_i\rangle$ est évidemment reliée aux propriétés de dualité de la transformation de Fourier. La bande d'états impliqués autour de ω_i a une largeur d'ordre $1/T$: elle est très large aux temps courts puisque, à ces instants, la perturbation est vue comme une fonction variant très vite et possède donc un spectre de Fourier très large. À l'inverse, aux temps très longs, la perturbation est vue comme une fonction constante et a donc un spectre qui, sur l'échelle de fréquence du système, est très pointu, centré sur la fréquence nulle.

7.3.2 Perturbation sinusoïdale

Considérons maintenant le cas où $V(t)$ a l'une de deux formes :

$$V(t) = V \sin \omega t , \quad V(t) = V \cos \omega t \quad (\omega > 0) , \quad (7.58)$$

où V est un opérateur indépendant du temps ; il est toujours licite de convenir que la pulsation ω a un signe donné, positif par exemple. Partant de (7.47) et en posant $\langle \psi_f | V | \psi_i \rangle = V_{\text{fi}}$, on a, pour le sinus :

$$|A_{i \rightarrow f}^{(1)}| = \hbar^{-1} |V_{\text{fi}}| \left| \frac{1}{2i} \int_0^t dt_1 e^{i\omega_{\text{fi}} t_1} (e^{i\omega t_1} - e^{-i\omega t_1}) \right| . \quad (7.59)$$

D'où, au premier ordre :

$$P_{i \rightarrow f}^{(1)}(t) = \frac{|V_{\text{fi}}|^2}{4\hbar^2} \left| \frac{e^{i(\omega_{\text{fi}} + \omega)t} - 1}{i(\omega_{\text{fi}} + \omega)} - \frac{e^{i(\omega_{\text{fi}} - \omega)t} - 1}{i(\omega_{\text{fi}} - \omega)} \right|^2 . \quad (7.60)$$

Avec le cosinus, on obtient la même expression – au signe central près²⁴ :

$$P_{i \rightarrow f}^{(1)}(t) = \frac{|V_{\text{fi}}|^2}{4\hbar^2} \left| \frac{e^{i(\omega_{\text{fi}} + \omega)t} - 1}{i(\omega_{\text{fi}} + \omega)} + \frac{e^{i(\omega_{\text{fi}} - \omega)t} - 1}{i(\omega_{\text{fi}} - \omega)} \right|^2 . \quad (7.61)$$

Le signe central \pm dans (7.60) et (7.61) ne joue que pour les termes croisés venant du développement du module carré ; ce double produit est toujours très petit si on se place dans le cas où $\omega_{\text{fi}} \gg 1$; en effet, le module carré de chacun des deux termes contribuant à l'amplitude est une fonction du genre $f(t, \omega \pm \omega_{\text{fi}})$, centrée en $\pm \omega_{\text{fi}}$ et de largeur $\Delta\omega = 1/t$. Dans la limite $\omega_{\text{fi}} \gg \Delta\omega$, le module au carré de l'amplitude se réduit quasiment à la somme des modules au carré de chaque contribution ; de la sorte, pour le sinus comme pour le cosinus on obtient la même expression approchée de la probabilité :

$$P_{i \rightarrow f}^{(1)}(t) \simeq \frac{|V_{\text{fi}}|^2}{4\hbar^2} \left[\left| \frac{e^{i(\omega_{\text{fi}} + \omega)t} - 1}{i(\omega_{\text{fi}} + \omega)} \right|^2 + \left| \frac{e^{i(\omega_{\text{fi}} - \omega)t} - 1}{i(\omega_{\text{fi}} - \omega)} \right|^2 \right] \equiv \frac{|V_{\text{fi}}|^2}{4\hbar^2} [f(t, \omega + \omega_{\text{fi}}) + f(t, \omega - \omega_{\text{fi}})] . \quad (7.62)$$

²²On ne peut donc pas en la circonstance définir une probabilité de transition par unité de temps qui est indépendante du temps.

²³Pour tous les couples d'états bien séparés en énergie (tels que $\omega_{ki} \gg 1$), f est quasiment nulle. En quelque sorte, la fonction f découple tout état initial donné $|\psi_i\rangle$ de tous les autres états éloignés en énergie et se comporte vis-à-vis de ces derniers un peu à la manière d'une fonction δ :

$$\forall \omega_i, \omega_k, \forall t \gg |\omega_i - \omega_k|^{-1} : f(t, \omega_{ki}) \simeq 2\pi t \delta(\omega_{ki}) . \quad (7.55)$$

On est tenté d'en déduire :

$$P_{i \rightarrow f}(t) \simeq 2\pi \hbar^{-2} |V_{\text{fi}}|^2 t \delta(\omega_{\text{fi}}) . \quad (7.56)$$

L'écriture à l'aide de la fonction δ ne s'impose nullement ici et est plutôt source de confusion : autant l'éviter.

²⁴En faisant $\omega = 0$ dans (7.61), on retrouve bien la probabilité pour une perturbation constante donnée par (7.49).

La probabilité de transition présente alors un maximum très pointu pour les deux valeurs de la pulsation ω , conventionnellement prise positive :

$$\omega = \omega_{\text{fi}} \quad \text{ou} \quad \omega = -\omega_{\text{fi}} . \quad (7.63)$$

Ainsi, la probabilité est résonnante pour les états d'énergie E_{f} telle que :

$$E_{\text{f}} = E_{\text{i}} + \hbar\omega \quad \text{ou} \quad E_{\text{f}} = E_{\text{i}} - \hbar\omega . \quad (7.64)$$

Le premier cas représente l'absorption, le second l'émission. Pour un couple d'états donnés, la résonance se produit lorsque la fréquence du champ est égale à la fréquence de Bohr associée à ce couple. Par ailleurs, pour un ω_{fi} donné et puisque ω est supposé positif, seul l'un des deux termes dans (7.62) est important : c'est le terme dit résonnant, l'autre étant appelé antirésonnant. Finalement :

$$P_{\text{i} \rightarrow \text{f}}^{(1)}(t) \simeq \frac{|V_{\text{fi}}|^2}{4\hbar^2} \times \begin{cases} f(t, \omega + \omega_{\text{fi}}) & \text{si } \omega_{\text{fi}} < 0 \\ f(t, \omega - \omega_{\text{fi}}) & \text{si } \omega_{\text{fi}} > 0 \end{cases} . \quad (7.65)$$

Cette approximation est correcte si la largeur $\Delta\omega = 1/t$ de chaque fonction f est très petite devant la distance entre les deux résonances, soit $2|\omega_{\text{fi}}|$. Comme seules les valeurs de ω proches de $|\omega_{\text{fi}}|$ sont pertinentes, ceci signifie au total que l'on doit avoir :

$$t \gg \frac{1}{\omega} . \quad (7.66)$$

Le sens physique de cette condition est clair : la résonance en absorption ou en émission ne se produit que si la durée de la perturbation est grande par rapport à la période du champ appliqué et si $\omega \sim \omega_{\text{fi}}$; en d'autres termes, le champ externe doit apparaître comme quasi-monochromatique au système – lequel doit avoir le temps de compter un grand nombre de périodes – et avoir la bonne fréquence. C'est seulement lorsque la largeur de Fourier de la perturbation est très petite devant la fréquence propre du système que celui-ci peut présenter un comportement très sélectif se traduisant par une résonance. Pour une perturbation très brève à la même échelle, la réponse du système est évidemment à peu près plate en fréquence.

Les calculs explicites ci-dessus ne concernent que le premier ordre, et impliquent de façon décisive l'élément de matrice V_{fi} . Si ce dernier est nul, la probabilité est nulle au premier ordre – d'où la notion de *règles de sélection* (voir plus loin). En pareil cas, il est utile d'aller à l'ordre suivant et d'utiliser l'expression (7.44), qui contient *deux* interactions et où, de ce fait, figure une somme interne sur tous les états du systèmes (processus *virtuels*). Il est bien clair que deux états $|\psi_{\text{i}}\rangle$ et $|\psi_{\text{f}}\rangle$ non-couplés au premier ordre parce que $V_{\text{fi}} = 0$, peuvent être couplés au second ordre : il suffit pour cela qu'il existe au moins un état $|E_{k_0}\rangle$ tel que les deux éléments de matrice $\langle E_{k_0}|V|\psi_{\text{i}}\rangle$ et $\langle \psi_{\text{f}}|V|E_{k_0}\rangle$ sont différents de zéro ; dès lors le couplage entre $|\psi_{\text{i}}\rangle$ et $|\psi_{\text{f}}\rangle$ se fait *via* l'état $|E_{k_0}\rangle$ et la transition $|\psi_{\text{i}}\rangle \rightarrow |\psi_{\text{f}}\rangle$ devient possible. Avec une onde quasi-monochromatique de pulsation ω , la probabilité sera sensiblement non-nulle seulement si²⁵ $2\omega \simeq \hbar^{-1}|E_{\text{i}} - E_{\text{f}}|$. On parle alors de transition à deux photons ; de façon imagée, on peut dire que ceux-ci se font la courte échelle pour faire passer le système de $|\psi_{\text{i}}\rangle$ vers $|\psi_{\text{f}}\rangle$ en transitant par $|E_{k_0}\rangle$. Il s'agit vraiment d'une image : à aucune étape de ce scénario par pas élémentaires l'énergie n'est conservée (il faut seulement que l'énergie soit conservée entre l'état initial et l'état final, soit $|E_{\text{i}} - E_{\text{f}}| = 2\hbar\omega$) – en particulier, l'état intermédiaire peut avoir une énergie absolument quelconque.

7.4 Amplitudes de transition entre un état discret et un continuum d'états finals. Règle d'or de Fermi

Les résultats précédents portent sur les transitions entre deux états discrets d'énergie parfaitement déterminée ou au moins bien séparés de leurs plus proches voisins²⁶. Cette situation n'est pas la seule possible, loin de là ; en pratique, il est aussi important de savoir traiter la situation où la perturbation peut conduire le système dans une partie continue – ou quasi-continue – du spectre d'énergie. De nombreuses situations physiques correspondent

²⁵L'intervention formelle du *carré* de la perturbation fait apparaître la fréquence *double* 2ω .

²⁶Dans ces conditions, le champ quasi-monochromatique "isole" les deux niveaux en harmonie avec la fréquence excitatrice ; c'est une condition *nécessaire* qui permet le cas échéant – moyennant quelques précautions – de faire l'approximation dite de l'atome à deux niveaux.

à ce cas ; par exemple, si l'on considère l'ionisation d'un atome, l'électron éjecté passe d'un état discret (lié) à un état du continuum (non-lié). L'émission spontanée d'un photon par un atome ou un noyau excité est un autre exemple de transition vers un continuum à partir d'un état discret, ou en tout cas d'énergie très bien définie^{27, 28}.

D'après les postulats, lorsqu'une variable prend des valeurs continues, on définit une *densité* de probabilité, avec laquelle il est possible de calculer des probabilités pour que la variable aléatoire prenne des valeurs situées dans un intervalle donné. Par exemple, si $\Psi(x, t)$ est la fonction d'onde ordinaire, $|\Psi(x, t)|^2$ est la densité de probabilité de présence ; alors, la probabilité $P_{[ab]}(t)$ de trouver la particule entre les deux points d'abscisses a et b est :

$$P_{[ab]}(t) = \int_a^b dx |\Psi(x, t)|^2 . \quad (7.67)$$

C'est bien l'intégrale (c'est-à-dire une *somme*) du module carré – non le module carré de l'intégrale – représentant la probabilité de trouver la particule entre a et b , obtenue en faisant la somme d'événements exclusifs (la particule est dans $[a, a + \delta x]$ ou dans $[a + \delta x, a + 2\delta x]$, ou dans $[a + 2\delta x, a + 3\delta x]$... ou dans $[b - \delta x, b]$).

De la même façon, pour mesurer l'impulsion \vec{p} d'une particule, il faut se référer à la représentation- p ; la fonction d'onde correspondante, $\Phi(\vec{p}, t)$, permet de définir une densité de probabilité pour les mesures de l'impulsion, soit $|\Phi(\vec{p}, t)|^2$. Le détecteur utilisé pour la mesure a une ouverture forcément finie, et donne un signal à chaque fois que l'impulsion est dans un domaine \mathcal{D} centré sur l'extrémité du vecteur \vec{p} choisi. Par exemple, on peut définir \mathcal{D} par un petit angle solide $\Delta\Omega$ autour d'une direction donnée, ce qui définit l'orientation du détecteur, et par un ΔE qui définit sa résolution en énergie autour d'une valeur moyenne nominale – ou, de façon équivalente, le module de l'impulsion. En tout état de cause, pour une valeur nominale p_f , la probabilité $P_{\vec{p}_f}$ d'obtenir un signal sur le détecteur s'obtient en sommant (intégrant) dans le domaine \mathcal{D} :

$$P_{\mathcal{D}}(t) \equiv P_{\vec{p}_f}(t) = \int_{\vec{p}-\vec{p}_f \in \mathcal{D}} d^3p |\Phi(\vec{p}, t)|^2 . \quad (7.68)$$

\mathcal{D} est un (petit) volume dans l'espace des impulsions, déterminé par le protocole expérimental, fixé dès que l'orientation du détecteur, son ouverture angulaire et sa fenêtre en énergie ont été choisies. Bien évidemment, l'intégrale sur le vecteur \vec{p} peut être remplacée par une intégration sur sa direction et son module, ce dernier étant relié à la seule énergie cinétique $E = \vec{p}^2/2m$ pour une particule libre dans l'état final. Autrement dit, on effectue un simple changement de variables en écrivant :

$$d^3p = p^2 dp d\Omega = (2mE) d(2mE)^{1/2} d\Omega = (2m^3 E)^{1/2} dE d\Omega \equiv \rho(E) dE d\Omega . \quad (7.69)$$

$\rho(E)$ s'appelle densité d'états en énergie ; $\rho(E)dE$ donne le nombre d'états dans l'intervalle $[E, E + dE]$. Dès lors, la probabilité précédente prend la forme :

$$P_{\mathcal{D}}(t) \equiv P_{E_f, \Delta\Omega}(t) = \int_{|E-E_f| \leq \Delta E, |\Omega-\Omega_f| \leq \Delta\Omega} \rho(E) dE d\Omega |\Phi(\vec{p}, t)|^2 . \quad (7.70)$$

Ces considérations se généralisent au cas où les états finals sont propres d'un ECOC contenant, outre l'énergie, des observables à spectre continu collectivement notées α (ci-dessus, α se réduit aux angles précisant la direction de \vec{p}_{p_f}). Dans ces conditions, en continuant à désigner le vecteur d'état par $|\Psi(t)\rangle$ – la fonction d'onde étant $\Psi(\alpha, E, t) = \langle \alpha, E | \Psi(t) \rangle$ –, la probabilité d'obtenir l'énergie E_f et les valeurs α_f à ΔE et $\Delta\alpha$ près peut toujours s'écrire comme l'intégrale d'un module carré :

$$P_{E_f \alpha_f}(t) = \int_{|E-E_f| \leq \Delta E, |\alpha-\alpha_f| \leq \Delta\alpha} dE d\alpha \rho(\alpha, E) |\langle \alpha, E | \Psi(t) \rangle|^2 ; \quad (7.71)$$

$|\langle \alpha, E | \Psi(t) \rangle|^2$ joue le rôle d'une *densité* de probabilité de transition vers l'état $|\alpha, E\rangle$, cependant que $\rho(\alpha, E)$ assure le décompte les états de même α et E . Ceci étant, le calcul de l'amplitude de transition, au premier

²⁷ On sait qu'il existe une relation du genre $\tau\delta\omega \sim 1$ entre la durée de vie τ de l'état excité et l'incertitude en énergie $\hbar\delta\omega$. Dans un cadre plus général (pour les particules élémentaires par exemple), un pic bien prononcé dans un spectre d'énergie est appelé *résonance*.

²⁸ Dans la situation la plus simple, le *continuum* est formé des états tensoriels de l'atome au fondamental par ceux d'un photon d'impulsion \vec{k} et d'énergie $\varepsilon_k = \hbar kc$. Si le champ est enfermé dans une boîte de dimension arbitrairement large, la quantification des modes du champ a lieu sur une échelle d'énergie "infiniment" petite.

ordre, entre un état initial discret $|\psi_i\rangle$ et un état du continuum $|\alpha, E\rangle$ démarre comme antérieurement : alors, aucune hypothèse quant à la nature (discrète ou continue) des états impliqués n'a été faite – sauf qu'ils sont propres de H_0 . Ainsi, dans les notations en cours, il faut trouver $\langle\alpha, E|\Psi(t)\rangle$, soit $\langle\alpha, E|U(t)|\psi_i\rangle$; au premier ordre, (7.43) donne :

$$|\langle\alpha, E|\Psi(t)\rangle| \equiv \langle\alpha, E|U(t)|\psi_i\rangle \simeq |A_{i \rightarrow f \equiv (\alpha, E)}^{(1)}| = \hbar^{-1} \left| \int_0^t dt_1 e^{\frac{i}{\hbar}(E-E_i)t_1} \langle\alpha, E|V(t_1)|\psi_i\rangle \right|. \quad (7.72)$$

Par exemple, pour une perturbation constante qui agit à partir de l'instant $t = 0$, on a :

$$|\langle\alpha, E|\Psi(t)\rangle| \simeq \hbar^{-1} \left| \langle\alpha, E|V|\psi_i\rangle \int_0^t dt_1 e^{\frac{i}{\hbar}(E-E_i)t_1} \right|. \quad (7.73)$$

Il en résulte que la densité de probabilité de transition est :

$$\hbar^{-2} |\langle\alpha, E|V|\psi_i\rangle|^2 \left| \int_0^t dt_1 e^{\frac{i}{\hbar}(E-E_i)t_1} \right|^2 \equiv |\langle\alpha, E|V|\psi_i\rangle|^2 f(t, \frac{E-E_i}{\hbar}), \quad (7.74)$$

où la fonction f a été définie en (7.49). La probabilité de transition entre $|\psi_i\rangle$ et le groupe d'états finals définis par les valeurs E_f et α_f à ΔE et $\Delta\alpha$ près est alors donnée par l'intégrale sommante sur toutes les valeurs de la fenêtre du détecteur :

$$P_{E_f \alpha_f}(t) = \hbar^{-2} \int_{|E-E_f| \leq \Delta E, |\alpha-\alpha_f| \leq \Delta\alpha} dE d\alpha \rho(\alpha, E) |\langle\alpha, E|V|\psi_i\rangle|^2 f(t, \frac{E-E_i}{\hbar}). \quad (7.75)$$

La fonction f varie très rapidement au voisinage de E_i et son intégrale en énergie vaut $2\pi\hbar t$. À l'inverse, puisque l'on est en présence d'un continuum d'états finals, la densité ρ et l'élément de matrice de V varient lentement, comparés à f : c'est précisément ce qui différencie un *continuum* d'énergie d'un spectre discret, où au contraire la densité a l'allure d'un peigne (irrégulier) de Dirac²⁹ (en toute rigueur d'ailleurs, d'un point de vue physique, il n'existe jamais de vrai *continuum* : il s'agit toujours d'une idéalisation (modélisation) d'une situation physique où les états sont si proches en énergie les uns des autres qu'il est exclu de les séparer en pratique³⁰). On peut donc, sur un strict plan opératoire, remplacer f par une fonction de Dirac³¹, soit précisément par $2\pi t \delta(E-E_i)$. En effet, dans un premier temps, on peut extraire ρ et l'élément de matrice de l'intégrale d'énergie, en les prenant à la valeur E_i où est centrée la fonction f :

$$P_{E_f \alpha_f}(t) \simeq \hbar^{-2} \int_{|\alpha-\alpha_f| \leq \Delta\alpha} d\alpha \rho(\alpha, E_i) |\langle\alpha, E_i|V|\psi_i\rangle|^2 \int_{|E-E_f| \leq \Delta E} dE f(t, \frac{E-E_i}{\hbar}). \quad (7.76)$$

Si l'énergie E_f n'est pas égale à E_i , l'intégrale de droite est quasiment nulle (schématiquement : on intègre un Dirac en dehors de son point de concentration). Au contraire, si $E_f \simeq E_i$ et si la largeur du détecteur ΔE est grande par rapport à \hbar/t , on peut rejeter les bornes sur l'énergie en $\pm\infty$ et on a à peu près :

$$\int_{|E-E_f| \leq \Delta E} dE f(t, \frac{E-E_i}{\hbar}) \simeq \int_{-\infty}^{+\infty} dE f(t, \frac{E-E_i}{\hbar}) = 2\pi\hbar t, \quad |E_f - E_i| \lesssim \mathcal{O}(\frac{\hbar}{t}), \quad \Delta E \gg \frac{\hbar}{t}, \quad (7.77)$$

d'où, dans ces conditions et avec $E_f \simeq E_i$:

$$P_{E_f \alpha_f}(t) \simeq \frac{2\pi}{\hbar} t \int_{|\alpha-\alpha_f| \leq \Delta\alpha} d\alpha \rho(\alpha, E_i) |\langle\alpha, E_i|V|\psi_i\rangle|^2. \quad (7.78)$$

Si $E_f - E_i$ est plus grand que \hbar/t , l'intégrale en énergie est pratiquement nulle. Dans la limite $t \rightarrow +\infty$, $P_{E_f \alpha_f}$ n'est strictement non nulle que si $E_f = E_i$: c'est la conservation de l'énergie.

²⁹ puisqu'il n'y a d'états que pour des valeurs infiniment précises de l'énergie.

³⁰ Voir par exemple le traitement ci-après de l'effet photoélectrique. La quantification de l'impulsion par application des conditions cycliques de Born - von Karman introduit une *discrétisation* des énergies que l'on peut rendre aussi petite que l'on veut, comparée à la résolution du détecteur le plus performant conceptuellement possible.

³¹ Dirac a précisément introduit sa "fonction" comme un objet purement opératoire : associé à une fonction lentement variable à l'échelle de ce qui va être modélisé par un δ , il sélectionne la valeur de la fonction lentement variable au point de concentration de δ .

De fait, l'expression (7.78) permet de définir une probabilité de transition par unité de temps, p , qui est cette fois indépendante du temps :

$$p_{E_f \alpha_f} \equiv \frac{P_{E_f \alpha_f}}{t} = \frac{2\pi}{\hbar} \int_{|\alpha - \alpha_f| \leq \Delta\alpha} d\alpha \rho(\alpha, E_i) |\langle \alpha, E_i | V | \psi_i \rangle|^2 . \quad (7.79)$$

Si on ne se préoccupe pas de la valeur de α dans l'état final (si par exemple on regarde seulement l'énergie), il faut intégrer sur α sans restriction (*i. e.* sur toutes les valeurs possibles de α) :

$$p_{E_f} = \frac{2\pi}{\hbar} \int_{\forall \alpha} d\alpha \rho(\alpha, E_i) |\langle \alpha, E_i | V | \psi_i \rangle|^2 , \quad |E_f - E_i| \lesssim \mathcal{O}\left(\frac{\hbar}{t}\right) . \quad (7.80)$$

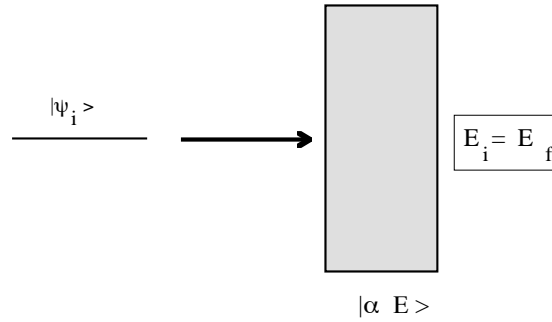


Figure 7.3: Illustration de la Règle d'or de Fermi avec une perturbation constante.

En général, il n'y a aucune raison pour que, après intégration, le résultat soit de la forme : une densité effective ρ multipliée par un élément de matrice. D'un autre côté, dans le cas où l'état final est caractérisé par sa seule énergie, il n'y a pas de grandeur du type α , donc pas d'intégration en α . Il reste alors :

$$p_{E_f} = \frac{2\pi}{\hbar} \rho(E_i) |\langle E_i | V | \psi_i \rangle|^2 , \quad |E_f - E_i| \lesssim \mathcal{O}\left(\frac{\hbar}{t}\right) , \quad (7.81)$$

d'une part ; d'autre part, si la condition sur l'énergie n'est pas satisfaite, la probabilité de transition est nulle dans la limite $t\Delta E \gg \hbar$. Le résultat (7.81) porte le nom de Règle d'or de Fermi et décrit – sur une grande échelle de temps à préciser dans chaque cas – ce qui se passe quand un état discret est couplé à un continuum d'états finals par une perturbation constante, “allumée” à partir d'un certain instant. Elle exprime d'une certaine façon la conservation de l'énergie : la probabilité est donnée par l'expression (7.81) quand l'état final a une énergie E_f qui est *égale* à celle de l'état de départ, E_i , et est nulle autrement ; en définitive, ceci peut se traduire en introduisant un *symbole* de Kronecker :

$$p_{E_f} = \frac{2\pi}{\hbar} \rho(E_i) |\langle \alpha_f, E_f | V | \psi_i \rangle|^2 \delta_{E_f E_i} . \quad (7.82)$$

Bien évidemment, le caractère *isoénergétique* de la transition possible – exprimé par le symbole de Kronecker dans (7.82) – résulte du fait que la perturbation V a été prise *constante*. Avec une perturbation monochromatique de pulsation ω , on trouve $\delta_{E_f E_i \pm \hbar\omega}$.

7.5 Applications exemplaires

Il s'agit ici de donner deux applications du formalisme précédent dans deux situations physiques importantes et constituant deux exemples de transitions soit entre états discrets, soit entre un état discret et un *continuum*.

7.5.1 Interaction d'un atome avec un champ électromagnétique (description semi-classique)

On va calculer³² au premier ordre les probabilités de transition induites dans un atome (à un seul électron, de masse μ) par un champ électromagnétique considéré comme extérieur³³ 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 zéro³⁴. Les deux champs sont :

$$\vec{\mathcal{E}} = \mathcal{E} \vec{a} \cos(\vec{k} \cdot \vec{r} - \omega t) , \quad \vec{\mathcal{B}} = \mathcal{B} \vec{b} \cos(\vec{k} \cdot \vec{r} - \omega t) , \quad (7.83)$$

\vec{a} et \vec{b} sont les polarisations des deux champs, $(\vec{b}, \vec{n} = \vec{k}/k, \vec{a})$ 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 :

$$\mathcal{E} = \mathcal{B}c , \quad (7.84)$$

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

$$\vec{A} = A_0 \vec{a} \sin(\vec{k} \cdot \vec{r} - \omega t) , \quad A_0 = \frac{\mathcal{E}}{\omega} = \frac{\mathcal{B}}{k} , \quad \mathcal{E} = \mathcal{B}c \quad (7.85)$$

restitue bien les deux champs (7.83) par les relations $\vec{\mathcal{E}} = -\partial \vec{A} / \partial t$ 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 :

$$H = \frac{1}{2\mu} (\vec{p} - e\vec{A})^2 + V(\vec{r}) - g_e \frac{e}{2\mu} \vec{S} \cdot \vec{\mathcal{B}} . \quad (7.86)$$

$V(\vec{r})$ est le potentiel intra-atomique et $g_e \simeq 2$. En développant l'expression (7.86) et en négligeant le terme diamagnétique quadratique en \vec{A} , on trouve :

$$H = \left[\frac{\vec{p}^2}{2\mu} + V(\vec{r}) \right] - \left[\frac{e}{2\mu} (\vec{A} \cdot \vec{p} + \vec{p} \cdot \vec{A}) + g_e \frac{e}{2\mu} \vec{S} \cdot \vec{\mathcal{B}} \right] \equiv H_0 + V(t) . \quad (7.87)$$

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} et \vec{p} commutent et on peut écrire :

$$V = V_1 + V_2 , \quad \text{avec : } V_1 = -\frac{e}{\mu} \vec{A} \cdot \vec{p} , \quad V_2 = -g_e \frac{e}{2\mu} \vec{S} \cdot \vec{\mathcal{B}} . \quad (7.88)$$

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

$$\frac{\langle V_2 \rangle}{\langle V_1 \rangle} \sim \frac{\langle \vec{S} \rangle \cdot \vec{\mathcal{B}}}{\vec{A} \cdot \langle \vec{p} \rangle} \sim \frac{\hbar k A_0}{A_0 p} = \frac{\hbar k}{p} \sim k a_0 \quad (7.89)$$

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 (7.85), on a :

$$V_1 = -\frac{e}{\mu} A_0 \vec{p} \cdot \vec{a} \sin(\vec{k} \cdot \vec{r} - \omega t) . \quad (7.90)$$

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 ky venant du produit scalaire $\vec{k} \cdot \vec{r}$ est

³² Voir [4], complément A_{XIII}.

³³ donc non-quantifié.

³⁴ [4], Appendice III, 4-b- α .

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 :

$$V_1 \simeq +\frac{e}{\mu} A_0 \vec{p} \cdot \vec{a} \sin \omega t \equiv V_{E1} . \quad (7.91)$$

Ceci constitue l'approximation dite dipolaire électrique³⁵ 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 :

$$\frac{d}{dt} \vec{r}_H = \frac{\vec{p}_H}{\mu} + \frac{e\mathcal{E}}{\mu\omega} \vec{a} \sin \omega t , \quad \frac{d}{dt} \vec{p}_H = -(\vec{\nabla}V)_H . \quad (7.92)$$

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 (7.91) 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, (7.91) résulte directement du choix de la jauge (7.85) (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 :

$$\langle \psi_f | V_{E1} | \psi_i \rangle = +\frac{e\mathcal{E}}{\mu\omega} \langle \psi_f | \vec{a} \cdot \vec{p} | \psi_i \rangle \sin \omega t \equiv +\frac{e\mathcal{E}}{\mu\omega} \langle \psi_f | p_z | \psi_i \rangle \sin \omega t . \quad (7.93)$$

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 :

$$i\hbar \frac{\vec{p}}{\mu} = [\vec{r}, H_0] , \quad (7.94)$$

on obtient³⁶ :

$$\langle \psi_f | V_{E1} | \psi_i \rangle = ie\mathcal{E} \frac{\omega_{fi}}{\omega} \sin \omega t \langle \psi_f | \vec{r} \cdot \vec{a} | \psi_i \rangle . \quad (7.95)$$

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 suffisant³⁷ 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 $R_{nl} Y_{lm}$; 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} :

$$z = r \sqrt{\frac{4\pi}{3}} Y_{10} , \quad x \pm iy = \mp r \sqrt{\frac{8\pi}{3}} Y_{1\pm 1} . \quad (7.96)$$

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

$$\int d\Omega Y_{l_f m_f}^* Y_{1m} Y_{l_i m_i} \quad (m = 0, \pm 1) . \quad (7.97)$$

³⁵approximation symbolisée par E1.

³⁶Par ailleurs, l'atome absorbe ou émet quand $\omega \simeq \omega_{fi}$ (résonance). En pratique, on récrit souvent l'expression (7.95) sous la forme simplifiée $\langle \psi_f | V_{E1} | \psi_i \rangle \simeq ie\mathcal{E} \sin \omega t \langle \psi_f | \vec{r} \cdot \vec{a} | \psi_i \rangle$. $e\vec{r}$ étant le moment dipolaire électrique \vec{d} , $\langle \psi_f | V_{E1} | \psi_i \rangle$ est finalement proportionnel à $\langle \psi_f | \vec{d} | \psi_i \rangle$, d'où la terminologie *transition dipolaire électrique*.

³⁷tant que l'on reste au premier ordre de la théorie des perturbations.

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 :

$$\Delta l = \pm 1, \quad \Delta m = 0, \pm 1. \quad (7.98)$$

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 (7.98) 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 (7.98) 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^{38, 39} :

$$V_1 \simeq V_{E1} - \frac{e\mathcal{B}}{\mu} (\vec{p} \cdot \vec{a}) (\vec{n} \cdot \vec{r}) \cos \omega t. \quad (7.99)$$

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 (7.99), l'écart à V_{E1} est d'ordre $(e\mathcal{B}/\mu)pr \sim (e\mathcal{B}/\mu)\hbar$, alors que $V_2 \sim (e/\mu)\hbar\mathcal{B}$. 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 (7.99) font notamment apparaître le moment cinétique orbital \vec{L} . Finalement, après regroupement de tous les termes, l'opérateur $V(t)$ complet (7.88) apparaît sous la forme approchée :

$$V(t) \simeq V_{E1} - \frac{e}{2\mu} (\vec{L} + g_e \vec{S}) \cdot \vec{\mathcal{B}} \cos \omega t - \frac{e}{2\mu\omega} [(\vec{k} \cdot \vec{r})(\vec{\mathcal{E}} \cdot \vec{p}) + (\vec{k} \cdot \vec{p})(\vec{\mathcal{E}} \cdot \vec{r})] \cos \omega t \equiv V_{E1} + V_{M1} + V_{E2}. \quad (7.100)$$

Le terme contenant les moments cinétiques est appelé dipolaire magnétique⁴⁰, 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 coordonnées⁴¹ 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 :

$$\Delta l = 0, \pm 2, \quad \Delta m = 0, \pm 1, \pm 2. \quad (7.101)$$

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.

7.5.2 Théorie élémentaire de l'effet photoélectrique

Il s'agit de calculer, au premier ordre, la probabilité de transition entre l'état fondamental $|\psi_i\rangle = |1s\rangle = |\psi_{n=1}\rangle$ et un état du *continuum* de l'atome d'hydrogène, sous l'effet d'un champ électromagnétique traité à nouveau

³⁸Au passage, on utilise la relation $kA_0 = B$.

³⁹ \vec{n} désigne un vecteur unitaire le long de \vec{k} .

⁴⁰Pour dipolaire magnétique, l'acronyme consacré est M1.

⁴¹On s'en convainc en utilisant (7.94).

classiquement[35]. Ceci constitue sans doute l'exemple le plus simple d'une transition entre un état discret (lié) et un état du *continuum* (non-lié), et constitue l'archétype de la fragmentation d'un système rendu instable par une action extérieure dynamique. Dans l'état final, l'électron n'est pas rigoureusement "libre" (il reste soumis au potentiel Coulombien), même s'il se trouve arbitrairement loin du noyau ; cependant, si l'énergie finale est grande à l'échelle de l'eV, on peut admettre que l'état Coulombien non-lié est en fait très voisin d'une onde plane. Selon Einstein, la réaction peut s'écrire :

$$\text{photon} + \text{électron lié} \rightarrow \text{électron non-lié} . \quad (7.102)$$

Le photon est absorbé par l'atome (il n'apparaît plus dans l'état final) ; en considérant l'électron à l'infini dans l'état final, la conservation de l'énergie se traduit par :

$$\hbar q c + E_I = \frac{1}{2}\mu\vec{v}_f^2 \equiv E_f \quad (q = \|\vec{q}\|) , \quad (7.103)$$

$\hbar\vec{q}$ est l'impulsion du photon incident, $\hbar q c \equiv \hbar\omega$ son énergie, E_I est l'énergie de l'électron quand il est lié dans l'état fondamental ($E_I = E_{n=1} \simeq -13.6$ eV). Pour un électron dans un métal, le seul changement à effectuer dans (7.103) consiste à remplacer E_I par $-W_S$, puisque le travail de sortie est traditionnellement défini comme une quantité positive.

On choisit à nouveau le potentiel-vecteur sous la forme (7.85) et on retient l'interaction la plus simple, V_1 , donnée par (7.90), que l'on réécrit sous la forme :

$$V_1(t) = w e^{i\omega t} + w^\dagger e^{-i\omega t} , \quad w = \frac{e}{2i\mu} A_0 \vec{a} \cdot \vec{p} e^{-i\vec{q} \cdot \vec{r}} , \quad w^\dagger = \frac{ie}{2\mu} A_0 \vec{a} \cdot \vec{p} e^{+i\vec{q} \cdot \vec{r}} , \quad \omega > 0 . \quad (7.104)$$

En toute rigueur, l'amplitude de transition (à un ordre donné de la théorie de perturbation) s'exprime à l'aide des éléments de matrice entre l'état $|1s\rangle$ et un état du *continuum* du potentiel Coulombien. Ces états sont connus, évidemment ; afin de s'en tenir à une description élémentaire, on va supposer que l'énergie des photons incidents est très grande devant l'énergie de liaison de l'électron, $E_{n=1}$. Alors, l'énergie finale de celui-ci est très voisine de $\hbar\omega$:

$$E_f \simeq \hbar\omega , \quad (7.105)$$

et, comme anticipé ci-dessus, il est alors licite de prendre pour état final de l'électron, $|\psi_f\rangle$, un état strictement libre, simple combinaison linéaire d'ondes planes dont le module du vecteur d'onde \vec{k} est fixé par⁴² :

$$E_f = \frac{\hbar^2 \vec{k}^2}{2\mu} . \quad (7.106)$$

La légitimité de ce choix repose sur le fait que les états non-liés fortement excités de l'atome d'hydrogène, s'ils ne sont pas strictement des ondes planes, doivent y ressembler fortement – pour des raisons physiques évidentes.

Le choix des ondes planes étant fait, une petite difficulté technique se présente, comme à chaque fois qu'il s'agit de manipuler des états non-normalisables. Pour traiter cette difficulté, différents trucs peuvent être utilisés. L'un d'entre eux consiste à imaginer que l'électron libéré reste toutefois confiné dans une boîte de dimension arbitrairement grande comparée aux autres longueurs déjà présentes dans le problème physique⁴³. Il est clair que les prévisions physiques ne sauraient dépendre de la taille de cette boîte⁴⁴, supposée en général cubique pour simplifier, de côté L .

On a alors une particule dans un puits (infini) de taille linéaire L , dont on sait que les états sont quantifiés ; leur énergie s'exprime à l'aide de trois entiers strictement positifs n_x , n_y et n_z :

$$E_{n_x n_y n_z} = \frac{\hbar^2 \vec{k}^2}{2\mu} = \frac{\hbar^2}{2\mu} (n_x^2 + n_y^2 + n_z^2) \frac{\pi^2}{L^2} \quad (n_u = 1, 2, 3, \dots) . \quad (7.107)$$

⁴²Si l'énergie du photon est supposée grande devant E_I , on suppose néanmoins que le photoélectron n'est pas relativiste.

⁴³Pour fixer les idées, cette boîte peut être choisie comme étant la pièce où se déroule l'expérience !

⁴⁴Il peut arriver que la véracité de cette affirmation repose sur des subtilités mises ici sous le tapis, notamment quand il existe des interactions à longue portée comme par exemple l'interaction de Coulomb nue, qui décroît comme $1/r$. Alors, il n'est pas forcément facile de trouver une échelle de longueur propre au problème traité, ou, plus précisément, cette échelle peut être formellement infinie. En pratique, ou bien on tient compte d'effets physiques ignorés jusqu'alors – et fabriquant automatiquement des effets d'écran –, ou bien on introduit une fonction de coupure formelle (pour le Coulombien, on prend souvent un potentiel de Yukawa, $e^{-r/\xi}/r$), et on prend astucieusement la limite (ξ infini, L infini).

L étant aussi grand que l'on veut par rapport à une distance typique du problème (ici, a_0 par exemple), il est bien clair que sur une échelle d'énergie atomique, ces énergies – formellement quantifiées – sont extrêmement rapprochées les unes des autres et constituent un quasi-*continuum* d'énergie⁴⁵. Par ailleurs, pour le puits cubique infini situé dans l'octant positif de \mathbb{R}^3 et dont l'un des sommets est l'origine du repère, on sait que les états propres sont :

$$\psi_{n_x n_y n_z} = \left(\frac{2}{L}\right)^{3/2} \sin n_x \frac{\pi x}{L} \sin n_y \frac{\pi y}{L} \sin n_z \frac{\pi z}{L} . \quad (7.108)$$

La pratique montre que ces fonctions sont d'un usage peu agréable et conduisent à des calculs inutilement compliqués – puisque de toute façon, on s'attend à ce que ce qui se passe aux bords de la boîte de confinement n'ait aucune pertinence physique. Pour cette raison, on utilise le plus souvent des ondes planes⁴⁶ $\psi_{\vec{k}}(\vec{r}) = C e^{i\vec{k}\cdot\vec{r}}$, définies dans une boîte cubique de côté L et auxquelles on impose d'autres conditions aux limites, dites de Born - von Karman. Ces conditions s'énoncent comme suit :

$$e^{-ik_x L/2} = e^{+ik_x L/2} , \quad e^{-ik_y L/2} = e^{+ik_y L/2} , \quad e^{-ik_z L/2} = e^{+ik_z L/2} . \quad (7.109)$$

Chacune d'entre elles identifie un point d'abscisse $-L/2$ avec le point d'abscisse $+L/2$ situé sur la face opposée du cube, comme si les deux faces parallèles étaient juxtaposées (pour cette raison, les conditions (7.109) sont dites *cycliques*). À une dimension, ceci revient, au lieu de considérer un grand segment de droite s'étendant entre les abscisses $\pm L/2$, à refermer ce segment en boucle ; si L est assez grand devant toutes les autres longueurs, la courbure artificielle ainsi introduite est sans effet et l'espace est localement “droit”.

Il résulte immédiatement de (7.109) qu'il existe trois entiers quelconques (pris maintenant dans \mathbb{Z} et non plus dans \mathbb{N}) donnant les composantes (quantifiées !) du vecteur \vec{k} :

$$k_x = n_x \frac{2\pi}{L} , \quad k_y = n_y \frac{2\pi}{L} , \quad k_z = n_z \frac{2\pi}{L} , \quad (7.110)$$

de sorte que l'espace du vecteur d'onde est discrétisé en petits cubes, tout comme dans le cas du puits infini. La quantification de \vec{k} entraîne *ipso facto* celle de l'énergie, donnée maintenant par l'expression suivante :

$$E_{n_x n_y n_z} = \frac{\hbar^2 \vec{k}^2}{2\mu} = \frac{\hbar^2}{2\mu} (n_x^2 + n_y^2 + n_z^2) \frac{(2\pi)^2}{L^2} \quad (n_u = 0, \pm 1, \pm 2, \pm 3, \dots) . \quad (7.111)$$

Maintenant, le vecteur d'onde n'est plus restreint à l'octant positif. La comparaison des deux expressions (7.107) et (7.111) montre un facteur 4 d'écart, mais cette différence n'en induit aucune au niveau des résultats. En effet, dans le cas des sinus, le nombre d'états $n_{\text{sinus}}(k)$ dont le vecteur d'onde a un module inférieur k et contenus dans l'octant positif s'obtient divisant le volume de la sphère de rayon k par le volume d'un petit cube de côté π/L , soit $\delta^3 k = (\pi/L)^3$. On a donc :

$$n_{\text{sinus}}(k) = \frac{1}{8} \frac{(4\pi/3) k^3}{(\pi/L)^3} = \frac{(kL)^3}{6\pi^2} . \quad (7.112)$$

Pour les ondes planes, il faut prendre la sphère entière, mais le volume élémentaire $\delta^3 k$ est maintenant $(2\pi/L)^3$ et on récupère le facteur 1/8. Au total, dans les sommations qui deviennent des intégrales, rien ne sera changé. Enfin, alors que la valeur $n_u = 0$ est exclue pour les sinus (la fonction d'onde serait identiquement nulle), rien ne l'interdit *a priori* pour les ondes planes. À la limite, l'adjonction d'un état “de mesure nulle” ne modifie pas les résultats. En définitive, les sinus et les ondes planes produisent les mêmes prédictions physiques.

Par construction l'espace \vec{k} est alors discrétisé en petits cubes $\delta^3 k$ de volume $(2\pi/L)^3$, formant chacun ce que l'on peut appeler un “gros point”⁴⁷. Ceci veut dire que toutes les sommations sur le vecteur \vec{k} seront en temps utile remplacées par des intégrales, la somme discrète à L fini tendant vers une somme de Darboux ordinaire, point de départ servant à définir, par passage à la limite, l'intégrale de Riemann. En pratique :

$$\sum_{\vec{k}} (\dots) \equiv \frac{1}{\delta^3 k} \sum_{\vec{k}} (\dots) \delta k \rightarrow \left(\frac{L}{2\pi}\right)^3 \int_{\mathbb{R}^3} (\dots) d^3 k . \quad (7.113)$$

⁴⁵Si on prend $L = 1$ m, l'écart entre les premiers niveaux du puits infini (qui sont déjà les moins denses) est de l'ordre de 10^{-20} eV !

⁴⁶Le choix de la constante de normalisation sera fait d'ici peu.

⁴⁷Ce type de pavage a été utilisé à propos du rayonnement thermique (voir cours de Licence).

Le facteur $[L/(2\pi)]^3$, inverse de $\delta^3 k$, est nécessaire ne serait-ce que pour des raisons d'homogénéité (les (...) dans les deux membres ont la même dimension, $d^3 k$ a la dimension (longueur)⁻³); schématiquement, le rapport $d^3 k/\delta^3 k$ tend vers 1 à la limite $L \rightarrow +\infty$. Bien évidemment, si tous les calculs sont faits convenablement, la longueur L doit disparaître tôt ou tard⁴⁸, pour ne produire finalement que des prévisions indépendantes des conditions aux limites précises, choisies uniquement en fonction de la commodité. Enfin, et c'est le plus important, la densité d'états en énergie, $\rho(E)$, est la même, que l'on prenne des sinus comme en (7.108) ou des ondes planes. Ce point est essentiel puisque, en temps utile, on remplacera les sommations discrètes par des intégrales où apparaît ρ .

Cette discrétisation étant admise, il n'y a plus de difficulté technique. En particulier, chaque onde plane peut maintenant être normalisée sans problème :

$$\int_{\text{boîte}} d^3 r |\psi_{\vec{k}}(\vec{r})|^2 = 1 \iff C = L^{-3/2} \implies \psi_{\vec{k}}(\vec{r}) = L^{-3/2} e^{i\vec{k}\cdot\vec{r}} . \quad (7.114)$$

Au premier ordre, la probabilité de transition de $|\psi_1\rangle$ vers $|\psi_{\vec{k}}\rangle$ est :

$$P_{\vec{k}} = |A_{1\rightarrow\vec{k}}^{(1)}|^2 , \quad (7.115)$$

où $A_{1\rightarrow\vec{k}}^{(1)}$ est donné par (7.43), d'où :

$$|A_{1\rightarrow\vec{k}}^{(1)}| = \hbar^{-1} \left| \int_0^t dt_1 e^{\frac{i}{\hbar}(E_{\vec{k}}-E_1)t_1} \langle \psi_{\vec{k}} | (w e^{i\omega t} + w^\dagger e^{-i\omega t}) | \psi_1 \rangle \right| . \quad (7.116)$$

Après intégration sur t_1 et élévation du module au carré, deux fonctions du type $f(T, \Omega)$ apparaissent ; la première vient du terme w et introduit $\Omega = \hbar^{-1}(E_{\vec{k}} - E_1) + \omega$, soit $f(t, \hbar^{-1}(E_{\vec{k}} - E_1) + \omega)$, la seconde vient de w^\dagger et donne $f(t, \hbar^{-1}(E_{\vec{k}} - E_1) - \omega)$. Comme $\hbar\omega > 0$ et que $E_{\vec{k}} \gg E_1$, seul le terme contenant la deuxième fonction va contribuer en pratique au module carré de l'amplitude, d'où :

$$P_{\vec{k}} \simeq |A_{1\rightarrow\vec{k}}^{(1)}|^2 = \hbar^{-2} |\langle \psi_{\vec{k}} | w^\dagger | \psi_1 \rangle|^2 f\left(t, \frac{E_{\vec{k}} - E_1}{\hbar} - \omega\right) . \quad (7.117)$$

Ceci est la probabilité élémentaire de trouver l'électron éjecté avec une impulsion \vec{k} dont l'extrémité appartient au volume "infiniment petit" $\delta^3 k$. En pratique, le détecteur a une ouverture $\Delta^3 k$ assez petite pour avoir une bonne résolution, mais suffisamment grande pour produire un signal significatif et en tout cas certainement beaucoup plus grande que $\delta^3 k$ – qui peut être rendu petit *ad libitum*. Alors, le signal enregistré sur le détecteur est la somme ΔP des probabilités élémentaires associées aux événements satisfaisant $\|\vec{k} - \vec{k}_f\| \leq \Delta^3 k$, \vec{k}_f étant la valeur nominale de réglage du détecteur. P_f est :

$$P_f = \sum_{\vec{k}, \|\vec{k} - \vec{k}_f\| \leq \Delta^3 k} |A_{1\rightarrow\vec{k}}^{(1)}|^2 . \quad (7.118)$$

La somme sur \vec{k} court par petits pas – ce sont des infiniment petits au sens du Physicien – et peut être remplacée par une intégrale, conformément à (7.113). Il vient ainsi :

$$P_f = \left(\frac{L}{2\pi}\right)^3 \int_{\|\vec{k} - \vec{k}_f\| \leq \Delta^3 k} d^3 k \hbar^{-2} |\langle \psi_{\vec{k}} | w^\dagger | \psi_1 \rangle|^2 f\left(t, \frac{E_{\vec{k}} - E_1}{\hbar} - \omega\right) . \quad (7.119)$$

P_f est bien un nombre pur : f est homogène au carré d'un temps, l'élément de matrice est une énergie au carré. Selon les expressions (7.114) et (7.104), cet élément de matrice s'écrit précisément :

$$\langle \psi_{\vec{k}} | w^\dagger | \psi_1 \rangle = \int_{\text{boîte}} d^3 r L^{-3/2} e^{-i\vec{k}\cdot\vec{r}} \frac{ie}{2\mu} A_0 \vec{a}\cdot\vec{p} e^{i\vec{q}\cdot\vec{r}} \psi_{100}(\vec{r}) , \quad (7.120)$$

⁴⁸sauf, bien sûr, si on calcule des grandeurs extensives.

soit, en remplaçant \vec{p} par $-i\hbar\vec{\nabla}$ et compte tenu du fait⁴⁹ que $\vec{a}\cdot\vec{q} = 0$:

$$\langle \psi_{\vec{k}} | w^\dagger | \psi_1 \rangle = \frac{e\hbar}{2\mu} A_0 L^{-3/2} \int_{\text{boîte}} d^3r e^{+i(\vec{q}-\vec{k})\cdot\vec{r}} \vec{a}\cdot\vec{\nabla} \psi_{100}(\vec{r}) . \quad (7.121)$$

En rapprochant mentalement (7.119) et (7.121), on voit qu'à cet endroit précis, tous les facteurs contenant L se compensent : on peut donc déjà anticiper la limite $\lim_{L \rightarrow +\infty}$ ici et là. Une intégration par parties (le terme tout intégré est nul) permet d'obtenir :

$$\langle \psi_{\vec{k}} | w^\dagger | \psi_1 \rangle = \frac{e\hbar}{2i\mu} A_0 L^{-3/2} (2\pi\hbar)^{3/2} (\vec{q} - \vec{k})\cdot\vec{a} \phi_{100}(\vec{p} = \hbar\vec{K}) , \quad (7.122)$$

où ϕ_{100} est la représentation- p de l'état $1s$:

$$\phi_{100}(\vec{p}) = (2\pi\hbar)^{-3/2} \int_{\mathbb{R}^3} d^3r e^{+(\vec{p}\cdot\vec{r})/(i\hbar)} \psi_{100}(\vec{r}) \equiv \chi_{10}(p) Y_{00} , \quad (7.123)$$

et où :

$$\vec{K} = \vec{k} - \vec{q} \quad (7.124)$$

est le transfert de moment du photon vers l'électron éjecté de l'atome. $\chi_{10}(p)$ se calcule facilement et est donné par :

$$\chi_{10}(p) = \frac{16\pi}{(2\pi\hbar)^{3/2}} \frac{a_0^{3/2}}{[1 + (pa_0/\hbar)^2]^2} . \quad (7.125)$$

Avec tout ceci, (7.119) devient :

$$P_f = \frac{8}{\pi} \left(\frac{eA_0}{\mu} \vec{a}\cdot\vec{K} \right)^2 a_0^3 \int_{\|\vec{k}-\vec{k}_f\| \leq \Delta^3 k} k^2 dk d\Omega \frac{1}{(1 + a_0^2 K^2)^4} f(t, \omega_k - \omega_i - \omega) . \quad (7.126)$$

Compte tenu de $\omega_k = \hbar\vec{k}^2/(2\mu)$ et de $\omega_k \gg \omega_i$, on a :

$$f(t, \omega_k - \omega_i - \omega) \simeq f(t, \omega_k - \omega) \simeq 2\pi t \delta(\omega_k - \omega) = \frac{2\pi\mu}{\hbar k} t \delta\left(k - \sqrt{2\mu\omega/\hbar}\right) . \quad (7.127)$$

Au passage, on a supposé t suffisamment grand pour que $\delta\omega \equiv 1/t$ soit aussi petit que l'on veut, ce qui permet (comme on l'a fait pour établir la Règle d'or de Fermi) de remplacer le filtre approximatif f par une fonction δ convenablement ajustée (c'est-à-dire conformément à (7.52)). En reportant (7.127) dans (7.126), il vient :

$$P_f = 16 t \frac{[eA_0 \vec{a}\cdot(\vec{k}_f - \vec{q})]^2}{\mu\hbar} a_0^3 \int_{\|\vec{k}-\vec{k}_f\| \leq \Delta^3 k} \frac{k dk d\Omega}{[1 + a_0^2(k - q)^2]^4} \delta\left(k - \sqrt{2\mu\omega/\hbar}\right) . \quad (7.128)$$

Maintenant, de deux choses l'une :

- ou bien l'énergie du photon $\hbar\omega$ est différente de $\hbar^2 k_f^2/(2\mu)$ – énergie nominale de la fenêtre du détecteur – et alors le domaine d'intégration de k est en-dehors du point de concentration de la fonction δ ; dans ces conditions, l'intégrale est nulle et $P_f = 0$.
- ou bien l'énergie du photon est égale à la valeur affichée sur le détecteur et la fonction δ agit comme d'habitude.

Ceci traduit la conservation de l'énergie. Dans le second cas, la probabilité est :

$$P_f = 16 t \frac{[eA_0 \vec{a}\cdot(\vec{k}_f - \vec{q})]^2}{\mu\hbar} a_0^3 \frac{k_f}{[1 + a_0^2(k_f - q)^2]^4} \Delta\Omega . \quad (7.129)$$

⁴⁹ce qui exprime la commutation entre \vec{p} et \vec{A} , aussi traduite par $\vec{\nabla}\cdot\vec{A} = 0$.

Il est agréable d'éliminer le potentiel-vecteur, pour ne faire apparaître que des quantités ayant un sens physique direct. L'énergie du champ dans la boîte de volume V est⁵⁰ :

$$E_{\text{champ}} = \frac{\epsilon_0}{2} \int_V d^3r \left(\vec{\mathcal{E}}^2 + \frac{1}{\epsilon_0 \mu_0} \vec{\mathcal{B}}^2 \right) = \epsilon_0 \int_V d^3r \vec{\mathcal{E}}^2 = N \hbar \omega . \quad (7.130)$$

La dernière égalité introduit N , nombre de photons d'énergie $\hbar \omega$ dans le (grand) volume $V = L^3$. Par ailleurs :

$$\vec{\mathcal{E}} = -\frac{\partial \vec{A}}{\partial t} = \omega A_0 \vec{a} \cos(\vec{k} \cdot \vec{r} - \omega t) , \quad \bar{\vec{\mathcal{E}}}^2 = \frac{1}{2} \omega^2 A_0^2 , \quad (7.131)$$

où la barre horizontale désigne une moyenne dans le temps sur une période. La dernière égalité de (7.130) devient alors :

$$\epsilon_0 V \frac{1}{2} \omega^2 A_0^2 = N \hbar \omega \quad (7.132)$$

et permet d'exprimer A_0^2 en fonction de la densité volumique de photons, N/V :

$$A_0^2 = \frac{2 \hbar}{\epsilon_0 \omega} \frac{N}{V} . \quad (7.133)$$

En reportant ceci dans (7.129), il vient finalement :

$$P_{\text{f}} = 32 t \frac{e^2 a_0^3 k_{\text{f}} \left[\vec{a} \cdot (\vec{k}_{\text{f}} - \vec{q}) \right]^2}{\epsilon_0 \mu \omega [1 + a_0^2 (k_{\text{f}} - q)^2]^4} \frac{N}{V} \Delta \Omega . \quad (7.134)$$

On note que le numérateur apporte une dépendance angulaire pour les photoélectrons. Il est d'ailleurs possible d'oublier \vec{q} , qui est très petit comparé à \vec{k}_{f} ; en effet :

$$\frac{\hbar^2 k_{\text{f}}^2}{2\mu} = \hbar \omega = \hbar q c \iff \frac{q}{k_{\text{f}}} = \frac{\hbar k_{\text{f}}}{2\mu c} = \frac{v_{\text{f}}}{2c} \ll 1 , \quad (7.135)$$

puisque les électrons sont supposés non-relativistes. D'où la forme simplifiée :

$$P_{\text{f}} \simeq 32 t \frac{e^2 a_0^3 k_{\text{f}} (\vec{a} \cdot \vec{k}_{\text{f}})^2}{\epsilon_0 \mu \omega [1 + a_0^2 k_{\text{f}}^2]^4} \frac{N}{V} \Delta \Omega . \quad (7.136)$$

La probabilité de transition par unité de temps est simplement $p_{\text{f}} = P_{\text{f}}/t$. En utilisant le fait que :

$$(a_0 k_{\text{f}})^2 = \left(\frac{\hbar^2}{\mu e^2} \right)^2 \frac{2\mu E_{\text{f}}}{\hbar^2} = \frac{E_{\text{f}}}{|E_{n=1}|} \gg 1 , \quad \hbar \omega \simeq E_{\text{f}} , \quad (7.137)$$

on peut mettre p_{f} sous la forme approchée :

$$p_{\text{f}} \simeq \frac{32}{\pi^2} \frac{c}{r_0} \lambda_{\text{C}}^3 \cos^2 \theta \left(\frac{|E_1|}{E_{\text{f}}} \right)^{5/2} \frac{N}{V} \Delta \Omega , \quad (7.138)$$

où r_0 et λ_{C} sont le rayon classique et la longueur d'onde Compton de l'électron :

$$r_0 = \frac{e^2}{\mu c^2} \simeq 3 \text{ F} , \quad \lambda_{\text{C}} = \frac{\hbar}{\mu c} \simeq 0.024 \text{ \AA} . \quad (7.139)$$

En définitive, (7.138) est l'expression de la probabilité de transition photoélectrique par unité de temps pour un \vec{k}_{f} à $\Delta \Omega$ près. θ est l'angle entre le vecteur d'onde final \vec{k}_{f} du photoélectron et la polarisation du champ électrique \vec{E} . Sans surprise, les électrons émis dans la direction du champ sont les plus nombreux : ce sont ceux qui vibrent le plus fortement sous l'effet de ce dernier ; quantitativement, le diagramme polaire de rayonnement est donné par $\cos^2 \theta$.

⁵⁰Tenir compte de (7.85).

7.6 Rudiments sur la description purement quantique de l'interaction champ - matière

Jusqu'à présent, le champ électromagnétique a été considéré comme un champ purement classique agissant de l'extérieur sur un système quantique (atome, par exemple). Cette description semi-classique est suffisante pour certaines applications mais est incapable de rendre compte de phénomènes importants, comme par exemple l'émission spontanée⁵¹. Un atome, un noyau, tout système porté dans un état excité n'y reste pas indéfiniment : même si le champ qui a excité le système a disparu entre temps, le système redescend au bout d'un certain temps à l'état fondamental, éventuellement en plusieurs étapes, c'est-à-dire en transitant par d'autres niveaux excités. Ce temps est foncièrement aléatoire ; sa moyenne⁵² est conventionnellement appelée durée de vie de l'état excité.

D'ailleurs, sur le plan des principes, on se doute qu'il va être tôt ou tard nécessaire de quantifier le champ électromagnétique, afin de construire un cadre théorique où champ et matière sont traités sur un même pied d'égalité. Il s'agit même d'un juste retour des choses : après tout, le rayonnement a joué un rôle crucial pour la genèse de la Théorie Quantique, qu'il s'agisse du rayonnement du corps noir, de l'effet photoélectrique, de l'effet Compton, des spectres atomiques, *etc.*

Il s'agit ici de présenter quelques idées permettant d'ébaucher une théorie complète du système couplé photons - matière. Le système quantique ainsi défini, supposé isolé, est décrit par un Hamiltonien que l'on écrit sous la forme⁵³ :

$$H = H_{\text{at}} + H_{\text{champ}} + V \equiv H_0 + V . \quad (7.140)$$

H_{at} est le Hamiltonien pris en compte pour la description semi-classique⁵⁴, alors noté H_0 , H_{champ} est le Hamiltonien du champ libre ; quant à V , c'est le terme qui décrit l'interaction entre les deux parties du système. Par construction, l'atome et le champ forment un système isolé : H est indépendant du temps. De fait, l'analyse suivante fera apparaître des expressions réminiscentes de la théorie des perturbations stationnaires (par exemple (7.178)). Toutefois, l'objectif final est d'aller au-delà d'un schéma perturbatif afin d'obtenir, notamment, la décroissance exponentielle de la population d'un niveau atomique excité.

La première tâche à accomplir est d'écrire H_{champ} , procédant comme à chaque fois que la quantification est en chantier : en s'inspirant de la description Hamiltonienne classique et en posant des relations de commutation après avoir défini les opérateurs représentant les grandeurs physiques. On s'attend à voir émerger les photons, grains de lumière se déplaçant à la vitesse c ; il n'existe pas de référentiel où le photon est au repos et ce fait permet de suspecter que le formalisme en fonction d'onde définie dans l'espace des coordonnées (représentation- q) risque d'être mal adapté – à supposer d'ailleurs qu'il existe. En revanche, cela a un sens de parler d'un photon d'énergie ou d'impulsion bien définie. On pressent ainsi qu'il convient d'abord de reformuler les équations classiques non dans l'espace réel mais dans l'espace "réciproque" des impulsions. Cette "relecture quantique" des équations de Maxwell [38] – qui, au passage, montre qu'il n'est pas possible en effet de définir une fonction d'onde en représentation- q pour le photon – permet d'écrire la composante électrique du champ (classique) sous la forme d'une intégrale sur le vecteur d'onde⁵⁵ \vec{k} :

$$\vec{\mathcal{E}}(\vec{r}, t) = i \int d^3k \sqrt{\frac{\hbar\omega}{2\epsilon_0(2\pi)^3}} \left[\vec{\alpha}(\vec{k}, t) e^{i\vec{k}\cdot\vec{r}} - \vec{\alpha}^*(\vec{k}, t) e^{-i\vec{k}\cdot\vec{r}} \right] . \quad (7.141)$$

En dépit de la présence de \hbar , il s'agit bien d'une expression purement classique ; \hbar apparaît dans la dérivation de (7.141) par l'usage de la relation entre énergie et pulsation, $E = \hbar\omega$. $\vec{\alpha}(\vec{k}, t)$ est un vecteur qui s'exprime en

⁵¹La description semi-classique rend évidemment bien compte de l'émission *induite*. Dans le domaine optique, l'émission induite est en général beaucoup plus lente que l'émission spontanée et est souvent négligée de ce fait. La situation est inversée dans la gamme des microondes.

⁵²On suppose qu'elle est finie. Évidemment, il existe des variables aléatoires distribuées suivant des lois larges et qui sont dépourvues de moyenne (et *a fortiori* d'écart-type).

⁵³Attention : dans (7.140), H_0 est le Hamiltonien de l'atome *et* du champ, sans interaction.

⁵⁴à des effets de "renormalisation" près.

⁵⁵Dans cette section, $\hbar\vec{k}$ désigne l'impulsion du photon.

combinaison linéaire de $\vec{\mathcal{E}}$ et $\dot{\vec{\mathcal{E}}}$ et qui satisfait l'équation :

$$i \frac{\partial}{\partial t} \vec{\alpha}(\vec{k}, t) = \omega \vec{\alpha}(\vec{k}, t) \iff i \hbar \frac{\partial}{\partial t} \vec{\alpha}(\vec{k}, t) = \hbar \omega \vec{\alpha}(\vec{k}, t) . \quad (7.142)$$

Ces équations rappellent les équations de Heisenberg pour les opérateurs de création et d'annihilation a et a^\dagger d'un oscillateur harmonique à une dimension, dont le Hamiltonien s'écrit :

$$H_{\text{osc.harm.}} = \hbar \omega \left(a^\dagger a + \frac{1}{2} \right) \equiv \hbar \omega \left(N + \frac{1}{2} \right) . \quad (7.143)$$

N est l'opérateur nombre de particules ; dans un état propre $|n\rangle$ de $H_{\text{osc.harm.}}$, on a :

$$\langle n | H_{\text{osc.harm.}} | n \rangle = n \hbar \omega , \quad (7.144)$$

à la constante additive $\hbar \omega / 2$ près⁵⁶. En outre, l'opérateur d'annihilation a agissant sur l'état fondamental $n = 0$ (état "vide") donne zéro :

$$a | n = 0 \rangle = 0 . \quad (7.145)$$

De fait, les vecteurs classiques $\vec{\alpha}(\vec{k}, t)$ et $\vec{\alpha}^*(\vec{k}, t)$ vont devenir en Mécanique Quantique les opérateurs de création et d'annihilation d'un photon transverse ; celui-ci est caractérisé par son impulsion $\hbar \vec{k}$ et sa polarisation $\vec{\varepsilon}$ et on notera précisément $a_{\vec{k}\vec{\varepsilon}}$ et $a_{\vec{k}\vec{\varepsilon}}^\dagger$ les opérateurs correspondants. Dès lors, le Hamiltonien (quantique) du champ se met sous la forme :

$$H_{\text{champ}} = \sum_{\vec{k}\vec{\varepsilon}} \hbar \omega_{\vec{k}\vec{\varepsilon}} a_{\vec{k}\vec{\varepsilon}}^\dagger a_{\vec{k}\vec{\varepsilon}} . \quad (7.146)$$

Deux différences sautent aux yeux entre (7.143) et (7.146). La première est l'adjonction d'une sommation sur les modes $(\vec{k}, \vec{\varepsilon})$, qui épuise tous les types de photons possibles ; la seconde, très mineure, est la suppression de la constante additive $\hbar \omega / 2$ dans l'énergie de chaque mode, donnant au total l'un de ces infinis dont on peut se débarrasser puisque seules les différences d'énergie comptent. D'un autre point de vue, l'expression (7.146) ne doit pas surprendre : on sait bien que le décompte des modes du champ électromagnétique (classique) fait apparaître celui-ci comme un ensemble d'oscillateurs harmoniques indépendants⁵⁷ ; (7.146) n'exprime pas autre chose : c'est une somme sur l'énergie de *chaque* oscillateur (mode), caractéristique de l'absence de termes à deux corps. Notons que, en raison de l'isotropie, l'énergie $\hbar \omega_{\vec{k}\vec{\varepsilon}}$ est indépendante de la direction de \vec{k} et de la polarisation ε ; elle sera désormais notée $\hbar \omega_k$ et contient la loi de dispersion du photon $\omega_k = kc$, soit $E_k = \hbar kc$, typique d'une particule relativiste de masse nulle. Enfin, les opérateurs $a_{\vec{k}\vec{\varepsilon}}$ et $a_{\vec{k}\vec{\varepsilon}}^\dagger$ satisfont les relations de commutation :

$$[a_{\vec{k}\vec{\varepsilon}}, a_{\vec{k}'\vec{\varepsilon}'}^\dagger] = \delta_{\vec{k}\vec{k}'} \delta_{\vec{\varepsilon}\vec{\varepsilon}'} \mathbf{1} \quad (7.147)$$

et les équations de Heisenberg :

$$i \frac{d}{dt} a_{\vec{k}\vec{\varepsilon}\text{H}} = \omega_k a_{\vec{k}\vec{\varepsilon}\text{H}} , \quad i \frac{d}{dt} a_{\vec{k}\vec{\varepsilon}\text{H}}^\dagger = -\omega_k a_{\vec{k}\vec{\varepsilon}\text{H}}^\dagger . \quad (7.148)$$

Ces deux équations s'intègrent en :

$$a_{\vec{k}\vec{\varepsilon}\text{H}}(t) = e^{-i\omega_k t} a_{\vec{k}\vec{\varepsilon}} , \quad a_{\vec{k}\vec{\varepsilon}\text{H}}^\dagger(t) = e^{+i\omega_k t} a_{\vec{k}\vec{\varepsilon}}^\dagger . \quad (7.149)$$

En revenant maintenant à (7.141), il est alors possible d'écrire l'expression quantique du champ électrique au sens de Heisenberg :

$$\vec{\mathcal{E}}_{\text{H}}(\vec{r}, t) = i \sum_{\vec{\varepsilon}} \int d^3k \sqrt{\frac{\hbar \omega_k}{2\epsilon_0 (2\pi)^3}} \left[a_{\vec{k}\vec{\varepsilon}} e^{i\vec{k}\cdot\vec{r}} e^{-i\omega_k t} - a_{\vec{k}\vec{\varepsilon}}^\dagger e^{-i\vec{k}\cdot\vec{r}} e^{i\omega_k t} \right] \vec{\varepsilon} . \quad (7.150)$$

⁵⁶On peut toujours oublier une constante additive, puisque ce sont les *différences* d'énergie qui ont un sens physique. On admet que cette pratique demeure légitime même quand la constante en question est infinie (quand on considère une infinité de modes harmoniques)...

⁵⁷Deux modes du champ classique sont sans interaction directe et forment un "gaz parfait". La matière est un intermédiaire obligé pour que deux photons puissent se voir ou avoir une filiation – d'où la nécessité d'un peu de matière pour assurer l'équilibre thermodynamique du champ, tout comme les collisions sont essentielles au maintien de l'équilibre d'un gaz parfait usuel (ou à sa relaxation).

Par $\vec{\mathcal{E}} = -\partial\vec{A}/\partial t$, on en déduit le potentiel-vecteur transverse :

$$\vec{A}_H(\vec{r}, t) = \sum_{\vec{\varepsilon}} \int d^3k \sqrt{\frac{\hbar}{2\epsilon_0(2\pi)^3\omega_k}} \left[a_{\vec{k}\vec{\varepsilon}}^\dagger e^{-i\vec{k}\cdot\vec{r}} e^{i\omega_k t} + a_{\vec{k}\vec{\varepsilon}} e^{i\vec{k}\cdot\vec{r}} e^{-i\omega_k t} \right] \vec{\varepsilon} ; \quad (7.151)$$

de cette expression, on trouve immédiatement le champ magnétique par $\vec{\mathcal{B}} = \vec{\nabla} \times \vec{A}$.

L'expression (7.146) permet d'affirmer que les états propres du champ libre (en l'absence de matière) est le produit tensoriel des états du type oscillateur harmonique, chaque oscillateur pouvant avoir une infinité de fréquences propres ω_k , selon la valeur de k ($\omega_k = kc$) :

$$|\text{état propre du champ}\rangle = |n_{\vec{k}_1\vec{\varepsilon}_1}\rangle \otimes |n_{\vec{k}_2\vec{\varepsilon}_2}\rangle \otimes \dots = |n_{\vec{k}_1\vec{\varepsilon}_1} n_{\vec{k}_2\vec{\varepsilon}_2} \dots\rangle, \quad n_{\vec{k}_i\vec{\varepsilon}_i} \in \mathbb{N}. \quad (7.152)$$

Un état du champ est donc complètement déterminé par l'énoncé du nombre d'occupation $n_{\vec{k}_i\vec{\varepsilon}_i}$ du mode i ; les opérateurs de création et d'annihilation font varier d'une unité le nombre d'occupation relatif à leur mode :

$$a_{\vec{k}_i\vec{\varepsilon}_i} |\dots n_{\vec{k}_i\vec{\varepsilon}_i} \dots\rangle = \sqrt{n_{\vec{k}_i\vec{\varepsilon}_i}} |\dots n_{\vec{k}_i\vec{\varepsilon}_i} - 1 \dots\rangle, \quad a_{\vec{k}_i\vec{\varepsilon}_i}^\dagger |\dots n_{\vec{k}_i\vec{\varepsilon}_i} \dots\rangle = \sqrt{n_{\vec{k}_i\vec{\varepsilon}_i} + 1} |\dots n_{\vec{k}_i\vec{\varepsilon}_i} + 1 \dots\rangle, \quad (7.153)$$

où tous les nombres cachés dans les \dots sont inchangés. Enfin, conformément à (7.146), l'énergie propre correspondante est la simple somme :

$$H_{\text{champ}} |\dots n_{\vec{k}_i\vec{\varepsilon}_i} \dots\rangle = E_{\text{champ}} |\dots n_{\vec{k}_i\vec{\varepsilon}_i} \dots\rangle, \quad E_{\text{champ}} = \sum_i n_{\vec{k}_i\vec{\varepsilon}_i} \hbar\omega_{\vec{k}_i\vec{\varepsilon}_i}. \quad (7.154)$$

L'état fondamental du champ, appelé état vide et noté $|\text{vac}\rangle$, est celui où tous les nombres d'occupation sont nuls : c'est bien l'état de plus basse énergie, puisque tous les $n_{\vec{k}_i\vec{\varepsilon}_i}$ sont des entiers positifs ou nuls. Clairement, justement parce que ces nombres sont entiers, la plus petite variation de l'état du champ est finie : ce *quantum* de champ n'est autre que le photon d'Einstein. Les états du champ définis en (7.152) sont des états où les nombres de photons sont fixés (ce sont de bons nombres quantiques) et ne varient pas dans le temps : ce sont bien les états du champ en l'absence de matière.

À ce stade, il est possible de saisir l'importance de la quantification du champ et notamment le rôle crucial de ce que l'on appelle "fluctuations du vide", en conséquence de la non-commutation entre les opérateurs du champ :

$$a_{\vec{k}\vec{\varepsilon}} a_{\vec{k}\vec{\varepsilon}}^\dagger \neq a_{\vec{k}\vec{\varepsilon}}^\dagger a_{\vec{k}\vec{\varepsilon}}. \quad (7.155)$$

Soit un état du champ à nombre donné de photons d'impulsion et de polarisation bien définies, noté $|n_{\vec{k}\vec{\varepsilon}}\rangle$; la moyenne de la relation de commutation (7.147) dans cet état donne :

$$\langle n_{\vec{k}\vec{\varepsilon}} | a_{\vec{k}\vec{\varepsilon}} a_{\vec{k}\vec{\varepsilon}}^\dagger | n_{\vec{k}\vec{\varepsilon}} \rangle = 1 + \langle n_{\vec{k}\vec{\varepsilon}} | a_{\vec{k}\vec{\varepsilon}}^\dagger a_{\vec{k}\vec{\varepsilon}} | n_{\vec{k}\vec{\varepsilon}} \rangle = 1 + n_{\vec{k}\vec{\varepsilon}}. \quad (7.156)$$

On peut deviner que le nombre de photons est directement relié à l'intensité des champs, laquelle contribue proportionnellement au nombre de transitions par unité de temps, autant en absorption induite qu'en émission induite⁵⁸. C'est pourquoi le 1 au second membre de (7.156) fait toute la différence alors que l'autre terme, $n_{\vec{k}\vec{\varepsilon}}$, donne une parfaite symétrie entre absorption et émission induites. Le retour à l'équilibre thermique ne pouvant se faire que si le taux d'émission est *supérieur* au taux d'absorption⁵⁹, le 1 – signature de la non-commutation – est donc crucial pour le maintien de l'équilibre ou la relaxation vers l'équilibre ; son apparition dans le formalisme constitue la justification de l'introduction phénoménologique par Einstein du coefficient A . Bien sûr, dans la limite des très grands nombres d'occupation ($n_{\vec{k}\vec{\varepsilon}} \gg 1$), la non-commutation est de peu d'importance.

Dans l'état vide de photons, on a évidemment :

$$\langle \text{vac} | a_{\vec{k}\vec{\varepsilon}}^\dagger a_{\vec{k}\vec{\varepsilon}} | \text{vac} \rangle = 0, \quad (7.157)$$

puisque $a_{\vec{k}\vec{\varepsilon}} | \text{vac} \rangle = 0$, en généralisation de (7.145). L'équation (7.157) exprime le fait que dans l'état vide, le nombre *moyen* de photons est nul, quelles que soient l'impulsion et la polarisation (tous les modes sont vides

⁵⁸Le qualificatif *induite* rappelle que c'est le champ qui provoque les transitions entre niveaux atomiques.

⁵⁹Puisque la population thermique d'un état est une fonction décroissante de l'énergie.

en moyenne). Il en va d'ailleurs de même⁶⁰ de la *fluctuation* du nombre de photons, puisque $|\text{vac}\rangle$ est propre de l'opérateur $N_{\vec{k}, \vec{\varepsilon}}$ et de son carré :

$$\langle \text{vac} | (a_{\vec{k}, \vec{\varepsilon}}^\dagger a_{\vec{k}, \vec{\varepsilon}}) (a_{\vec{k}, \vec{\varepsilon}}^\dagger a_{\vec{k}, \vec{\varepsilon}}) | \text{vac} \rangle = 0 . \quad (7.158)$$

Les champs $\vec{\mathcal{E}}$ et $\vec{\mathcal{B}}$ sont également nuls en moyenne dans tout état (7.152) et en particulier pour l'état vide (ils s'expriment en combinaison linéaire des $a_{\vec{k}, \vec{\varepsilon}}^\dagger$ et des $a_{\vec{k}, \vec{\varepsilon}}$ dont tous les éléments diagonaux sont nuls). En revanche, même dans l'état vide, les *carrés* de ces champs ne le sont pas ; en effet, l'élévation au carré d'une expression telle que (7.150) produit quatre types de termes :

$$a_{\vec{k}, \vec{\varepsilon}} a_{\vec{k}, \vec{\varepsilon}} , \quad a_{\vec{k}, \vec{\varepsilon}}^\dagger a_{\vec{k}, \vec{\varepsilon}}^\dagger , \quad a_{\vec{k}, \vec{\varepsilon}} a_{\vec{k}, \vec{\varepsilon}}^\dagger , \quad a_{\vec{k}, \vec{\varepsilon}}^\dagger a_{\vec{k}, \vec{\varepsilon}} . \quad (7.159)$$

Tous ces termes, sauf ceux du 3^{ème} type, ont une valeur moyenne nulle dans l'état vide ; en vertu de la relation de commutation (7.147) :

$$\langle \text{vac} | a_{\vec{k}, \vec{\varepsilon}} a_{\vec{k}, \vec{\varepsilon}}^\dagger | \text{vac} \rangle = \langle \text{vac} | (\mathbf{1} + a_{\vec{k}, \vec{\varepsilon}}^\dagger a_{\vec{k}, \vec{\varepsilon}}) | \text{vac} \rangle = 1 + 0 = 1 . \quad (7.160)$$

Ainsi, même dans l'état vide, les carrés des champs ne sont pas nuls ; ce sont ces fluctuations qui sont responsables, notamment, de l'émission spontanée, émission d'un photon alors que le champ est identiquement nul en moyenne – situation que l'on décrirait classiquement en disant “*le champ est nul*”. Ces fluctuations sont aussi responsables du *Lamb shift* – que l'on peut se représenter comme l'émission - absorption incessante de photons (virtuels) par un atome (voir ci-dessous).

Ayant trouvé le potentiel-vecteur quantique (7.151), l'interaction atome - champ réduite à sa plus simple et dominante expression V_1 (voir (7.88)) est :

$$V = -\frac{e}{\mu} \sum_{\vec{\varepsilon}} \int d^3k \sqrt{\frac{\hbar}{16\pi^3\epsilon_0\omega}} \left(a_{\vec{k}, \vec{\varepsilon}}^\dagger e^{-i\vec{k}\cdot\vec{r}} + a_{\vec{k}, \vec{\varepsilon}} e^{i\vec{k}\cdot\vec{r}} \right) \vec{p}\cdot\vec{\varepsilon} ; \quad (7.161)$$

ainsi, l'interaction V dans (7.140) est une combinaison *linéaire* des $a_{\vec{k}, \vec{\varepsilon}}$ et des $a_{\vec{k}, \vec{\varepsilon}}^\dagger$. Le processus élémentaire⁶¹ entre états du champ à nombre donné de photons est donc la variation d'une unité du nombre d'occupation – ce qui confirme que le photon est bien le *quantum* de champ –, cependant que l'atome lui-même change ou ne change pas⁶² d'état (bien sûr, en allant à l'ordre n de la théorie de perturbation, on met en évidence des transitions possibles entre états du champ différant exactement de n photons, mais ces transitions procèdent par une succession de sauts élémentaires $\Delta n = \pm 1$). Ces mécanismes élémentaires peuvent être représentés par des dessins appelés diagrammes de Feynman.

Il s'agit maintenant de définir un modèle simple et de diagonaliser aussi bien que possible le Hamiltonien correspondant, c'est-à-dire de trouver ses états propres. Ceci fait, il sera possible de trouver l'évolution temporelle d'un état non-stationnaire : la dépendance en temps qui réapparaîtra dans la suite n'est pas la conséquence de l'application d'une perturbation dépendant du temps mais de la préparation d'un état qui n'est pas propre du Hamiltonien *complet* H . Les états propres de H_0 s'obtiennent par produit tensoriel des états atomiques et des états du champ libre donnés en (7.153).

À titre d'exemple, fixons-nous comme objectif de construire une description (très simplifiée) de l'émission spontanée. On admet de prime abord que l'atome possède seulement deux états, notés⁶³ $|g\rangle$ et $|e\rangle$, en supposant bien sûr que l'élément de matrice de \vec{p} entre ces deux états n'est pas nul⁶⁴. En ce qui concerne le champ, on n'introduit que deux types d'états, l'état vide $|\text{vac}\rangle$, et les états à un seul photon, $|\vec{k}, \vec{\varepsilon}\rangle$. Au total, le système atome + champ possède 4 types d'états, obtenus par produit tensoriel ($|g\rangle \otimes |\vec{k}, \vec{\varepsilon}\rangle \equiv |g, \vec{k}, \vec{\varepsilon}\rangle$, *etc.*) :

$$\{ |g, \vec{k}, \vec{\varepsilon}\rangle \}_{\vec{k}, \vec{\varepsilon}} , \quad |e, \text{vac}\rangle , \quad |g, \text{vac}\rangle , \quad \{ |e, \vec{k}, \vec{\varepsilon}\rangle \}_{\vec{k}, \vec{\varepsilon}} , \quad (7.162)$$

⁶⁰La fluctuation du nombre de photons est d'ailleurs nulle dans tout état du type (7.152), pas seulement pour le vide.

⁶¹au sens où *une seule* interaction est en jeu.

⁶²Si le photon est récupéré dans le milieu extérieur (processus réel), alors la conservation d'énergie exige que l'atome change d'état.

⁶³ g pour *ground*, e pour *excited*.

⁶⁴Les éléments diagonaux de \vec{p} , pris sur les états atomiques, sont évidemment nuls puisqu'il s'agit d'états *liés*. Typiquement, g désigne un état de type s (le fondamental par exemple) et e un état de type p ; ce couple satisfait bien les règles de sélection E1 obtenues plus haut.

d'énergies respectives⁶⁵ :

$$\hbar kc, \quad E_e, \quad 0, \quad E_e + \hbar kc. \quad (7.163)$$

Dans le sous-espace engendré par les vecteurs (7.162), V est diagonal par blocs : il ne peut y avoir d'élément de matrice non-nul qu'entre deux états différant par un photon ; en outre, il faut en plus que \vec{p} ait un élément de matrice également non-nul, ce qui exclut tous les éléments de matrice diagonaux vis-à-vis des variables atomiques. Si on définit les projecteurs :

$$P = |e, \text{vac}\rangle\langle e, \text{vac}| + \sum_{\vec{k}, \vec{\varepsilon}} |g, \vec{k}, \vec{\varepsilon}\rangle\langle g, \vec{k}, \vec{\varepsilon}|, \quad Q = \mathbf{1} - P, \quad (7.164)$$

on voit que $PVQ = QVP = 0$. En outre, dans chacun des deux sous-espaces définis par P et Q , V a tous ses éléments diagonaux nuls.

L'émission spontanée est manifestement décrite par les processus mettant en jeu les états qui apparaissent dans P , qui sont donc exclusivement considérés dans la suite. Parmi ces états, certains sont quasi-dégénérés, voire strictement dégénérés : ils peuvent donc être les points de départ et d'arrivée d'une transition réelle. L'autre sous-espace contient deux types d'états forcément séparés par une grande différence d'énergie (au moins égale à $E_e - E_g$), qui ne peuvent de ce fait apparaître qu'en tant qu'intermédiaire dans des processus virtuels. En définitive, s'agissant de décrire le plus simplement possible l'émission spontanée – émission réelle d'un photon vers l'extérieur –, on ne retient que les états⁶⁶ :

$$\{|g, \vec{k}, \vec{\varepsilon}\}_{\vec{k}, \vec{\varepsilon}}, \quad |e, \text{vac}\rangle. \quad (7.165)$$

À cette approximation, le spectre de H_0 se compose de la demi-droite réelle s'étendant de 0 à $+\infty$ (le photon libre, avec k variant de 0 à $+\infty$) et du point isolé E_e . Reprenant le langage utilisé à propos de la résolvante, G_0 a donc un pôle simple en E_e , noyé dans le *continuum* de photons libres.

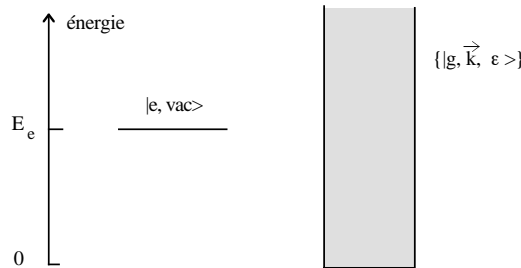


Figure 7.4: Schéma des niveaux d'énergie de H_0 dans l'espace effectif défini par P (voir 7.164)).

La question que l'on se pose peut se formuler comme suit : le système étant supposé préparé à $t = 0$ dans l'état $|e, \text{vac}\rangle$ (pas de photon et l'atome à l'état excité), quelle est la probabilité pour l'atome d'être encore à l'état excité au temps t ? Cette probabilité dépend du temps puisque l'état préparé $|e, \text{vac}\rangle$ n'est pas propre de H (mais seulement de H_0) : pour le Hamiltonien complet, c'est donc un état non-stationnaire et la probabilité cherchée est bien une fonction explicite du temps ; évidemment, la réponse à cette question donnera notamment la durée de vie du niveau excité. En d'autres termes, ce qu'il faut trouver c'est la variation au cours du temps de l'amplitude $A(t) = \langle e, \text{vac} | \Psi(t) \rangle$, dont le module au carré donnera la probabilité cherchée. En notant simplement $|E_e\rangle \equiv |e, \text{vac}\rangle$ et en reprenant les notations de la théorie des perturbations dépendant du temps, l'expression de l'amplitude cherchée est :

$$A(t) = \langle E_e | U_0(t) U_1(t) | E_e \rangle = e^{\frac{i}{\hbar} E_e t} \langle E_e | U_1(t) | E_e \rangle \equiv e^{\frac{i}{\hbar} E_e t} a(t). \quad (7.166)$$

Dans ces conditions, la probabilité de trouver l'atome encore excité à l'instant $t > 0$ est $P(t)$:

$$P(t) = |A(t)|^2 = |a(t)|^2. \quad (7.167)$$

Il s'agit clairement d'une probabilité *conditionnelle*.

⁶⁵L'énergie du fondamental de l'atome est prise comme origine.

⁶⁶L'espace effectif d'états ainsi défini a une dimension infinie non-dénombrable puisque \vec{k} prend ses valeurs dans \mathbb{R}^3 .

Calcul perturbatif

La quantité essentielle est donc $a(t) = \langle E_e | U_1(t) | E_e \rangle$, dont il est facile d'exprimer la dérivée ; en effet :

$$\frac{da}{dt} = \langle E_e | \frac{\partial}{\partial t} U_1(t) | E_e \rangle , \quad (7.168)$$

et, compte tenu de (7.22), il vient :

$$\frac{da}{dt} = \frac{1}{i\hbar} \langle E_e | V_I(t) U_1(t) | E_e \rangle . \quad (7.169)$$

Faisons d'abord le calcul au plus bas ordre non-nul, soit en remplaçant U_1 par son expression au premier ordre⁶⁷ :

$$\frac{da}{dt} = \frac{1}{i\hbar} \langle E_e | V_I(t) \left[\mathbf{1} + \frac{1}{i\hbar} \int_0^t dt' V_I(t') \right] | E_e \rangle . \quad (7.170)$$

Dans le sous-espace associé à P , V_I , tout comme V , est purement non-diagonal ; le terme $\mathbf{1}$ du crochet a donc une contribution nulle et il reste⁶⁸ :

$$\frac{da}{dt} = \frac{1}{(i\hbar)^2} \langle E_e | V_I(t) \int_0^t dt' V_I(t') | E_e \rangle = -\hbar^{-2} \int_0^t dt' \langle E_e | V_I(t) \sum_{\vec{k}, \vec{\varepsilon}} |g, \vec{k}, \vec{\varepsilon}\rangle \langle g, \vec{k}, \vec{\varepsilon}| V_I(t') | E_e \rangle . \quad (7.171)$$

En supposant le champ enfermé dans une très grande boîte, la somme discrète sur \vec{k} porte sur des niveaux formant un *quasi-continuum* et peut être remplacée par une intégrale. En introduisant $E = \hbar kc$, la densité d'états $\rho(E)$ (qui incorpore la sommation sur les polarisations, soit $\sum_{\vec{k}, \vec{\varepsilon}} \bullet \rightarrow \int dE \rho(E) \bullet$) et en posant :

$$\langle g, \vec{k}, \vec{\varepsilon} | V | E_e \rangle = v(E) \quad (E = \hbar kc) , \quad (7.172)$$

on voit que l'amplitude $a(t)$ satisfait :

$$\frac{da}{dt} = -\frac{1}{2\pi\hbar} \int_0^{+\infty} dE \int_0^t dt' \Gamma(E) e^{-(i/\hbar)(E-E_e)(t-t')} , \quad (7.173)$$

où la fonction Γ , homogène à l'inverse d'un temps, est définie par :

$$\Gamma(E) = \frac{2\pi}{\hbar} \rho(E) |v(E)|^2 . \quad (7.174)$$

Cette fonction est nulle pour $E < 0$ (le spectre de H_0 commence en $E = 0$), et tend vers zéro aux hautes énergies ; en effet, $v(E)$ contient en facteur l'élément de matrice $\langle e | e^{i\vec{k}\cdot\vec{r}} \vec{p} | g \rangle \equiv \phi(\vec{k})$, qui est la transformée de Fourier d'un produit de deux fonctions représentant des états liés : leur extension spatiale linéaire est donc d'ordre a_0 . Il en résulte que la fonction $\phi(\vec{k})$ est sensiblement non nulle seulement si $k \lesssim a_0^{-1}$, soit si $E \lesssim a_0^{-1}/(\hbar c)$, ou encore⁶⁹ $E \lesssim \alpha\mu c^2$. En d'autres termes, $\Gamma(E)$ a vaguement l'allure d'une cloche de largeur $\hbar\Omega \sim \alpha\mu c^2$. Par ailleurs, l'intégrale sur le temps t' fournit une partie réelle et une partie imaginaire, caractérisable chacune par une largeur d'ordre \hbar/t ; si :

$$\frac{\hbar}{t} \ll \hbar\Omega , \quad (7.175)$$

alors on peut rejeter la borne t à $+\infty$ et on obtient⁷⁰ :

$$\frac{da}{dt} = -\frac{1}{2\pi} \int_0^{+\infty} dE \Gamma(E) \left[\pi \delta(E_e - E) + i\mathcal{P} \frac{1}{E_e - E} \right] \quad (7.177)$$

⁶⁷Au total, il s'agit d'un calcul au second ordre en V puisqu'une interaction est déjà présente en facteur de U_1 dans (7.169). Comme le couplage atome - champ fait varier d'une unité le nombre de photons (V est une combinaison *linéaire* des a et des a^\dagger), le premier terme $\mathbf{1}$ dans (7.170) donne une contribution nulle à l'amplitude de transition.

⁶⁸Compte tenu de la forme de la perturbation, il faut bien deux interactions pour que l'élément de matrice diagonal sur $|E_e\rangle$ soit différent de zéro.

⁶⁹ α est la constante de structure fine.

⁷⁰Se souvenir des relations :

$$\int_0^{+\infty} dt e^{-(\eta+ix)t} = \frac{1}{\eta+ix} = \frac{\eta}{x^2+\eta^2} - i \frac{x}{x^2+\eta^2} \quad (\eta > 0) , \quad \lim_{\eta \rightarrow 0} \frac{1}{\eta+ix} = \pi \delta(x) - i\mathcal{P} \frac{1}{x} . \quad (7.176)$$

où \mathcal{P} désigne la partie principale de Cauchy. En posant maintenant⁷¹ :

$$\Delta(E) = \frac{1}{2\pi} \mathcal{P} \int_0^{+\infty} dE' \frac{\Gamma(E')}{E - E'} , \quad (7.178)$$

on obtient finalement :

$$\frac{da}{dt} = - \left[\frac{1}{2} \Gamma(E_e) + i \Delta(E_e) \right] \quad (7.179)$$

qui s'intègre membre à membre en :

$$a(t) = 1 - \left[\frac{1}{2} \Gamma(E_e) + i \Delta(E_e) \right] t . \quad (7.180)$$

Il s'ensuit immédiatement que la probabilité $P(t) = |A(t)|^2 = |a(t)|^2 = a^*(t)a(t)$ est donnée au même ordre par :

$$P(t) = \left[1 - \frac{1}{2} \Gamma(E_e) t \right]^2 + [\Delta(E_e) t]^2 \simeq 1 - \Gamma(E_e) t + \mathcal{O}(t^2) . \quad (7.181)$$

Ce résultat, qui clôt le traitement perturbatif, n'est rien d'autre que la règle d'or de Fermi, l'état $|e, \text{vac}\rangle$ jouant le rôle de l'état noté $|\psi_i\rangle$ dans la section 7.4, le *continuum* étant formé ici par les états $|g, \vec{k}, \vec{\varepsilon}\rangle$. Clairement, (7.181) n'a de sens que si $\Gamma(E_e)t \ll 1$. Compte tenu de (7.175), il faut au total :

$$\Omega^{-1} \ll t \ll \Gamma(E_e)^{-1} . \quad (7.182)$$

Ceci n'est possible que si :

$$\Omega^{-1} \ll \Gamma(E_e)^{-1} \iff \Gamma(E_e) \ll \Omega ; \quad (7.183)$$

cette condition signifie que la largeur spectrale de l'état atomique "renormalisé" doit être petite devant la largeur du *continuum* effectivement couplé à l'état atomique⁷². Ce que l'on vient d'obtenir a une portée générale et donne la description (approchée), valable aux temps courts, pour un niveau *étroit* (ici, infiniment étroit) couplé à un *continuum large*.

Par ailleurs, un calcul complémentaire permet vérifier que l'énergie du photon final est égale à E_e , ce qui exprime la conservation de l'énergie pour un processus réel.

Calcul non perturbatif

Il est possible de dépasser le traitement perturbatif et, notamment, de montrer que l'expression (7.181) est bien le début d'une exponentielle – alors $\Gamma(E_e)^{-1}$ pourra indiscutablement être identifié avec la durée de vie de l'état excité. En tout état de cause, le développement (7.181) n'a évidemment de sens que si $\Gamma(E_e)t \ll 1$, c'est-à-dire aux temps très courts, donc juste après l'excitation de l'atome.

La résolvante $G(z)$:

$$G(z) = \frac{1}{z\mathbf{1} - H} \quad (7.184)$$

est un outil très commode pour dépasser le traitement perturbatif, comme on l'a déjà vu à propos de la théorie des perturbations indépendantes du temps. On connaît le spectre de H_0 , qui se compose de la demi-droite réelle \mathbb{R}_+ et du point isolé E_e ; en revanche, celui de H est inconnu à ce stade, mais il est bien évidemment situé dans la droite réelle. Considérons maintenant l'intégrale :

$$U_+(t) = \frac{1}{2i\pi} \int_{C_+} \frac{e^{\frac{1}{i\hbar} zt}}{z\mathbf{1} - H} dz \quad (7.185)$$

⁷¹La structure de $\Delta(E)$ est celle d'une correction d'énergie au second ordre de la théorie des perturbations indépendantes du temps, écrite pour des états formant un continuum : $\sum_l \bullet \rightarrow \int dE \rho(E) \bullet$.

⁷²Comme $\Gamma(E_e)$ contient l'interaction au carré, il est plausible, en général, que $\Gamma(E_e) \sim \alpha^2 \mu c^2 \sim \alpha \Omega \ll \Omega$.

où C_+ est une ligne parallèle à l'axe réel, située juste au-dessus et parcourue de $+\infty + i0$ vers $-\infty + i0$. Si $t < 0$, on referme le contour par un grand demi-cercle de rayon R du côté⁷³ $\Im z > 0$, qui n'apporte aucune contribution dans la limite $R \rightarrow +\infty$. Il n'y a aucune singularité dans le demi-plan $\Im z > 0$: par le théorème des résidus, l'intégrale est nulle. Au contraire, pour $t > 0$, on referme par le bas et on ramasse toutes les singularités de la résolvante, qui constituent le spectre de H . Au total :

$$U_+(t) = \begin{cases} 0 & \text{si } t < 0 \\ e^{\frac{1}{i\hbar} Ht} & \text{si } t > 0 \end{cases} . \quad (7.186)$$

$U_+(t)$ est donc nul pour $t < 0$ et coïncide avec l'opérateur d'évolution⁷⁴ $U(t)$ pour $t > 0$. En terme d'intégrale sur l'axe réel, l'égalité suivante en résulte :

$$U_+(t) = \lim_{\varepsilon \rightarrow 0} \frac{1}{2i\pi} \int_{+\infty}^{-\infty} e^{\frac{1}{i\hbar}(E+i\varepsilon)t} G(E+i\varepsilon) dE , \quad (7.187)$$

où ε est un réel positif. Muni de ces résultats, on voit que l'amplitude $A(t > 0) = \langle E_e | U(t) | E_e \rangle$ possède la représentation intégrale :

$$A(t) = \frac{1}{2i\pi} \int_{C_+} G_e(z) e^{\frac{1}{i\hbar}zt} dz = \lim_{\varepsilon \rightarrow 0} \frac{1}{2i\pi} \int_{+\infty}^{-\infty} G_e(E+i\varepsilon) e^{\frac{1}{i\hbar}(E+i\varepsilon)t} dE \quad (7.188)$$

où :

$$G_e(z) = \langle E_e | G(z) | E_e \rangle . \quad (7.189)$$

Toute la question est alors de calculer $G_e(z)$, que l'on va déduire de l'équation de Dyson :

$$G(z) = G_0(z) + G_0(z)VG(z) \quad (7.190)$$

itérée une fois :

$$G(z) = G_0(z) + G_0(z)VG_0(z) + G_0(z)VG_0(z)VG(z) . \quad (7.191)$$

G_0 est diagonal ; la perturbation V , au contraire, est purement non-diagonale et ne couple entre eux que des états différant d'un photon ; en prenant l'élément de matrice de (7.191) et en se référant à (7.172), il vient :

$$G_e(z) = \frac{1}{z - E_e} + 0 + \frac{1}{z - E_e} \int_0^{+\infty} dE \rho(E) v^*(E) \frac{1}{z - E} v(E) G_e(z) . \quad (7.192)$$

On en déduit (voir (7.174)) :

$$G_e(z) = \frac{1}{z - E_e - \frac{\hbar}{2\pi} \int_0^{+\infty} dE' \frac{\Gamma(E')}{z - E'}} . \quad (7.193)$$

Compte tenu de la structure algébrique de cette expression – et du fait que $\Gamma(E) > 0$ (voir (7.174)) –, cette fonction ne peut avoir de pôle : par exemple, si $\Im z > 0$, l'intégrale du dénominateur a une partie imaginaire *négative* ; comme l'intégrale est précédée du signe –, au total la partie imaginaire du dénominateur est la somme de deux quantités positives et ne peut donc s'annuler ; la conclusion est la même si $\Im z < 0$. Le dénominateur ne peut donc jamais être nul⁷⁵.

Les équations (7.193) et (7.178) donnent maintenant⁷⁶ :

$$G_e(E + i0) = \frac{1}{E + i0 - E_e - \hbar\Delta(E) + i\frac{\hbar}{2}\Gamma(E)} ; \quad (7.194)$$

de la même façon, on trouve :

$$G_e(E - i0) = \frac{1}{E - i0 - E_e - \hbar\Delta(E) - i\frac{\hbar}{2}\Gamma(E)} . \quad (7.195)$$

⁷³ $\Im z$ désigne la partie imaginaire de z .

⁷⁴ Cet opérateur, appelé propagateur avancé (ou causal), a été rencontré dans le chapitre 1, section 1.5.

⁷⁵ Bien évidemment, $G(z)$ a heureusement des singularités – voir le théorème de Liouville ! Ces singularités sont denses et définissent la coupure de 0 à $+\infty$ sur l'axe réel.

⁷⁶ On peut évidemment oublier $+i0$ dans $E+i0$ au dénominateur ; cette précaution d'écriture ne fait que rappeler par quel chemin on obtient le second membre de (7.194) et (7.195).

Les deux équations (7.194) et (7.195) confirment que la fonction $G(z)$ a une coupure s'étendant sur l'axe réel de 0 à $+\infty$ (la fonction $\Gamma(E)$ est identiquement nulle pour $E < 0$, de sorte que $G_e(E - i0) = G_e(E + i0)$ si $E < 0$) :

$$G_{e\pm}(E) \equiv \lim_{\varepsilon \rightarrow 0} G_e(E \pm i\varepsilon) = \frac{1}{E - E_e - \hbar\Delta(E) \pm i\frac{\hbar}{2}\Gamma(E)} . \quad (7.196)$$

Avant de continuer le calcul, il est utile d'interpréter ce résultat et d'élucider le contenu physique de la fonction $G_{e\pm}(E)$. Imaginons que le couplage avec le champ tende vers zéro. Alors :

$$G_{e\pm}(E) \rightarrow G_{e\pm}^{(0)}(E) = \frac{1}{E - E_e \pm i0} . \quad (7.197)$$

En vertu de $\Im \lim_{\eta \rightarrow 0} \frac{1}{x+i\eta} = -\pi\delta(x)$, on en déduit :

$$\mp \frac{1}{\pi} \Im G_{e\pm}^{(0)}(E) = \delta(E - E_e) . \quad (7.198)$$

$\delta(E - E_e)$ n'est autre que la densité d'états de l'atome près de l'énergie E_e : quand le couplage atome - champ est nul, le seul état possible de l'atome (dit "nu") est celui d'énergie exactement égale à E_e . Cette constatation permet de comprendre que la partie imaginaire de $G_{e\pm}(E)$ fournit la densité d'états⁷⁷ de l'atome en présence du champ (atome "habillé"⁷⁸) :

$$\rho_{\text{at}}(E) = \mp \frac{1}{\pi} \Im G_{e\pm}(E) = \frac{1}{2\pi} \frac{\hbar\Gamma(E)}{[E - E_e - \hbar\Delta(E)]^2 + \frac{[\Gamma(E)]^2}{4}} \quad (E \sim E_e) . \quad (7.199)$$

Essentiellement – en négligeant la variation (lente) de $\Delta(E)$ et de $\Gamma(E)$ – $\rho_{\text{at}}(E)$ a une variation lorentzienne.

Une fois l'élément de matrice $G_{e+}(E)$ déterminé, il suffit d'effectuer l'intégration (7.188) pour trouver l'amplitude $A(t)$, le long de C_+ ou de tout autre contour équivalent – un préalable à cette opération est l'identification des singularités de $G_e(z)$. En se plaçant toujours dans l'hypothèse d'un couplage faible, ce calcul est grandement facilité en remplaçant la vraie fonction $G_e(z)$ par une autre, dont les propriétés analytiques sont très différentes mais qui donnera à peu près la même valeur pour l'intégrale – au moins pour les temps pas trop longs. Pour ceci, on remarque que les fonctions Δ et Γ étant à variation lente, on ne commet pas une erreur grossière en remplaçant leur argument E par E_e . Ainsi apparaît la fonction⁷⁹ $G_{e\text{ap}+}(E)$:

$$G_{e\text{ap}+}(E) = \frac{1}{E - E_e - \hbar\Delta_e + i\frac{\hbar}{2}\Gamma_e} , \quad \Delta_e = \Delta(E_e) , \quad \Gamma_e = \Gamma(E_e) . \quad (7.200)$$

C'est ce que l'on appelle souvent l'approximation du pôle : visiblement, une fois effectué son prolongement analytique immédiat, $G_{e\text{ap}+}(z)$ a une et une seule singularité, un pôle en $E_e + \hbar\Delta_e - i\frac{\hbar}{2}\Gamma_e$ et a de ce fait des propriétés analytiques très différentes de celles de la fonction "exacte" $G_e(z)$, qui a une coupure (et rien d'autre tant qu'elle n'est pas prolongée par continuité sur sa surface de Riemann). En tout cas, l'inversion de Fourier - Laplace (7.188) avec $G_{e\text{ap}+}(z)$ est alors immédiate et donne :

$$A(t) \simeq e^{\frac{1}{\hbar}[E_e + \hbar\Delta_e - i\frac{\hbar}{2}\Gamma_e]t} = e^{\frac{1}{\hbar}(E_e + \hbar\Delta_e)t} e^{-\frac{1}{2}\Gamma_e t} \quad (t > 0) . \quad (7.201)$$

Cette expression est transparente ; elle montre d'abord que, sous l'effet du couplage avec le champ, l'énergie de l'atome "nu" est déplacée de la quantité $\hbar\Delta_e$, c'est ce que l'on appelle le Lamb *shift* que l'on peut se représenter comme le résultat de l'émission-absorption de photons virtuels par l'atome produisant une "renormalisation" de son énergie ; ces processus sont bien virtuels au sens où ils impliquent des états d'énergie arbitraire (voir la définition (7.178) de Δ_e , qui contient une somme sur l'énergie). Ce déplacement est loin d'être négligeable et joue un rôle important en Physique atomique⁸⁰ ; en particulier, il est responsable de la levée de dégénérescence des niveaux $2S_{1/2}$ et $2P_{1/2}$ de l'atome d'hydrogène, une fois "habillé" par les photons. Chacun de ces niveaux

⁷⁷Il n'est pas difficile d'étendre ce résultat : d'une façon générale, la partie imaginaire de la trace de la "bonne" résolvante fournit la densité d'états.

⁷⁸expression due à Cl. Cohen-Tannoudji ; en anglais, on dit *dressed atom*.

⁷⁹La densité d'états correspondante, $\rho_{\text{at ap}}(E)$ est alors strictement lorentzienne.

⁸⁰Son calcul précis exige toutefois d'étendre le petit modèle présenté ici en détail.

a sa propre correction d'énergie mais il se trouve que cette correction est beaucoup plus grande pour $2P_{1/2}$ que pour $2S_{1/2}$ ⁸¹ ; au total, l'écart est manifeste et on trouve :

$$E_{2S_{1/2}} - E_{2P_{1/2}} \simeq 1060 \text{ MHz} . \quad (7.202)$$

Cette valeur est tout à fait comparable à l'élargissement Doppler à l'ambiante pour une raie optique et donne un effet très visible.

En outre, (7.201) donne :

$$P(t) = |A(t)|^2 \simeq e^{-\Gamma_e t} . \quad (7.203)$$

Dans l'approximation considérée, le déclin est effectivement purement exponentiel ; la durée de vie est Γ_e^{-1} , la largeur naturelle est Γ_e .

Dans l'approximation du pôle, tout se passe finalement comme si l'interaction atome - champ se bornait à remplacer l'énergie réelle E_e de l'atome sans champ par l'énergie *complexe* $\tilde{E}_e = E_e + \hbar\Delta_e - i\hbar\Gamma_e/2$; le pôle réel de G_0 a donné naissance à un pôle situé à proximité mais ayant une partie imaginaire finie, et négative. Ceci est une vision simplifiée de ce qui se passe pour la fonction $G_e(z)$; on sait qu'elle a une coupure, mais il est possible de la prolonger par continuité dans le 2^{ème} feuillet de Riemann ; dans ce feuillet, le prolongement analytique a effectivement un pôle en \tilde{E}_e . En tout cas, l'apparition, dans l'approximation du pôle, d'une énergie complexe fournit la justification théorique de l'introduction de hamiltoniens phénoménologiques non-hermitiques.

Le traitement précédent repose sur un certain nombre d'hypothèses autorisant certaines approximations, notamment celle du couplage faible et d'un niveau étroit couplé à un *continuum* large. Dans le cas contraire, des phénomènes intéressants se produisent (exemple : l'oscillation de Rabi, résonance au sens large – elle se produit lorsque la fréquence du champ appliqué est strictement égale à la fréquence propre du système à deux niveaux – l'oscillation s'effectuant à une pulsation donnée par l'élément de matrice de couplage atome - champ, ce qui est fort inusuel, donc très lentement à l'aune des échelles usuelles de la dynamique d'un atome). Notons toutefois que même en restant à l'intérieur d'une hypothèse de couplage faible, le déclin n'est pas exponentiel jusqu'au bout : des arguments généraux⁸² permettent de montrer que, pour des temps très grands devant Γ_e^{-1} , la décroissance suit en fait une loi puissance du genre $t^{-\lambda}$. Même si ce comportement non-exponentiel intervient très tard – au moment où il reste très peu d'atomes excités –, il n'en est pas pour autant sans intérêt : après tout, il constitue un exemple de curiosité⁸³ peut-être un peu inattendue, obtenue finalement dans un cadre assez banal.

Une dernière remarque, pour terminer. L'analyse précédente a certes été exprimée dans le langage de la Physique atomique, mais la méthode, les concepts et les outils utilisés ont une portée tout à fait générale et valent tout autant à chaque fois qu'il s'agit de décrire l'état instable d'un système (particule en Physique des hautes énergies – on parle souvent alors de *résonance* –, atome excité, quasi-excitations en Matière condensée, complexe transitoire en Chimie, *etc.*).

⁸¹L'état $2S_{1/2}$ est métastable – sa durée de vie est gigantesque, de l'ordre de la seconde – traduisant le fait que son couplage effectif au rayonnement est très faible (le retour vers le fondamental $1S_{1/2}$ est une transition à deux photons).

⁸²À cause de la coupure, l'intégrale portant sur le prolongement analytique de $G_e(z)$ s'écrit comme une somme de deux termes : le résidu en \tilde{E}_e qui produit une composante de déclin exponentiel comme ci-dessus, et une intégrale ("lacet") sur les deux bords de la coupure qui donne la décroissance algébrique en loi-puissance, $t^{-\lambda}$. Aux grands temps, l'exponentielle est quasiment éteinte et seule subsiste en pratique la décroissance en $t^{-\lambda}$.

⁸³Par exemple, si l'exposant λ est trop faible, on peut se trouver embarrassé pour définir une durée de vie au sens ordinaire (l'intégrale de $P(t)$ diverge si λ n'est pas supérieur à 1). En tout cas, il y a maintenant au moins deux échelles de temps : la durée de vie définie par Γ_e^{-1} , c'est l'échelle courte, et une autre immensément longue par rapport à la précédente.

Chapitre 8

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.*

8.1 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 (par exemple en terme 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 cohabitant 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 – et restent à demeure – dans deux régions

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 $N^{\frac{7}{5}}$ en fonction du nombre N de particules et la limite thermodynamique (qui exige $E \propto N$) tout simplement n'existerait pas ! ([39], § 2. 2. 1).

distinctes de l'espace³. Aucune propriété d'un système contenant des particules identiques n'est modifiée si l'on procède à l'échange de deux d'entre elles.

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. À 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" individuel.

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 condition 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 temps⁴.

L'impossibilité de distinguer deux particules identiques disparaît dans la limite classique : alors, les paquets d'ondes restent très bien localisés et ne se mélangent pas, leur centre suit en gros la trajectoire classique, qui reprend alors tout son sens – le non-recouvrement entre les paquets d'onde restitue *de facto* la discernabilité classique et l'impossibilité de distinguer deux particules identiques disparaît.

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 "frou" introduit par la collision des paquets d'ondes rend impossible toute autre identification que les grandeurs physiques intrinsèques de chaque particule (masse, charge, spin, *etc.*), qui sont identiques puisque précisément les particules le sont.

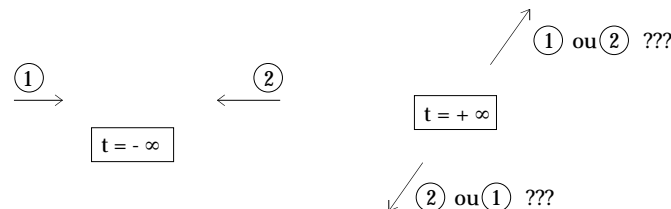


Figure 8.1: Illustration de la collision de deux spins $\frac{1}{2}$.

Un exemple ([4], § 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 $\frac{1}{2}$, dont on ignore pour simplifier les degrés de liberté orbitaux. L'espace des états est de dimension 4, et est engendré par les vecteurs $|m_1, m_2\rangle$ 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. 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 [20] "Faut-il toujours antisymétriser la fonction d'onde ?"

⁴En revanche, si à tout instant les paquets d'ondes restent bien séparés et ne se mélangent pas, la trajectoire classique reprend alors tout son sens : s'ils ne s'étalent pas trop et si le potentiel ne varie pas trop vite dans l'espace, leur centre suit en gros la trajectoire classique (Théorème d'Ehrenfest). En définitive, il existe heureusement une sorte de continuité vers la situation classique : le non-recouvrement entre les paquets d'ondes restitue *de facto* la discernabilité classique et l'impossibilité de distinguer deux particules identiques disparaît.

l'autre⁵ et, dans la formulation théorique présentée jusqu'ici, constitue la description maximale possible. 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 :

$$|\psi\rangle = a|+\frac{1}{2}, -\frac{1}{2}\rangle + b|-\frac{1}{2}, +\frac{1}{2}\rangle, \quad |a|^2 + |b|^2 = 1. \quad (8.1)$$

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 (8.1) 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 nettes et précises sur le 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 , $|\pm\frac{1}{2}\rangle$:

$$|\frac{1}{2}\rangle_x = \frac{1}{\sqrt{2}}[|+\frac{1}{2}\rangle + |-\frac{1}{2}\rangle]. \quad (8.2)$$

Pour les deux spins, le produit tensoriel se développe suivant :

$$|\frac{1}{2}\rangle_{1x} \otimes |\frac{1}{2}\rangle_{2x} = \frac{1}{2}[|+\frac{1}{2}, +\frac{1}{2}\rangle + |+\frac{1}{2}, -\frac{1}{2}\rangle + |-\frac{1}{2}, +\frac{1}{2}\rangle + |-\frac{1}{2}, -\frac{1}{2}\rangle]. \quad (8.3)$$

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 $|\psi\rangle$ (8.1) est le module au carré du produit scalaire :

$$P(a, b) = |(\langle\psi|, |\frac{1}{2}\rangle_{1x} \otimes |\frac{1}{2}\rangle_{2x})|^2. \quad (8.4)$$

Le calcul à partir de (8.3) et (8.1) 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, et en l'état actuel des choses, 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 propres 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) et les bosons (spin entier) ; 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.

8.2 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 :

$$\Psi(\dots, i, \dots, j, \dots) = + \Psi(\dots, j, \dots, i, \dots) , \quad (8.5)$$

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 :

$$\Psi(\dots, i, \dots, j, \dots) = - \Psi(\dots, j, \dots, i, \dots) , \quad (8.6)$$

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 note 2.

¹⁰Il en va de même pour le 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.

¹²Il faut évidemment un nombre *pair* de fermions pour former un spin entier.

- les particules de spin entier sont des bosons, et sont donc décrits par des états symétriques satisfaisant (8.5)
- les particules de spin demi-entier sont des fermions, et sont donc décrits par des états antisymétriques satisfaisant (8.6).

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. À 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'échange de leurs constituants élémentaires donnera globalement un signe $-$ (nombre impair de fermions)¹³.

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 celle-ci ; il faut donc que l'on ait :

$$|\Psi(\dots, i, \dots, j, \dots)|^2 = |\Psi(\dots, j, \dots, i, \dots)|^2 . \quad (8.7)$$

soit :

$$\Psi(\dots, i, \dots, j, \dots) = e^{i\alpha} \Psi(\dots, j, \dots, i, \dots) . \quad (8.8)$$

En définitive, l'affirmation du postulat de symétrisation est plus forte que celle rappelant l'invariance des probabilités : ce postulat dit que la phase α ne peut être 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 :

$$|\chi_1\rangle = \frac{1}{\sqrt{2}} \left[\left| +\frac{1}{2}, -\frac{1}{2} \right\rangle + \left| -\frac{1}{2}, +\frac{1}{2} \right\rangle \right] . \quad (8.9)$$

- (b) $a = +1$ et $b = -1$ correspondant à l'état singulet (spin total $S = 0$) et donnant la combinaison :

$$|\chi_0\rangle = \frac{1}{\sqrt{2}} \left[\left| +\frac{1}{2}, -\frac{1}{2} \right\rangle - \left| -\frac{1}{2}, +\frac{1}{2} \right\rangle \right] . \quad (8.10)$$

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. Il n'y a pas de difficulté : 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 une symétrie¹⁴ très différente : elle n'est pas quelconque mais entièrement déterminée par la valeur de

¹³C'est pourquoi les deux isotopes de l'Hélium ${}^3\text{He}$ et ${}^4\text{He}$ ont des propriétés radicalement différentes. ${}^3\text{He}$ est un fermion, alors que ${}^4\text{He}$ est un boson ; la superfluidité de ce dernier à basse température est la manifestation de la condensation de Bose.

¹⁴En 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 $\frac{1}{2}$.

M_S , et peut s'obtenir à l'aide de graphiques appelés *tableaux d'Young* [20]). La probabilité $P(a, b)$ (8.4) 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 :

$$\Psi_A(1, 2, 3, 4) = (-1)^2 \Psi_A(2, 1, 4, 3) . \quad (8.11)$$

Plus généralement, pour une permutation quelconque, Ψ acquiert un signe qui est, par définition, la *signature* de la permutation (voir ci-dessous, section 8.3).

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 identiques¹⁵. 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 :

$$\vec{S} = \sum_{i=1}^N \vec{S}_i \quad (8.12)$$

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 :

$$H = \sum_{i=1}^N \left[\frac{\vec{p}_i^2}{2m} + V(\vec{r}_i) \right] + \frac{1}{2} \sum_{i=1}^N \sum_{j=1, j \neq i}^N W(r_{ij}) ; \quad (8.13)$$

le facteur $\frac{1}{2}$ évite de compter deux fois l'interaction entre deux particules données¹⁶. À nouveau, H est visiblement symétrique dans toute permutation.

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 :

$$\Psi(1, 2, \dots, N ; t + dt) = \Psi(1, 2, \dots, N ; t) + \frac{dt}{i\hbar} H(1, \dots, N) \Psi(1, 2, \dots, N ; t) + \mathcal{O}(dt^2) . \quad (8.15)$$

¹⁵Notons 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.

¹⁶De ce point de vue, le terme à deux corps dans (8.13) s'écrit tout autant :

$$\sum_{i=1}^N \sum_{j=i+1}^N W(r_{ij}) . \quad (8.14)$$

Effectuons l'échange de deux quelconques des particules dans les deux membres. Il vient :

$$\Psi(\dots, j, \dots, i, \dots; t + dt) = \Psi(\dots, j, \dots, i, \dots; t) + \frac{dt}{i\hbar} H(1, \dots, N) \Psi(\dots, j, \dots, i, \dots; t) + \mathcal{O}(dt^2) . \quad (8.16)$$

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 :

$$\Psi_S(\dots, j, \dots, i, \dots; t + dt) = \Psi_S(\dots, i, \dots, j, \dots; t + dt) \quad (8.17)$$

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 :

$$\Psi_A(\dots, j, \dots, i, \dots; t + dt) = -\Psi_A(\dots, i, \dots, j, \dots; t) - \frac{dt}{i\hbar} H(1, \dots, N) \Psi_A(\dots, i, \dots, j, \dots; t) + \mathcal{O}(dt^2) , \quad (8.18)$$

puisque les deux fonctions au second membre changent de signe par hypothèse mais pas H . Il en résulte :

$$\Psi_A(\dots, j, \dots, i, \dots; t + dt) = -\Psi_A(\dots, i, \dots, j, \dots; t + dt) . \quad (8.19)$$

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.

8.3 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 :

$$\begin{pmatrix} 1 & 2 & \dots & N \\ i_1 & i_2 & \dots & i_n \end{pmatrix} , \quad (8.20)$$

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 :

$$\begin{pmatrix} 1 & 2 & \dots & N \\ i_1 & i_2 & \dots & i_n \end{pmatrix} \begin{pmatrix} 1 & 2 & \dots & N \\ j_1 & j_2 & \dots & j_n \end{pmatrix} = \begin{pmatrix} j_1 & j_2 & \dots & j_N \\ k_1 & k_2 & \dots & k_n \end{pmatrix} \begin{pmatrix} 1 & 2 & \dots & N \\ j_1 & j_2 & \dots & j_n \end{pmatrix} . \quad (8.21)$$

Le produit au second membre n'est autre que :

$$\begin{pmatrix} 1 & 2 & \dots & N \\ k_1 & k_2 & \dots & k_n \end{pmatrix} . \quad (8.22)$$

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 :

$$P_\lambda = \begin{pmatrix} 1 & 2 & \dots & N \\ i_1 & i_2 & \dots & i_N \end{pmatrix} , \quad (8.23)$$

alors :

$$P_\lambda \Psi(1, 2, \dots, N) = \Psi(i_1, i_2, \dots, i_N) , \quad (8.24)$$

Les P_λ sont visiblement des opérateurs unitaires ; en effet, un produit scalaire tel que :

$$(P_\lambda \Psi, P_\lambda \Phi) \quad (8.25)$$

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 :

$$(P_\lambda \Psi, P_\lambda \Phi) = (\Psi, \Phi) \quad (8.26)$$

qui est la relation caractéristique d'un opérateur unitaire. Toute permutation étant unitaire, son adjoint est son inverse :

$$P_\lambda^\dagger = P_\lambda^{-1} . \quad (8.27)$$

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 :

$$P_\mu P_{\lambda_1} = P_\nu \quad P_\mu P_{\lambda_2} = P_\nu \iff P_{\lambda_1} = P_\mu^{-1} P_\nu \quad P_{\lambda_2} = P_\mu^{-1} P_\nu \quad (8.28)$$

d'où $P_{\lambda_1} = P_{\lambda_2}$. Le lemme de réarrangement peut s'écrire formellement :

$$P_\lambda \mathcal{S}_N = \mathcal{S}_N . \quad (8.29)$$

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 :

$$T_{ij} = \begin{pmatrix} 1 & 2 & 3 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & N \\ 1 & 2 & 3 & \dots & i-1 & j & i+1 & \dots & j-1 & i & j+1 & \dots & N \end{pmatrix} , \quad (8.30)$$

que l'on peut noter plus simplement :

$$T_{ij} = \begin{pmatrix} i & j \\ j & i \end{pmatrix} . \quad (8.31)$$

Comme toute permutation, les transpositions sont unitaires, mais comme leur carré est l'identité, elles sont également hermitiques :

$$T_{ij}^2 = \mathbf{1} \iff T_{ij} = T_{ij}^{-1} = T_{ij}^\dagger . \quad (8.32)$$

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 a une parité parfaitement définie, que l'on appelle précisément 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 :

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad (8.33)$$

peut se décomposer comme suit :

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = T_{12} T_{23} . \quad (8.34)$$

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 :

$$P_\lambda \Psi_A(1, 2, \dots, N) = (-1)^{\sigma_\lambda} \Psi_A(1, 2, \dots, N) \quad (\text{fermions}) . \quad (8.35)$$

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 :

$$P_\lambda \Psi_S(1, 2, \dots, N) = \Psi_S(1, 2, \dots, N) \quad (\text{bosons}) . \quad (8.36)$$

Ces définitions étant données, on peut maintenant montrer comment construire les fonctions physiquement acceptables, c'est-à-dire celles ayant la symétrie requise vis-à-vis des permutations. L'idée est d'engendrer, à partir d'une fonction Φ donnée mais quelconque, 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'écrire¹⁷ :

$$\Phi_S = (\mathbf{1} + T_{12}) \Phi(1, 2) , \quad \Phi_A = (\mathbf{1} - T_{12}) \Phi(1, 2) \quad (8.37)$$

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 définissant¹⁸ les deux projecteurs Y_S et Y_A :

$$Y_S = \frac{1}{N!} \sum_{\lambda=1}^{N!} P_\lambda , \quad Y_A = \frac{1}{N!} \sum_{\lambda=1}^{N!} (-1)^{\sigma_\lambda} P_\lambda \quad (8.38)$$

que l'on peut collectivement dénoter Y_I :

$$Y_I = \frac{1}{N!} \sum_{\lambda=1}^{N!} \varepsilon_\lambda P_\lambda , \quad \varepsilon_\lambda = \begin{cases} 1 & (\text{I=S, bosons}) \\ (-1)^{\sigma_\lambda} & (\text{I=A, fermions}) \end{cases} . \quad (8.39)$$

Il est facile de voir que

$$Y_S Y_A = Y_A Y_S = 0 . \quad (8.40)$$

De plus, Y_I est bien un projecteur ; en effet, notons d'abord que :

$$P_\nu = P_\lambda P_\mu \implies \sigma_\nu = \sigma_\lambda + \sigma_\mu \iff \varepsilon_\nu = \varepsilon_\lambda \varepsilon_\mu \quad (8.41)$$

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 . Puis :

$$Y_I^2 = \frac{1}{N!^2} \sum_{\lambda=1}^{N!} \sum_{\mu=1}^{N!} \varepsilon_\lambda \varepsilon_\mu P_\lambda P_\mu = \frac{1}{N!^2} \sum_{\nu=1}^{N!} \varepsilon_\nu P_\nu \sum_{\lambda=1}^{N!} 1 . \quad (8.42)$$

La dernière forme est obtenue à partir du lemme de réarrangement. La deuxième sommation donne le simple facteur $N!$, de sorte que :

$$Y_I^2 = Y_I . \quad (8.43)$$

¹⁷Ces combinaisons linéaires ne sont pas normalisées.

¹⁸Le facteur $\frac{1}{N!}$ est introduit pour l'idempotence – voir (8.43).

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 évidemment 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 :

$$P_\mu Y_I = \frac{1}{N!} \sum_{\lambda=1}^{N!} \varepsilon_\lambda P_\mu P_\lambda = \frac{1}{N!} \sum_{\nu=1}^{N!} \frac{\varepsilon_\nu}{\varepsilon_\mu} P_\nu , \quad (8.44)$$

où l'on a à nouveau utilisé (8.41). Comme les ε_λ valent ± 1 , $\frac{\varepsilon_\nu}{\varepsilon_\mu} = \varepsilon_\nu \varepsilon_\mu \forall \mu, \nu$, (8.44) s'écrit tout autant :

$$P_\mu Y_I = \frac{1}{N!} \sum_{\nu=1}^{N!} \varepsilon_\nu \varepsilon_\mu P_\nu \equiv \varepsilon_\mu Y_I . \quad (8.45)$$

Soit maintenant une fonction quelconque Φ ; le résultat de l'action de Y_I sur Φ est soit une fonction de symétrie I, soit une fonction identiquement nulle. En effet, par exemple avec Y_A , on trouve, compte tenu de (8.45) :

$$P_\mu Y_A \Phi = (-1)^{\sigma_\mu} Y_A \Phi . \quad (8.46)$$

La fonction $Y_A \Phi$ est donc une fonction antisymétrique dans toute transposition. De la même façon, on a :

$$P_\mu Y_S \Phi = Y_S \Phi . \quad (8.47)$$

Ainsi, Y_I projette toute fonction donnée 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 zéro¹⁹. Usuellement, on voit d'emblée si l'on est dans un tel cas particulier. Notons enfin que deux fonctions de symétrie différente A et S sont orthogonales d'après (8.40).

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 :

$$\Psi_A = C Y_A \Phi , \quad \Psi_S = C' Y_S \Phi , \quad (8.48)$$

où C et C' sont des constantes de normalisation.

Traitons un exemple avec $N = 3$. Soit $\Phi(1, 2, 3)$ une fonction quelconque. La fonction symétrique (mais non normalisée) engendrée à partir de Φ est :

$$\begin{aligned} \Psi_S &= C' \left[\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right. \\ &\quad \left. + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right] \Phi(1, 2, 3) \\ &= C' \frac{1}{3!} [\Phi(1, 2, 3) + \Phi(2, 1, 3) + \Phi(3, 2, 1) + \Phi(1, 3, 2) + \Phi(2, 3, 1) + \Phi(3, 1, 2)] . \end{aligned} \quad (8.49)$$

Il est évident que cette dernière fonction est complètement symétrique dans toute permutation. De même :

$$\begin{aligned} \Psi_A &= C \left[\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} - \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right. \\ &\quad \left. + \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right] \Phi(1, 2, 3) \\ &= C \frac{1}{3!} [\Phi(1, 2, 3) - \Phi(2, 1, 3) - \Phi(3, 2, 1) - \Phi(1, 3, 2) + \Phi(2, 3, 1) + \Phi(3, 1, 2)] . \end{aligned} \quad (8.50)$$

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

¹⁹ Autre façon de voir : si Φ a déjà une symétrie I ou S, alors $Y_I \Phi = \Phi$. Dès lors $Y_I \Phi = Y_I Y_I \Phi = \delta_{I I'} \Phi$ en vertu de (8.40).

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 orthogonales²⁰.

Pour illustrer 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 distinctes ψ et ϕ supposées orthonormalisées. Compte tenu du postulat de symétrisation, les seuls états physiquement acceptables sont de la forme :

$$\Psi(1, 2) = C [\mathbf{1} + \varepsilon T_{12}] \psi(1)\phi(2) \quad (8.51)$$

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 :

$$|\Psi\rangle = \frac{1}{\sqrt{2}} [\mathbf{1} + \varepsilon T_{12}] |\psi \phi\rangle \quad (8.52)$$

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 :

$$a |a_n\rangle = a_n |a_n\rangle, \quad A |a_n a_m\rangle = (a_n + a_m) |a_n a_m\rangle. \quad (8.53)$$

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 et est donc de la forme :

$$|\Psi'\rangle = \frac{1}{\sqrt{2}} [\mathbf{1} + \varepsilon T_{12}] |a_n a_m\rangle. \quad (8.54)$$

La probabilité de trouver ce résultat est donc :

$$P = \left| \frac{1}{2} \langle a_n a_m | [\mathbf{1} + \varepsilon T_{12}^\dagger] [\mathbf{1} + \varepsilon T_{12}] |\psi \phi\rangle \right|^2. \quad (8.55)$$

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 :

$$P = |\langle a_n a_m | [\mathbf{1} + \varepsilon T_{12}] |\psi \phi\rangle|^2 = |\langle a_n a_m | \psi \phi\rangle + \varepsilon \langle a_n a_m | \phi \psi\rangle|^2. \quad (8.56)$$

Ainsi, dans l'amplitude donnant la probabilité P , les bosons interfèrent avec le signe +, les fermions avec le signe -.

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 :

$$Y_I H \Phi = Y_I E \Phi = E Y_I \Phi, \quad (8.57)$$

d'une part ; d'autre part, en vertu de $Y_I H = H Y_I$, il vient :

$$Y_I H \Phi = H (Y_I \Phi) = E (Y_I \Phi), \quad (8.58)$$

où les parenthèses ne sont là que pour la clarté. (8.58) montre que si Φ est propre de H , alors $Y_I \Phi$ est encore propre²¹ 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 caractère d'état propre.

²⁰Dans 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.

²¹D'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 les deux combinaisons particulières satisfaisant à l'exigence du postulat de symétrisation.

8.4 Applications et exemples

8.4.1 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 particules. 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 :

$$H(1, 2, \dots, N) = \sum_{i=1}^N \left[\frac{\vec{p}_i^2}{2m} + V(\vec{r}_i) \right] \equiv \sum_{i=1}^N H_1(i) . \quad (8.59)$$

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 :

$$H_1 \psi_k(\rho) = E_k \psi_k(\rho) \quad (8.60)$$

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 :

$$\psi_k(\rho) \equiv \phi_k(\vec{r}) \chi(m_s) , \quad H_1 \phi_k(\vec{r}) = E_k \phi_k(\vec{r}) . \quad (8.61)$$

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.

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 :

$$\begin{aligned} H \prod_{j=1}^N \psi_{k_j}(\rho_j) &= \sum_{i=1}^N \prod_{j=1}^N H_1(\rho_i) \psi_{k_j}(\rho_j) = \sum_{i=1}^N \left[\prod_{j=1, j \neq i}^N \psi_{k_j}(\rho_j) \right] H_1(\rho_i) \psi_{k_i}(\rho_i) \\ &= \sum_{i=1}^N \left[\prod_{j=1, j \neq i}^N \psi_{k_j}(\rho_j) \right] E_{k_i} \psi_{k_i}(\rho_i) = \sum_{i=1}^N E_{k_i} \prod_{j=1}^N \psi_{k_j}(\rho_j) \end{aligned} \quad (8.62)$$

On a donc :

$$H \prod_{j=1}^N \psi_{k_j}(\rho_j) = E_{k_1, k_2, \dots, k_N} \prod_{j=1}^N \psi_{k_j}(\rho_j) , \quad E_{k_1, k_2, \dots, k_N} = \sum_{j=1}^N E_{k_j} . \quad (8.63)$$

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 ; il convient maintenant de lui appliquer l'un des opérateurs Y_{Γ} (8.39) afin d'obtenir une fonction (non normalisée) de symétrie convenable. Considérons d'abord le cas des bosons. Une bonne fonction est :

$$\Psi_S(1, 2, \dots, N) = C' Y_S \prod_{i=1}^N \psi_{k_i}(\rho_i) \equiv C' Y_S \Phi(1, 2, \dots, N) \quad (8.64)$$

l'énergie associée à Ψ_S étant la somme des E_{k_i} , conformément à (8.63). 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 fondamental²² 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

²²Il s'agit bien sûr de l'état à température nulle.

à nombre de bosons fixé, au phénomène appelé condensation de Bose²³. L'état fondamental des N bosons est alors :

$$\Psi_{\text{fond. bosons}} = \psi_1(\rho_1) \psi_1(\rho_2) \dots \psi_1(\rho_N) . \quad (8.65)$$

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 :

$$Y_A = \frac{1}{2!} \left[\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right] . \quad (8.66)$$

D'où il résulte :

$$\Psi_A(1, 2) = C Y_A \psi_k(1) \psi_k(2) = \frac{C}{2} [\psi_k(1) \psi_k(2) - \psi_k(2) \psi_k(1)] \equiv 0 . \quad (8.67)$$

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 ailleurs²⁴, 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 :

$$\Psi_A(1, 2, \dots, N) = C Y_A [\psi_1(\rho_1) \psi_2(\rho_2) \dots \psi_N(\rho_N)] \equiv C Y_A \Phi(1, 2, \dots, N) . \quad (8.68)$$

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 ont été introduites dans le chapitre 5. On peut donc utiliser une fonction ϕ_k (orbitale) deux fois au plus, l'associant soit avec α , soit avec β . Ainsi, pour un nombre pair d'électrons, tous ceux-ci sont appariés dans l'état fondamental et une bonne fonction pour cet état s'écrira :

$$\Psi_A(1, 2, \dots, N) = C Y_A \prod_{i=1}^{\frac{N}{2}} \phi_i(\vec{r}_{2i-1}) \alpha(m_{s\ 2i-1}) \phi_i(\vec{r}_{2i}) \beta(m_{s\ 2i}) . \quad (8.69)$$

²³À température finie, le jeu des fluctuations thermiques peut supprimer la condensation de Bose si la dimensionnalité d est faible : l'argument *standard* strict – qui n'est pas exempt de critiques (Yvan Simon, communication privée) – montre qu'il n'y a pas de condensation de Bose si $d \leq 2$ (il donne $T_C \propto [\zeta(\frac{d}{2})]^{-\frac{2}{d}}$ où ζ est la fonction de Riemann : pour $d \in \mathbb{R}$, $d \leq 2$, la fonction de Riemann est infinie et la température critique est *nulle*). Pour $d > 2$, la condensation se produit à une certaine température finie 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) ; d'ailleurs, quand on examine pour un gaz parfait presque classique (densité \ll (longueur d'onde de de Broglie) $^{-d} \propto T^{\frac{d}{2}}$) la première correction quantique à la fonction de partition, celle-ci peut s'interpréter en considérant une interaction effective (dépendant de la température...), attractive pour les bosons, répulsive pour les fermions. Notons enfin que la question de la condensation de Bose en présence d'interactions entre les bosons (au sens ordinaire du terme) est encore actuellement un problème largement ouvert.

²⁴notamment si la loi de dispersion est la même.

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 seul électron non apparié, responsable d'un moment cinétique de spin total pour le système égal à $S = \frac{1}{2}$. 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 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 :

$$\Psi_{S=1, M_S=1}(1, 2) = \frac{1}{\sqrt{2}} [\phi_1(\vec{r}_1)\phi_2(\vec{r}_2) - \phi_2(\vec{r}_1)\phi_1(\vec{r}_2)] \alpha(m_{s_1})\alpha(m_{s_2}) , \quad (8.70)$$

et la probabilité de trouver les deux fermions l'un près de \vec{r}_1 , l'autre près de \vec{r}_2 est :

$$dP_{S=1}(\vec{r}_1, \vec{r}_2) = \frac{1}{2} |\phi_1(\vec{r}_1)\phi_2(\vec{r}_2) - \phi_2(\vec{r}_1)\phi_1(\vec{r}_2)|^2 d^3r_1 d^3r_2 . \quad (8.71)$$

En particulier, la probabilité élémentaire est nulle de trouver les deux électrons au même point :

$$dP_{S=1}(\vec{r}_1 = \vec{r}_2) = 0 . \quad (8.72)$$

C'est ce que l'on appelle le "trou de Fermi" : 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) de spin donné en un point interdit l'accès de ce point à un autre électron de même spin. 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 :

$$\Psi_{S=0, M_S=0}(1, 2) = \frac{1}{\sqrt{2}} [\phi_1(\vec{r}_1)\phi_2(\vec{r}_2) + \phi_2(\vec{r}_1)\phi_1(\vec{r}_2)] \frac{1}{\sqrt{2}} [\alpha(m_{s_1})\beta(m_{s_2}) - \beta(m_{s_1})\alpha(m_{s_2})] . \quad (8.73)$$

et la densité $dP_{S=0}(\vec{r}_1, \vec{r}_2)$ ne s'annule pas en $\vec{r}_1 = \vec{r}_2$.

Ainsi, 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. 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. 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.

À 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 :

$$\Psi_A(1, 2) = \frac{1}{\sqrt{2!}} \begin{vmatrix} \psi_{k_1}(1) & \psi_{k_2}(1) \\ \psi_{k_1}(2) & \psi_{k_2}(2) \end{vmatrix} . \quad (8.74)$$

Il est facile de se convaincre que la fonction d'onde antisymétrique (8.68) construite sur un produit de N fonctions²⁵ à une particule $\{\psi_{k_i}\}$ peut de même s'écrire sous la forme d'un déterminant, historiquement appelé déterminant de Slater :

$$\Psi_A(1, 2, \dots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{k_1}(1) & \psi_{k_2}(1) & \dots & \psi_{k_N}(1) \\ \psi_{k_1}(2) & \psi_{k_2}(2) & \dots & \psi_{k_N}(2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{k_1}(N) & \psi_{k_2}(N) & \dots & \psi_{k_N}(N) \end{vmatrix} . \quad (8.75)$$

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 (8.68) et en posant $\Phi = \prod_{i=1}^N \psi_{k_i}$, on a :

$$\langle \Psi_A | \Psi_A \rangle = |C|^2 \langle Y_A \Phi | Y_A \Phi \rangle = |C|^2 \langle \Phi | Y_A^\dagger Y_A | \Phi \rangle . \quad (8.76)$$

Puisque Y_A est hermitique et idempotent, ceci est égal à :

$$\langle \Psi_A | \Psi_A \rangle = |C|^2 \langle \Phi | Y_A \Phi \rangle = \frac{|C|^2}{N!} \langle \Phi | \sum_{\lambda=1}^{N!} P_\lambda | \Phi \rangle . \quad (8.77)$$

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 :

$$\langle \Psi_A | \Psi_A \rangle = \frac{|C|^2}{N!} \prod_{i=1}^N \langle \psi_{k_i} | \psi_{k_i} \rangle = \frac{|C|^2}{N!} = 1 \iff |C| = \sqrt{N!} . \quad (8.78)$$

Finalement :

$$|\Psi_A\rangle = \frac{\sqrt{N!}}{N!} \sum_{\lambda=1}^{N!} P_\lambda |\Phi\rangle = \frac{1}{\sqrt{N!}} \sum_{\lambda=1}^{N!} P_\lambda \left| \prod_{i=1}^N \psi_{k_i} \right\rangle , \quad (8.79)$$

qui reproduit bien l'expression (8.75).

Sur la forme (8.75), on voit immédiatement que si deux fonctions ψ_{k_i} coïncident (c'est-à-dire si deux *spin-orbitales* sont les mêmes), la fonction d'onde est identiquement nulle : alors, le déterminant a deux colonnes identiques. Enfin, signalons la notation abrégée courante pour l'expression (8.75) :

$$\Psi_A(1, 2, \dots, N) = \frac{1}{\sqrt{N!}} \text{Det} [\psi_{k_1} \psi_{k_2} \dots \psi_{k_N}] . \quad (8.80)$$

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. Dans une telle approche, les ψ_{k_i} sont au contraire inconnues *a priori* 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 pratique – 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

²⁵Bien noter que le déterminant est construit avec des spin-orbitales, pas avec des orbitales !

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 eux²⁶. En pareil 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 (8.70) 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 $S_- = S_{1-} + S_{2-}$ produit²⁷ :

$$(S_{1-} + S_{2-}) \frac{1}{\sqrt{2}} \text{Det}[\phi_1\alpha \phi_2\beta] = \frac{\hbar}{\sqrt{2}} [\text{Det}[\phi_1\beta \phi_2\alpha] + \text{Det}[\phi_1\alpha \phi_2\beta]] . \quad (8.81)$$

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 :

$$\Psi_{S=1 M_S=0} = \frac{1}{2} [\text{Det}[\phi_1\beta \phi_2\alpha] + \text{Det}[\phi_1\alpha \phi_2\beta]] . \quad (8.82)$$

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 (8.82) est donc :

$$\Psi_{S=0 M_S=0} = \frac{1}{2} [\text{Det}[\phi_1\beta \phi_2\alpha] - \text{Det}[\phi_1\alpha \phi_2\beta]] , \quad (8.83)$$

obtenue par un simple changement de signe dans la combinaison linéaire des deux déterminants.

8.4.2 L'atome d'hélium

L'atome d'hélium ($Z = 2$) est l'atome le plus simple contenant déjà la complexité foncière 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 que les interactions de type électrostatique de Coulomb – étant entendu que les corrections relativistes ne peuvent être ignorées, surtout dans le cas des atomes lourds²⁸. Dans ces conditions, le Hamiltonien s'écrit avec des notations évidentes²⁹ :

$$H(\vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2) = \frac{\vec{p}_1^2}{2m} - \frac{2e'^2}{r_1} + \frac{\vec{p}_2^2}{2m} - \frac{2e'^2}{r_2} + \frac{e'^2}{r_{12}} , \quad (8.84)$$

ou, sous une forme plus concise :

$$H(\vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2) = [H_1(\vec{r}_1, \vec{p}_1) + H_1(\vec{r}_2, \vec{p}_2)] + \frac{e'^2}{r_{12}} \equiv H_0 + V . \quad (8.85)$$

$H_1(\vec{r}, \vec{p}) = \frac{p^2}{2m} - 2\frac{e'^2}{r}$ est le Hamiltonien hydrogénoïde pour $Z = 2$. On note au passage que H est bien invariant dans toute permutation, comme ils se doit pour des particules identiques, ici les deux électrons.

²⁶De 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.

²⁷étant entendu que $S_- \alpha = \hbar\sqrt{\frac{1}{2}(\frac{1}{2}+1) - \frac{1}{2}(\frac{1}{2}-1)} \beta \equiv \hbar\beta$.

²⁸Il reste que la structure fine de l'hélium est aisément observable.

²⁹On fait ici l'approximation du noyau infiniment massif.

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, qui joue un rôle très 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 $\frac{e'^2}{a_0}$, soit plusieurs eV. Ceci revient à considérer d'abord H_0 défini en (8.85), qui est une somme d'opérateurs à une particule dont on connaît les solutions exactes ; H_0 définit l'approximation dite des électrons indépendants. Si ϕ_n est une fonction propre hydrogénoïde avec $Z = 2$, d'énergie³⁰ :

$$E_n = -\frac{m(2e')^2}{2n^2\hbar^2} \simeq -\frac{54.45}{n^2} \text{ eV} , \quad (8.86)$$

alors tout produit $\phi_n\phi_m$ est fonction propre de H_0 :

$$H_0 \phi_n(\vec{r}_1)\phi_m(\vec{r}_2) = (E_n + E_m) \phi_n(\vec{r}_1)\phi_m(\vec{r}_2) \equiv E_{nm}\phi_n(\vec{r}_1)\phi_m(\vec{r}_2) . \quad (8.87)$$

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 :

$$H_0 \phi_n(\vec{r}_1)\chi(m_{s_1})\phi_m(\vec{r}_2)\chi(m_{s_2}) = E_{nm}\phi_n(\vec{r}_1)\chi(m_{s_1})\phi_m(\vec{r}_2)\chi(m_{s_2}) . \quad (8.88)$$

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 :

$$X_{1,+1} = \alpha(m_{s_1})\alpha(m_{s_2}) , \quad X_{1,-1} = \beta(m_{s_1})\beta(m_{s_2}) \quad (8.89)$$

$$X_{1,0} = \frac{1}{\sqrt{2}} [\alpha(m_{s_1})\beta(m_{s_2}) + \beta(m_{s_1})\alpha(m_{s_2})] \quad (8.90)$$

2. état singulet, dont la composante est antisymétrique :

$$X_{0,0} = \frac{1}{\sqrt{2}} [\alpha(m_{s_1})\beta(m_{s_2}) - \beta(m_{s_1})\alpha(m_{s_2})] . \quad (8.91)$$

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 :

$$\Phi_1(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2}} [\phi_n(\vec{r}_1)\phi_m(\vec{r}_2) - \phi_m(\vec{r}_1)\phi_n(\vec{r}_2)] , \quad (8.92)$$

tandis que la partie spatiale du singulet est :

$$\Phi_0(\vec{r}_1, \vec{r}_2) = \frac{1}{\sqrt{2}} [\phi_n(\vec{r}_1)\phi_m(\vec{r}_2) + \phi_m(\vec{r}_1)\phi_n(\vec{r}_2)] . \quad (8.93)$$

En définitive, les états construits sur un produit d'orbitales sont :

$$\Psi_{1, M_S} = \frac{1}{\sqrt{2}} [\phi_n\phi_m - \phi_m\phi_n] X_{1, M_S} , \quad M_S = 0, \pm 1 \quad (\text{triplet}) . \quad (8.94)$$

$$\Psi_{0,0} = \frac{1}{\sqrt{2}} [\phi_n\phi_m + \phi_m\phi_n] X_{0,0} , \quad M_S = 0 \quad (\text{singulet}) . \quad (8.95)$$

³⁰(8.86) montre aussi que l'état fondamental de l'ion He^+ a pour énergie -54.45 eV.

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 cas³¹ $N = 2$. 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-hélium³².

Les états de spin étant déterminés, ainsi que la forme de la fonction d'onde, 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 :

$$\begin{aligned} & (1s)^2 \\ & (1s) (2s) \quad (1s) (2p) \\ & (1s) (3s) \quad (1s) (3p) \quad (1s) (3d) \quad \textit{etc.} \end{aligned} \quad (8.96)$$

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 :

$$\begin{aligned} (1s)^2 & : \phi_{100} \phi_{100} \\ (1s) (2s) & : \phi_{100} \phi_{200} \quad (1s) (2p) : \phi_{100} \phi_{21m_l} \\ (1s) (3s) & : \phi_{100} \phi_{300} \quad (1s) (3p) : \phi_{100} \phi_{31m_l} \quad (1s) (3d) : \phi_{100} \phi_{32m_l} \quad \textit{etc.} \end{aligned} \quad (8.97)$$

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 :

$$\Psi_{1^1S}(1, 2) = \phi_{100}(\vec{r}_1) \phi_{100}(\vec{r}_2) \frac{1}{\sqrt{2}} [\alpha(m_{s_1})\beta(m_{s_2}) - \beta(m_{s_1})\alpha(m_{s_2})] , \quad (8.98)$$

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 numériquement³³ que l'énergie de tout état où les deux électrons sont excités est supérieure à l'énergie de l'état fondamental de l'ion 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 :

$$\text{triplet} : \frac{1}{\sqrt{2}} [\phi_{100} \phi_{2lm_l} - \phi_{2lm_l} \phi_{100}] X_{1, M_S} \quad (l = 0, 1, -l \leq m_l \leq l) \quad (8.99)$$

$$\text{singulet} : \frac{1}{\sqrt{2}} [\phi_{100} \phi_{2lm_l} + \phi_{2lm_l} \phi_{100}] X_{0, 0} \quad (l = 0, 1, -l \leq m_l \leq l) . \quad (8.100)$$

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 (voir (8.86)) 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 : la fonction $\phi_{100}(\vec{r}_1) \phi_{2lm_l}(\vec{r}_2) - \phi_{2lm_l}(\vec{r}_1) \phi_{100}(\vec{r}_2)$ est visiblement nulle si $\vec{r}_1 = \vec{r}_2$.

³¹et aussi, pour N quelconque, pour les composantes $M_S = \pm S$ de l'état $S = N \frac{\hbar}{2}$ (état "ferromagnétique").

³²Seules des transitions Mq , en particulier dipolaires magnétiques ($q = 1$), permettent de "convertir" une espèce dans l'autre.

³³Si les deux électrons sont dans l'état 2s, l'énergie (associée à H_0) est $2 \times \frac{1}{4}(-54.45)$ eV > -54.45 eV. Physiquement, un tel atome ainsi préparé s'ionise spontanément.

Si on choisit la configuration (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 1^3S , 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 1^3P 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 $V = \frac{e'^2}{r_{12}}$ comme une perturbation. On verra après coup que la répulsion est loin d'être une petite correction (ce qui n'est guère surprenant physiquement) – d'où quelque suspicion sur la validité d'un traitement perturbatif –, mais la prise en compte, même grossière, de la répulsion de Coulomb lève le voile sur un phénomène inattendu, déjà mentionné : en dépit de l'absence d'interactions magnétiques (le Hamiltonien est purement électrostatique), les états atomiques 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 :

$$\Delta E_{1^1S} = \langle \Psi_{1^1S} | \frac{e'^2}{r_{12}} | \Psi_{1^1S} \rangle . \quad (8.101)$$

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 :

$$\Delta E_{1^1S} = \langle \phi_{100}\phi_{100} | \frac{e'^2}{r_{12}} | \phi_{100}\phi_{100} \rangle , \quad (8.102)$$

où l'élément de matrice est l'intégrale :

$$\int d^3r_1 \int d^3r_2 \phi_{100}(r_1)\phi_{100}(r_2) \frac{e'^2}{r_{12}} \phi_{100}(r_1)\phi_{100}(r_2) \equiv J_{1s1s} . \quad (8.103)$$

L'intégrale J_{1s1s} s'appelle intégrale directe³⁴ et représente la répulsion électrostatique des deux distributions de charges de densités respectives $|\phi_{100}(\vec{r}_1)|^2$ et $|\phi_{100}(\vec{r}_2)|^2$; elle vaut³⁵ $\frac{5}{4} \frac{e'^2}{a_0} \simeq 34.01$ eV [17] ; c'est visiblement une quantité positive – c'est également une correction très importante ! À cet ordre, l'énergie du fondamental est :

$$E_{1^1S} = 2E_1 + J_{1s1s} = -74.89 \text{ eV} . \quad (8.104)$$

La répulsion de Coulomb entre électrons déstabilise considérablement le fondamental, ce qui n'est pas surprenant.

Le premier niveau excité d'énergie $E_1 + E_2$ (à l'ordre zéro) est dégénéré 16 fois ; il convient donc d'une part d'écrire la matrice de la perturbation dans le sous-espace dégénéré et d'autre part 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 :

$$[V, \vec{L}] = 0 , \quad [V, \vec{S}] = 0 , \quad (8.105)$$

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 :

$$\langle \Psi_{LSM_L M_S} | \frac{e'^2}{r_{12}} | \Psi_{L'S'M'_L M'_S} \rangle \propto \delta_{LL'} \delta_{SS'} \delta_{M_L M'_L} \delta_{M_S M'_S} , \quad (8.106)$$

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 :

³⁴ *Attention !* Ceci est la terminologie usuelle en physique atomique et moléculaire (voir par exemple [40]). En revanche, Cohen-Tannoudji *et al.* [4] 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).

³⁵ Les intégrales du genre J et K se calculent en développant $\frac{1}{r_{12}}$ en produit (tensoriel) d'harmoniques sphériques, ce qui a le bon goût de séparer les variables 1 et 2 (voir [16], Appendice B).

la répulsion de Coulomb va lever partiellement la dégénérescence présente dans l'approximation grossière des électrons indépendants³⁶.

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 :

$$\begin{aligned} \Delta E_{2^1S} &= \frac{1}{2} \langle \phi_{100}\phi_{200} + \phi_{200}\phi_{100} | \frac{e'^2}{r_{12}} | \phi_{100}\phi_{200} + \phi_{200}\phi_{100} \rangle \\ &= \langle \phi_{100}\phi_{200} | \frac{e'^2}{r_{12}} | \phi_{100}\phi_{200} \rangle + \langle \phi_{100}\phi_{200} | \frac{e'^2}{r_{12}} | \phi_{200}\phi_{100} \rangle \equiv J_{1s2s} + K_{1s2s} . \end{aligned} \quad (8.107)$$

On retrouve une intégrale du type direct, J_{1s2s} (répulsion électrostatique entre les deux distributions ϕ_{100} et ϕ_{200} , égale à $\frac{34}{81} \frac{e'^2}{a_0}$, soit environ 11.42 eV) et une autre intégrale, K_{1s2s} , appelée intégrale d'échange³⁷, égale 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'échange³⁸ est toujours positive ([41], Appendice 19) :

$$K_{nm} = \int d^3r_1 \int d^3r_2 \phi_n(\vec{r}_1)\phi_m(\vec{r}_2) \frac{e'^2}{r_{12}} \phi_m(\vec{r}_1)\phi_n(\vec{r}_2) > 0 \quad \forall \phi_n, \phi_m . \quad (8.108)$$

À cet ordre, on a maintenant :

$$E_{2^1S} = E_1 + E_2 + J_{1s2s} + K_{1s2s} \simeq -55.45 \text{ eV} . \quad (8.109)$$

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 :

$$E_{1^3S} = E_1 + E_2 + J_{1s2s} - K_{1s2s} \simeq -57.84 \text{ eV} , \quad (8.110)$$

et, en raison du signe de l'intégrale d'échange, (8.108), on obtient l'inégalité importante :

$$E_{1^3S} < E_{2^1S} \quad (8.111)$$

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 incapable d'expliquer les ordres de grandeur.

Pour les états P on trouve de façon analogue³⁹ :

$$E_{1^1P} = E_1 + E_2 + J_{1s2p} + K_{1s2p} \simeq -53.91 \text{ eV} \quad (8.112)$$

$$E_{1^3P} = E_1 + E_2 + J_{1s2p} - K_{1s2p} \simeq -55.77 \text{ eV} , \quad (8.113)$$

avec à nouveau :

$$E_{1^3P} < E_{1^1P} . \quad (8.114)$$

³⁶Bien sûr, les énergies perturbées ne peuvent dépendre ni de M_L , ni de M_S , puisqu'il n'y a pas de direction privilégiée dans l'espace (physique).

³⁷L'apparition du terme d'échange est clairement une conséquence du postulat de symétrisation et n'a pas d'interprétation classique.

³⁸L'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 $100 a_0$ [40]. K provient du postulat de symétrisation : la symétrisation est donc inutile si les fonctions d'onde ne se recouvrent pas [20].

³⁹ $J_{1s2p} \simeq 13.22$ eV, $K_{1s2p} \simeq 0.93$ eV.

En définitive, la succession des niveaux classés par énergie croissante est obtenue comme :

$$E_{1^1S} < E_{1^3S} < E_{1^3P} < E_{2^1S} < E_{1^1P} . \quad (8.115)$$

Au total, cette approximation assez grossière⁴⁰ 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 l'ordre des niveaux n'est pas conforme à l'expérience ; du spectre observé, et en raisonnant avec les règles de sélection, on peut en effet conclure que l'ordre réel est le suivant :

$$E_{1^1S} < E_{1^3S} < E_{2^1S} < E_{1^3P} < E_{1^1P} . \quad (8.116)$$

Le *splitting* entre 1^3S et 2^1S vaut environ 0.8 eV, celui entre 1^3P et 1^1P vaut 0.74 eV. L'écart entre E_{2^1S} et E_{1^3P} est égal à 0.34 eV ; la relative faiblesse de cet écart permet d'accepter la nécessité d'un traitement nettement plus précis pour rendre compte du bon ordonnancement.

8.4.3 Modèle simpliste d'un solide. Conducteurs et isolants

La matière condensée à l'état solide possède des propriétés très variables d'un corps à l'autre ; cette extrême variété suggère de prime abord une classification selon des caractéristiques macroscopiques, aisément mises en évidence au moins au plan qualitatif. Elle est en grande partie fondée sur des arguments physiques exploitant la robustesse des constituants "élémentaires" (très précisément : atomes, molécules, ions, ...) : dans la phase condensée, ces derniers se retrouvent profondément altérés ou au contraire presque intacts par rapport à leur état libre en phase diluée. Par ailleurs, la disposition dans l'espace des constituants une fois rassemblés permet de distinguer deux grandes classes de solides : les cristaux, caractérisés par un arrangement spatial ayant une symétrie de translation discrète (existence d'un réseau géométrique noté \mathcal{B}) et présentant un ordre à longue portée (en toute rigueur infinie⁴¹), et les verres, qui ne possèdent qu'un ordre à courte distance et que l'on peut se représenter comme des liquides figés.

Dans le cas de la matière ordonnée, il existe de nombreux cas où ce sont des ions positifs qui sont disposés aux nœuds du réseau⁴². Comme usuellement, la condensation se fait à partir d'objets électriquement neutres, la bonne vision du solide est dans ce cas un réseau statique⁴³ d'ions ; pour les cristaux ioniques (exemple : NaCl), les atomes d'une espèce cèdent chacun un électron à ceux de l'autre espèce, et il n'y a pas de charge plus ou moins libre⁴⁴. Au contraire, pour les métaux (exemple le sodium), chaque atome perd un électron lors de la condensation solide et la bonne image du solide est celle d'un réseau d'ions positifs donne immergé dans une mer d'électrons presque libres. Dans ce dernier cas, le réseau d'ions positifs constitue le potentiel *périodique* au sein duquel se meuvent les électrons libérés lors de la constitution du solide formé à partir d'un système globalement neutre. C'est de ce type de solide qu'il est question dans la suite.

Les électrons libérés forment un fluide – un fluide de fermions – dont on peut négliger dans un premier temps les interactions internes⁴⁵. On se retrouve alors dans le cadre d'un modèle de particules indépendantes, chacune d'entre elles étant soumise au potentiel périodique du réseau d'ions. Les états du solide – quand on

⁴⁰On 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.

⁴¹Il s'agit bien sûr d'une vision idéale, ignorant délibérément les défauts de toute nature (impuretés, lacunes, dislocations, etc.)

⁴²C'est le cas pour les cristaux ioniques et pour les métaux.

⁴³On laisse de côté les vibrations du réseau (phonons). Celles-ci sont cependant cruciales pour comprendre l'origine de la loi d'Ohm : sans les collisions électron-phonon, le gaz d'électrons constitue un système purement mécanique, qui n'est le siège d'aucune dissipation d'énergie. Dans un modèle classique sans frottement, les électrons soumis à un champ électrique $\vec{\mathcal{E}}$ ont un mouvement uniformément accéléré ; le courant est $\vec{j} = \rho \vec{v}$ et ne peut entrer dans la loi d'Ohm selon $\vec{j} = \sigma \vec{\mathcal{E}}$ puisque la vitesse n'est pas constante ; il en va de même en Mécanique Quantique. La description du transport quantique puise dans l'arsenal de la Mécanique Statistique hors d'équilibre (équation de Boltzmann, ou l'un de ses avatars, et théorie de la réponse linéaire, notamment).

⁴⁴Si un cristal ionique a néanmoins une conductivité finie (non-nulle), c'est grâce aux défauts ; quoi qu'il en soit les porteurs sont alors des ions, de masse énorme par rapport à celle d'un électron et la mobilité est très petite.

⁴⁵Le traitement de l'atome d'hélium donné dans la section 8.4.2 montre que négliger la répulsion de Coulomb entre électrons est une grossière approximation. Pour un solide (et notamment un métal), les choses sont plus subtiles. L'approximation des électrons indépendants y est réhabilitée par le fait que le gaz d'électrons, en raison de sa densité, est le siège du phénomène d'écrantage. De ce fait, la répulsion de Coulomb est fortement réduite et se trouve (au moins pour les densités électroniques assez élevées) reléguée à un rôle secondaire pour la description qualitative en cours.

néglige les vibrations du réseau, ce sont en fait les états du gaz d'électrons – seront donc des déterminants de Slater construits sur des spin-orbitales obtenues comme états propres d'un Hamiltonien à un électron contenant le potentiel périodique ci-dessus. Les modes propres (fonctions et valeurs propres) d'un tel Hamiltonien présentent des caractéristiques dépendant fortement de la géométrie du réseau. Pour la discussion en cours, le seul point important est l'existence de bandes d'énergie permises ou interdites : les domaines d'énergie autorisée à un électron sont *non-connexes* ; ce fait majeur est conséquence directe de la symétrie discrète de translation⁴⁶.

Pour un électron libre, la loi de dispersion est :

$$\varepsilon^{(0)}(\vec{k}) = \frac{\hbar^2 \vec{k}^2}{2m} . \quad (8.117)$$

En conséquence, il existe des états pour *toutes* les valeurs réelles positives de l'énergie : toutes les énergies sont permises et sont données par une loi parabolique en fonction du vecteur d'onde \vec{k} ; exprimé autrement, la densité d'états en énergie⁴⁷, $\rho(\varepsilon)$, est une fonction non-nulle de 0 à $+\infty$, ne présentant aucun trou (*gap*). Au contraire, en présence d'un potentiel périodique, une réflexion (dite de Bragg) se produit à chaque fois que⁴⁸ \vec{k} appartient à un certain réseau défini dans l'espace des \vec{k} , appelé réseau réciproque $\tilde{\mathcal{B}}$. Au voisinage d'un tel point la parabole libre s'incurve et un *gap* s'ouvre, d'autant plus grand que le potentiel périodique est fort : la parabole se fragmente et la loi de dispersion est une succession d'arcs disjoints $\varepsilon_n(k)$. Par exemple, à une dimension, pour un réseau de paramètre a , un *gap* s'ouvre pour les valeurs de k telles que $k = \frac{\pi}{a} + n\frac{2\pi}{a} \equiv k_n$, $n \in \mathbb{Z}$. Le *gap* en énergie est donné par (voir fig. 8.2) :

$$\Delta E_{g_n} = \lim_{k \rightarrow k_n+0} \varepsilon_{n+1}(k) - \lim_{k \rightarrow k_n-0} \varepsilon_n(k) . \quad (8.120)$$

Au total, quand l'énergie varie de 0 à $+\infty$, on rencontre une succession de bandes alternativement permises et interdites, c'est-à-dire qu'il n'existe d'états pour l'électron que si l'énergie propre appartient à certains segments *disjoints* de l'axe réel, complètement déterminés par la nature géométrique du réseau de Bravais et l'intensité du potentiel périodique.

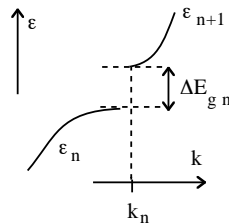


Figure 8.2: Représentation schématique du *gap* entre deux bandes permises.

En toute rigueur, pour un solide *infini* dans toutes les directions, l'énergie varie *continûment* à l'intérieur d'une bande permise. Pour des raisons physiques – et pour éviter des complications techniques inessentiels – on raisonne sur un échantillon de taille linéaire L arbitrairement grande mais finie et on adopte des conditions cycliques de Born - von Karman ; dès lors, les états permis forment un quasi-*continuum*, deux états consécutifs

⁴⁶Se souvenir du modèle de Kronig - Penney.

⁴⁷Dans \mathbb{R}^d , $\rho(\varepsilon)$ se définit comme :

$$\rho(\varepsilon)d\varepsilon = \Omega_d k^{d-1} dk , \quad (8.118)$$

\vec{k} étant défini par inversion de la loi de dispersion ; d est la dimensionnalité et Ω_d la surface de la sphère de rayon unité dans \mathbb{R}^d ($\Omega_d = \frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})}$, Γ étant la fonction d'Euler de deuxième espèce). Ainsi, par exemple, de (8.117) et (8.118) on déduit la densité d'états pour un électron libre :

$$\rho^{(0)}(\varepsilon) = \frac{1}{2} \Omega_d \left(\frac{2m}{\hbar^2} \right)^{\frac{d}{2}} \varepsilon^{\frac{d}{2}-1} . \quad (8.119)$$

On remarque que la densité d'états libre est constante à $d = 2$ et diverge à énergie nulle pour $d < 2$.

⁴⁸La quantité $\hbar\vec{k}$ a tous les aspects d'une impulsion pour l'électron dans le cristal. On verra par la suite pourquoi il vaut mieux l'appeler *pseudo-impulsion* de l'électron.

étant séparés en énergie par $\delta\varepsilon = \|\vec{\nabla}\varepsilon\| \delta k$, avec $\delta k = \frac{2\pi}{L}$ (qui est aussi petit que l'on veut). Cet échantillon fini contient par construction un nombre fini d'électrons, N .

Il reste maintenant à construire les états du solide, à commencer par l'état fondamental⁴⁹. Celui-ci s'obtient en mettant les électrons un à un dans chaque état disponible (une spin-orbitale), conformément au principe de Pauli, par ordre d'énergie croissante. Chaque état propre du Hamiltonien (orbitale) peut, comme d'habitude, contenir au plus deux électrons. Plusieurs cas peuvent se présenter :

1. la 1^{ère} bande d'énergie est pleine avant épuisement du stock d'électrons
2. on épuise le stock d'électrons avant d'avoir complètement rempli la 1^{ère} bande
3. on arrive pile, avec les deux derniers électrons, au dernier état disponible de la bande considérée.

Dans le cas 1., on prend la bande suivante et on continue le remplissage pour se retrouver finalement soit dans le cas 2. soit dans le cas 3., seuls cas qu'il suffit donc de considérer dans la suite. En définitive, la plus haute bande pertinente (utilisée) est donc soit pleine, soit partiellement remplie⁵⁰.

- Quand la plus haute bande est complètement remplie, on ne peut donner de l'énergie au gaz d'électrons qu'en lui fournissant au moins ΔE_g , valeur du *gap* séparant la bande pleine de la prochaine bande permise (vide). ΔE_g vaut quelques eV pour le diamant et une fraction d'eV pour les semi-conducteurs.

Appliquer un champ électrique au solide, éventuellement aussi petit que l'on veut – mais de toute façon très petit par rapport aux champs microscopiques – c'est vouloir accélérer les électrons en leur donnant de l'énergie cinétique. Dans le cas présent, c'est strictement impossible, le champ appliqué ne pouvant donner aux électrons les plus hauts en énergie une énergie supplémentaire au moins égale à ΔE_g leur permettant de franchir le *gap*. Une estimation de l'énergie gagnée par un électron sous l'effet d'un champ \mathcal{E} donne $\Delta E_{\text{cin}} \sim |e|\mathcal{E}l$, où l est le libre parcours moyen, soit $\Delta E_{\text{cin,eV}} \sim \mathcal{E}l$. Pour un champ de 1 V/cm et avec $l \sim 100 \text{ \AA}$, on a :

$$\Delta E_{\text{cin,eV}} \sim 1 \times 10^{22} \times 100 \times 10^{-10} = 10^{-6} \text{ eV} \ll \Delta E_g . \quad (8.121)$$

En pareil cas, le solide est un isolant : aucun champ appliqué ne peut donner lieu, à température nulle, au passage d'un courant électrique⁵¹. En réalité, c'est en considérant la possibilité de saut du *gap* par activation thermique que la notion d'isolant apparaît de façon pertinente. S'en tenant pour simplifier à des facteurs de Boltzmann, on voit que $\exp[-\frac{\Delta E_g}{k_B T}]$ reste un nombre incroyablement petit à l'ambiante pour un solide comme le diamant, dont l'énergie de *gap* est de l'ordre de 5 eV. En revanche, pour un semi-conducteur intrinsèque ($\Delta E_g \sim$ quelques $\frac{1}{10}$ eV), la résistivité n'est pas si grande – et *décroît* fortement avec la température – au contraire d'un métal. En fait, ce qui importe c'est à la fois le facteur de Boltzmann (qui reste très petit à l'ambiante) et la densité de charge pseudo-libres (en gros, le produit des deux).

- Quand, au contraire, la dernière bande est partiellement remplie, il existe des niveaux d'énergie vides, juste au-dessus du dernier niveau occupé⁵² ; l'application d'un champ électrique peut alors devenir efficace et faire grimper les électrons de plus haute énergie vers ces états très proches et initialement vides, donnant lieu au passage d'un courant : le solide est un métal.

Ainsi, les propriétés de transport électrique dépendent crucialement du remplissage de la dernière bande requise pour placer tous les électrons disponibles – un concept qui d'ailleurs n'a de sens que dans la mesure

⁴⁹Le raisonnement est fait à température nulle. À température finie, la fonction de Fermi - Dirac est une marche arrondie mais, compte tenu de la forte dégénérescence du gaz d'électrons pour les métaux usuels, les résultats qualitatifs énoncés ci-après demeurent inchangés (à l'ambiante, $k_B T \simeq 25 \text{ meV}$, alors que l'énergie de Fermi ε_F est usuellement de l'ordre de 10^4 K). D'une façon générale, les fluctuations thermiques concernent essentiellement les électrons situés au voisinage de ε_F .

⁵⁰Les éventuelles bandes profondes complètement remplies sont appelées bandes de valence et ne jouent aucun rôle pour le transport électrique. La première bande incomplète est appelée bande de conduction.

⁵¹En pratique, l'affirmation doit évidemment être nuancée, notamment parce que le solide peut contenir des impuretés faisant apparaître des états dans les *gaps*. C'est d'ailleurs ce que l'on fait délibérément pour constituer un semi-conducteur extrinsèque.

⁵²La quantification de \vec{k} résultant des conditions cycliques n'est pas en jeu ; elle est, elle, arbitrairement petite.

où le principe de Pauli joue un rôle essentiel : en définitive, c'est le caractère fermionique des porteurs (les électrons) qui permet de comprendre pourquoi certains solides sont conducteurs et d'autres sont isolants⁵³. À vrai dire, l'explication de la conduction électrique repose sur des considérations fort éloignées de ce qui se passe dans l'espace réel, contrairement à ce que l'intuition naturelle aurait permis d'espérer.

Le dernier état occupé est caractérisé par un vecteur d'onde, k_F , appelé vecteur d'onde de Fermi ; pour $d = 3$ et pour des électrons libres, il est déterminé en module par :

$$2 \frac{\frac{4\pi}{3} k_F^3}{\left(\frac{2\pi}{L}\right)^3} = N \iff k_F = \left(3\pi^2 \frac{N}{L^3}\right)^{1/3} \equiv (3\pi^2 \rho)^{1/3}, \quad (8.122)$$

où $\rho = \frac{N}{V}$ est la densité spatiale moyenne d'électrons. Le facteur 2 au premier membre prend en compte les deux possibilités pour le spin, une orbitale étant donnée. L'énergie correspondante, ε_F , est appelée énergie de Fermi. La mer de Fermi désigne le réservoir d'électrons d'énergie $\varepsilon \leq \varepsilon_F$; elle est délimitée par sa surface (surface de Fermi), définie géométriquement dans l'espace du vecteur d'onde \vec{k} par la relation $\varepsilon(\vec{k}) = \varepsilon_F$. La forme de la surface de Fermi reflète la structure du réseau réciproque, elle-même liée à la nature du réseau de Bravais et au nombre d'électrons disponibles ; si le potentiel périodique est faible, la surface de Fermi est proche de la sphère libre ; il se peut toutefois que des petites variations topologiques aient des conséquences physiques spectaculaires (en un sens, la surface de Fermi peut non seulement se déformer, mais également développer des protubérances qui, finalement, la "trouent").

Remarques

1. En toute généralité, le niveau de Fermi se définit comme la limite du potentiel chimique à température nulle :

$$\varepsilon_F = \lim_{T \rightarrow 0} \mu(T, V, N). \quad (8.123)$$

Pour un métal, ε_F ainsi défini est effectivement l'énergie du plus haut état occupé. Dans le cas d'un semi-conducteur ou d'un isolant, la limite $\mu(T = 0, V, N)$ est située *dans* le *gap*, entre le sommet de la bande de valence et le bas de la bande de conduction, à un certain endroit qui dépend de leur symétrie, c'est-à-dire de leur dégénérescence. Si les deux bandes ont même dégénérescence, $\varepsilon_F = \mu(T = 0, V, N)$ est en milieu de *gap*.

2. Il est souvent commode de considérer l'état fondamental comme un état "vide" (d'excitations). En ce sens, la mer de Fermi est la cousine de la mer de Dirac – que l'on a supposée complètement remplie pour éviter les ennuis avec les états d'énergie négative. À partir de cet état vide, l'excitation de basse énergie (par des photons, par exemple) crée des paires électron-trou, analogues (sur une autre échelle d'énergie !) des paires électron-positron de la théorie de Dirac. Pour un isolant ou un semi-conducteur, ces transitions bande à bande se produisent pour une fréquence du champ électromagnétique appliqué supérieure ou égale à $\hbar^{-1} \Delta E_g$, appelé seuil de photoconductivité⁵⁴. La mesure de cette fréquence-seuil est un moyen courant pour obtenir la valeur numérique du *gap*.

8.5 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

⁵³On apprendra par la suite que cette explication élémentaire doit être raffinée dans certains cas (semi-métaux).

⁵⁴Il s'agit bien de *photoconductivité* : l'exposition à la lumière – de fréquence suffisante, encore une preuve expérimentale du photon ! – porte des électrons dans la bande de conduction initialement vide : l'isolant, ou le semi-conducteur, devient un vrai conducteur, la conductivité augmentant avec le flux lumineux, toutes choses égales par ailleurs.

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 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 :

$$|n_1, n_2, \dots, n_k, \dots\rangle \quad (8.124)$$

et on introduit des opérateurs de création a_k^\dagger et d'annihilation a_k tels que :

$$a_k^\dagger |n_1, n_2, \dots, n_k, \dots\rangle \propto |n_1, n_2, \dots, n_{k-1}, n_k + 1, n_{k+1}, \dots\rangle, \quad (8.125)$$

$$a_k |n_1, n_2, \dots, n_k, \dots\rangle \propto |n_1, n_2, \dots, n_{k-1}, n_k - 1, n_{k+1}, \dots\rangle. \quad (8.126)$$

Ce sont ces opérateurs qui permettent de passer d'un secteur à l'autre de l'espace de Fock. Chaque opérateur a_k et a_k^\dagger ne modifie que le nombre d'occupation n_k , et le fait varier exactement d'une unité, dans un sens ou dans l'autre – sans doute la traduction technique la plus simple de la notion de *quantum*.

Ceci étant, il est possible de distinguer notamment :

- l'état vide :

$$|\text{vide}\rangle \equiv |n_1 = 0, n_2 = 0, \dots, n_k = 0, \dots\rangle. \quad (8.127)$$

Pour l'état vide, on a :

$$a_k |\text{vide}\rangle = 0. \quad (8.128)$$

Cette relation exprime le fait que l'on ne peut détruire une particule qui n'existe pas.

- les états à une particule :

$$|\Psi_k\rangle \equiv |n_1 = 0, n_2 = 0, \dots, n_{k-1} = 0, n_k = 1, n_{k+1} = 0, \dots\rangle, \quad (8.129)$$

que l'on peut construire en faisant agir une seule fois un seul opérateur de création sur l'état vide, conformément à (8.125) :

$$|\Psi_k\rangle = a_k^\dagger |\text{vide}\rangle. \quad (8.130)$$

Par ailleurs, toujours grâce à l'argument suivant lequel il est impossible de détruire une particule qui n'existe pas, on a :

$$a_k |0, 0, \dots, n_{k'}, 0, \dots\rangle = 0 \quad \forall k \neq k', \quad (8.131)$$

- les états à deux particules, $|n_k = 1, n_l = 1\rangle$. Ils sont engendrés à partir du vide par action d'un produit tel que $a_k^\dagger a_l^\dagger$:

$$|\Psi_{kl}\rangle = a_k^\dagger a_l^\dagger |\text{vide}\rangle. \quad (8.132)$$

Si les particules sont des bosons, l'état doit être symétrique ; l'égalité suivante doit donc être vraie :

$$a_k^\dagger a_l^\dagger |\text{vide}\rangle = a_l^\dagger a_k^\dagger |\text{vide}\rangle \quad \forall k, l. \quad (8.133)$$

Elle doit l'être d'ailleurs quel que soit l'état sur lequel agit le produit $a_k^\dagger a_l^\dagger$, puisque qu'un état de bosons, quel qu'il soit, doit être symétrique dans toute transposition. On en déduit, pour des bosons :

$$a_k^\dagger a_l^\dagger = a_l^\dagger a_k^\dagger \quad \forall k, l \iff [a_k^\dagger, a_l^\dagger] = 0 \quad (\text{bosons}) . \quad (8.134)$$

La relation hermitique conjuguée est aussi vraie :

$$a_k a_l = a_l a_k \quad \forall k, l \iff [a_k, a_l] = 0 \quad (\text{bosons}) . \quad (8.135)$$

Au contraire, s'il s'agit de fermions, l'état doit être antisymétrique dans toute transposition ; il faut que :

$$a_k^\dagger a_l^\dagger |\text{vide}\rangle = -a_l^\dagger a_k^\dagger |\text{vide}\rangle \quad \forall k, l \quad (8.136)$$

et, plus généralement⁵⁵ :

$$a_k^\dagger a_l^\dagger = -a_l^\dagger a_k^\dagger \quad \forall k, l \iff \{a_k^\dagger, a_l^\dagger\} = 0 \quad (\text{fermions}) , \quad (8.137)$$

qui entraîne à son tour :

$$a_k a_l = -a_l a_k \quad \forall k, l \iff \{a_k, a_l\} = 0 \quad (\text{fermions}) . \quad (8.138)$$

Ainsi, le produit d'un nombre quelconque d'opérateurs de création agissant sur le vide permet de construire les états à nombre quelconque de particules. Tous ces états constituent l'espace de Fock.

Il convient de déterminer les constantes de proportionnalité dans (8.125) et (8.126), relations que l'on peut écrire plus précisément (compte tenu du fait que a_k et a_k^\dagger sont hermitiques conjugués l'un de l'autre) :

$$a_k^\dagger |\dots, n_k, \dots\rangle = C_{n_k+1} |\dots, n_k+1, \dots\rangle , \quad (8.139)$$

$$a_k |\dots, n_k, \dots\rangle = C_{n_k}^* |\dots, n_k-1, \dots\rangle . \quad (8.140)$$

Tous les nombres $n_{k' \neq k}$ sont sous-entendus dans les *kets*, 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 :

$$a_k^\dagger a_k |\dots, n_k, \dots\rangle = C_{n_k}^* a_k^\dagger |\dots, n_k-1, \dots\rangle = |C_{n_k}|^2 |\dots, n_k, \dots\rangle . \quad (8.141)$$

Ainsi, les états de Fock sont propres de $N_k \equiv a_k^\dagger a_k$. L'opérateur N_k généralise pour le mode k l'opérateur $a^\dagger a$ introduit à propos de la description de Dirac d'un oscillateur harmonique à une dimension ; pour la même raison que dans ce cadre élémentaire, N_k est appelé *opérateur 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 (*quanta*) effectivement placées dans l'état k :

$$N_k = a_k^\dagger a_k, \quad N_k |\dots, n_k, \dots\rangle = n_k |\dots, n_k, \dots\rangle . \quad (8.142)$$

Les relations (8.139), (8.140) et (8.141) 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 de prendre cette dernière réelle positive :

$$C_{n_k} = \sqrt{n_k} , \quad (8.143)$$

d'où il résulte, pour les bosons ($n_k \in \mathbb{N}$) (voir (8.139) et (8.140)) :

$$a_k^\dagger |\dots, n_k, \dots\rangle = \sqrt{1+n_k} |\dots, n_k+1, \dots\rangle \quad (8.144)$$

$$a_k |\dots, n_k, \dots\rangle = \sqrt{n_k} |\dots, n_k-1, \dots\rangle . \quad (8.145)$$

La combinaison de ces deux équations redonne bien (8.141), compte tenu de (8.143) :

$$a_k^\dagger a_k |\dots, n_k, \dots\rangle = \sqrt{n_k} a_k^\dagger |\dots, n_k-1, \dots\rangle = \sqrt{n_k} \sqrt{1+(n_k-1)} |\dots, (n_k-1)+1, \dots\rangle . \quad (8.146)$$

⁵⁵ Les accolades désignent l'anticommutateur : $\{A, B\} \stackrel{\text{d\'ef}}{=} AB + BA$.

- Pour les fermions, il reste vrai que $|C_{n_k}|^2$ est une fonction linéaire de n_k , en accord avec (8.141), 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 ($n_k = 0, 1$) :

$$a_k^\dagger |\dots, n_k, \dots\rangle = \sqrt{1 - n_k} |\dots, n_k + 1, \dots\rangle \quad (8.147)$$

$$a_k |\dots, n_k, \dots\rangle = \sqrt{n_k} |\dots, n_k - 1, \dots\rangle . \quad (8.148)$$

Il est possible de rassembler ces équations sous la forme :

$$a_k^\dagger |\dots, n_k, \dots\rangle = \sqrt{1 + \varepsilon n_k} |\dots, n_k + 1, \dots\rangle \quad (8.149)$$

$$a_k |\dots, n_k, \dots\rangle = \sqrt{n_k} |\dots, n_k - 1, \dots\rangle \quad (8.150)$$

avec :

$$\varepsilon = \begin{cases} +1 & (\text{bosons}) \\ -1 & (\text{fermions}) \end{cases} . \quad (8.151)$$

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 :

$$[a_k, a_l^\dagger] = \delta_{kl} \quad (\text{bosons}) , \quad (8.152)$$

$$\{a_k, a_l^\dagger\} = \delta_{kl} \quad (\text{fermions}) , \quad (8.153)$$

soit :

$$a_k a_l^\dagger - \varepsilon a_l^\dagger a_k = \delta_{kl} . \quad (8.154)$$

Dans tous les cas, $N_k = a_k^\dagger a_k$ est l'opérateur nombre de particules dans le mode k ; ses valeurs propres sont des entiers non-négatifs :

$$N_k = \begin{cases} 0, 1, 2, \dots & (\text{bosons}) \\ 0, 1 & (\text{fermions}) \end{cases} . \quad (8.155)$$

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 :

$$\{c_k, c_l^\dagger\} = \delta_{kl} , \quad \{c_k, c_l\} = \{c_k^\dagger, c_l^\dagger\} = 0 \quad \forall k, l \quad (8.156)$$

$$c_k |\dots, n_k, \dots\rangle = \sqrt{n_k} |\dots, n_k - 1, \dots\rangle , \quad c_k^\dagger |\dots, n_k, \dots\rangle = \sqrt{1 - n_k} |\dots, n_k + 1, \dots\rangle \quad (8.157)$$

- Bosons :

$$[b_k, b_l^\dagger] = \delta_{kl} , \quad [b_k, b_l] = [b_k^\dagger, b_l^\dagger] = 0 \quad \forall k, l \quad (8.158)$$

$$b_k |\dots, n_k, \dots\rangle = \sqrt{n_k} |n_k - 1\rangle , \quad b_k^\dagger |\dots, n_k, \dots\rangle = \sqrt{1 + n_k} |\dots, n_k + 1, \dots\rangle . \quad (8.159)$$

Le point essentiel à retenir est que la symétrie de permutation est *incluse* dans l'algèbre des opérateurs de création et d'annihilation, qui commutent pour les bosons, anticommulent pour les fermions. 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 toujours d'emblée 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 maintenant un état propre pour une particule, de Hamiltonien H_1 :

$$H_1 |\Psi_k\rangle = \varepsilon_k |\Psi_k\rangle \iff H_1 = \sum_k |\Psi_k\rangle \varepsilon_k \langle \Psi_k| . \quad (8.160)$$

On a, par définition :

$$|\Psi_k\rangle = a_k^\dagger |\text{vac}\rangle . \quad (8.161)$$

Considérons maintenant un ensemble de telles particules sans interaction, dont le Hamiltonien est la grandeur additive⁵⁶ $H = \sum_i H_1(i)$; 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 :

$$E = \sum_k n_k \varepsilon_k . \quad (8.162)$$

Le Hamiltonien peut donc s'écrire :

$$H = \sum_k \varepsilon_k a_k^\dagger a_k \equiv \sum_k \varepsilon_k N_k . \quad (8.163)$$

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 permettant d'écrire :

$$|\Psi_k\rangle = \sum_\kappa U_{\kappa k} |\Phi_\kappa\rangle , \quad U_{\kappa k} = \langle \Phi_\kappa | \Psi_k \rangle . \quad (8.164)$$

En considérant (8.161), on voit que la même transformation existe entre les opérateurs associés aux deux bases ; avec des notations évidentes :

$$a_k^\dagger = \sum_\kappa U_{\kappa k} A_\kappa^\dagger . \quad (8.165)$$

Ceci étant, H (8.163) se transforme comme suit :

$$H = \sum_{\kappa \lambda} \sum_k \varepsilon_k U_{\kappa k} U_{\lambda k}^* A_\kappa^\dagger A_\lambda ; \quad (8.166)$$

la somme sur k est :

$$\sum_k \varepsilon_k \langle \Phi_\kappa | \Psi_k \rangle \langle \Phi_\lambda | \Psi_k \rangle^* = \sum_k \langle \Phi_\kappa | \Psi_k \rangle \varepsilon_k \langle \Psi_k | \Phi_\lambda \rangle \equiv \langle \Phi_\kappa | H_1 | \Phi_\lambda \rangle , \quad (8.167)$$

de sorte que H s'écrit :

$$H = \sum_{\kappa \lambda} \langle \Phi_\kappa | H_1 | \Phi_\lambda \rangle A_\kappa^\dagger A_\lambda ; \quad (8.168)$$

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 :

$$\Omega = \sum_{kl} \langle \Psi_k | \omega | \Psi_l \rangle a_k^\dagger a_l \equiv \sum_{kl} \omega_{kl} a_k^\dagger a_l . \quad (8.169)$$

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 (8.163).

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 (8.130). On peut de ce fait écrire, tenant compte de (8.169) :

$$\langle \Psi_k | \Omega | \Psi_l \rangle = \langle \text{vac} | a_k \Omega a_l^\dagger | \text{vac} \rangle = \sum_{k' l'} \langle \text{vac} | a_k a_{k'}^\dagger a_{l'} a_l^\dagger | \text{vac} \rangle \omega_{k' l'} . \quad (8.170)$$

L'idée consiste à jouer avec les relations de commutation⁵⁷ afin de mettre les opérateurs d'annihilation à droite dans le produit, juste devant $|\text{vac}\rangle$: en vertu de (8.128), le terme correspondant sera nul. Le produit des quatre opérateurs dans (8.170) est aussi $a_k a_{k'}^\dagger (\delta_{ll'} + \varepsilon a_l^\dagger a_{l'})$ (voir (8.154)) ; comme $a_{l'} |\text{vac}\rangle = 0$, il reste :

$$\langle \Psi_k | \Omega | \Psi_l \rangle = \sum_{k' l'} \langle \text{vac} | a_k a_{k'}^\dagger \delta_{ll'} | \text{vac} \rangle \omega_{k' l'} ; \quad (8.171)$$

⁵⁶Le Hamiltonien H_0 dans (8.85) est de cette forme, avec $N = 2$.

⁵⁷Le jeu qui consiste à faire commuter les opérateurs d'un produit pour réduire de deux unités le nombre de facteurs s'appelle une *contraction*.

en recommençant la même opération avec les opérateurs restants, il vient :

$$\langle \Psi_k | \Omega | \Psi_l \rangle = \sum_{k'l'} \langle \text{vac} | (\delta_{kk'} + \varepsilon a_{k'}^\dagger a_k) | \text{vac} \rangle \delta_{ll'} \omega_{k'l'} = \sum_{k'l'} \delta_{kk'} \delta_{ll'} \omega_{k'l'} = \omega_{kl} ; \quad (8.172)$$

Quant au premier membre, c'est très précisément :

$$\langle 0, 0, \dots, 0, 1_k, 0, \dots | \sum_i \omega(i) | 0, 0, \dots, 0, 1_l, 0, \dots \rangle ; \quad (8.173)$$

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 (8.172).

Considérons maintenant une observable à deux corps :

$$V = \frac{1}{2} \sum_{i \neq j} v(i, j) . \quad (8.174)$$

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 (8.169), on écrit⁵⁸ :

$$V = \frac{1}{2} \sum_{k, l, k', l'} v_{k k' l l'} a_{k'}^\dagger a_k^\dagger a_l a_{l'} \quad (8.175)$$

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 :

$$v_{k k' l l'} = \sum_{\text{spins}} \int d^3 r d^3 r' \Psi_k^*(\vec{r}) \Psi_{k'}^*(\vec{r}') v(\vec{r} - \vec{r}') \Psi_l(\vec{r}) \Psi_{l'}(\vec{r}') \equiv \langle \Psi_k \Psi_{k'} | v | \Psi_l \Psi_{l'} \rangle \quad (8.176)$$

Au total, un Hamiltonien ne contenant que des termes à un et à deux corps :

$$H = \sum_{i=1}^N H_1(i) + \frac{1}{2} \sum_{i=1}^N \sum_{j=1, j \neq i}^N H_2(i, j) \quad (8.177)$$

s'écrit comme suit :

$$H = \sum_{k l} \varepsilon_{kl} a_k^\dagger a_l + \frac{1}{2} \sum_{k, l, k', l'} v_{k k' l l'} a_{k'}^\dagger a_k^\dagger a_l a_{l'} , \quad (8.178)$$

où :

$$\varepsilon_{kl} = \langle \Psi_k | H_1 | \Psi_l \rangle , \quad v_{k k' l l'} = \langle \Psi_k \Psi_{k'} | H_2 | \Psi_l \Psi_{l'} \rangle . \quad (8.179)$$

Le point remarquable est que l'expression (8.178) est indépendante du nombre (fixe ou variable) de particules, N .

⁵⁸ Attention à l'ordre des indices !

Chapitre 9

Introduction à la théorie de la diffusion

Il s'agit de présenter les bases de la diffusion élastique de particules par un potentiel.

9.1 Présentation

On connaît l'importance des expériences de diffusion¹ de particules pour l'étude de la matière : c'est en faisant interagir une "pluie" de particules (faisceau incident) sur une cible et en analysant les caractéristiques du faisceau diffusé que l'on peut espérer obtenir des informations sur les interactions fondamentales, sur la dynamique interne de la cible, sur sa structure, *etc.* Le faisceau peut être de la lumière (photons) de fréquence *a priori* quelconque, mais évidemment adaptée au but poursuivi (rayons γ , rayons X, UV, IR, *etc.*), soit des particules massives (neutrons, électrons, atomes). La figure 9.1 donne le schéma de principe d'une expérience de collision.

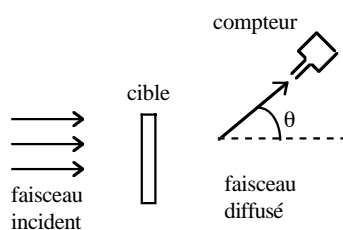


Figure 9.1: Schéma d'une expérience de diffusion.

Les phénomènes de diffusion sont en général fort complexes ; notamment, les particules du faisceau peuvent être composites et se fragmenter sous l'effet de la collision – tout comme les constituants de la cible, d'ailleurs. Afin de s'en tenir à l'essentiel et aux idées premières, on adopte pour la suite les hypothèses simplificatrices suivantes :

1. On ignore le spin des particules. Ceci n'est pas toujours possible : la diffusion de neutrons, outre une interaction de contact à l'échelle atomique², procède également d'une interaction entre le spin du neu-

¹On parlera indifféremment de théorie de la diffusion ou de théorie des collisions.

²Cette interaction a une portée de l'ordre du rayon nucléaire et peut de ce fait être considérée comme ponctuelle à l'échelle de l'Å.

tron et les spins nucléaires ou électroniques. Dans ce dernier cas, la diffusion correspondante donne des informations précieuses sur les fonctions d'onde des électrons (extension, anisotropie, *etc.*).

2. Toutes les particules en jeu (projectiles et constituants de la cible) sont supposées sans structure interne : seules les collisions *élastiques* seront considérées, où notamment l'énergie cinétique des projectiles est la même avant et après la collision. Physiquement, ceci signifie que les éventuels degrés de liberté internes ne peuvent être excités qu'avec des énergies très supérieures à celles mises en jeu lors de la collision.
3. Les diffuseurs de la cible sont supposés infiniment massifs (et sans structure, comme déjà admis). Ceci permet de remplacer chacun d'entre eux par le champ qu'il crée, donnant un potentiel³ $V(\vec{r})$. En outre, la cible est supposée assez mince pour que chaque projectile ne subisse qu'une collision
4. On néglige toute corrélation entre les particules diffusées, ce qui exige notamment que le faisceau soit de faible densité⁴. Cette hypothèse est presque une nécessité méthodologique : on veut observer la *signature* de la cible, ce qui exige que le faisceau de projectiles constitue une *petite* perturbation.

Ces hypothèses étant admises, la première étape consiste à trouver les états quantiques d'un projectile dans le champ d'un diffuseur ; il faut finalement résoudre l'équation aux valeurs propres pour un certain Hamiltonien H :

$$H = \frac{\vec{p}^2}{2m} + V(\vec{r}) . \quad (9.1)$$

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 :

$$dn = j_i \sigma_d(\theta, \phi) d\Omega . \quad (9.2)$$

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

$$\sigma = \int d\Omega \sigma_d(\theta, \phi) = \int_0^{2\pi} d\phi \int_0^\pi \sin \theta d\theta \sigma_d(\theta, \phi) . \quad (9.3)$$

9.2 Nature des états de diffusion

Les états stationnaires de diffusion se construisent à partir des états propres ψ du Hamiltonien (9.1) :

$$\left[\frac{\vec{p}^2}{2m} + V(\vec{r}) \right] \psi(\vec{r}) = E\psi(\vec{r}) . \quad (9.4)$$

E est l'énergie d'un projectile (la même avant et après la diffusion), m est sa masse. V est l'énergie potentielle du projectile dans le champ créé par un diffuseur⁵. En posant :

$$E = \frac{\hbar^2 k^2}{2m} \quad (k \in \mathbb{R}) , \quad V = \frac{\hbar^2}{2m} U , \quad (9.5)$$

on obtient :

$$[\Delta + k^2 - U(\vec{r})] \psi(\vec{r}) = 0 . \quad (9.6)$$

On sait résoudre exactement cette équation pour quelques potentiels remarquables, les fonctions propres diffèrent 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

³plus précisément, une énergie potentielle.

⁴La faible densité (schématiquement : le non-recouvrement spatial des fonctions d'onde des projectiles) autorise également à délaissier toute complication liée à l'indiscernabilité des projectiles.

⁵ $V(+\infty) = 0$ comme d'habitude, de sorte que $E > 0$ et $k \in \mathbb{R}$.

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. Dans ce qui suit, on raisonne avec l'axe Oz parallèle au faisceau incident et passant par le centre diffuseur (voir fig. 9.2)

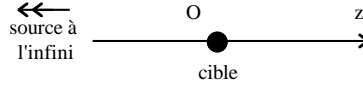


Figure 9.2: Géométrie utilisée dans le texte. La source est située à l'infini en $z = -\infty$.

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 à symétrie sphérique du genre $\frac{1}{r} e^{ikr}$, représentant l'onde effectivement diffusée ; de fait, on vérifie facilement que cette fonction satisfait⁶ l'équation dans \mathbb{R}^3 :

$$(\Delta + k^2) \frac{1}{r} e^{ikr} = 0 \quad (r \rightarrow +\infty), \quad (9.7)$$

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. Très précisément, on pose (en délaissant la question de la normalisation), dans la géométrie précisée :

$$\psi_{\vec{k}}(\vec{r}) \simeq C \left(e^{ikz} + f_k(\theta, \phi) \frac{e^{ikr}}{r} \right) \quad (r \text{ grand}), \quad (9.8)$$

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 située en $z = -\infty$ (il est de toute façon essentiel de se souvenir que l'expression de $\psi_{\vec{k}}$ (9.8) prend explicitement en compte la position de la source dans l'espace). Finalement, l'expression (9.8) est une combinaison linéaire de deux solutions (de même énergie) : c'est bien la solution (asymptotique) de l'équation aux valeurs propres (9.4), compte tenu de la position de la source (laquelle constitue *ipso facto* une condition aux limites).

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 ($1/r = 0$) et le courant associé⁷ à e^{ikz} est $\hbar k/m$. 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 :

$$j_i = |C|^2 \frac{\hbar k}{m}. \quad (9.10)$$

⁶Le facteur $\frac{1}{r}$ provient des trois dimensions de l'espace. Dans \mathbb{R}^d , l'onde sphérique solution de (9.7) est $r^{-(d-1)/2} e^{ikr}$ (voir l'expression de $\Delta \equiv -\hbar^{-2} p_r^2$ dans le ch. 4). L'exposant $(d-1)/2$ assure d'ailleurs que le flux du courant radial à travers une sphère de rayon R est indépendant de R ; ceci ne fait que traduire la conservation du nombre de projectiles quelle que soit la distance où on les compte tous, en particulier à l'infini.

⁷Le courant est généralement donné par :

$$\vec{j} = \frac{\hbar}{2im} (\psi^* \vec{\nabla} \psi - \psi \vec{\nabla} \psi^*) = \frac{\hbar}{m} \Im(\psi^* \vec{\nabla} \psi) \equiv \frac{\hbar}{m} \Im(\psi^* \vec{\nabla} \psi). \quad (9.9)$$

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 :

$$\vec{\nabla} = \left(\frac{\partial}{\partial r}, \frac{1}{r} \frac{\partial}{\partial \theta}, \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \right). \quad (9.11)$$

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

$$\vec{j}_d = |C|^2 \left(\frac{\hbar k}{m r^2} |f_k|^2, \frac{\hbar}{m r^3} \Im \left(f_k^* \frac{\partial f_k}{\partial \theta} \right), \frac{\hbar}{m r^3 \sin \theta} \Im \left(f_k^* \frac{\partial f_k}{\partial \phi} \right) \right). \quad (9.12)$$

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 :

$$\vec{j}_d \simeq \vec{j}_{\text{radial}} = |C|^2 \frac{\hbar k}{m r^2} |f_k|^2 \frac{\vec{r}}{r}. \quad (9.13)$$

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 détecteur⁸. Il en résulte, compte tenu de (9.10) :

$$dn = j_d r^2 d\Omega = |C|^2 \frac{\hbar k}{m r^2} |f_k|^2 r^2 d\Omega \equiv j_i |f_k|^2 d\Omega, \quad (9.14)$$

d'où, en se référant à (9.2), la relation simple entre amplitude de diffusion et section efficace différentielle :

$$\sigma_d(\theta, \phi) = |f_k(\theta, \phi)|^2. \quad (9.15)$$

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 bel et bien de trouver f_k . Ceci permettra de comparer $|f_k|^2$ à la section efficace mesurée et de valider un modèle ou un autre en vertu de (9.15).

9.3 Méthode des déphasages

Cette méthode est construite par comparaison systématique entre la diffusion par un potentiel $V(r)$, supposé à symétrie sphérique⁹, et la résolution du problème de la particule libre dans \mathbb{R}^3 en coordonnées sphériques ; comme on l'a noté au chapitre 4, il s'agit d'un cas particulier de champ central. Alors, on a notamment rencontré une relation importante :

$$e^{i k z} = \sum_{l=0}^{+\infty} (2l+1) i^l j_l(kr) P_l(\cos \theta), \quad (9.16)$$

et comme :

$$P_l(\cos \theta) = \sqrt{\frac{4\pi}{2l+1}} Y_{l0}(\theta), \quad (9.17)$$

on a aussi :

$$e^{i k z} = \sum_{l=0}^{+\infty} i^l \sqrt{4\pi(2l+1)} j_l(kr) Y_{l0}(\theta). \quad (9.18)$$

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

$$\psi_{k l m}^{(0)}(\vec{r}) = C_k j_l(kr) Y_{l m}(\theta, \phi), \quad (9.19)$$

⁸puisque $d\Omega$ est par définition la surface de sphère de rayon unité délimitée par le petit cône d'axe (θ, ϕ) .

⁹Dès lors, l'amplitude f_k , et donc aussi la section efficace, ne dépend pas de l'angle azimutal ϕ .

où C_k est une constante de normalisation¹⁰, j_l une fonction de Bessel sphérique et Y_{lm} une harmonique sphérique. k est le label quantique représentant l'énergie :

$$E = \frac{\hbar^2 k^2}{2m} . \quad (9.21)$$

Les j_l dépendent de E par l'intermédiaire du vecteur d'onde k

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 les développements asymptotiques de j_l :

$$j_l(kr) \simeq \frac{1}{kr} \sin\left(kr - l\frac{\pi}{2}\right) \quad (r \gg k^{-1}) , \quad (9.22)$$

ce qui permet, toujours dans le cas libre, de récrire (9.18) :

$$e^{ikz} \simeq \frac{1}{kr} \sum_{l=0}^{+\infty} i^l \sqrt{4\pi(2l+1)} \sin\left(kr - l\frac{\pi}{2}\right) Y_{l0}(\theta) , \quad (9.23)$$

pour une onde plane issue de $z = -\infty$ ($k > 0$). En définitive, l'expression (9.23) 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, (9.23) : 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 :

$$\psi_k(\vec{r}) \simeq C' \sum_{l=0}^{+\infty} i^l \sqrt{4\pi(2l+1)} \frac{1}{kr} \sin\left(kr - l\frac{\pi}{2} + \delta_l\right) Y_{l0}(\theta) . \quad (9.24)$$

Bien évidemment, il ne s'agit pour l'instant que d'un *ansatz* ; s'il est correct, on doit retrouver¹¹ le comportement asymptotique deviné pour ψ_k , (9.8). 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 (9.24) est correcte – ce que confirme un calcul explicite ([4], VIII-4-a-β). On a :

$$\begin{aligned} \sum_{l=0}^{+\infty} i^l \sqrt{4\pi(2l+1)} \frac{1}{kr} \sin\left(kr - l\frac{\pi}{2} + \delta_l\right) Y_{l0}(\theta) = \\ e^{-i\delta_l} \sum_{l=0}^{+\infty} i^l \sqrt{4\pi(2l+1)} \frac{e^{ikr} e^{-il\pi/2} e^{2i\delta_l} - e^{-ikr} e^{+il\pi/2}}{2ikr} Y_{l0}(\theta) . \end{aligned} \quad (9.25)$$

En utilisant :

$$e^{2i\delta_l} = 1 + 2i e^{i\delta_l} \sin \delta_l , \quad (9.26)$$

le terme au numérateur de la grande fraction dans (9.25) est :

$$e^{ikr} e^{-il\pi/2} (1 + 2i e^{i\delta_l} \sin \delta_l) - e^{-ikr} e^{+il\pi/2} = 2i e^{i\delta_l} \sin \delta_l + e^{i(kr-il\pi/2)} - e^{-i(kr-il\pi/2)} , \quad (9.27)$$

d'où :

$$\psi_k(\vec{r}) \simeq C' e^{-i\delta_l} \sum_{l=0}^{+\infty} i^l \sqrt{4\pi(2l+1)} \left[\frac{e^{ikr} e^{-il\pi/2} e^{i\delta_l} \sin \delta_l}{kr} + \frac{e^{i(kr-l\pi/2)} - e^{-i(kr-l\pi/2)}}{2ikr} \right] Y_{l0}(\theta) . \quad (9.28)$$

¹⁰ Avec $C_k = \sqrt{2k^2/\pi}$, on a :

$$\langle \psi_{klm}^{(0)} | \psi_{k'l'm'}^{(0)} \rangle = \delta(k-k') \delta_{ll'} \delta_{mm'} . \quad (9.20)$$

¹¹ Dans tous les cas, on a bien une onde plane issue de $z = -\infty$.

Le deuxième terme du crochet donne très exactement, d'après (9.23), l'expression asymptotique de l'onde plane e^{ikz} ; d'où :

$$\psi_k(\vec{r}) \simeq C' e^{-i\delta_l} \left[\sum_{l=0}^{+\infty} i^l \sqrt{4\pi(2l+1)} e^{-il\pi/2} \frac{e^{ikr}}{kr} e^{i\delta_l} \sin \delta_l + e^{ikz} \right] Y_{l0}(\theta) . \quad (9.29)$$

Ceci est bien du type (9.8) (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 déphasages¹² – donc aussi, par (9.15), la section efficace en fonction des δ_l :

$$f_k(\theta) = \frac{1}{k} \sum_{l=0}^{+\infty} \sqrt{4\pi(2l+1)} e^{i\delta_l} \sin \delta_l Y_{l0}(\theta) . \quad (9.30)$$

D'après (9.15) et par intégration sur les angles selon (9.3), on en déduit la section efficace totale σ ; compte tenu de l'orthonormalisation des Y_{lm} , il vient :

$$\sigma = \int d\Omega |f_k(\theta)|^2 = \frac{4\pi}{k^2} \sum_{l=0}^{+\infty} (2l+1) \sin^2 \delta_l(k) . \quad (9.31)$$

En pratique, un tel développement n'est vraiment utile que pour les potentiels donnant un petit nombre de déphasages sensiblement différents de zéro¹³. Les δ_l peuvent en principe se trouver par identification, si le potentiel $V(r)$ est connu, en résolvant l'équation aux valeurs propres correspondante pour chaque valeur de l et en écrivant l'expression asymptotique des fonctions propres.

Il se trouve que le nombre de déphasages pertinents peut être semi-quantitativement 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 pour établir ce fait de considérer le comportement de ces fonctions près de l'origine. Sachant que :

$$j_l(kr) \simeq \frac{(kr)^l}{(2l+1)!!} \quad (kr \ll 1) , \quad (9.32)$$

on voit que 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 seul déphasage sensiblement non-nul, δ_0 . Plus généralement, les δ_l significatifs auront un indice l satisfaisant :

$$l \lesssim k\xi . \quad (9.33)$$

Sans surprise, le nombre de déphasages importants 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 :

$$V(r) = \begin{cases} 0 & \text{si } r > \xi \\ +\infty & \text{si } r < \xi \end{cases} . \quad (9.34)$$

D'après ce qui précède, seuls les déphasages satisfaisant (9.33) 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 (9.30) :

$$f_k(\theta) = \frac{\sqrt{4\pi}}{k} e^{i\delta_0} \sin \delta_0 Y_{00} = \frac{1}{k} e^{i\delta_0(k)} \sin \delta_0(k) . \quad (9.35)$$

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 :

$$\sigma = \frac{4\pi}{k^2} \sin^2 \delta_0(k) . \quad (9.36)$$

¹²Il n'y a pas ici de dépendance par rapport à l'angle azimutal ϕ (la cible est ponctuelle et il y a visiblement une symétrie de rotation autour de l'axe Oz).

¹³sauf bien sûr si, par bonne fortune, on sait resommer la série.

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 :

$$u'' + k^2 u = 0 \quad (\forall r > \xi) , \quad (9.37)$$

puisque $u(r)$ satisfait une équation aux valeurs propres à une dimension (réduite à \mathbb{R}_+) avec le potentiel effectif $V_{\text{eff}} = V(r) + \hbar^2 l(l+1)/(2mr^2) = 0$ si $l = 0$ et $r > \xi$. En outre, par la continuité de la fonction d'onde – identiquement nulle pour $r < 0$ –, on doit avoir :

$$u(r = \xi) = 0 . \quad (9.38)$$

La solution est donc :

$$u(r) = \begin{cases} A \sin[k(r - \xi)] & \text{si } r > \xi \\ 0 & \text{si } r < \xi \end{cases} . \quad (9.39)$$

Par ailleurs, le déphasage est identifié en écrivant le comportement asymptotique de la fonction radiale $R(r)$ (voir (9.24)) :

$$R(r) \simeq \sqrt{4\pi} \frac{C'}{kr} \sin(kr + \delta_0) \iff u(r) \simeq C'' \sin(kr + \delta_0) \quad (r \rightarrow +\infty) . \quad (9.40)$$

Par comparaison avec (9.39), il vient $\delta_0 = -k\xi$ et donc, suivant (9.31) :

$$\sigma = \frac{4\pi}{k^2} \sin^2 k\xi . \quad (9.41)$$

Comme on a supposé $k\xi \ll 1$, le résultat significatif est en réalité :

$$\sigma \simeq 4\pi \xi^2 , \quad (9.42)$$

obtenu après développement du sinus. Ce résultat, valide notamment à très basse énergie, montre que dans une telle limite la section efficace est indépendante de l'énergie. L'expression (9.42) n'est en fait rien d'autre que la section efficace totale classique ($\hbar = 0$) d'une sphère dure de rayon ξ pour un projectile ponctuel... 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 longueur¹⁴, 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, elle a bien le temps de sonder le potentiel. Ces résonances existent aussi dans \mathbb{R}^3 et sont visibles dès l'onde p ($l = 1$).

À l'inverse, à haute énergie, les longueurs d'onde deviennent très petites et les déphasages importants sont en grand nombre ; dans ce cas, on trouve ([4], C_{VIII}, eq. (22)) :

$$\sigma \simeq 2\pi \xi^2 . \quad (9.43)$$

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 encore lié à l'existence d'une discontinuité du potentiel. Dans tous les cas, on a :

$$\sigma > \sigma_{\hbar=0} , \quad (9.44)$$

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 classiquement¹⁵. Il en résulte une *augmentation* de la section efficace de diffusion.

¹⁴comme par exemple la longueur d'onde de de Broglie $\hbar/\sqrt{2mE} = k^{-1}$.

¹⁵faute de quoi, leur position ne saurait être parfaitement définie.

Bibliographie

- [1] E. Schrödinger, *Mémoires sur la Mécanique Ondulatoire* (J. Gabay, Paris, 1988)
- [2] H. Bacry, *Leçons sur la théorie des groupes et les symétries des particules élémentaires* (Gordon & Breach, distribué par Dunod, Paris, 1967)
- [3] D. Bohm, *Quantum Theory* (Prentice Hall, Englewood Cliffs, 1951)
- [4] Cl. Cohen-Tannoudji, B. Diu et F. Laloë, *Mécanique Quantique* (Hermann, Paris, 1973)
- [5] Jean Dieudonné, *Abrégé d'histoire des mathématiques* Hermann (1986)
- [6] <http://aleph0.clarku.edu/~djoyce/hilbert/problems.html>

- [7] Richard P. Feynman *Lumière et Matière* (Collection "Points Sciences", InterEditions, Paris, 1987).
- [8] M. Jammer, *The Conceptual developments of Quantum Mechanics* (McGraw - Hill, New York, 1966)
- [9] E. Klein, *Conversations avec le Sphinx*, p. 176 (Albin Michel, 1991)
- [10] R. Penrose, *The Emperor's New Mind* (Oxford University Press, 1989)
- [11] A. Einstein, B. Podolsky et N. Rosen, Phys. Rev. **47**, 777, 1935.
- [12] A. Aspect, J. Dalibard et G. Roger, Phys. Rev. Letters **4**, 91, 1982.
- [13] A. Aspect, dans *Le Monde Quantique*, édité par S. Deligeorge (Points Sciences, Seuil, 1984).
- [14] L. E. Ballentine, *Quantum Mechanics* (Prentice - Hall, Englewood Cliffs, 1990)
- [15] L. Landau et E. Lifchitz, *Mécanique Quantique* (Editions Mir, Moscou, 1967)
- [16] A. Messiah, *Mécanique Quantique* (Dunod, Paris, 1965)
- [17] E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1970)
- [18] L. I. Schiff, *Quantum Mechanics* (McGraw - Hill, New York, 1968)
- [19] L. Landau et I. Lifchitz, *Mécanique* (Mir, Moscou, 1966)
- [20] A. Messsiah, *Mécanique Quantique*, tome II (Dunod, Paris 1965)
- [21] E. Merzbacher, Amer. J. Phys., **30**, 237 (1962)
- [22] R. Courant et D. Hilbert *Methods of Mathematical Physics*, vol. I, (Interscience Pub., New York, 1953)
- [23] M. Abramowitz et I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970)
- [24] S. Flügge, *Practical Quantum Mechanics* (Springer, Berlin, 1974)
- [25] R. Jackiw, in *M. A. Bég Memorial Volume*, M. C. Gutzwiller, A. Inomata, J. R. Klauder et L. Steit éditeurs (World Scientific, Singapour, 1991)

- [26] J. Fernando Perez & F. A. B. Coutinho, *Amer. J. Phys.*, **59**, 52, 1991
- [27] H. V. McIntosh, *Amer. J. Phys.*, **27**, 620, 1959
- [28] M. Bander & C. Itzykson, *Rev. Mod. Phys.*, **38**, 330, 1966
- [29] L. Landau & I. Lifchitz, *Mécanique Quantique* (Mir, Moscou, 1967)
- [30] Cl. Cohen-Tannoudji, *Cours du Collège de France, 1973-1974*
- [31] A. Sokal et J. Bricmont, *Impostures intellectuelles* (Odile Jacob, Paris, 1997)
- [32] *Encyclopædia Universalis*, **16**, 300.
- [33] L. Ballentine, *Quantum Mechanics* (Prentice Hall, Englewood Cliffs, 1990)
- [34] A. W. Overhauser, A.R. Collola et S. A. Werner, *Phys. Rev. Let.*, **33**, 1237 (1974)
- [35] A. S. Davydov, *Quantum Mechanics* (Pergamon Press, 1965)
- [36] E. Schrödinger, *Ann. Physik*, **79**, 361 (1926)
- [37] P. Pechukas et J. C. Light, *J. Chem. Phys.*, **44**, 3897 (1966)
- [38] Cl. Cohen-Tannoudji, *Cours du Collège de France (1973/1974)*
- [39] N. Goldenfeld, *Lectures on Phase Transitions and the Renormalization Group* (Addison - Wesley, New York, 1997)
- [40] H. A. Bethe et E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Springer, Berlin, 1957)
- [41] J. C. Slater, *Quantum Theory of Atomic Structure, I* (McGraw - Hill, New York, 1960)