

Physique Quantique (II)

Table des Matières

10 L'avènement de la Mécanique Quantique	7
10.1 Problèmes de l'Ancienne Théorie des Quanta	7
10.2 La Mécanique des Matrices	8
10.3 Les ondes de matière (de Broglie, 1923)	13
10.4 L'équation de Schrödinger	16
10.5 Vitesse de groupe	21
10.6 Diffraction des particules matérielles	24
10.7 Limite classique	24
10.7.1 Longueur d'onde pour un objet macroscopique	25
10.7.2 Limite classique de la fonction d'onde	25
11 Fonction d'onde	29
11.1 Fentes d'Young	29
11.2 Interprétation probabiliste de la fonction d'onde et conséquences	34
11.2.1 L'interprétation de Born et Jordan	35
11.2.2 Calcul des valeurs moyennes	37
11.2.3 Le déterminisme quantique	41
11.3 Principe d'incertitude de Heisenberg	44
11.3.1 Principe d'incertitude spatial	44
11.3.2 Principe d'incertitude temporel	47
12 Magnétisme atomique	49
12.1 Magnétisme classique	49
12.1.1 Moment magnétique	49
12.1.2 Précession de Larmor	51

12.1.3	Paramagnétisme classique	52
12.1.4	Expériences de Einstein - de Haas et de Barnett	53
12.2	Expérience de Stern et Gerlach	55
13	Postulats et structure formelle de la Mécanique Quantique	61
13.1	Enoncé des postulats	61
13.1.1	Notion d'état	61
13.1.2	Notion d'observable	63
13.1.3	Résultats possibles de la mesure d'une grandeur physique	66
13.1.4	La réduction du paquet d'ondes	69
13.1.5	Evolution des systèmes dans le temps	71
13.2	Illustration des postulats : retour sur l'expérience de Stern et Gerlach	73
13.3	Les bases du formalisme de la Mécanique Quantique	73
13.3.1	Changement de base	73
13.3.2	Produit scalaire	75
13.3.3	Généralisations	79
14	Opérateurs	85
14.1	Propriété fondamentale des observables : hermiticité	85
14.2	Valeur moyenne d'une observable : utilisation de sa base propre	88
14.3	Représentation des opérateurs hermitiques et des opérateurs unitaires	90
14.4	Retour sur la notation de Dirac	92
14.5	Combinaisons d'opérateurs	93
14.6	Représentation- \vec{r}	95
14.7	Représentation- \vec{p}	100
15	Evolution temporelle d'un système quantique	103
15.1	Description de l'évolution dans le temps	103
15.1.1	La description de Schrödinger	105
15.1.2	La description de Heisenberg	109
15.1.3	Le théorème d'Ehrenfest	113
15.2	Propagateur	114
15.3	Séparation espace - temps et états stationnaires	117

16 Problèmes à une dimension	121
16.1 Propriétés générales des problèmes à une dimension	121
16.2 La quantification comme conséquence des conditions imposées à Ψ	127
16.2.1 $\varepsilon < 0 \iff E < 0$	127
16.2.2 $\varepsilon > 0 \iff E > 0$	128
16.3 Le puits carré	130
16.3.1 Le puits fini	131
16.3.2 Le puits infini	144
16.4 La marche de potentiel	147
16.5 La barrière de potentiel	149
17 L'oscillateur harmonique	153
17.1 L'importance de l'oscillateur harmonique	153
17.2 Résolution de l'équation aux valeurs propres	156
17.3 Quantification canonique, opérateurs de création et d'annihilation	165
17.4 Le propagateur de l'oscillateur harmonique	169

Chapitre 10

L'avènement de la Mécanique Quantique

L'élaboration précise de la Mécanique Quantique a été accomplie lors de deux étapes décisives constituées par les travaux de Heisenberg (Mécanique des Matrices, 1925) et de Schrödinger (Mécanique Ondulatoire, 1926). Ces deux formalismes, à première vue très différents, sont en fait équivalents et – comme l'a montré Schrödinger – ne sont que deux écritures distinctes d'une seule et même théorie, la Mécanique Quantique. Ces travaux sont presque contemporains de l'apport de de Broglie (1923), qui est d'une tout autre nature, au sens où aucune proposition opérationnelle n'accompagne l'idée fondamentale – associer une onde à toute particule matérielle –, n'autorisant de ce fait aucune prévision théorique permettant une validation immédiate. En fait, la justesse de la proposition de de Broglie fut démontrée par Davisson et Germer en 1927 (tout comme les rayons X, les électrons sont diffractés par un cristal).

Le but de ce chapitre est d'exposer de façon succincte les démarches fondamentales des pères fondateurs d'une théorie à la robustesse exemplaire et probablement unique : plus de trois quarts de siècle après sa construction, aucune expérience n'autorise à mettre en doute les bases d'une théorie pourtant déroutante, aux conclusions souvent “paradoxaux” quand elles sont examinées dans une optique classique, et en tout cas en violente opposition avec le sens commun.

10.1 Problèmes de l'Ancienne Théorie des Quanta

Avant d'en venir aux propositions de Heisenberg et de Schrödinger, il est utile de faire le point sur la nécessité d'une nouvelle théorie. L'Ancienne Théorie des Quanta, construite par un aller-retour entre formalisme théorique et expérience (par exemple, le modèle de Bohr pour rendre compte de la série de Balmer), a connu de grands succès, le plus spectaculaire étant sans doute l'explication de la structure fine de la raie H_α par Sommerfeld. Cependant, des difficultés subsistaient :

- l'aspect arbitraire des règles de quantification, consistant à passer au peigne fin les trajectoires classiques pour ne retenir que celles satisfaisant l'exigence de Bohr - Wilson - Sommerfeld, traduite par :

$$\oint p_j dq_j = n_j h . \quad (10.1)$$

Cette règle, sans aucune autre justification qu'empirique, possédait un caractère arbitraire eu égard à la définition ordinaire d'une théorie physique. Si le pragmatisme est de règle lorsqu'il s'agit de construire une nouvelle théorie, la forme finale de celle-ci ne doit reposer que sur des principes de portée universelle (le principe de causalité, le principe de relativité, les symétries fondamentales, *etc.*).

- la cohabitation forcée du classique et du néo-quantique donnait à la théorie un aspect un peu bâtard : notamment, rien n'était remis en cause concernant l'Électromagnétisme. Sa conséquence : une charge

accélérée rayonne – et donc l'instabilité de l'atome – devait être mise de côté dès qu'il s'agissait des cercles de Bohr ! L'affirmation de principe de Bohr (“quand l'électron est sur un tel cercle, il ne rayonne pas”) relève plutôt de la formule magique (le démon de Bohr n'existe pas).

- la nécessité de recourir peu à peu à des “recettes” de plus en plus nombreuses pour rendre compte de la richesse des spectres atomiques, révélée au fur et à mesure que les expériences se faisaient de plus en plus fines et précises. De ce point de vue, l'introduction à *la main* du nombre quantique de spin par Uhlenbeck et Goudsmit (1925) est exemplaire¹.

Au stade de son évolution où elle allait être délaissée, l'Ancienne Théorie des Quanta était finalement une collection volumineuse de règles – dont les rapports mutuels étaient loin d'être clairs – et dont certaines étaient d'ailleurs en stricte contradiction avec d'autres théories physiques². Vouée à disparaître, cette théorie a toutefois joué un rôle majeur dans le cheminement vers la bonne théorie³. De surcroît et avec le recul, il est possible de démontrer qu'elle est une théorie approximant la vraie Mécanique Quantique dans la limite dite quasi-classique, où les effets quantiques sont petits (mais restent singuliers). Examinée indépendamment, l'Ancienne Théorie des Quanta n'a, à proprement parler, aucune justification théorique ; au contraire, revisitée dans un cadre plus vaste, elle apparaît comme une approximation valide dans une certaine limite⁴.

10.2 La Mécanique des Matrices

En 1925, Heisenberg (à 24 ans !) publia un article [1] qui doit être considéré comme celui où furent formalisées pour la première fois les règles fondamentales de la nouvelle mécanique, et notamment la *non-commutativité* du produit des objets mathématiques représentant les grandeurs physiques. La démarche de Heisenberg, assez complexe, mérite indéniablement d'être discutée un peu dans ses détails, ne serait-ce que pour percevoir en quoi elle est un mélange extraordinaire de connaissances maîtrisées et d'audace intellectuelle.

Le point de départ est une forme de renoncement : Heisenberg décide de ne plus tenter de décrire directement la position et la vitesse d'une particule comme l'électron gravitant autour du proton dans l'atome d'hydrogène. Ce parti pris est inspiré par une impossibilité expérimentale qu'un physicien ne peut décider d'ignorer : aucune expérience ne permettait (et c'est toujours le cas !) d'*observer* la trajectoire de l'électron, reléguée *de facto* dans l'imaginaire du physicien et, de ce point de vue, étant plutôt d'ordre métaphysique. D'emblée, Heisenberg adopte le point de vue suivant lequel la description théorique de la dynamique atomique doit se fixer comme seul objectif d'expliquer ce qui est *observable*, c'est-à-dire en l'espèce les spectres atomiques.

L'une des premières observations de Heisenberg est de rappeler que, dans la limite des très grands nombres quantiques (au sens de l'Ancienne Théorie des Quanta), la fréquence d'une transition d'un état n à un état m voisin ($|n - m| \ll n, m$) coïncide à peu près avec la fréquence de rotation de l'électron sur son orbite ; ceci est l'expression fondamentale du Principe de Correspondance de Bohr : quand n est très grand, la quantité nh varie en fait, et relativement, par toutes petites quantités h , “infinitésimales” dans cette limite, reproduisant asymptotiquement un comportement quasi-continu érigé en principe implicite par la Physique Classique⁵.

D'un point de vue purement classique, toute grandeur dynamique relative à un mouvement périodique peut se décomposer en série de Fourier. Par exemple, pour une particule libre mais confinée entre deux murs distants de a , le mouvement est périodique ($x(t)$ est une fonction en dents de scie) de période $T = 2a/v$ où v est la vitesse (constante en module) de la particule rebondissant élastiquement sur les murs ; la pulsation correspondante est ω_E , c'est une fonction de l'énergie⁶ $E = mv^2/2$:

$$\omega_E = \frac{\pi}{a} \sqrt{\frac{2E}{m}} . \quad (10.2)$$

¹ainsi que sa conséquence empirique appelée “Principe de Pauli”.

²Sans aborder d'autres difficultés, comme celle de l'impossibilité de fournir une explication précise des spectres des atomes complexes.

³Notamment, Heisenberg s'y est constamment référé pour élaborer sa Mécanique des Matrices.

⁴Il n'est d'ailleurs pas rare de l'utiliser encore pour obtenir rapidement des estimations d'ordre de grandeur.

⁵Sur la notion de continu en Physique, voir les réflexions de Schrödinger dans *Physique Quantique et représentation du monde* [13].

⁶Sur ce point, l'oscillateur harmonique est très particulier : la pulsation du mouvement n'est autre que ω_0 , la pulsation propre de l'oscillateur, et est *indépendante* de la valeur de l'énergie (la période d'un pendule est indépendante de l'énergie pour les oscillations de petite amplitude). Toutes les pulsations $\omega(n, 1)$ définies ci-dessous en (10.4) sont égales entre elles dans le cas de l'oscillateur harmonique.

Il en résulte que l'abscisse $x(t)$ peut être décomposée comme suit :

$$x(t) = \sum_{k \in \mathbb{Z}} x_k e^{ik\omega_E t} , \quad (10.3)$$

où les x_k sont les composantes de Fourier (aussi appelées amplitudes) d'une fonction en dents de scie, aisément calculables et paramétrées par l'énergie E , dont elles dépendent *ipso facto*.

Si maintenant l'on procède, pour ce mouvement classique, à la quantification selon (10.1), E se trouve quantifiée : $E \in \{E_n\}$, de sorte que la composante de Fourier x_k est complètement définie par deux entiers, k et n , c'est en un sens une fonction de n et de k que l'on peut noter $x(n, k)$ (et l'abscisse peut être notée $x_n(t)$). De même, la pulsation fondamentale ω_{E_n} peut être plus simplement notée ω_n ; on peut aussi la désigner par $\omega(n, 1)$, toutes les harmoniques $\omega_k = k\omega_E$ se notant alors logiquement $\omega(n, k)$ et sont telles que :

$$\omega(n, k) = k\omega(n, 1) . \quad (10.4)$$

Ainsi, la quantification "ancienne" étant effectuée et ces nouvelles notations adoptées, le développement (10.3) devient :

$$x_n(t) = \sum_{k \in \mathbb{Z}} x(n, k) e^{ik\omega(n, k)t} . \quad (10.5)$$

L'argumentation précédente et ces définitions ne sont nullement tributaires de l'exemple choisi pour fixer les idées et ont une portée générale. Pour un mouvement périodique quelconque à un seul degré de liberté (pour simplifier) ayant subi la quantification de Bohr - Wilson - Sommerfeld, on peut écrire un développement du genre⁷ :

$$q_n(t) = \sum_{k \in \mathbb{Z}} q(n, k) e^{i\omega(n, k)t} . \quad (10.6)$$

Il est bien clair que la connaissance de $q_n(t)$ ou celle, conjointe, des $q(n, k)$ et des $\omega(n, k)$, sont strictement équivalentes : connaître une fonction ou connaître ses coefficients de Fourier (et les fréquences correspondantes) constituent deux *représentations* distinctes d'une même grandeur dont l'une permet de trouver l'autre et réciproquement. En définitive, disposer la loi horaire ou de sa série de Fourier revient strictement au même quoique, bien sûr, $q(t)$ donne une représentation plus imagée que son développement de Fourier. On va voir que Heisenberg a dû se livrer à une manipulation formelle d'apparence anodine sur les indices (opération appelée "bricolage" dans la suite), inoffensive pour les très grands nombres quantiques mais aux conséquences spectaculaires pour les petits nombres quantiques ; il n'est pas exagéré de dire que c'est très précisément à ce point que la notion même de trajectoire disparaît du formalisme quantique.

Par construction (voir (10.4)), les pulsations classiques du développement de Fourier satisfont :

$$\omega(n, k) = -\omega(n, -k) , \quad (10.7)$$

et :

$$\omega(n, k) + \omega(n, k') = \omega(n, k + k') ; \quad (10.8)$$

en outre, comme $q(t)$ est une grandeur physique, donc une fonction à valeurs réelles, on a :

$$q(n, k) = q^*(n, -k) . \quad (10.9)$$

Les considérations précédentes ne sont rien d'autre qu'un examen de l'analyse de Fourier classique à travers le filtre des relations de quantification. C'est à ce stade que Heisenberg effectue un premier pas important. Les fréquences de Bohr sont définies en soi : ce sont simplement (à \hbar près) les différences d'énergies ($\hbar\omega = \Delta E$) ; en termes plus précis, considérant un couple d'énergies E_n et E_{n-k} , on introduit la pulsation correspondante, notée $\omega_{n, n-k}$:

$$\omega_{n, n-k} = \frac{1}{\hbar} (E_n - E_{n-k}) . \quad (10.10)$$

⁷Revenant aussi à la notation traditionnelle de la Mécanique Analytique, q pour une coordonnée, p pour son moment conjugué.

C'est cette fréquence qui doit, dans la limite des grands nombres quantiques, coïncider⁸ avec la pulsation classique $\omega(n, k)$. Heisenberg postule que cette coïncidence vaut tout autant au niveau des *amplitudes*, ce qui conduit à la nécessité préalable de postuler l'existence de quantités $q_{n, n-k}$ soumise à la même condition aux limites. Dès lors, on doit avoir simultanément :

$$\omega_{n, n-k} \rightarrow \omega(n, k) , \quad q_{n, n-k} \rightarrow q(n, k) \quad (n \gg 1) . \quad (10.11)$$

D'un autre côté, par leur définition en terme de différence de deux énergies, (10.10), les fréquences de Bohr satisfont :

$$\omega_{n, n-k'} + \omega_{n-k', n-k} = \omega_{n, n-k} . \quad (10.12)$$

Au total, Heisenberg pose finalement que la quantité :

$$q_{n, n-k} e^{i\omega_{n, n-k} t} \equiv Q_{n, n-k}(t) \quad (10.13)$$

est la description quantique du terme classique :

$$q(n, k) e^{i\omega(n, k)t} \quad (10.14)$$

et que l'ensemble des objets $Q_{n, n-k}(t)$ constitue la représentation quantique de la coordonnée⁹.

Heisenberg examine ensuite la question de la représentation du *carré* de la coordonnée. A partir de (10.6), on trouve :

$$[q_n(t)]^2 = \sum_{k, k' \in \mathbb{Z}^2} q(n, k)q(n, k') e^{i[\omega(n, k) + \omega(n, k')]t} , \quad (10.15)$$

ce qui, compte tenu de (10.8), s'écrit :

$$[q_n(t)]^2 = \sum_{k, k'' \in \mathbb{Z}^2} q(n, k)q(n, k'' - k) e^{i\omega(n, k'')t} . \quad (10.16)$$

Moyennant un réarrangement des indices de sommation, on obtient finalement :

$$[q_n(t)]^2 = \sum_{k \in \mathbb{Z}} \left[\sum_{k' \in \mathbb{Z}} q(n, k')q(n, k - k') \right] e^{i\omega(n, k)t} . \quad (10.17)$$

Si donc on note $q^{(2)}(n, k)$ les composantes de Fourier du carré de la coordonnée $q^2(t)$, (10.17) montre que la règle de composition classique est :

$$q^{(2)}(n, k) = \sum_{k' \in \mathbb{Z}} q(n, k')q(n, k - k') . \quad (10.18)$$

La question se pose maintenant de représenter quantiquement $q^2(t)$. *A priori*, le plus naturel consiste à injecter dans (10.15) les représentants quantiques définis en (10.13) ; en s'y prenant de la sorte, on obtient :

$$\sum_{k, k' \in \mathbb{Z}^2} q_{n, n-k'} q_{n, n-k} e^{i\omega_{n, n-k'} t} e^{i\omega_{n, n-k} t} = \sum_{k, k' \in \mathbb{Z}^2} q_{n, n-k'} q_{n, n-k} e^{i(\omega_{n, n-k'} + \omega_{n, n-k})t} . \quad (10.19)$$

Le facteur de phase de chaque terme, $e^{i(\omega_{n, n-k'} + \omega_{n, n-k})t}$, ne convient pas. En effet, tout comme on représente $q_n(t)$ – dont le développement fait apparaître les $q(n, k) e^{i\omega(n, k)t}$ – par les quantités $Q_{n, n-k}$ définies en (10.13), il convient par cohérence de représenter $[q(t)]^2$ par un ensemble de quantités de la même forme, à savoir $q_{n, n-k}^{(2)} e^{i\omega_{n, n-k} t}$, les amplitudes $q_{n, n-k}^{(2)}$ et les pulsations de Bohr $\omega_{n, n-k}$ étant toujours astreintes à redonner leurs équivalents classiques dans la limite des très grands nombres quantiques : tout ce qui a été affirmé à ce sujet doit valoir tout autant pour n'importe quelle grandeur dynamique (par exemple le carré de la coordonnée)

⁸Il serait tout aussi acceptable d'imposer à $\omega_{n+k, n}$ de tendre vers $\omega(n, k)$. Cette tolérance vis-à-vis d'un (petit) glissement des indices est non-pertinente si $k \ll n$; d'ailleurs ... (voir la suite).

⁹Bien sûr, la notation à deux indices évoque une matrice mais, à ce stade, procéder brutalement à l'assimilation stricte des $Q_{nn'}$ avec une matrice est prématuré : une matrice est un objet obéissant à des règles algébriques bien précises et, pour l'instant, on ne connaît pas le statut des $Q_{nn'}$ vis-à-vis de ces règles. À l'époque, la théorie des matrices était fort peu familière aux physiciens – à l'exception notable de Born qui l'avait utilisée pour l'étude des vibrations mécaniques d'un solide cristallin.

que pour la coordonnée elle-même. Or, visiblement, $e^{i\omega_{n,n-k}t}$ – que l’on attend pour q^2 – est différent du facteur de phase apparaissant dans (10.19) puisque :

$$\omega_{n,n-k'} + \omega_{n,n-k} \neq \omega_{n,n-k} . \quad (10.20)$$

Pour que tout “marche” bien, il faudrait, au premier membre de (10.20), obtenir la somme :

$$\omega_{n,n-k'} + \omega_{n-k',n-k} . \quad (10.21)$$

Évoquant une “conclusion irréprouvable” (... *an almost cogent conclusion*, cité par Jammer [2], p. 201), Heisenberg arrange à la main les indices et se livre à la substitution¹⁰ :

$$\omega_{n,n-k'} + \omega_{n,n-k} \mapsto \omega_{n,n-k'} + \omega_{n-k',n-k} , \quad (10.22)$$

tant pour les pulsations que pour les amplitudes :

$$q_{n,n-k'}q_{n,n-k} \mapsto q_{n,n-k'}q_{n-k',n-k} . \quad (10.23)$$

Au total, Heisenberg remplace (10.19) par :

$$\sum_{k,k' \in \mathbb{Z}^2} q_{n,n-k'}q_{n-k',n-k} e^{i\omega_{n,n-k'}t} e^{i\omega_{n-k',n-k}t} \equiv \sum_{k,k' \in \mathbb{Z}^2} Q_{n,n-k'}Q_{n-k',n-k} , \quad (10.24)$$

et c’est ceci qu’il retient pour représenter q^2 dans sa théorie quantique :

$$q_{n,n-k}^{(2)} = \sum_{k' \in \mathbb{Z}} q_{n,n-k'}q_{n-k',n-k} . \quad (10.25)$$

Le point-clé est le suivant : le bricolage des indices est négligeable si les entiers n sont grands¹¹, de sorte que la substitution intuitive (10.19) ou sa version modifiée par Heisenberg, (10.24), sont quasi-indiscernables dans la limite des très grands nombres quantiques. Au contraire, si n est de l’ordre de l’unité, les deux expressions seront très différentes. Ainsi, avec $n \gg 1$ (donc dans la vision quasi-classique), la construction (10.24) – qui respecte la règle d’association des fréquences de Bohr – est tout aussi légitime que (10.19) – et c’est elle qu’il faut retenir, justement pour cette dernière raison. Dès lors, il devient évident que la théorie qui se dessine tendra d’une façon ou d’une autre vers la Mécanique Classique si les nombres quantiques sont grands, mais s’en démarquera nettement dans le cas contraire.

Clairement, la démarche de Heisenberg n’a à ce stade, aucune justification réelle¹² et son point d’arrivée provisoire est le fruit d’une déduction jalonnée par des étapes qui peuvent apparaître *ad hoc*¹³.

En résumé, Heisenberg propose de représenter la coordonnée d’une particule (inobservable en tant que telle) non par une loi horaire du genre $q(t)$, mais par une collection de nombres $Q_{n,n'}(t)$ assujettis à une algèbre non-commutative, un fait surprenant contenu dans la représentation du carré de la coordonnée par des nombres $Q_{n,n'}^2$ tels que :

$$Q_{n,n'}^{(2)} = \sum_{m \in \mathbb{Z}} q_{n,m}q_{m,n'} . \quad (10.26)$$

Ici apparaît la nature matricielle des tableaux de nombres $Q_{n,n'}$, puisque l’on reconnaît au second membre la multiplication des matrices. Ainsi, la coordonnée est représentée par une matrice Q d’éléments $Q_{n,n'}$, cependant

¹⁰C’est le “bricolage” des indices annoncé plus haut.

¹¹Schématiquement, $q_{100000,100001}$ doit pouvoir être confondu avec $q_{99999,100000} \dots$

¹²De la même façon, on verra que l’équation de Schrödinger ne se démontre pas, mais se construit d’une manière *inductive* (non purement déductive). Dans un cas comme dans l’autre, il reste à la nouvelle théorie à faire ses preuves.

¹³Dans une lettre à Einstein (datée du 15 juillet 1925), Born écrivait ([3], p. 100) :

“La nouvelle étude de Heisenberg, qui va bientôt paraître, a une allure très mystique, mais elle est sûrement exacte et profonde”

L’enthousiasme contenu de Born est à comparer à ce que lui disait Einstein, quelque temps plus tard (4 décembre 1926) :

“La mécanique quantique force le respect. Mais une voix intérieure me dit que ce n’est pas encore le nec plus ultra. La théorie nous apporte beaucoup de choses, mais elle nous rapproche à peine du secret du Vieux. De toute façon, je suis convaincu que lui, au moins, ne joue pas aux dés.” (ibid., p. 107).

que le carré de la coordonnée est représentée par le carré de la matrice Q^2 , d'éléments $Q_{n,n'}^{(2)}$, donnés par la règle de multiplication des matrices, (10.26). Ces matrices possèdent une propriété remarquable résultant de la réalité de $q(t)$, (10.9). Si on y fait la substitution brutale, on obtient :

$$q_{n,n-k} = q_{n,n+k}^* . \quad (10.27)$$

Cette relation n'est quère plaisante ; en réaménageant à nouveau les indices (en les faisant glisser de k à droite), on adopte plutôt :

$$q_{n,n-k} = q_{n-k,n}^* \iff q_{n,n'} = q_{n',n}^* . \quad (10.28)$$

Autrement dit, la matrice transposée est égale à la matrice initiale complexe conjuguée. Cette symétrie remarquable caractérise les matrices dites *hermitiques*, dont le rôle de tout premier plan sera abondamment rediscuté dans la suite¹⁴.

Une fois admise la loi de représentation pour la coordonnée et son carré, Heisenberg l'a généralisée de proche en proche pour une puissance quelconque de la coordonnée et également pour le produit de deux coordonnées distinctes, comme x et y . Ceci permet de construire formellement la représentation quantique de n'importe quelle fonction développable en série entière des coordonnées.

La démarche de Heisenberg a été guidée par le souci premier de renoncer à donner une description théorique ce qui n'est pas directement accessible à l'observation, c'est-à-dire fondamentalement les trajectoires (par exemple, celle de l'électron dans l'atome d'hydrogène). Dans son formalisme, au contraire, apparaissent de façon cruciale les fréquences qui, elles, sont directement observables puisqu'elles précisent les caractéristiques de la lumière émise par l'atome (pour ne citer que cet exemple). Il reste bien sûr à élucider le sens physique des quantités $q_{nn'}$: on verra notamment que certaines d'entre elles sont reliées à l'intensité des raies observées et fournissent les règles de sélection¹⁵. Plus généralement, la matrice $Q_{nn'}(t)$ contient toute la dynamique de la coordonnée, au sens de la Mécanique Quantique.

Il appartient à Born d'avoir généralisé la correspondance proposée par Heisenberg entre variables dynamiques classiques et représentations quantiques, en posant que ce qui était admis pour les coordonnées, devait aussi s'appliquer aux impulsions p_u ($u = x, y, z$). L'ensemble des règles ainsi adoptées permettait alors de construire le représentant quantique de n'importe quelle fonction des variables coordonnées et impulsion, pourvu que cette fonction soit développable en série entière. On savait ainsi désormais représenter par exemple le Lagrangien ou le Hamiltonien d'un système en partant de son expression classique.

L'aspect le plus surprenant des résultats de Heisenberg et Born est qu'elle introduit une algèbre non commutative : la position q et l'impulsion p sont finalement représentées par des matrices. La théorie des matrices avait été fondée soixante-dix ans plus tôt, en 1855, par Cayley et on sait bien, qu'en général, le produit de deux matrices A et B dépend de l'ordre des facteurs :

$$AB \neq BA \quad (10.29)$$

ce qui s'écrit, en définissant le commutateur $[,]$:

$$[A, B] \equiv AB - BA \neq 0 . \quad (10.30)$$

À n'en pas douter, il convenait d'examiner le commutateur fondamental de q et de p , les deux grandeurs centrales de la dynamique classique. Born et Jordan généralisèrent formellement les équations de Hamilton en termes de matrices ; cette extension implique clairement la notion de dérivée d'une matrice, ici une dérivée en temps. La définition adoptée est la suivante ; la dérivée d'une matrice d'éléments $\alpha_{nm}(t)$ est la matrice admettant comme éléments les quantités $\dot{\alpha}_{nm}$ – ceci est naturel compte tenu de la règle d'addition des matrices et du fait que la dérivée implique la différence de deux objets pris l'un à t , l'autre à $t + dt$. Munis de cette définition, Born et Jordan démontrèrent que l'extrémalisation de la trace¹⁶ de la matrice :

$$L = p\dot{q} - H(q, p) \quad (10.31)$$

¹⁴On verra que toute grandeur physique est nécessairement représentée par un opérateur dont la matrice, sur une base orthonormée, est hermitique.

¹⁵dans l'approximation dite dipolaire électrique, en abrégé E1.

¹⁶La trace d'une matrice est la somme de ses éléments diagonaux. C'est un invariant par rapport à tout changement de base : si les éléments diagonaux changent individuellement d'une base à l'autre, leur somme reste au contraire inaltérée.

conduit aux équations en termes de matrices :

$$\dot{q} = \frac{\partial H}{\partial p} \quad \dot{p} = -\frac{\partial H}{\partial q} , \quad (10.32)$$

formellement semblables aux équations canoniques de Hamilton. C'est alors que Born trouva l'expression du commutateur fondamental des matrices de q et p :

$$qp - pq \equiv [q, p] = i\hbar \mathbf{1} , \quad (10.33)$$

où $\mathbf{1}$ désigne la matrice identité (qui est diagonale sur toute base et dont tous les éléments diagonaux sont égaux à 1). La relation de commutation (10.33) est la relation fondamentale de la nouvelle mécanique, de laquelle tout (ou presque) est déductible. Elle est en remarquable analogie avec le crochet de Poisson de la Mécanique Analytique ; en effet :

$$\{q, p\} \equiv \frac{\partial q}{\partial q} \frac{\partial p}{\partial p} - \frac{\partial p}{\partial q} \frac{\partial q}{\partial p} = 1 \times 0 - 0 \times 0 = 1 . \quad (10.34)$$

Par comparaison avec (10.33), ce résultat élucide la correspondance formelle entre le commutateur quantique fondamental et son crochet de Poisson classique :

$$\{q, p\} \leftrightarrow \frac{1}{i\hbar} [q, p] . \quad (10.35)$$

Cette relation tenant au niveau de la coordonnée et de l'impulsion, elle est par conséquence vraie de proche en proche pour tous les couples de grandeurs dynamiques :

$$\{A, B\} \leftrightarrow \frac{1}{i\hbar} [A, B] , \quad (10.36)$$

puisque – comme on peut s'en convaincre sur des exemples –, l'algèbre des crochets de Poisson et celle des commutateurs sont identiques. C'est ainsi que les deux composantes du moment cinétique $\vec{L} = \vec{r} \times \vec{p}$ satisfont :

$$[L_x, L_y] = i\hbar L_z , \quad (10.37)$$

et bien d'autres relations remarquables, toutes déductibles du crochet classique en utilisant la correspondance (10.35).

À ce stade, il n'est pas nécessaire de poursuivre l'exposé de la Mécanique des Matrices : comme mentionné plus haut, elle et l'approche de Schrödinger seront par la suite englobées dans un même formalisme qui sera au contraire l'objet d'une discussion très détaillée. Il faut toutefois signaler que sa première application par Heisenberg fut le traitement de l'oscillateur harmonique. Le résultat majeur ainsi obtenu est la valeur des énergies possibles (quantifiées !) :

$$E \in \{E_n\} \quad E_n = \left(n + \frac{1}{2}\right) h\nu \quad (n \in \mathbb{N}) . \quad (10.38)$$

Ainsi, dans la limite $n \gg 1$ on a :

$$E_n \simeq n h\nu . \quad (10.39)$$

Ce résultat issu de la Mécanique des Matrices n'est autre que l'hypothèse formulée par Planck en 1900 !

Il convient enfin de signaler que bien d'autres résultats ont été obtenus avec le formalisme de Heisenberg ; en particulier, la généralisation aux systèmes non périodiques ([2], § 5.2) a été effectuée par Born et Wiener.

10.3 Les ondes de matière (de Broglie, 1923)

Juste avant la Mécanique des Matrices, une autre approche était proposée, reprenant les travaux de Hamilton publiés entre 1828 et 1837 (et tombés aux oubliettes). La formulation de la Mécanique proposée par Hamilton trouva son origine dans ses recherches en l'Optique. Le point fondamental est de remarquer l'analogie entre

le Principe de Moindre Action de Maupertuis (PMA), le plus ancien principe variationnel de la dynamique ordinaire, et le Principe de Moindre Durée de Fermat (PMD) conduisant aux lois de l'optique géométrique et en particulier à la loi de Snell - Descartes. Le premier peut se retrouver à partir de la forme moderne du principe variationnel conduisant aux équations de Lagrange. S'agissant de ces dernières, le point de départ est, pour un système conservatif d'énergie constante E :

$$S = \int_0^t L dt' = \int_0^t (p\dot{q} - H) dt' = \int_0^t p\dot{q} dt' - Et = \int_{M_0}^{M_t} pdq - Et \equiv \mathcal{J} - Et . \quad (10.40)$$

Pour un ensemble de trajectoires d'énergie donnée, la variation porte sur la seule intégrale \mathcal{J} ; V étant l'énergie potentielle de la particule de masse m , le principe variationnel de la Mécanique s'écrit :

$$\delta \int_{M_0}^{M_t} pdq = 0 \iff \delta \int_{M_0}^{M_t} \sqrt{2m(E - V)} dq = 0 . \quad (10.41)$$

Quant au Principe de Fermat, affirmant que le temps de parcours de la lumière doit être le plus court, il s'énonce (n est l'indice du milieu, s le chemin curviligne de la lumière) :

$$\delta \int \frac{ds}{c/n} = 0 \iff \delta \int_{M_0}^{M_t} n ds = 0 . \quad (10.42)$$

La comparaison des équations (10.41) et (10.42) conduisit Hamilton à affirmer que la quantité $p = \sqrt{2m(E - V)}$ joue en Mécanique le même rôle¹⁷ que l'indice n ; mais l'analogie optico-mécanique de Hamilton va encore plus loin.

En effet, on sait d'une part que l'énergie est l'opposé de la dérivée de S par rapport au temps ; d'autre part, les relations $p_i = \partial S / \partial q_i$ dans \mathbb{R}^3 peuvent s'écrire avec $\vec{\nabla}$. Au total :

$$\frac{\partial S}{\partial t} = -E , \quad \vec{p} = \vec{\nabla} S . \quad (10.43)$$

La deuxième équation montre que le vecteur impulsion, qui définit en tout point la tangente à la trajectoire, est normal aux surfaces d'action constante. En d'autres termes, sans considérer l'aspect cinématique du mouvement (la loi horaire), les trajectoires sont, d'un strict point de vue géométrique, les "rayons" des surface d'équi-action.

De l'autre côté, pour tout mode propre du champ électromagnétique dans le vide, le vecteur d'onde \vec{k} et la pulsation $\omega = 2\pi\nu$ définissent la phase $\phi = \vec{k} \cdot \vec{r} - \omega t$ par :

$$\frac{\partial \phi}{\partial t} = -\omega , \quad \vec{k} = \vec{\nabla} \phi . \quad (10.44)$$

\vec{k} est normal aux surfaces d'équi-phase, exprimant l'orthogonalité entre ces surfaces et les rayons lumineux. L'analogie entre \vec{p} et \vec{k} est manifeste.

Ces conclusions remarquables de Hamilton ne furent considérées que comme des analogies formelles dénuées d'implications profondes ; compte tenu des succès de la Physique d'alors, nul ne ressentait le besoin d'une révolution des concepts. Cette révolution aurait consisté à admettre que, tout comme l'optique géométrique n'est qu'une approximation (les rayons lumineux ne sont que les traces ("trajectoires") des ondes dans la limite où la longueur d'onde est très petite), la Mécanique Classique n'est elle aussi qu'une approximation d'une autre théorie plus générale. Admettre ceci eut été admettre que des ondes doivent être associées aux particules, les trajectoires de ces dernières n'étant que les signes visibles à l'œil d'une réalité bien plus complexe – tout comme les rayons lumineux de l'Optique Géométrique ne sont qu'une visualisation schématique des ondes de l'Optique Physique. À l'époque, aucune expérience cruciale ne justifiait un tel saut conceptuel¹⁸.

La relation de Planck - Einstein $E = h\nu$ donne une nouvelle dimension à l'analogie optico-mécanique de Hamilton ; elle établit en effet un lien strict entre deux grandeurs, l'une ayant une signification mécanique – l'énergie, l'autre une signification ondulatoire – la fréquence :

$$E = h\nu = \hbar\omega . \quad (10.45)$$

¹⁷à un facteur près.

¹⁸Toutefois, certains penseurs hétérodoxes, pour des raisons surtout philosophiques, franchirent le pas avant l'heure, en-dehors de toute méthodologie propre aux sciences "dures" – voir [2], p. 241.

En utilisant les premières relations de (10.43) et (10.44), (10.45) se réécrit :

$$\frac{\partial S}{\partial t} = \hbar \frac{\partial \phi}{\partial t} . \quad (10.46)$$

À partir du moment où l'on admet la proportionnalité entre S et ϕ *via* la constante de Planck h , la même proportionnalité doit tenir entre \vec{p} et \vec{k} :

$$\vec{p} = \vec{\nabla} S = \vec{\nabla}(\hbar\phi) = \hbar\vec{\nabla}\phi \equiv \hbar\vec{k} . \quad (10.47)$$

Avec $k = 2\pi/\lambda$ et¹⁹ $p = mv$, (10.47) prend la forme historique proposée par de Broglie :

$$\lambda = \frac{h}{mv} . \quad (10.48)$$

Par cette relation, de Broglie associe une onde (“onde associée”) à toute particule de masse (propre) m et de vitesse v ; dans le cas relativiste (10.48) devient^{20, 21} :

$$\lambda = \frac{h}{\sqrt{(E/c)^2 - m^2c^2}} = \frac{h}{p} = \frac{h}{\gamma mv} . \quad (10.49)$$

La longueur d’onde est d’autant plus petite que l’énergie est grande ; pour une énergie donnée, elle est d’autant plus grande que la masse est petite. Les démarches respectives de Hamilton et de de Broglie peuvent être résumées comme indiqué sur la figure (10.1).

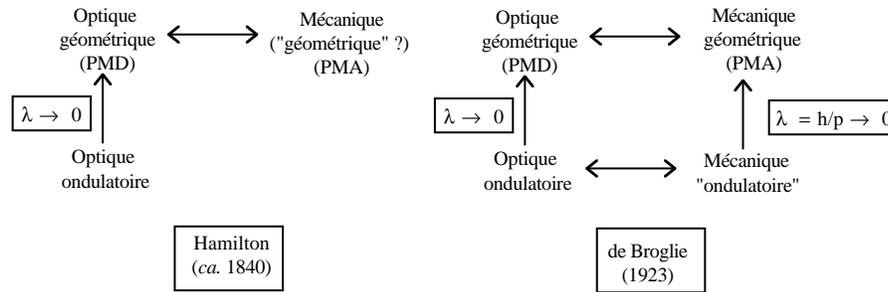


Figure 10.1: Démarches comparées de Hamilton et de Broglie.

À ce stade, il est naturel de se poser la question de la validité de la Mécanique ordinaire (“géométrique”) : on peut y répondre par analogie avec le critère précisant celle de l’Optique géométrique. Cette dernière est applicable lorsque l’indice varie lentement à l’échelle de la longueur d’onde, et est donc formellement exacte à la limite $\lambda \rightarrow 0$. Transposée à la Mécanique, cette condition s’énonce (voir (10.41) et (10.42)) : les variations d’énergie potentielle doivent être lentes à l’échelle de la longueur d’onde associée, soit :

$$\left| \frac{V(l+\lambda) - V(l)}{V(l)} \right| \ll 1 , \quad (10.50)$$

où l est une longueur typique du problème. A titre d’exemple, soit l’atome d’hydrogène dans l’optique de Bohr²² ; par un simple argument dimensionnel, on voit que la longueur d’onde associée est en ordre de grandeur donnée par :

$$\lambda = \frac{h}{md/T} , \quad (10.51)$$

¹⁹La relation de de Broglie est en général énoncée sous l’une autre forme, $\lambda = h/(mv)$ ou $\lambda = h/p$; on peut donc se demander comment elle s’écrit en présence d’un champ magnétique. Ceci ressemble à un faux-problème : l’essentiel est que, dans un cas comme dans l’autre, on sait construire le Hamiltonien du système, d’où tout est déductible ; de ce fait, écrire la relation de de Broglie n’est pas une étape obligée pour avancer.

²⁰L’expression (10.49) de λ a été anticipée lors de la description de la diffusion d’électrons relativistes par un noyau – voir Ch. 9. Dans ces relations, m désigne la masse propre (“au repos”) de l’électron.

²¹Bien évidemment, pour le photon on retrouve $\lambda = hc/E = c/\nu$.

²²Il est facile de vérifier que la longueur des cercles de Bohr est un multiple *entier* de λ .

où d est le périmètre du cercle et T la période du mouvement circulaire uniforme. Pour une fréquence de rotation optique ($\sim 10^{14}$ Hz) et une longueur atomique, on trouve λ de l'ordre de quelques dizaines d'Å ; il est clair que, sur une telle distance, V varie énormément puisque la taille de l'atome est de l'ordre de 1 Å : selon ce critère, la mécanique de l'atome est donc *nécessairement* ondulatoire !

Ce critère n'est pas infaillible (il pourrait faire croire que pour une particule libre – V est alors constant –, la Mécanique Classique est toujours valide, ce qui n'est pas le cas). En réalité, le bon critère repose sur la considération de l'action typique du problème : si celle-ci est grande par rapport à h , les effets quantiques sont petits, si elle d'ordre h (ou plus petite), ceux-ci seront importants. Pour un mouvement circulaire uniforme, l'action la plus immédiate est le moment cinétique J :

$$J = mvR = mR^2\omega ; \quad (10.52)$$

considérons d'abord la révolution diurne de la Terre. On a²³ :

$$J_{\text{Terre}} = 6 \times 10^{24} \times (150 \times 10^9)^2 \frac{2\pi}{86400 \times 365} \simeq 2.7 \times 10^{40} \text{ J s} \simeq 0.4 \times 10^{74} h \text{ (!!!)} . \quad (10.53)$$

Par contraste, pour un électron orbitant autour du proton à une fréquence optique, on a :

$$J_e = 9 \times 10^{-31} \times (2 \times 10^{-10})^2 \times 4 \times 10^{15} \simeq 1.5 \times 10^{-34} \text{ J s} \simeq 0.25 h . \quad (10.54)$$

10.4 L'équation de Schrödinger

Cette section est consacrée à l'établissement de l'équation de Schrödinger ; cette construction procède par induction (en fin de course, Schrödinger précise bien qu'il admet par *postulat* l'équation qu'il est parvenue à écrire) et ne doit donc être considérée comme une "démonstration". L'exposé ci-dessous reprend l'essentiel de la démarche originelle, publiée dans quatre articles en 1926, constituant ce que Born désigna en 1961 comme une "*contribution insurpassée en Physique Théorique*".

Dans le deuxième article [4], Schrödinger exploite à fond l'analogie optico-mécanique de Hamilton, à la lumière des idées de de Broglie. Le point de départ est l'équation de Hamilton - Jacobi pour une particule de masse m et d'énergie potentielle V (se souvenir que la relation générale $p_k = \partial S / \partial q_k$ s'écrit vectoriellement $\vec{p} = \vec{\nabla} S$ dans \mathbb{R}^3) :

$$\frac{\partial S}{\partial t} = -H(\vec{r}, \vec{\nabla} S) . \quad (10.55)$$

Pour une trajectoire d'énergie E , on a :

$$\frac{\partial S}{\partial t} = -E , \quad |\vec{\nabla} S|^2 = 2m[E - V(\vec{r})] = p . \quad (10.56)$$

Soit une surface d'action constante Σ , définie par l'équation $S(\vec{r}, t) = \text{C}^{\text{ste}} = S_0$. Au fil du temps, cette surface se déplace : x , y et z varient, tout comme t , mais de telle sorte que la fonction S garde la même valeur S_0 . Un point M donné de cette surface décrit une ligne normale à la surface²⁴ et se déplace de M_t à M_{t+dt} entre t et $t + dt$; pour ces deux points séparés²⁵ de dl le long de la normale à Σ , la variation de S est formellement :

$$dS = |\vec{\nabla} S| dl + \frac{\partial S}{\partial t} dt = \sqrt{2m(E - V)} dl - E dt = p dl - E dt . \quad (10.57)$$

Par construction, dS est nul, ce qui donne la vitesse de déplacement v_ϕ d'un point d'une surface d'équi-action :

$$v_\phi \equiv \frac{dl}{dt} = \frac{E}{\sqrt{2m(E - V)}} = \frac{E}{p} , \quad (10.58)$$

²³La masse de la Terre est environ 6×10^{24} kg et la distance Terre - Soleil vaut (en moyenne) à peu près 150 millions de kilomètres.

²⁴Tout comme les lignes de champ (électrique) sont les normales aux équipotentiels (électrostatiques).

²⁵ dl est donc le déplacement normal.

un résultat déjà obtenu au ch. 9. Ayant fait apparaître cette vitesse, Schrödinger lui fait jouer le rôle d'une vitesse de phase dans une certaine équation d'onde pour une certaine fonction, notée Ψ et qu'il appelle "fonction d'onde", équation recopiée de l'équation d'onde de l'Électromagnétisme :

$$\Delta\Psi - \frac{1}{v_\phi^2} \frac{\partial^2\Psi}{\partial t^2} = 0 . \quad (10.59)$$

En posant une dépendance temporelle sous la forme $e^{-2i\pi\nu t}$, soit $e^{-2i\pi(E/h)t}$:

$$\Psi(\vec{r}, t) = e^{-2i\pi(E/h)t} \psi(\vec{r}) , \quad (10.60)$$

il obtient :

$$\Delta\psi - \frac{1}{v_\phi^2} \left(\frac{-2i\pi E}{h} \right)^2 \psi = 0 . \quad (10.61)$$

Tenant alors compte de (10.58), Schrödinger obtient la forme historique :

$$\Delta\psi + \frac{8\pi^2 m}{h^2} (E - V)\psi = 0 . \quad (10.62)$$

Schrödinger appliqua immédiatement son équation à l'atome d'hydrogène en choisissant de retenir les solutions partout monovaluées, finies et tendant vers zéro à l'infini ; l'application de ces conditions aux limites, pour l'instant quelque peu arbitraires²⁶, induit immédiatement la *quantification* des valeurs possibles de E , un phénomène universel à propos des équations de ce type (penser à une corde vibrante). Ce résultat est en soi remarquable, mais, encore plus extraordinaire, les valeurs possibles E_n se trouvent être strictement celles obtenues par Bohr en 1913 pour retrouver la série de Balmer !

Après ce premier succès, Schrödinger résolut son équation dans le cas de l'oscillateur harmonique à une dimension et obtint à nouveau la quantification de l'énergie ; les valeurs ainsi trouvées, $E_n = [n + (1/2)]h\nu$, sont exactement les mêmes que celles obtenues par Heisenberg dans sa Mécanique des Matrices, un formalisme en apparence situé aux antipodes (en outre, Schrödinger obtint, (avec l'aide de Weyl), les solutions ψ elles-mêmes, inédites et apparemment absentes (en fait implicites) dans la formulation de Heisenberg). En tout cas, l'identité de ces résultats précoces est un argument très fort en faveur de la justesse des deux démarches, parce qu'elles dévoilent finalement la même réalité.

L'équation (10.62) peut s'écrire sous une autre forme, devenue habituelle :

$$-\frac{\hbar^2}{2m}\Delta\psi + V\psi = E\psi , \quad (10.63)$$

et est en fait une équation aux valeurs propres pour un certain opérateur ; en effet, le premier membre désigne l'action de l'opérateur

$$-\frac{\hbar^2}{2m}\Delta + V \quad (10.64)$$

sur la fonction ψ . Le résultat, lu au second membre de (10.63), est la *même* fonction, simplement multipliée par le scalaire E . Les solutions de (10.63) sont donc les *fonctions propres* de l'opérateur (10.64), et E est la valeur propre correspondante²⁷. La fonction $\Psi(\vec{r}, t)$ définie en (10.60) satisfait aussi cette équation aux valeurs propres puisque le facteur de phase ne dépend que du temps ; on a donc aussi :

$$-\frac{\hbar^2}{2m}\Delta\Psi + V\Psi = E\Psi . \quad (10.65)$$

La valeur propre E varie d'une solution à l'autre. De plus, persuadé que la forme la plus générale de l'équation de sa nouvelle mécanique devait à la fois dépendre du temps et ne pas contenir cette valeur propre, Schrödinger entreprit d'écrire une équation satisfaisant ces (ses !) exigences. Il faut bien dire que l'étape cruciale suivante

²⁶C'est la signification physique ultérieurement attribuée à Ψ qui justifiera le choix de ces conditions aux limites.

²⁷Une fonction propre et la valeur propre correspondante forment un couple propre indissociable ; pour rappeler ce fait, on peut noter ψ_E la fonction propre associée à la valeur propre E .

reste quelque peu mystérieuse, comme le montre clairement la lecture des mémoires originaux²⁸. En fin de compte, postulant une dépendance fondamentale “élémentaire” en $e^{-2i\pi\nu t} = e^{Et/(i\hbar)}$ et en éliminant l'énergie “à l'envers”²⁹, Schrödinger obtient finalement :

$$-\frac{\hbar^2}{2m}\Delta\Psi + V\Psi = i\hbar\frac{\partial\Psi}{\partial t} . \quad (10.67)$$

C'est l'équation de Schrödinger ; la forme réduite aux seules variables d'espace (10.63) est l'équation aux valeurs propres correspondante.

L'apparition du i provient de la dépendance temporelle supposée, (10.60). Ici, au contraire de ce que l'on fait en Optique ou en Électricité, il ne s'agit pas d'une astuce technique fondée sur l'usage commode des nombres complexes et où, en fin de compte on prend une partie réelle puisque l'on s'intéresse à des grandeurs physiques, réelles par essence. Comme l'écrit Schrödinger³⁰ :

... Ψ ne doit plus être limitée à sa partie réelle mais prise dans son sens mathématique propre, la fonction d'onde étant considérée maintenant comme essentiellement complexe.

Le formalisme développé par Schrödinger lui permet non seulement d'obtenir les énergies d'un système quantique, E_n , mais également les fonctions propres ψ_n de l'opérateur (10.64) et les fonctions Ψ_n en tant que solutions *particulères* de l'équation (10.67). Par ailleurs, se posait de façon assez énigmatique le problème de l'équivalence entre la Mécanique des Matrices de Heisenberg - Born - Jordan et la Mécanique Ondulatoire de Schrödinger, les deux approches donnant les mêmes résultats pour tous les exemples traités. Convaincu de l'existence de cette équivalence, Schrödinger obtint lui aussi la relation fondamentale³¹ (10.33) dans le langage des fonctions d'onde, en exploitant la remarque suivante. Si on introduit l'opérateur différentiel $-i\hbar\partial/\partial q$, on vérifie trivialement l'égalité suivante :

$$q\left(-i\hbar\frac{\partial}{\partial q}\psi(q)\right) - \left(-i\hbar\frac{\partial}{\partial q}(q\psi(q))\right) = i\hbar\psi(q) . \quad (10.68)$$

En effet, la dérivation du produit $q\psi(q)$ dans la deuxième grande parenthèse produit deux termes dont l'un se compense avec ce qui vient de la première. Cette égalité est vraie pour toute fonction $\psi(q)$ et, par conséquent, tient au niveau des opérateurs eux-mêmes. On peut donc écrire :

$$q\left(-i\hbar\frac{\partial}{\partial q}\bullet\right) - \left(-i\hbar\frac{\partial}{\partial q}(q\bullet)\right) = i\hbar\bullet , \quad (10.69)$$

soit, avec la notation en commutateur :

$$[q, -i\hbar\frac{\partial}{\partial q}] = i\hbar\mathbf{1} , \quad (10.70)$$

²⁸A propos du i ($i^2 = -1$) présent en facteur de la dérivée en temps de (10.67), le mieux est de laisser la parole à Schrödinger lui-même qui, en 1932, commentait ainsi son équation fondamentale :

“... Mais comment le $\sqrt{-1}$ a-t-il pu s'introduire dans cette équation ? Une réponse, dont je n'ose indiquer ici que le sens général, a été donnée à cette question par un physicien, qui a autrefois quitté l'Autriche, mais qui, malgré de longues années passées l'étranger, n'a pas complètement perdu son humeur mordante de Viennois et qui, en outre, est connu pour sa faculté de trouver toujours le mot juste : Le $\sqrt{-1}$ s'est glissé dans l'équation comme quelque chose que nous laissons échapper par hasard, en éprouvant toutefois un soulagement inappréciable après lui avoir donné naissance involontairement”. [5], p. 166 .

²⁹Avec la fonction $\bullet \equiv e^{Et/(i\hbar)}$, on a :

$$\frac{E\bullet}{i\hbar} = \frac{\partial\bullet}{\partial t} \iff E\bullet = i\hbar\frac{\partial\bullet}{\partial t} . \quad (10.66)$$

³⁰En 1932, Schrödinger commente comme suit sa propre remarque :

“Les mots essentiellement complexes cherchent à dissimuler ici une grande difficulté. Dans son désir d'envisager à tout prix le phénomène de la propagation des ondes Ψ comme quelque chose de réel dans le sens classique du terme, l'auteur s'était refusé de reconnaître franchement que tout le développement de la théorie mettait de plus en plus clairement en évidence le caractère essentiellement complexe de la fonction d'onde.”

³¹On peut même la qualifier de *fondatrice*.

où $\mathbf{1}$ désigne l'opérateur identité. Cette relation est en tout point analogue à la relation fondamentale (10.33) de la Théorie des Matrices, ce qui suggère de représenter l'impulsion p par l'opérateur différentiel :

$$p \mapsto -i\hbar \frac{\partial}{\partial q} . \quad (10.71)$$

Plus généralement, à la composante cartésienne p_u ($u = x, y, z$) du moment conjugué on associe la dérivation $-i\hbar\partial/\partial u$; alors, la relation fondamentale se généralise en :

$$[u, -i\hbar \frac{\partial}{\partial v}] = i\hbar\delta_{uv} \mathbf{1} \quad (u, v = x, y, z) , \quad (10.72)$$

où δ_{uv} est le symbole de Kronecker (x commute avec $\partial/\partial y$, etc.).

Cette règle étant admise, on réalise que l'équation de Schrödinger (10.67) peut s'écrire *automatiquement* de la façon suivante. On part de l'expression du Hamiltonien de la Mécanique Analytique (plus tard appelé "classique" quand il faudra le distinguer du Hamiltonien quantique H qui va s'en déduire) en terme des variables \vec{r} et \vec{p} exprimées en coordonnées *cartésiennes*³², soit $H_{\text{class}}(\vec{r}, \vec{p})$; ceci étant fait, on remplace dans cette expression \vec{r} par l'opérateur multiplication par \vec{r} , \vec{p} par l'opérateur différentiel $-i\hbar\vec{\nabla}$; ceci fabrique de fait le premier membre de (10.67) (le carré du gradient est le Laplacien). En résumé, le mode de construction de l'équation de Schrödinger passe par la substitution :

$$H_{\text{class}}(\vec{r}, \vec{p}) , \quad \vec{r} \mapsto \times\vec{r} , \quad \vec{p} \mapsto -i\hbar\vec{\nabla} \implies H(\vec{r}, -i\hbar\vec{\nabla}) . \quad (10.73)$$

Avec le H ainsi construit, (10.67) s'écrit de manière particulièrement concise :

$$i\hbar \frac{\partial \Psi}{\partial t} = H \Psi . \quad (10.74)$$

La "recette" n'est pas toujours suffisante, puisqu'il existe des concepts quantiques (comme le spin), qui s'évanouissent totalement à la limite classique et que l'on ne peut donc obtenir par un processus de limite effectué en sens inverse ; dans certaines situations, il convient donc de compléter le Hamiltonien (quantique) ainsi construit en y rajoutant *à la main* des termes spécifiques (comme le couplage magnétique entre le moment magnétique associé au spin et un champ magnétique). En outre, il est intéressant de remarquer que la Mécanique Quantique se trouve parée d'un statut épistémologique très particulier, unique même : puisque le Hamiltonien est formé, dans un premier temps, par référence à la Mécanique Classique, on peut dire que la Mécanique Quantique a besoin explicitement de sa théorie limite (la Mécanique Classique) pour pouvoir être formulée en détail³³. Usuellement, il n'en va pas ainsi : toute nouvelle théorie physique généralisant une théorie antérieure est formulée intrinsèquement et, bien sûr, dans la bonne limite, redonne la "vieille" théorie. Par exemple, la Relativité Restreinte se formule à partir de deux postulats (l'invariance des lois physiques et la constante de la vitesse de la lumière dans tous les référentiels d'inertie) ; de ces deux postulats, on déduit tout et, bien sûr, prenant la limite c infini, on retombe sur la bonne vieille Mécanique de Galilée.

Retrouver (10.33) dans le langage de Schrödinger constitue l'étape décisive pour établir l'équivalence entre l'approche de celui-ci et celle de Heisenberg. Eckart et Pauli, puis finalement Dirac, ont établi en détail la stricte biunivocité des deux théories. Techniquement, le point central est le suivant. L'équation aux dérivées partielles de Schrödinger admet un ensemble de fonctions propres, ψ_n ; comme l'équation est linéaire, toute combinaison linéaire est encore une solution³⁴. S'introduit alors naturellement la notion d'espace vectoriel (c'est un espace de fonctions) et il est bien connu que dans un tel espace on sait définir et faire agir des opérateurs – en général linéaires³⁵ – représentables par des matrices dès qu'une base de l'espace est choisie. C'est ainsi que l'équation d'onde de Schrödinger, assortie des conditions aux limites imposées par le sens physique attribué à la fonction d'onde³⁶, ouvre la voie vers les matrices de Heisenberg.

Il n'est pas utile d'en dire plus pour l'instant : une grande partie du jeu ultérieur sera constitué d'un aller-retour entre une équation aux dérivées partielles et une représentation matricielle ; par la suite, on construira

³²Point important ! Pour des coordonnées non rectangulaires, la substitution n'est pas aussi simple.

³³La relation fondatrice (10.33) a été obtenue en Mécanique des Matrices par considération de la relation de quantification de l'ATQ, laquelle était le critère de choix parmi les *trajectoires* classiques.

³⁴C'est ce que l'on appelle parfois le principe de superposition linéaire ; il ne s'agit pas d'un principe mais bien d'une conséquence de la nature de l'équation de Schrödinger.

³⁵L'opération de renversement du temps exige l'introduction d'opérateurs *antilinéaires*.

³⁶et qui produisent spontanément la quantification de l'énergie, du moment cinétique, ...

d'ailleurs le formalisme de façon plus abstraite (*à la Dirac*), et on verra comment il s'incarne dans l'une ou l'autre de ces deux formes principales.

Le principe d'une résolution systématique de l'équation (10.67) est le suivant. On commence par résoudre³⁷ l'équation aux valeurs propres (10.65), ce qui fournit un ensemble de couples propres $(E_\alpha, \psi_\alpha(\vec{r}))$. Chacun d'entre eux permet de construire une solution particulière³⁸ de (10.67) :

$$\Psi_\alpha(\vec{r}, t) = \psi_\alpha(\vec{r}) e^{-i\omega_\alpha t} \quad (\omega_\alpha = \hbar^{-1} E_\alpha) . \quad (10.75)$$

Puis, en vertu de la linéarité de (10.67), on écrit la solution la plus générale Ψ sous la forme d'une combinaison linéaire :

$$\Psi(\vec{r}, t) = \sum_{\alpha} A_{\alpha} \Psi_{\alpha}(\vec{r}, t) , \quad (10.76)$$

où les coefficients A_{α} sont à déterminer. \sum_{α} désigne une opération de sommation en général : ce peut être une somme discrète et/ou une somme continue (intégrale)³⁹. Pour trouver les coefficients A_{α} du développement (10.76), il faut faire référence à une condition initiale donnée indépendamment⁴⁰ ; à défaut d'une telle information, le problème ne peut être résolu. Se donnant la fonction d'onde au départ :

$$\Psi(\vec{r}, t = 0) = a(\vec{r}) \quad \forall \vec{r} \in \mathbb{R}^3 , \quad (10.77)$$

où a est une fonction connue, on est en mesure d'écrire une équation pour les A_{α} ; faisant $t = 0$ dans l'expression de la solution la plus générale⁴¹, (10.76), on obtient l'équation :

$$\sum_{\alpha} A_{\alpha} \psi_{\alpha}(\vec{r}) = a(\vec{r}) . \quad (10.78)$$

La fonction $a(\vec{r})$ est donnée, les $\psi_{\alpha}(\vec{r})$ sont connues par la résolution de l'équation aux valeurs propres : l'inversion du système (10.78) fournit les A_{α} , de façon unique.

Ces éléments techniques étant précisés, la question du sens physique de la fonction d'onde de Schrödinger demeure. C'est évidemment un point capital dont l'élucidation complète ne sera faite qu'un peu plus tard (Born, 1927). Bien sûr, Schrödinger s'est interrogé sur ce point, mais n'a fourni que des éléments de réponse ; sa conviction était que le produit⁴² $\Psi\Psi^*$ représente les fluctuations électriques de charge dans l'espace physique : on voit qu'il n'était pas loin de la notion de probabilité de présence d'une charge. Cette croyance était fondée essentiellement sur le fait suivant ; à partir de l'équation (10.67), il est facile de montrer que si on pose (pour une particule de charge e et de masse m) :

$$\rho_Q = e\Psi\Psi^* , \quad \vec{j}_Q = \frac{e\hbar}{2im} [\Psi^* \vec{\nabla}\Psi - \Psi \vec{\nabla}\Psi^*] , \quad (10.79)$$

alors les quantités ρ_Q et \vec{j}_Q satisfont l'équation de conservation :

$$\frac{\partial \rho_Q}{\partial t} + \text{div } \vec{j}_Q = 0 ; \quad (10.80)$$

Il faut bien retenir que (10.80) est une *conséquence* de l'équation fondamentale (10.67). ρ_Q et \vec{j}_Q ont tout ce qu'il faut pour être respectivement une densité de charge et un courant de charge. Écrite pour la charge e , cette équation est manifestement valide pour tout attribut scalaire⁴³ et aussi évidemment sans ce facteur scalaire. Très généralement, on a donc :

$$\frac{\partial \rho}{\partial t} + \text{div } \vec{j} = 0 , \quad (10.81)$$

³⁷En pratique, les solutions ne sont connues exactement que pour un petit nombre de potentiels V .

³⁸Ces solutions particulières sont remarquables et sont appelées *états stationnaires* ; on verra par la suite la raison de cette terminologie.

³⁹Que la solution *la plus générale* de (10.67) puisse s'écrire en combinaison linéaire des fonctions propres n'a rien d'évident. Cette propriété repose sur la complétude du système $\{\psi_{\alpha}\}_{\alpha}$; par la suite, on verra à la réflexion que la complétude est assurée pour des raisons de cohérence interne de la théorie.

⁴⁰Tout comme dans un problème de Mécanique élémentaire, la résolution conduisant à une solution unique exige la donnée des valeurs initiales des coordonnées et des vitesses.

⁴¹Remarquer que, selon (10.75), $\Psi_{\alpha}(\vec{r}, t = 0) = \psi_{\alpha}(\vec{r})$.

⁴²Dans toute la suite, * désigne par défaut la quantité complexe conjuguée.

⁴³La masse, par exemple, on a alors une densité et un courant de masse.

avec :

$$\rho = \Psi\Psi^* , \quad \vec{j} = \frac{\hbar}{2im} [\Psi^* \vec{\nabla}\Psi - \Psi \vec{\nabla}\Psi^*] . \quad (10.82)$$

Fondamentalement, la densité $\rho = \Psi\Psi^*$ est donc conservée localement, au même titre que la densité de charge en Électricité. Il est facile de montrer, en raisonnant avec un état stationnaire, que la conservation de ρ impose à la grandeur E (c'est en fait l'énergie) d'être réelle. Ceci est un argument supplémentaire sur la voie de l'interprétation cohérente de la fonction d'onde. À la lumière de ce qui sera proposé par Born en 1927, on peut dire que Schrödinger était vraiment à deux doigts de formuler précisément l'interprétation retenue depuis lors.

Remarques

1. Il est intéressant de remarquer que si on pose $\Psi = \sqrt{\rho}e^{i\phi}$ (ϕ réel), le courant (10.82) s'exprime simplement comme :

$$\vec{j} = \frac{\hbar}{m} \rho \vec{\nabla}\phi , \quad (10.83)$$

On retiendra qu'une fonction d'onde réelle (ou ayant une phase constante dans l'espace) donne un courant nul.

2. Avec les grandeurs ρ et ϕ et en posant $\hbar\phi = S$ (se souvenir de la relation $E = h\nu$ plaquée sur l'analogie optico-mécanique de Hamilton), on peut récrire les équations de conservation, (10.81), et de Schrödinger, (10.67), sous les formes suivantes :

$$\frac{\partial\rho}{\partial t} + \text{div} \left(\frac{1}{m} \rho \vec{\nabla}S \right) = 0 , \quad (10.84)$$

$$\frac{\partial S}{\partial t} + \frac{1}{2m} (\vec{\nabla}S)^2 + V + \frac{\hbar^2}{4m} \left[\frac{(\vec{\nabla}\rho)^2}{2\rho^2} - \frac{\Delta\rho}{\rho} \right] = 0 , \quad (10.85)$$

Si on considère S comme l'action, ces deux équations peuvent être interprétées de façon "classique" ; la première exprime la conservation de la densité de particules ; la seconde est une équation du genre Hamilton-Jacobi, complétée par un terme de "potentiel quantique", corrigeant le V classique. Cette reformulation, séduisante, n'est pas pour autant satisfaisante : le principe de superposition linéaire joue un rôle fondamental en Mécanique Quantique (il fournit la règle de composition des amplitudes de probabilités sans laquelle les expériences fondamentales seraient incompréhensibles). Ce principe ne peut s'exprimer à l'aide de ρ et de S sans que l'on revienne à la fonction d'onde elle-même : on ne peut donc en fait se passer de cette dernière⁴⁴. De plus, d'autres aspects, comme les symétries de permutation pour un système de particules identiques, s'expriment très maladroitement en termes de ρ et S : la tentative de réinterpréter la théorie à l'aide de corrections quantiques conduit à une impasse, ce qui montre que la Mécanique Quantique doit être prise au sérieux dans son essence même.

10.5 Vitesse de groupe

Dans sa tentative de préciser le sens physique de Ψ , Schrödinger proposa que les aspects corpusculaires usuels observés pour les particules – donnant lieu à des effets localisés dans l'espace –, résultent d'une superposition linéaire de fonctions d'onde. Cette idée peut être l'objet de critiques, en particulier parce que les paquets d'ondes ont tendance à s'étaler au cours du temps (à l'exception notable de ceux attachés à l'oscillateur harmonique⁴⁵), alors qu'une particule est toujours susceptible de produire un effet localisé (un impact sur un écran, par exemple).

Quoi qu'il en soit, la notion de paquet d'ondes fait ici son apparition. Ce qui suit est d'abord un rappel de notions familières à propos des phénomènes oscillants, puis une transposition dans le cadre de la Mécanique Ondulatoire.

⁴⁴Quand on veut chasser le diable (la diabolique fonction d'onde [6] !), il rentre par la fenêtre.

⁴⁵Heisenberg a montré que cette exception résultait précisément du fait que l'oscillateur harmonique est un système "trop" simple ; étant caractérisé par une seule fréquence (et ses harmoniques), traduisant le fait que tous les niveaux d'énergie sont équidistants, toute superposition linéaire possède une évolution périodique en temps (c'est une série de Fourier standard).

Il est bien connu que toute fonction suffisamment régulière peut être écrite sous la forme d'une intégrale de Fourier. Pour fixer les idées, on considère d'abord le cas d'une fonction $f(\vec{r})$ dépendant des trois coordonnées d'espace. La transformée de Fourier de f est⁴⁶ :

$$f(\vec{r}) = (2\pi)^{-3/2} \int_{\mathbb{R}^3} d^3k F(\vec{k}) e^{i\vec{k}\cdot\vec{r}} . \quad (10.86)$$

La fonction F est la *représentation* de Fourier de la fonction f ; \vec{k} est appelé vecteur d'onde. La relation inverse est :

$$F(\vec{k}) = (2\pi)^{-3/2} \int_{\mathbb{R}^3} d^3r f(\vec{r}) e^{-i\vec{k}\cdot\vec{r}} . \quad (10.87)$$

Si f est à valeurs réelles, on a la symétrie :

$$F(-\vec{k}) = F^*(\vec{k}) . \quad (10.88)$$

La même analyse peut être faite pour une fonction de la variable temps, $h(t)$. On écrit alors :

$$h(t) = (2\pi)^{-1/2} \int_{\mathbb{R}} d\omega H(\omega) e^{-i\omega t} \iff H(\omega) = (2\pi)^{-1/2} \int_{\mathbb{R}} dt h(t) e^{i\omega t} . \quad (10.89)$$

On peut enfin combiner temps et espace et, pour une fonction dépendant de \vec{r} et t , écrire :

$$g(\vec{r}, t) = (2\pi)^{-2} \int_{\mathbb{R}^4} d^3k d\omega G(\vec{k}, \omega) e^{i(\vec{k}\cdot\vec{r} - \omega t)} \quad (10.90)$$

et :

$$G(\vec{k}, \omega) = (2\pi)^{-2} \int_{\mathbb{R}^4} d^3r dt g(\vec{r}, t) e^{-i(\vec{k}\cdot\vec{r} - \omega t)} . \quad (10.91)$$

Les fonctions rencontrées en Physique satisfont souvent une équation aux dérivées partielles ; alors, leurs caractéristiques de Fourier \vec{k} et ω ne sont pas indépendantes mais liées par une relation appelée loi de dispersion. Considérons par exemple l'équation de propagation :

$$\frac{1}{c^2} \frac{\partial^2 \bullet}{\partial t^2} - \Delta \bullet = 0 . \quad (10.92)$$

Cette équation a pour solutions (particulières) des ondes planes $e^{i(\vec{k}\cdot\vec{r} - \omega t)}$, à la condition expresse que l'égalité suivante soit vérifiée :

$$-\frac{1}{c^2} \omega^2 + \vec{k}^2 = 0 , \quad (10.93)$$

soit, en terme des modules :

$$\omega = kc . \quad (10.94)$$

Si c est la vitesse de la lumière, (10.94) n'est autre que la relation de dispersion des ondes électromagnétiques dans le vide. Elle est remarquablement simple (proportionnalité entre ω et k : le vide n'est pas dispersif, ce qui signifie que les paquets d'ondes ne s'y déforment pas ("ne se dispersent pas") en se propageant. Ceci est vrai de tout milieu où la relation de dispersion entre fréquence et vecteur d'onde est linéaire ; la grandeur assurant la proportionnalité s'appelle par définition vitesse de phase. Dans un milieu non dispersif, la vitesse de phase est indépendante de la fréquence : toutes les composantes de Fourier (différant par leur fréquence) avancent donc à la même vitesse de sorte que le profil spatial de l'onde ne varie pas lors de la propagation⁴⁷, lorsque le temps s'écoule : le paquet d'ondes est invariant en forme au cours du temps.

⁴⁶Le signe dans le facteur de phase est purement conventionnel, mais conditionne la suite de relations intégrales.

⁴⁷Cette propriété se voit d'ailleurs directement sur l'équation (10.92) ; un petit calcul montre que toutes les solutions g à une dimension satisfont $g(x, t) = g(x - ct, t = 0)$: le profil à l'instant t est simplement le profil à l'instant zéro, translaté (déplacé en bloc) de la longueur ct . Bien évidemment, cette propriété est perdue pour un milieu d'indice $n(k)$, pour lequel la relation de dispersion prend la forme $\omega = kc/n(k)$ et n'est plus linéaire en k .

Une équation linéaire telle que (10.92) admet encore pour solution toute combinaison linéaire de solutions particulières ; il en résulte que, $\forall G$, la fonction g définie en (10.90) obéit à (10.92). Toutefois, comme tout ceci ne tient que si la relation (10.93) est satisfaite (ω et \vec{k} ne sont pas indépendants), l'ordre de l'intégrale passe de 4 à 3 et on a très précisément :

$$g(\vec{r}, t) = (2\pi)^{-3/2} \int_{\mathbb{R}^3} d^3k G(\vec{k}, \omega(\vec{k})) e^{i(\vec{k} \cdot \vec{r} - \omega(k)t)} \equiv (2\pi)^{-3/2} \int_{\mathbb{R}^3} d^3k A(\vec{k}) e^{i(\vec{k} \cdot \vec{r} - \omega(k)t)} \quad (10.95)$$

où $\omega(k)$ est la fonction représentant la loi de dispersion. A une dimension d'espace, on a :

$$g(x, t) = (2\pi)^{-1/2} \int_{\mathbb{R}} dk G(k, \omega(k)) e^{i(kx - \omega(k)t)} \equiv (2\pi)^{-1/2} \int_{\mathbb{R}} dk A(k) e^{i(kx - \omega(k)t)} . \quad (10.96)$$

Les expressions (10.95) ou (10.96) sont des paquets d'ondes : ce sont des empilements d'ondes planes pondérées par les coefficients donnés par les valeurs successives de l'amplitude $A(k)$.

Pour introduire la vitesse de groupe simplement, on se place dans la suite à une seule dimension d'espace. Soit donc une fonction $f(x, t)$ définie comme :

$$f(x, t) = (2\pi)^{-1/2} \int_{\mathbb{R}} dk A(k) e^{i(kx - \omega(k)t)} , \quad (10.97)$$

la fonction $\omega(k)$ étant préalablement déterminée par considération de l'équation dynamique supposée satisfaite par f . La condition initiale s'écrit :

$$f_0(x) = (2\pi)^{-1/2} \int_{\mathbb{R}} dk A(k) e^{ikx} \equiv f(x, t=0) , \quad (10.98)$$

où $f_0(x)$ est une fonction connue dont la descendante à t est précisément $f(x, t)$. Il en résulte, par inversion de Fourier :

$$A(k) = (2\pi)^{-1/2} \int_{\mathbb{R}} dx' f_0(x') e^{-ikx'} , \quad (10.99)$$

ce qui donne la fonction inconnue $A(k)$. En définitive, la solution issue de $f_0(x)$ a pour expression :

$$f(x, t) = (2\pi)^{-1} \int_{\mathbb{R}} dk \int_{\mathbb{R}} dx' f_0(x') e^{i[k(x-x') - \omega(k)t]} . \quad (10.100)$$

Un paquet d'ondes peut évidemment avoir une forme quelconque. Il en existe dont le profil est une fonction à une seule bosse bien localisée dans l'espace : ce sont eux dont on espère qu'ils vont reproduire des effets localisés, réconciliant de ce fait la notion d'onde avec celle de corpuscule. Dans la suite, on suppose donc que $f_0(x)$ possède cette propriété et s'étend principalement sur un intervalle donné Δx . En vertu des propriétés de réciprocity des transformées de Fourier, $A(k)$ est une fonction présentant un maximum principal en une certaine valeur k_0 et sensiblement non nulle sur un intervalle $\Delta k \sim 1/\Delta x$ entourant k_0 . En examinant l'expression (10.97), on voit que, pour un t donné, $f(x, t)$ sera maximum au voisinage du point x où la phase est *stationnaire* ; c'est en effet lorsque la phase varie lentement – et c'est bien le cas près d'un *extremum* – que les interférences seront constructives. Par ailleurs, ce qui compte pour l'intégrale en k , ce sont les valeurs de k autour de k_0 et s'étendant en gros sur Δk , puisque c'est l'amplitude $A(k)$ qui pondère les différentes ondes, donnant une importance primordiale à celles de vecteur d'onde $k \simeq k_0$ à Δk près. L'abscisse x où le maximum de f survient à un instant donné, soit x_{\max} , est donc celui pour lequel la phase $kx - \omega(k)t$ est extrémale si $k = k_0$ (c'est bien lorsque à la fois la phase et l'amplitude varient lentement que les interférences sont constructives). Analytiquement, ceci s'écrit en annulant la dérivée par rapport à k de cette quantité :

$$\left[\frac{d}{dk} (kx_{\max} - \omega t) \right]_{k=k_0} = 0 \iff x_{\max}(t) = \left(\frac{d\omega}{dk} \right)_{k_0} t . \quad (10.101)$$

Cette équation⁴⁸ montre que le centre du paquet a un mouvement uniforme, de vitesse $(d\omega/dk)_{k=k_0}$, appelée vitesse de groupe et notée v_g :

$$v_g = \left(\frac{d\omega}{dk} \right)_{k=k_0} . \quad (10.102)$$

⁴⁸Ce résultat peut aussi être obtenu en écrivant l'intégrand de (10.97) sous la forme $e^{i[kx - \omega(k)t - i \ln A(k)]}$. La stationnarité de la phase s'écrit alors $x - \omega' t - i A'/A = 0$, qui est vrai si à la fois $x = \omega' t$ et si $A' = 0$.

v_g ne coïncide pas avec la vitesse de phase $v_\phi = E/p = \omega/k$, sauf si la relation de dispersion est linéaire, de la forme (10.94). Ces deux vitesses ont une interprétation graphique simple : la courbe représentant la fonction $\omega(k)$ étant tracée, v_g est donnée par la pente de la tangente au point d'abscisse k , alors que v_ϕ est la pente de la corde joignant l'origine à ce point.

Via les relations de Planck - Einstein et de de Broglie, la vitesse de groupe coïncide formellement⁴⁹ avec la vitesse de la particule. En effet :

$$v_g = \frac{d\omega}{dk} = \frac{d(\hbar\omega)}{d(\hbar k)} \equiv \frac{dE}{dp} ; \quad (10.103)$$

maintenant, que l'on prenne $E = p^2/(2m)$ ou $E = \sqrt{p^2c^2 + m^2c^4}$, on trouve $v_g = v$.

10.6 Diffraction des particules matérielles

A partir du moment où une onde est associée à toute particule, on se doute que tous les phénomènes essentiellement ondulatoires (diffraction, interférences, ...) observés avec le rayonnement électromagnétique vont également survenir avec un faisceau de particules matérielles. Cette idée, révolutionnaire, a été vérifiée expérimentalement pour la première fois par Davisson et Germer en 1927, qui ont observé la diffraction de Bragg⁵⁰ en envoyant sur un cristal de nickel un faisceau d'électrons d'énergie appropriée. Pour un électron, et en utilisant $E = p^2/(2m)$, la relation de de Broglie prend la forme spécifique :

$$\lambda_{\text{\AA}} \simeq \frac{12.26}{\sqrt{E_{\text{eV}}}} . \quad (10.104)$$

Pour observer la diffraction de Bragg, il faut que λ soit du même ordre de grandeur que les paramètres de maille, soit $\lambda \sim 1 \text{ \AA}$ pour la matière ordinaire (de basse énergie) ; l'application numérique montre que E doit être de quelques dizaines d'eV.

De fait, Davisson et Germer, prenant précisément $E = 54 \text{ eV}$ (soit $\lambda \simeq 1.67 \text{ \AA}$) mirent en évidence, en fonction de l'angle θ d'observation, plusieurs maxima d'intensité très fins (très bien résolus en angle). Le premier survient à $\theta = 50^\circ$, ce qui permet de trouver la distance réticulaire correspondante, $d = \lambda/\sin 50^\circ \simeq 2.18 \text{ \AA}$. C'est précisément la valeur que l'on connaissait d'expériences antérieures conventionnelles, effectuées avec des rayons X.

Cette expérience démontra la possibilité d'avoir des interférences avec des faisceaux de particules matérielles. La preuve expérimentale de la "dualité onde-corpuscule" était ainsi donnée et découvrait un champ immense de possibilités. On pouvait dès lors envisager d'abord de substituer aux rayons X des faisceaux électroniques de basse énergie ; le simple ajustement de la ddp du canon à électrons permet de produire le faisceau ayant la bonne longueur d'onde pour une expérience donnée. De plus, ce qui était vrai pour les électrons devait l'être également pour toutes les particules matérielles (en 1929, Estermann et Stern montrèrent que les atomes d'hélium et des molécules d'hydrogène étaient également diffractés suivant la règle de Bragg) ; puisque la longueur d'onde associée varie à énergie donnée comme l'inverse de la racine carrée de la masse, on peut envisager de couvrir un intervalle quasiment illimité de longueurs d'onde. Par exemple, des protons d'énergie 100 eV ont une longueur d'onde associée voisine de 0.008 Å, les neutrons thermiques eux redonnent des longueurs d'onde de l'ordre de 1 Å ; à l'inverse, revenant aux électrons mais les choisissant relativistes, on obtient des longueurs d'onde commensurables avec la dimension nucléaire, autorisant l'étude détaillée de la structure nucléaire (voir Ch. 9).

10.7 Limite classique

Il est utile d'examiner ce que deviennent les quantités introduites en Mécanique Ondulatoire lorsqu'il s'agit non plus de particules mais d'objets macroscopiques, pour lesquels la Mécanique Classique est valide, cela ne fait

⁴⁹Il s'agit bien d'une identité formelle. Les relations (10.103) n'acquiescent de contenu physique précis que si on se rapporte au centre du paquet, la quantité v étant plutôt une valeur moyenne puisque, en Mécanique Quantique, les grandeurs dynamiques sont représentées par des opérateurs – les valeurs physiques résultant d'une certaine opération de moyenne. D'ailleurs, la même remarque vaut pour la relation de de Broglie : dans $\lambda = h/(mv)$, v désigne une vitesse moyenne, à préciser dans chaque cas ...

⁵⁰bien connue avec les rayons X.

aucun doute. En premier lieu, on peut s'intéresser à la longueur d'onde associée à un tel objet.

10.7.1 Longueur d'onde pour un objet macroscopique

Soit un grain de poussière, de dimension typique $l \simeq 1 \mu$. Si sa masse volumique est $\sim 1 \text{ g/cm}^3$, sa masse M est d'ordre 10^{-15} kg et varie comme l^3 . Ce grain de poussière est supposé en équilibre thermique avec le milieu : sa vitesse est donc de l'ordre de $\sqrt{3k_B T/M}$. Au total :

$$\lambda \sim \frac{h}{\sqrt{3Mk_B T}} . \quad (10.105)$$

L'application numérique à l'ambiante donne $\lambda \sim 1 \text{ F}$. Cette longueur d'onde est incroyablement petite à l'échelle macroscopique (c'est l'ordre de grandeur de la dimension du noyau !) et peut donc être considérée comme quasi-nulle. On se retrouve alors dans la situation analogue à la limite de l'Optique géométrique et la mécanique ordinaire est évidemment applicable. Ceci est *a fortiori* vrai pour les objets plus gros, la longueur d'onde associée variant finalement comme $l^{-3/2}$, toutes choses égales par ailleurs. D'un autre côté, l'abaissement de la température fait croître λ , toutes choses égales par ailleurs : les effets quantiques peuvent toujours se manifester à condition de se placer à température ultra-basse.

On peut objecter que, comme v figure au dénominateur de la formule de de Broglie, il suffit, pour une particule de masse donnée, de lui attribuer une vitesse assez petite pour obtenir une longueur d'onde plus grande, ou en tout cas commensurable avec toute valeur prescrite à l'avance. Ceci est indiscutable, mais l'argument – qui tend à conduire vers une incohérence (une particule macroscopique de très faible vitesse serait quand même quantique) – ne résiste pas à l'application numérique. Reprenons le grain de poussière, sans faire référence à un milieu extérieur à la température T (pourtant inévitable⁵¹) pour estimer sa vitesse v ; celle-ci est donc reliée à λ par :

$$v = \frac{h}{M\lambda} \sim \frac{6.6 \times 10^{-34}}{10^{-15} \lambda_m} \sim \frac{10^{-18}}{\lambda_m} \text{ m/s} . \quad (10.106)$$

Si λ est très petit devant la dimension du grain – seule autre longueur disponible – il est clair que la Mécanique ordinaire suffit. Inversement, on attend (?) des effets quantiques lorsque $\lambda \sim l$ ou moins ; il faut donc que la vitesse soit inférieure ou de l'ordre de :

$$v \sim \frac{10^{-18}}{10^{-6}} \text{ m/s} = 10^{-12} \text{ m/s} = 10^{-2} \text{ \AA/s} . \quad (10.107)$$

Une telle vitesse n'a pas de sens pour un objet macroscopique – ne serait-ce que parce que sa rugosité de surface n'est certainement pas mieux définie qu'à 1 \AA près.

10.7.2 Limite classique de la fonction d'onde

La fonction d'onde est l'objet quantique par excellence, dont il vaut la peine d'examiner la limite classique. Par simplicité, on considère une particule de masse m dans une dimension spatiale ; pour un état (stationnaire) d'énergie E , le point de départ est l'équation aux valeurs propres (10.63). Quand le potentiel V est constant, (on peut le prendre nul), les solutions sont des ondes planes de la forme e^{ikx} dont le vecteur d'onde est relié à l'énergie par :

$$E = \frac{\hbar^2 k^2}{2m} . \quad (10.108)$$

Quand le potentiel n'est plus constant, mais varie lentement (en un sens à préciser), on se trouve dans des conditions où une approche quasi-classique devrait constituer une bonne approximation. Indépendamment de cette hypothèse d'ailleurs, il est toujours possible d'essayer une solution du genre $e^{i\phi(x)}$, ce qui revient simplement à effectuer un changement de fonction inconnue ($\phi(x)$ se réduit à une fonction linéaire de x dans le cas $V \equiv 0$) ;

⁵¹ ... sauf à considérer un grain unique dans l'Univers à température nulle, situation pour le moins extrême, où le caractère spécieux de l'objection saute aux yeux ... ; imaginer mettre le grain dans une enceinte vide pour le soustraire à l'agitation thermique de l'atmosphère ambiante est un leurre : dans l'enceinte "vide", il y a du rayonnement thermique !

si V varie lentement, on peut espérer obtenir ϕ par approximations successives, en commençant par négliger sa dérivée seconde, puis à l'incorporer dans un terme correctif. En posant donc :

$$\psi(x) = e^{i\phi(x)} , \quad p(x) = \sqrt{2m[E - V(x)]} \equiv \hbar k(x) , \quad (10.109)$$

l'équation aux valeurs propres (10.63) se transforme en une équation (non-linéaire) pour ϕ :

$$i\phi'' - \phi'^2 + [k(x)]^2 = 0 \iff \phi'^2 = [k(x)]^2 + i\phi'' . \quad (10.110)$$

La quantité $p(x)$ n'est autre que l'impulsion classique de la particule, exprimée à l'aide de la coordonnée x en utilisant la constante du mouvement E . Bien sûr, classiquement, cette quantité est positive : c'est donc ainsi qu'il faut interpréter le radical dans (10.109), dans la région où $E > V(x)$, c'est-à-dire dans la région où, classiquement, la particule peut seulement se trouver⁵². Inversement, en Mécanique Quantique, rien *a priori* n'empêche E d'être plus petit que $V(x)$ et il convient donc d'étendre le domaine de définition de la fonction qui coïncide avec $p(x)$ dans la région classique. On définit donc d'une façon générale une fonction $p(x)$, qui donne l'impulsion classique quand $E > V(x)$ (alors $p \in \mathbb{R}_+$) et qui est prolongée par continuité dans la région inaccessible classiquement, où elle est imaginaire pure ; il en va de même pour $k(x)$. Le cas échéant, on utilisera la fonction réelle positive $\kappa(x)$ telle que :

$$-i\hbar\kappa(x) = \sqrt{2m[E - V(x)]} \quad (E < V(x)) . \quad (10.111)$$

Le schéma itératif d'approximation part de l'idée que quand $V = 0$, la dérivée seconde ϕ'' est identiquement nulle. Avec un potentiel lentement variable, l'approximation la plus brutale consiste à oublier ϕ'' dans (10.110) ; ceci produit l'approximation ϕ_0 pour ϕ :

$$\phi_0'^2 = [k(x)]^2 \iff \phi_0(x) = \pm \int k(x) dx + C_0 \quad (10.112)$$

où C_0 est une constante. L'approximation suivante consiste à remplacer 0 par ϕ_0'' au second membre de (10.110), ce qui donne une nouvelle équation pour une fonction ϕ_1 :

$$\phi_1'^2 = [k(x)]^2 + i\phi_0'' . \quad (10.113)$$

La solution est, compte tenu de (10.112) :

$$\phi_1 = \pm \int \sqrt{k^2(x) + i\phi_0''} dx + C_1 = \pm \int \sqrt{k^2(x) \pm ik'(x)} dx + C_1 . \quad (10.114)$$

Et ainsi de suite. Les ϕ_0, ϕ_1, \dots constituent les approximations successives de la fonction ϕ cherchée ; si on s'arrête à ϕ_1 , on a donc $\phi(x) \simeq \phi_1(x)$. Dans l'expression (10.114), les signes sont solidaires (+ va avec +, - avec -).

Si l'on s'en tient à ϕ_1 , ceci implique que le rapport $|k'/k^2|$ est très petit devant 1 : garder la racine carrée n'aurait donc pas grand sens. Dans ces conditions, le moins que l'on puisse faire est un développement limité du radical ; à l'ordre le plus bas, on écrit donc :

$$\phi_1 \simeq \pm \int k(x) \left(1 + \frac{ik'(x)}{2k^2(x)} \right) dx + C_1 . \quad (10.115)$$

En effectuant l'intégration venant du second terme, on trouve :

$$\phi_1 \simeq \pm \int k(x) dx + \frac{i}{2} \ln \frac{k(x)}{K} , \quad (10.116)$$

où K est une constante d'intégration incorporant C_1 . Revenant à ψ par (10.109), on obtient les deux solutions :

$$\psi(x) \simeq \psi_{\pm}(x) \equiv \left[\frac{\hbar K}{p(x)} \right]^{1/2} e^{\pm i\Phi(x)} , \quad \Phi(x) = \hbar^{-1} \int p(x) dx . \quad (10.117)$$

⁵²Dit autrement : une particule classique ne peut se trouver dans une région où, par la constance de l'énergie, elle aurait une énergie cinétique négative.

Cette expression est valide si $|k'/k^2| \ll 1$, soit⁵³ :

$$|p'(x)| \ll \hbar^{-1}|p^2(x)| \iff \left| \frac{\hbar}{p} \right| \ll |p(x)| \iff |\lambda(x)p'(x)| \ll |p(x)| . \quad (10.118)$$

p' représente le taux de variation de p , piloté par celui du potentiel $V(x)$. En définitive, la dernière écriture signifie que le potentiel doit varier lentement à l'échelle de la longueur d'onde.

On dispose maintenant de deux solutions ψ_{\pm} ; la solution la plus générale de (10.63) (équation linéaire) est donc :

$$\psi(x) \simeq A_+ \psi_+(x) + A_- \psi_-(x) , \quad (10.119)$$

où les A_{\pm} sont des constantes à trouver. C'est ici que les difficultés techniques commencent réellement ([7], Ch. 7 et notamment p. 122).

En effet, considérons un point d'abscisse x_0 – dit point tournant – où la quantité $E - V(x)$ s'annule :

$$E - V(x_0) = 0 , \quad (10.120)$$

et, pour fixer les idées, supposons que dans le voisinage de ce point, on a :

$$E - V(x_0) \begin{cases} > 0 & \text{si } x < x_0 \\ < 0 & \text{si } x > x_0 \end{cases} . \quad (10.121)$$

La région de droite ($x > x_0$) est donc inaccessible classiquement (l'énergie cinétique serait négative) : un point tournant est un point où la vitesse classique change de sens, la particule classique faisant demi-tour. Il est clair que la rapidité du changement de sens de la vitesse est d'autant plus grande que le potentiel varie vite en x_0 : pour une paroi totalement réfléchissante, la vitesse classique passe instantanément de $+v_0$ à $-v_0$.

D'une façon générale, la condition de validité (10.118) est donc certainement violée près d'un point tournant ; ceci veut dire qu'une expression telle que (10.119) n'a de sens que *loin* d'un tel point, donc loin de x_0 , à gauche ou à droite. Dans chacune des régions ainsi précisées, l'expression (10.119) est correcte ; on peut donc de fait retenir :

$$\psi(x) \simeq \begin{cases} A_+ \psi_+(x) + A_- \psi_-(x) & \text{si } x \ll x_0 \\ A'_+ \psi_+(x) + A'_- \psi_-(x) & \text{si } x \gg x_0 \end{cases} . \quad (10.122)$$

Il y a maintenant 4 constantes à trouver ; elles ne sont pas indépendantes les unes des autres : il faut bien comprendre que les deux expressions (10.122) sont deux formes approchées d'une *seule et même fonction* : si on connaissait celle-ci, son développement à grand $|x|$ fournirait toutes les constantes. Pour relier explicitement les A_{\pm} aux A'_{\pm} , il faut trouver des conditions de raccordement ... qui ne peuvent être écrites en x_0 en considérant seulement les expressions (10.122), puisque celles-ci ne sont valables que *loin* de x_0 !

La clé consiste à contourner l'obstacle et, au lieu de rester sur l'axe réel de l'abscisse x , à se promener dans le plan complexe en contournant le point x_0 de suffisamment loin. Alors, moyennant un calcul de mathématiques appliquées très subtil (où les traquenards ne manquent pas), il est de fait possible d'établir les relations existant entre les A_{\pm} et les A'_{\pm} .

Ceci étant fait, on obtient la limite dite quasi-classique⁵⁴ de la fonction d'onde pour le problème considéré. Le point énigmatique est de trouver une fonction d'onde parfaitement définie dans la région inaccessible classiquement : quel peut en être le sens ? Dans cette région, la fonction d'onde ψ est petite (elle contient une exponentielle réelle, qui ne peut être physiquement qu'une exponentielle décroissante, du genre $e^{-\kappa|x|}$), mais elle n'est pas strictement nulle (dans un langage ondulatoire, on peut parler d'onde exponentiellement amortie). On verra par la suite que ce fait traduit un effet spectaculaire, strictement quantique, appelé effet-tunnel : la particule passe de l'autre côté de la montagne sans passer par dessus.

Terminons par l'expression complète de la fonction d'onde $\Psi(x, t)$, telle qu'elle ressort de l'approximation quasi-classique. En ne considérant que la solution ψ_+ , on a (voir (10.117)) :

$$\Psi(x, t) \simeq \left[\frac{\hbar K}{p(x)} \right]^{1/2} e^{i\Phi(x)} e^{-i\hbar^{-1}Et} . \quad (10.123)$$

⁵³Il convient de garder les modules : p peut fort bien être complexe.

⁵⁴On dit aussi : approximation WKB, pour Wentzel - Kramers - Brillouin (ou BKW dans les ouvrages en français !)

En reconnaissant l'action classique $S = \int p \, dq - Et$, ceci s'écrit aussi :

$$\Psi(x, t) \simeq \left[\frac{\hbar K}{p(x)} \right]^{1/2} \exp \left[\frac{i}{\hbar} S(x, t) \right] . \quad (10.124)$$

À partir de cette expression, le courant j défini en (10.82) ressort comme :

$$j = \rho(x) \frac{1}{m} \frac{\partial S}{\partial x} \quad \text{avec} \quad \rho = \Psi \Psi^* . \quad (10.125)$$

Puisque $\partial S / \partial x = p = mv$, on trouve finalement :

$$j = \rho(x) v(x) . \quad (10.126)$$

A nouveau, cette expression montre que $\rho = \Psi \Psi^*$ joue bien le rôle d'une densité de particules. Cette idée est le fondement du postulat de Born (1927), énoncé dans le chapitre suivant.

Chapitre 11

Fonction d'onde

La fonction d'onde Ψ est manifestement la quantité centrale de la Mécanique Ondulatoire de Schrödinger. Jusqu'à présent, cette fonction n'a pas encore reçu d'interprétation précise bien que Schrödinger lui-même ait commencé à poser quelques jalons sur la voie menant au sens et à l'utilisation de Ψ en tant qu'outil de prévision et de description du monde microscopique. Le but de ce chapitre est de donner l'interprétation usuellement acceptée¹, construite peu à peu entre 1926 et 1930 grâce aux travaux de Born [9], Heisenberg [10] et Bohr [11]. Cette théorie est non-relativiste. La généralisation relativiste a été initiée par Dirac [12]. Auparavant, pour préparer le terrain, il convient de discuter en détails certaines conséquences des idées admises jusqu'à présent ; on verra comment celles-ci imposent une révision complète de certains concepts considérés inébranlables par la Physique Classique.

11.1 Fentes d'Young

L'expérience des fentes d'Young est bien connue en Optique : c'est elle qui permet la mise en évidence des interférences lumineuses. Comme déjà noté, l'association d'une longueur d'onde à une particule matérielle permet de prévoir l'existence de phénomènes typiquement ondulatoires avec ces particules, un fait confirmé par Davisson et Germer en 1927 (voir chapitre précédent).

La présentation suivante est une analyse des interférences produites avec des particules qui, pour fixer les idées, seront des électrons. Elle se réfère à une expérience dont la mise en œuvre effective est bien plus complexe que la description schématique qui va en être donnée ne pourrait le faire croire. Ici, le but est de décrire une expérience mentale : c'est une expérience, en principe réalisable, destinée à illustrer des concepts établis afin d'édifier – éventuellement par essai et erreur – une nouvelle doctrine théorique.

Le principe du dispositif est représenté sur la figure 11.1. Des électrons accélérés par une ddp U passent à travers une grande ouverture S en direction d'un plan (plaque) percé(e) de deux trous (fentes) séparé(e)s par la distance d et numérotées 1 et 2 pour la commodité. La longueur d'onde associée aux électrons après leur accélération par la ddp U est :

$$\lambda = \frac{h}{\sqrt{2m|e|U}} . \quad (11.1)$$

Les franges d'interférences sont nettement visibles si $\lambda \simeq d$, ce qui peut être réalisé en ajustant U . Les électrons sont détectés dans un plan P situé à la distance L de la plaque ; un compteur mobile permet de savoir combien d'électrons arrivent à une abscisse x donnée.

¹Le débat sur l'interprétation de la Mécanique Quantique est de nos jours moins passionné qu'il ne l'a été. Devant les impressionnants succès de cette théorie, et l'effet révolutionnaire étant émoussé par l'habitude, l'immense majorité des physiciens a adopté une attitude pragmatique qui peut surprendre quand on en revient à quelques interrogations majeures sur les fondements mêmes de la Mécanique Quantique (selon Penrose [8], la Mécanique Quantique est "*la plus exacte et la plus mystérieuse de toutes les théories physiques.*"). Ce pragmatisme n'a rien d'épistologiquement nouveau : aucun physicien ne croit qu'une théorie donnée est un achèvement dans l'absolu, mais s'y réfère en confiance tant qu'aucune *expérience* ne vient l'ébranler sur ses bases. C'est le statut actuel de la Mécanique Quantique.

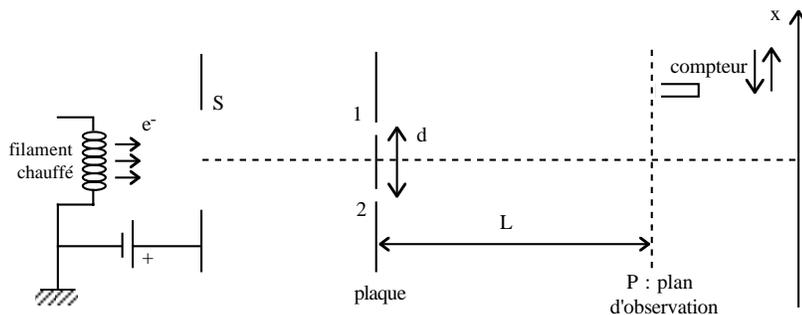


Figure 11.1: Schéma de l'expérience d'Young avec des électrons.

À condition d'utiliser un compteur suffisamment sensible et une source pas trop intense, le “courant” mesuré dans le plan d'observation n'est pas continu, mais est caractéristique d'une “pluie” de particules : le détecteur donne une suite de coups séparés dans le temps, entre lesquels aucun *bip* ne se produit. C'est bien pour cela que l'on parle par habitude de *particules*, sous-entendant par là des objets donnant lieu à des manifestations localisées, finalement assimilés à des points matériels. D'un autre côté, rien n'interdit d'utiliser une source suffisamment intense pour que le courant mesuré soit quasi constant – à des petites fluctuations près.

On sait bien ce que produit l'expérience ci-dessus si, au lieu d'une source d'électrons, on dispose d'une source lumineuse banale² et monochromatique, de longueur d'onde λ_0 . L'intensité lumineuse $I(x)$ possède un maximum central en $x = 0$ et des maxima secondaires décalés en angle de $2\lambda_0/d$ (voir 11.1). La figure vue sur l'écran est fixe dans le temps grâce à l'énormité du flux de photons. Bien évidemment, si l'on réduit considérablement l'intensité de la source, on va finir par mettre en évidence³ l'aspect granulaire de la lumière : on ne voit plus des franges permanentes mais une suite de petites taches lumineuses éphémères, *situées à un endroit ou un autre, apparemment au hasard* – la lumière est constituée de photons ! Bien sûr, à condition d'intégrer dans le temps sur un intervalle suffisamment long, la distribution des impacts finira par se confondre avec la distribution d'intensité de la figure d'interférences obtenue d'emblée avec une source conventionnelle.

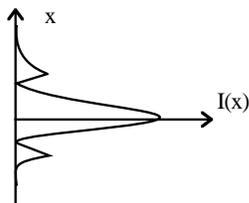


Figure 11.2: Représentation schématique de la figure d'interférences lumineuses.

Revenons à l'expérience conduite avec des électrons. Une longue exposition en un point d'abscisse x variable permet de reconstituer peu à peu le profil de distribution des impacts, $N(x)$. En divisant par le nombre total d'impacts, on obtient la distribution statistique des fréquences, $f(x)$, fonction qui tend vers un certain $P(x)$ à la limite d'un nombre “infini” de coups ; on a envie de considérer $P(x)$ comme la (densité) de probabilité des impacts en un point d'abscisse x : le produit $P(x)dx$ donne la proportion d'impacts reçus entre les points d'abscisses x et $x + dx$, normalisée l'unité. Le point remarquable est que cette fonction $P(x)$ est proportionnelle à la distribution d'intensité lumineuse quand on fait l'expérience avec de la lumière (voir figure (11.2)) :

$$P(x) \propto I(x) . \quad (11.2)$$

Si on est déjà persuadé de la “réalité” de l'onde associée de de Broglie, ce résultat ne devrait pas surprendre. Toutefois, pénétré de l'habitude mentale considérant les électrons “avant tout” comme des particules au sens

²On se souvient (voir l'effet photoélectrique ordinaire) qu'une source ordinaire donne un courant de photons gigantesque (à 1 m d'une lampe de 100 W émettant dans le jaune, on reçoit environ 10^{16} photons par seconde et par cm^2).

³à condition que le détecteur soit assez sensible pour mesurer un petit nombre de photons.

commun, on peut tenter de raisonner comme suit, adoptant *implicitement* une vision purement mécaniste. L'arrivée d'un électron en un point donné d'abscisse x est un événement résultant de l'un *ou* l'autre des deux événements élémentaires mutuellement exclusifs : l'électron passe par une fente *ou* passe par l'autre. Dans le même état d'esprit que ci-dessus, on devrait donc pouvoir introduire deux probabilités P_1 et P_2 définies comme :

$$P_i(x) = \text{Probabilité d'arriver en } x \text{ sur l'écran après passage par la fente } i \text{ (} i = 1, 2 \text{)} \quad (11.3)$$

(il s'agit bien sûr de *densités* de probabilité). Dans la vision mécaniste ainsi adoptée, il n'y a pas d'autres possibilités ; le "ou" ci-dessus se traduit, en probabilités, par l'addition (la probabilité associée à l'occurrence d'un même événement final (l'arrivée au point x) en conséquence de deux événements mutuellement exclusifs est la somme des probabilités). Dans cette vision des choses, on doit donc *logiquement* écrire que l'arrivée en x est distribuée suivant la somme :

$$P_1(x) + P_2(x) . \quad (11.4)$$

A quoi ressemblent les fonctions $P_i(x)$? Naïvement, on s'attend à deux courbes en cloche comme indiqué à gauche sur la figure 11.3. Les ayant, il est facile de tracer leur somme (à droite) : elle n'a rien à voir avec $P(x)$!

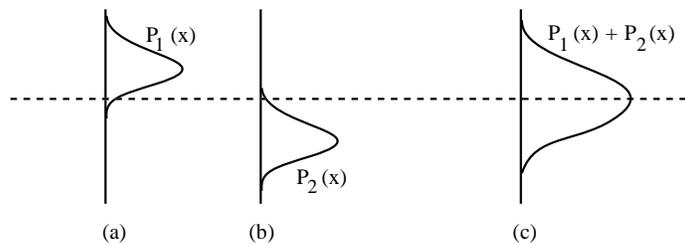


Figure 11.3: Probabilités de passer par un trou ou par l'autre ???

D'un autre côté, et pour en avoir le cœur net, on peut procéder comme suit, à titre de mise à l'épreuve de la validité des deux courbes P_1 et P_2 devinées intuitivement. On commence par fermer la fente 2 ; alors, les électrons ne peuvent passer que par 1 et on observe de fait sur l'écran une distribution du genre $P_1(x)$ ayant l'allure indiquée sur la partie (a) de la figure 11.3. Puis, on obture seulement la fente 1 ; les électrons ne peuvent passer que par 2 et on observe de fait la distribution $P_2(x)$ de la partie (b). L'addition graphique ces deux fonctions donne la courbe (c) ... qui n'a rien à voir avec $P(x)$, observée quand les deux trous sont *simultanément* ouverts ! Le paradoxe semble même atteindre son comble quand on réalise que, lorsque les deux fentes sont ouvertes, l'électron a plus de chemins possibles. Et pourtant, l'ouverture de la deuxième fente vient lui interdire d'arriver là où il le pouvait quand l'un des deux trous seulement est ouvert. Ceci est le fait quantique fondamental : *quand plusieurs potentialités se présentent, le résultat n'est pas la simple superposition des effets associés individuellement à chaque possibilité isolée.*

En définitive, quelque chose est faux dans le raisonnement ci-dessus ; il n'est pas question de remettre en cause la théorie des probabilités, qui est une théorie cohérente en soi, formulée indépendamment d'un champ d'application. L'erreur doit donc se tenir dans le *maniement* des probabilités pour la situation analysée ; il doit être faux d'additionner les probabilités comme on vient de le faire. Comme ce n'est pas la théorie des probabilités qui est en cause mais son usage, la seule conclusion logique est : lorsque les deux fentes sont ouvertes, il est faux de dire "l'électron est passé par une fente *ou* par l'autre".

On verra par la suite que c'est l'ensemble du raisonnement qui est incorrect : lorsque les deux fentes sont ouvertes, l'arrivée en un point x de l'écran est calculable en additionnant non des probabilités mais des *amplitudes de probabilités* (définies en temps utile par la suite), chacune étant associée à un "chemin" possible. Par ailleurs, on a essayé de corrélérer un événement (le passage par une fente) à un événement ultérieur (l'arrivée sur l'écran) ; la corrélation stricte n'est possible que dans la mesure où l'événement intermédiaire est unique : c'est bien le cas lorsqu'une seule une fente est ouverte et alors la distribution P_1 ou P_2 reflète bien le seul événement intermédiaire possible. Au contraire, lorsque les deux fentes sont ouvertes, la figure globale caractérisée par $P(x)$ ne permet pas de "remonter" à un événement intermédiaire : tous jouent un rôle comparable. Quand les deux fentes sont ouvertes, on ne peut pas dire si l'électron est passé par (1) ou par (2) au vu de la distribution $P(x)$. Ne pas pouvoir le dire n'est pas prendre position sur la réalité d'un tel événement ; cependant,

on imagine mal, si la trajectoire existe, qu'elle puisse dépendre de l'état respectif des fentes : si la trajectoire existe au sens commun et si l'électron passe par une fente, le même sens commun ne peut accepter l'idée que cette trajectoire est sensible au fait que l'autre fente est ouverte ou fermée.

La conclusion à ce stade se résume par la non-égalité suivante. Les deux fentes étant ouvertes, la distribution des impacts définie ci-dessus – identique à celle obtenue avec une lumière de même longueur d'onde, voir (11.2) – n'est pas égale à la somme $P_1 + P_2$:

$$P(x) \neq P_1(x) + P_2(x) . \quad (11.5)$$

Il s'agit d'un fait d'expérience, donc indiscutable ; seule son l'interprétation peut éventuellement être l'objet d'un débat. Il est important de noter que l'égalité $P = P_1 + P_2$ (qui est fautive !) a un côté quelque peu hybride : elle relie la probabilité $P(x)$ qui correspond à une situation expérimentale donnée (les deux fentes ouvertes) aux probabilités P_1 et P_2 qui correspondent à une autre situation expérimentale (une fente est fermée, l'autre est ouverte). Cette hybridité n'est pas suspecte s'il est possible de suivre pas à pas l'évolution spatio-temporelle d'une particule. Elle ne l'est pas non plus si certains détails du protocole expérimental n'affectent pas les phénomènes eux-mêmes, *i. e.* si la mesure ne perturbe pas le système étudié. Tout ceci est vrai dans le cadre classique, ou plus précisément y produit des effets indécélables⁴. Manifestement, s'agissant d'électrons, ces affirmations ne tiennent plus. La conclusion spectaculaire de l'expérience d'Young est donc qu'il n'est pas possible d'attribuer une trajectoire à l'électron, sauf à sombrer dans l'absurdité⁵ : présupposer que celle-ci existe serait projeter pour la représentation du monde microscopique une image qui s'est imposée à une autre échelle – projection qui n'est pas à rejeter dans l'absolu en l'absence d'éléments contradictoires⁶ – mais qui se révèle ici précisément incompatible avec l'observation. Il en résulte qu'une particule n'est pas seulement un petit objet, très petit à l'échelle macroscopique, auquel on peut appliquer par extrapolation les concepts usuels de la Mécanique Classique : il existe bel et bien une spécificité quantique⁷ à prendre au sérieux. Comme la trajectoire au sens classique suppose la donnée simultanée de la coordonnée et de la vitesse – qui fixent à tout instant le point et la tangente de la courbe –, la négation de son existence contient en essence le Principe d'Incertitude de Heisenberg qui sera formulé plus précisément par la suite.

S'il n'est plus possible de parler de particules au sens usuel, il n'est pas non plus possible de parler seulement d'ondes : les électrons qui peu à peu construisent la figure d'interférences produisent *aussi* des impacts localisés sur l'écran. La seule conclusion acceptable à ce stade est la suivante : comme il est incorrect de parler exclusivement en termes de particules ou en termes d'ondes, il faut admettre que les objets microscopiques sont *les deux à la fois*. Cette double facette n'introduit aucune contradiction logique, elle est au contraire nécessaire pour réconcilier les deux aspects qui s'imposent au vu des résultats de l'expérience. Admettre ceci est l'une des difficultés dans l'abord de la Mécanique Quantique : la pensée classique a défini une sorte de carcan mental légitimé par d'autres situations, dont il faut s'affranchir⁸.

Ainsi émerge une nouvelle notion, historiquement appelée dualité onde-corpuscule. Diverses propositions ont été faites pour donner un nom spécifique aux petits objets de la Physique (“particlondes”, “quants”, “particules”, ...) ; aucun d'entre eux ne s'est vraiment imposé : on utilise toujours le terme “particule”, tout en gardant à l'esprit cette dualité, notamment parce que dans la limite classique un électron apparaît bel et bien comme un tout petit point matériel : si l'on fait l'expérience d'Young avec des électrons de trop haute énergie, leur longueur d'onde étant très petite devant la distance entre les fentes⁹, on se retrouve dans la limite classique (c'est la limite “géométrique”) et on voit sur l'écran deux petites taches constituant les images géométriques des deux fentes. Dans ces conditions, obturer une fente produit un effet quasi-imperceptible sur l'image de l'autre.

La relation $P \neq P_1 + P_2$ lorsque les deux trous sont ouverts ne constitue pas à proprement parler une conclusion : c'est un fait d'expérience. Sur un plan logique, sa conséquence est que l'on ne peut analyser la

⁴Pour observer la collision entre deux boules de billard, il faut les éclairer. Il est certain que la lumière transporte de l'impulsion et donc affecte la collision. Toutefois, seul un physicien ayant perdu toute raison s'amuserait (?) à inclure cet effet dans la description théorique.

⁵Du genre : la trajectoire existe, l'électron passe par 1 et elle dépend de l'état de la fente 2.

⁶De la même façon, on peut penser que la loi de Coulomb est vraie à toute échelle, aussi petite que l'on veut. Il s'agit, au sens strict, d'une extrapolation mentale ; elle est acceptable tant qu'aucune expérience ne vient la malmenner. Il semble d'ailleurs que Schrödinger envisageait une révolution encore plus radicale, remettant notamment en cause les notions de continuité de l'espace et du temps [13].

⁷On ne doit pas penser que la trajectoire continue à exister et que le problème se réduit à l'impossibilité de l'observer en pratique.

⁸Selon Schrödinger, nos habitudes mentales ont forgé des images traduites par les vocables ondes et corpuscules “*images que nous sommes obligés de garder toutes les deux parce que nous ne savons pas encore comment nous en débarrasser*” ([5], p. XIV).

⁹et la dimension linéaire de chacune d'entre elles.

dynamique de l'électron en raisonnant seulement en terme de trajectoire à travers le trou 1 ou à travers le trou 2. Il n'en reste pas moins que, ne doutant pas que quand une seule fente est ouverte, l'électron passe par elle, il est possible d'essayer d'en savoir plus ("où passe l'électron ?") en analysant pour cela l'expérience suivante.

L'électron est capable de diffuser la lumière : c'est l'effet Compton. Une telle diffusion est décrite précisément comme une collision élastique entre un électron et un photon. Donc, si on dispose une source lumineuse derrière le plan des deux trous, l'observation d'une diffusion lumineuse au voisinage d'un trou est le signe indiscutable de la présence d'un électron dans le voisinage de ce trou¹⁰. Le résultat d'une telle expérience est le suivant : on observe effectivement des diffusions de photons près de chaque trou ! Chacune d'entre elles révèle indiscutablement la présence d'un électron (figure 11.4). En compliquant les choses, on peut imaginer qu'un appareil mesure la charge qui passe : cette charge est bien entière, ce n'est pas un "demi" électron . . . Arrivé à ce stade, il semble qu'un paradoxe surgit. D'une part l'expérience des fentes d'Young conventionnelle (sans la source lumineuse) impose la conclusion : on ne peut affirmer que l'électron passe par un trou ou passe par l'autre, puisque cette affirmation implique (et réciproquement) que $P = P_1 + P_2$ or l'expérience révèle $P \neq P_1 + P_2$. D'autre part, l'expérience conduite avec la source lumineuse derrière l'écran révèle au contraire indiscutablement la matérialisation d'un électron juste à droite d'un trou (en comptant le nombre de diffusions des photons près de chaque trou et en en faisant la somme, on reconstitue d'ailleurs la somme $P_1 + P_2$ schématisée sur la partie (c) de la figure 11.3).

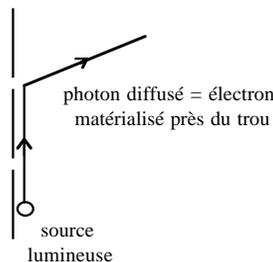


Figure 11.4: "Mesure" du trou où "passe un électron"

En somme, quand on met en évidence expérimentalement la fente près de laquelle se trouve un électron après passage de la plaque, on obtient¹¹ ($P(x)$ est toujours la distribution observée sur l'écran pour un type d'expérience donné) :

$$P(x) = P_1(x) + P_2(x) \quad (\text{avec observation de l'endroit où "passe l'électron"}) \quad (11.6)$$

alors que si on n'observe pas cette étape intermédiaire on a :

$$P(x) \neq P_1(x) + P_2(x) \quad (\text{sans observation de l'endroit où "passe l'électron"}) . \quad (11.7)$$

A la réflexion, ces deux résultats distincts peuvent se comprendre. Indéniablement, l'éclairage par de la lumière pour voir où passe l'électron perturbe l'évolution des phénomènes (le photon "tape" sur l'électron). Il faudrait donc pouvoir réduire cette perturbation au point d'en rendre les conséquences imperceptibles – mais c'est impossible, justement parce que la lumière est composée de photons indivisibles ! On pourrait envisager d'utiliser une lampe de faible puissance : réduire l'intensité de la lampe réduit en pratique le nombre de collisions, mais si une collision se produit, alors la perturbation est importante. Avec une source d'intensité très très faible, la figure d'interférences est en effet très peu modifiée, mais comme la probabilité pour qu'il se produise effectivement une diffusion est alors très faible, il ne s'en produit presque jamais et on ne peut donc presque jamais dire par quel trou est passé un électron. Globalement, la perturbation "de mesure" est bien devenue infiniment petite, mais elle ne fournit aucun résultat, aucune mesure à proprement parler, aucune réponse à la question posée : en réduisant à l'excès la perturbation de mesure, on rend celle-ci non-signifiante. Avec une source d'intensité moyenne, on observe une superposition des deux figures : la figure d'interférences résultant

¹⁰Tout électron traversant l'appareil ne donne pas forcément lieu à une diffusion Compton. Ici encore intervient un aspect probabiliste : la probabilité de diffusion Compton n'est pas égale à 1.

¹¹Les guillemets sont là pour rappeler que la diffusion Compton près d'un trou signe seulement la présence d'un électron en ce "point", rien de plus. Il ne faut pas pour autant affirmer que l'électron alors ainsi matérialisé est *passé* par le trou situé à proximité ; c'est la diffusion qui réduit le paquet d'ondes. Cette réduction est de nature irréversible : l'observation à un instant t ne permet aucune extrapolation concernant un instant antérieur à la réduction.

des impacts des électrons qui n'ont pas subi de collision avec un photon et les deux pics "non-cohérents" P_1 et P_2 associés aux électrons qui ont produit une diffusion Compton, les uns au voisinage d'une fente, les autres au voisinage de l'autre.

Enfin, on pourrait se dire que pour réduire la perturbation due à chaque collision il suffit d'utiliser des photons de grande longueur d'onde, puisque pour le photon $E = pc = hc/\lambda$. Mais il y a une limite-butoir : pour pouvoir distinguer les deux trous l'un de l'autre par la diffusion localisée du photon, il faut que la longueur d'onde utilisée soit plus petite que la distance entre les trous : pour distinguer des détails de dimension d (la distance entre les deux fentes) il faut $\lambda \leq d$. Par ailleurs, une source de longueur d'onde λ (la source est ici le point où la diffusion se produit) ne peut être localisée spatialement à une précision meilleure que λ . Il existe donc une borne supérieure pour la longueur d'onde des photons *utiles* et ceci fixe pour ces derniers une énergie *minimale* pour le but poursuivi. En définitive, la perturbation inévitable à chaque collision ne peut pas être rendue arbitrairement petite s'il s'agit d'effectuer une observation (mesure) signifiante. Ici apparaît une spécificité du monde quantique : toute expérience physique impliquant nécessairement un processus physique élémentaire, une inévitable et importante perturbation¹² se produit entre l'objet mesuré et la sonde physique utilisée pour la mesure elle-même.

Finalement, on voit que la différence cruciale entre les deux types d'expériences (11.6) et (11.7) est la notion de *mesure*. Pour l'expérience de type (11.6), on effectue de fait une *mesure de la position*, ce qui n'est pas le cas avec le protocole (11.7). En anticipant sur la suite : l'expérience (11.6) provoque une *réduction du paquet d'ondes*, ce que ne fait pas (11.7). Ce dernier processus sera rediscuté par la suite, lors de l'exposé des postulats de la Mécanique Quantique (Ch. 13) : il a été l'objet de vifs débats en raison des "paradoxes" auquel il semble conduire (le plus célèbre est celui dit du *Chat de Schrödinger*).

En conclusion de cette discussion de l'expérience d'Young, il est utile de rappeler les trois points suivants :

- il existe une incompatibilité entre les faits expérimentaux et l'attribution à une particule d'une trajectoire définie au sens classique¹³
- en général, toute opération de mesure s'accompagne d'une perturbation produisant un effet spectaculaire que l'on ne peut négliger
- la physique du monde microscopique ne peut être décrite que dans un cadre statistique, ou probabiliste.

11.2 Interprétation probabiliste de la fonction d'onde et conséquences

L'équation de Schrödinger ne se borne pas à fournir les énergies d'un système, mais produit également des fonctions propres, et plus généralement des combinaisons linéaires de fonctions propres, génériquement notées Ψ , dont il est temps de préciser le sens. En réalité, ces deux questions sont intimement mêlées : les énergies propres obtenues résultent du choix de conditions aux limites ; l'application de ces conditions fait le tri parmi toutes les solutions mathématiques de l'équation aux dérivées partielles ; seules restent les solutions "acceptables" d'un point de vue physique.

Or on ne peut pas déclarer qu'une solution donnée est acceptable ou non sans avoir préalablement dit ce que l'on entend par là, ce qui revient bien à avoir d'abord précisé le sens physique attribué à la fonction cherchée. Les premiers succès de l'équation aux valeurs propres par son aptitude à donner les bonnes énergies¹⁴ restent fragiles quant à leurs fondements théoriques tant que la boucle n'est pas bouclée, c'est-à-dire tant que l'on a pas élu rationnellement un critère de choix parmi toutes les solutions mathématiquement possibles. Après tout, l'Ancienne Théorie des Quanta a eu, elle aussi, ses succès ; la concordance entre les prévisions théoriques et les faits expérimentaux, obtenue à partir d'un ensemble de règles posées par affirmation, ne peut être pleinement

¹²Il existe toutefois des situations particulières où la mesure ne provoque aucune perturbation (par exemple : mesurer l'énergie quand le système est dans un état d'énergie donnée).

¹³Dans la reformulation à la Feynman de la Mécanique Quantique, on peut définir une trajectoire généralisée qui se démarque nettement – heureusement – de la définition classique.

¹⁴pour l'atome d'hydrogène et l'oscillateur harmonique.

satisfaisante tant que ces prescriptions n'ont pas reçu des fondements dont la solidité se mesure par la cohérence de la théorie qu'ils définissent.

Par ailleurs, l'énergie est certes une grandeur importante, mais la connaissance d'un système ne peut s'y réduire ; les solutions de l'équation de Schrödinger, les fonctions d'onde Ψ , doivent donc permettre d'aller plus loin dans la description précise d'un système. En fait – et c'est un *postulat* – la fonction d'onde d'un système est réputée contenir *toute* l'information escomptable sur ce système.

11.2.1 L'interprétation de Born et Jordan

L'interprétation acceptée par la majorité des physiciens est due principalement à Born [9] ; elle a été formulée en 1926, très peu de temps après les étapes décisives de Heisenberg, de Broglie et Schrödinger. Très vite, Schrödinger a suggéré que Ψ décrit des fluctuations de charge dans l'espace. L'idée de fluctuations ne tient que dans un contexte statistique, ou probabiliste. Selon l'idée première de Schrödinger, la fonction d'onde doit donc avoir une nature de ce type.

L'expérience des fentes d'Young confirme cette idée et aide à préciser le sens de la fonction d'onde, autant par les résultats troublants qu'elle confirme que par les analogies qu'elle autorise avec l'Optique. Par ailleurs, la discussion de cette expérience doit convaincre que la notion de probabilité de présence d'une particule en un point est un ingrédient essentiel et obligé de la Mécanique Ondulatoire : quand on compte les électrons un à un, on voit bien qu'il existe un aléatoire foncier autorisant seulement à décrire l'arrivée en un point par une (densité de) probabilité $P(x)$.

Partant de l'idée que la valeur de la fonction d'onde $\Psi(x)$ doit pouvoir être reliée à la probabilité de présence de (ou, à tout le moins, la probabilité d'observer) la particule au point x , on réalise tout de suite que $\Psi(x)$ ne peut convenir : $\Psi(x)$ est une fonction à valeurs complexes et une probabilité est un nombre positif (ou nul). D'ailleurs, même en oubliant cet aspect un peu particulier¹⁵, en se bornant à considérer seulement une fonction d'onde réelle, on voit que, si elle doit permettre d'expliquer les interférences, il faut bien qu'elle puisse au moins devenir négative et changer de signe d'un point à l'autre de l'espace ; c'est bien d'ailleurs ce que font en général les solutions de Schrödinger, comme on le verra par la suite¹⁶. Ceci montre bien à nouveau que $\Psi(x)$ elle-même ne saurait être une probabilité.

En Optique, les figures d'interférences résultent de la combinaison (la somme) de plusieurs ondes, représentant les champs électrique et magnétique \vec{E} et \vec{B} , dont les valeurs sont réelles, positives ou négatives ; mais les intensités des franges elles-mêmes sont données par les *carrés* des champs, \vec{E}^2 et \vec{B}^2 . Par analogie, et puisque $\Psi(x)$ est essentiellement complexe, on en vient à l'idée suivant laquelle la probabilité de présence de la particule au point x est proportionnelle au module carré de $\Psi(x)$, soit $|\Psi(x)|^2$. C'est ce qu'ont proposé Born et Jordan en admettant, par postulat, que la probabilité dP pour une particule de se trouver¹⁷ dans le petit volume d^3r autour du point défini par le vecteur \vec{r} est donnée par :

$$dP(\vec{r}, t) = |\Psi(\vec{r}, t)|^2 d^3r . \quad (11.8)$$

Il s'agit bel et bien d'un postulat, dont la véracité s'établit par la mise à l'épreuve de l'expérience. $|\Psi|^2$ est donc une *densité* de probabilité et a de ce fait la dimension L^{-d} dans \mathbb{R}^d . En ce qui concerne la fonction Ψ elle-même, et par analogie avec l'Électromagnétisme, on dit qu'il s'agit d'une amplitude, tout comme \vec{E} et \vec{B} sont les amplitudes des champs – précisant qu'il s'agit d'une amplitude de probabilité. Toutefois, cette amplitude est à valeurs complexes, une propriété qui démarque radicalement la Mécanique Quantique de l'Électromagnétisme, où l'usage des nombres complexes pour représenter les amplitudes des champs n'est qu'une commodité technique sans aucune contrepartie physique (les forces s'expriment à l'aide des champs et les forces sont réelles !).

Dès que l'on accepte (11.8), il faut disposer d'une équation de conservation : si $|\Psi(\vec{r}, t)|^2$ donne une probabilité, l'évolution de la probabilité dans l'espace au cours du temps est analogue à l'écoulement d'un

¹⁵Un temps, Schrödinger considéra – à tort – que seule la partie réelle de $\Psi(x)$ devait être considérée.

¹⁶À l'exception notable de celle décrivant l'état fondamental d'un système.

¹⁷Plus précisément, c'est la probabilité de trouver la particule lors d'une mesure de position.

fluide dans un circuit sans perte. Or cette équation de conservation, Schrödinger l'a déjà trouvée, en tant que *conséquence* de son équation – voir Ch. 10. On a vu qu'en posant (une particule de masse m) :

$$\rho(\vec{r}, t) = |\Psi(\vec{r}, t)|^2, \quad \vec{j}(\vec{r}, t) = \frac{\hbar}{2im} (\Psi^* \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^*), \quad (11.9)$$

une conséquence de l'équation de Schrödinger est :

$$\frac{\partial \rho}{\partial t} = -\text{div } \vec{j}. \quad (11.10)$$

Une autre conséquence du postulat fondamental est la nécessité pour la fonction d'onde de posséder des propriétés d'intégrabilité bien définies. Il faut que l'intégrale de $|\Psi|^2$ dans tout domaine \mathcal{D} fini soit finie, ce qui exclut déjà que Ψ possède des divergences trop violentes. En outre, $I_{\mathcal{D}} \equiv \int_{\mathcal{D}} |\Psi|^2 d^3r$ ne peut pas faire n'importe quoi : cela n'aurait pas de sens de trouver que cette intégrale diverge quand la distance du centre de \mathcal{D} à un point donné fixe augmente indéfiniment. Ceci interdit, par exemple, à Ψ de diverger exponentiellement à l'infini. En d'autres termes, $I_{\mathcal{D}}$ doit être finie y compris pour un domaine *fini* \mathcal{D} situé à l'infini :

$$\int_{\mathcal{D}} |\Psi|^2 d^3r < \infty \quad \forall \mathcal{D}. \quad (11.11)$$

Il est possible de préciser ces contraintes en faisant référence aux cas rencontrés en pratique. En général, une particule est soumise à un champ de forces lui conférant l'énergie potentielle¹⁸ $V(\vec{r})$; suivant la nature du champ de forces, il existe des états liés, ou des états non-liés, ou les deux à la fois. Les états liés sont les seuls à exister avec des potentiels très attractifs sur une longue distance : alors la particule se trouve préférentiellement près du point d'équilibre, sans pour autant être strictement confinée ; en tout cas, la probabilité de présence dans un domaine \mathcal{D} tend alors vers zéro si le domaine s'éloigne à l'infini du centre attractif. Si le potentiel est plus “mou” – tout en possédant un caractère attractif –, des états liés et des états non-liés sont simultanément possibles, selon la valeur de l'énergie de la particule ; dans un état non-lié, la probabilité mesurée par $I_{\mathcal{D}}$ ne tend pas vers zéro si \mathcal{D} part à l'infini. Enfin, pour les potentiels insuffisamment attractifs, ou pour les potentiels purement répulsifs, seuls des états non-liés existent.

S'agissant des états liés, on verra que, dans tous les cas rencontrés par la suite, l'interdiction faite à Ψ de diverger à l'infini produit automatiquement des solutions qui non seulement tendent vers zéro à l'infini, mais sont également de module carré sommable¹⁹. Alors, la probabilité de trouver la particule dans tout l'espace accessible est finie et, en conséquence de (11.8), il en va de même pour l'intégrale de $|\Psi|^2$ dans tout l'espace :

$$\int_{\mathbb{R}^3} |\Psi(\vec{r}, t)|^2 d^3r < +\infty. \quad (11.12)$$

Si (11.12) est satisfaite, on dit que Ψ est normalisable. Comme il est d'usage de compter les probabilités entre 0 et 1, on peut toujours en pareil cas remultiplier Ψ par le bon facteur pour avoir :

$$\int_{\mathbb{R}^3} |\Psi(\vec{r}, t)|^2 d^3r = 1. \quad (11.13)$$

Ceci constitue la condition de normalisation, satisfaite par défaut dans la suite sauf mention contraire.

La condition (11.12) est cruciale, comme le montre sa conséquence la plus spectaculaire : se traduisant en pratique par une condition aux limites²⁰, (11.12) permet de faire le tri parmi toutes les solutions de l'équation de Schrödinger et *fait apparaître spontanément la quantification*. En définitive, le postulat probabiliste à lui seul engendre la quantification, laquelle n'est plus dès lors le résultat d'une règle arbitraire comme c'était le cas dans l'Ancienne Théorie des Quanta. Cette “explication” théorique de la quantification observée est peut-être la preuve la plus éclatante de la justesse du postulat ci-dessus.

Les potentiels attractifs qui tendent vers une constante à l'infini, et les potentiels répulsifs, possèdent aussi des états propres non-normalisables (qui violent (11.12) mais satisfont (11.11)) et auxquels il est de ce fait

¹⁸L'usage en Mécanique Quantique est d'employer indifféremment les mots “énergie potentielle” ou “potentiel”.

¹⁹La “nature” est bien faite . . .

²⁰qui est l'expression technique du critère de rejet des solutions inacceptables physiquement.

possible de donner un sens physique. Il s'agit alors d'états de collision représentant la diffusion de particules par le potentiel V en question : ce sont clairement des états non-liés, pour lesquels $|\Psi|^2$ ne tend pas vers zéro à l'infini. Ces états donnent un courant \vec{j} non nul, et servent à construire des paquets d'ondes représentant une particule partant de l'infini en direction du centre du potentiel, avant d'y retourner après diffusion : il s'agit d'un phénomène dépendant du temps, au contraire de la description fournie par un état stationnaire (construit sur un seul état propre). En tout cas, les états non-liés ne sont en aucune façon astreints à satisfaire la condition de normalisabilité (11.12) ; en conséquence, l'énergie des états non-liés n'est pas quantifiée, toutes les valeurs sont possibles à partir d'une certaine valeur.

A titre d'exemple, soit l'atome d'hydrogène ; le potentiel est $-e'^2/r$, il est attractif mais assez mou et en tout cas à portée infiniment grande. Ce potentiel possède des états liés d'énergie négative²¹ et des états non-liés d'énergie positive, où l'énergie peut prendre n'importe quelle valeur entre 0 et $+\infty$. L'ensemble des valeurs propres de l'énergie peut alors être schématisé comme indiqué sur la figure (11.5).

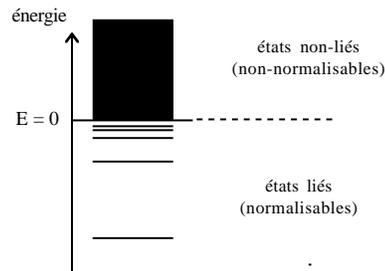


Figure 11.5: Spectre d'énergie du potentiel Coulombien

11.2.2 Calcul des valeurs moyennes

Une fois accepté le postulat de Born (11.8), d'une part la théorie acquiert définitivement un caractère probabiliste, d'autre part les moyens sont donnés de calculer la valeur moyenne (espérance mathématique) de n'importe quelle grandeur dynamique s'exprimant à l'aide de la coordonnée. Dans la suite, on se place dans \mathbb{R} (une seule dimension d'espace) pour simplifier les écritures, mais tous les résultats s'étendent immédiatement en dimension supérieure.

Etant donné une loi de probabilité $\Pi(\xi)$ – discrète ou continue – décrivant les probabilités pour qu'une variable aléatoire prenne les valeurs ξ , la valeur moyenne de cette aléatoire, notée $\langle \xi \rangle$, s'obtient en effectuant l'opération :

$$\langle \xi \rangle = \text{somme} (\xi \times \text{probabilité d'avoir la valeur } \xi) . \quad (11.14)$$

Si les valeurs possibles sont discrètes, $\{\xi_n\}_{n \in I}$, avec les probabilités $\{\Pi_n\}_{n \in I}$ et si la somme des probabilités est normalisée à l'unité, on a :

$$\langle \xi \rangle = \sum_{n \in I} \xi_n \Pi_n \quad \text{avec : } \sum_{n \in I} \Pi_n = 1 ; \quad (11.15)$$

au contraire, si ξ prend des valeurs continues dans un domaine \mathcal{D} avec la (densité) de probabilité $\Pi(\xi)$, on a :

$$\langle \xi \rangle = \int_{\xi \in \mathcal{D}} \xi \Pi(\xi) d\xi \quad \text{avec : } \int_{\xi \in \mathcal{D}} \Pi(\xi) d\xi = 1 . \quad (11.16)$$

Ce sont ces définitions usuelles que l'on transcrit ici avec la fonction d'onde. Compte tenu de (11.8), la valeur moyenne de la position d'une particule sur \mathbb{R} décrite par $\Psi(x, t)$ est donnée par :

$$\langle x \rangle(t) = \int_{x \in \mathbb{R}} x |\Psi(x, t)|^2 dx \quad \text{avec : } \int_{x \in \mathbb{R}} |\Psi(x, t)|^2 dx = 1 . \quad (11.17)$$

²¹ E_n est donnée par une expression qui coïncide avec celle de Bohr. C'est la valeur zéro qui sépare les valeurs discrètes et les valeurs continues, en conséquence du choix de l'origine du potentiel ($V(\infty) = 0$).

La notation au premier membre rappelle que la dépendance en temps (éventuelle) résulte de l'opération de moyenne : c'est Ψ qui (éventuellement) contient le temps. Au cas où la fonction d'onde n'est pas normalisée à l'unité, il faut écrire :

$$\langle x \rangle(t) = \frac{\int x |\Psi(x, t)|^2 dx}{\int |\Psi(x, t)|^2 dx} . \quad (11.18)$$

Dans la suite, on supposera Ψ normalisée à 1 et, pour simplifier les écritures, la variable t sera sous-entendue. D'une façon générale, la valeur moyenne d'une fonction quelconque de x , $f(x)$, est :

$$\langle f(x) \rangle = \int_{x \in \mathbb{R}} f(x) |\Psi(x)|^2 dx . \quad (11.19)$$

Ces règles permettent de trouver l'écart quadratique moyen de la coordonnée, qui est la mesure la plus simple des *fluctuations* de position :

$$(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2 = \langle (x - \langle x \rangle)^2 \rangle . \quad (11.20)$$

Comme tout écart quadratique, $(\Delta x)^2$ est une quantité positive, qui ne s'annule que si x prend une et une seule valeur, c'est-à-dire lorsque la variable aléatoire ne l'est pas (elle prend *une* certaine valeur avec probabilité 1). Explicitement, on a :

$$\Delta x^2 = \int_{x \in \mathbb{R}} x^2 |\Psi(x)|^2 dx - \left(\int_{x \in \mathbb{R}} x |\Psi(x)|^2 dx \right)^2 , \quad (11.21)$$

ou encore :

$$\Delta x^2 = \int_{x \in \mathbb{R}} \left(x - \int_{x \in \mathbb{R}} x |\Psi(x)|^2 dx \right)^2 |\Psi(x)|^2 dx . \quad (11.22)$$

À ce stade, on sait donc exprimer toutes les valeurs moyennes relatives à la position seule. Il faut maintenant trouver les règles de calcul pour l'impulsion p , représentée selon Schrödinger par l'opérateur différentiel :

$$p = -i\hbar \frac{\partial}{\partial x} . \quad (11.23)$$

Il faut associer p avec Ψ de la bonne façon pour construire la moyenne de l'impulsion. À la réflexion, il n'y a que trois possibilités :

$$\langle p \rangle_a = \int_{x \in \mathbb{R}} \left(-i\hbar \frac{\partial}{\partial x} \right) |\Psi(x)|^2 dx , \quad (11.24)$$

$$\langle p \rangle_b = \int_{x \in \mathbb{R}} \Psi^*(x) \left(-i\hbar \frac{\partial}{\partial x} \right) \Psi(x) dx , \quad (11.25)$$

$$\langle p \rangle_c = \int_{x \in \mathbb{R}} \Psi(x) \left(-i\hbar \frac{\partial}{\partial x} \right) \Psi^*(x) dx . \quad (11.26)$$

Dans ces expressions, l'opérateur différentiel agit sur ce qui est à sa droite. L'expression (11.24) est à rejeter d'emblée : elle donne pour la valeur moyenne de l'impulsion un nombre imaginaire pur ! Reste à choisir entre (11.25) et (11.26) ; on note d'abord que, pour une fonction d'onde qui s'annule à l'infini (c'est toujours vrai pour un état lié), on a :

$$\langle p \rangle_b = -\langle p \rangle_c , \quad (11.27)$$

comme le montre une intégration par parties. Il suffit donc choisir le bon signe pour avoir la bonne prescription – laquelle est universelle. Elle peut par conséquent être trouvée en raisonnant physiquement dans un cas particulier ; le plus simple est de considérer une particule libre. Le Hamiltonien H est alors $p^2/(2m)$ et l'équation de Schrödinger s'écrit :

$$i\hbar \frac{\partial \Psi}{\partial t} = \frac{1}{2m} \left(-i\hbar \frac{\partial}{\partial x} \right)^2 \Psi = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} . \quad (11.28)$$

Il en résulte $\Psi(x, t + dt) = \Psi(x, t) + [i\hbar/(2m)] \Psi'' dt + \dots$, où Ψ'' note simplement $\partial^2 \Psi / \partial x^2$. Il est alors facile d'écrire la valeur moyenne de la position à l'instant $t + dt$:

$$\langle x \rangle(t + dt) = \int_{\mathbb{R}} x \Psi^*(t + dt) \Psi(t + dt) dx \quad (11.29)$$

et de la comparer avec sa valeur à l'instant t . Un calcul simple donne :

$$\langle x \rangle(t + dt) = \langle x \rangle(t) + \left[\frac{i\hbar}{2m} \int_{\mathbb{R}} x (-\Psi''^* \Psi + \Psi^* \Psi'') dx \right] dt + \dots \quad (11.30)$$

Le terme du premier ordre en dt au second membre doit être la valeur moyenne de la vitesse, soit la valeur moyenne $\langle p/m \rangle$. Par identification :

$$\langle p \rangle = \frac{i\hbar}{2} \int_{\mathbb{R}} x (-\Psi''^* \Psi + \Psi^* \Psi'') dx \quad (11.31)$$

En effectuant à nouveau plusieurs intégrations par parties, on trouve que le second membre de (11.31) est égal à celui de (11.25). Ce résultat fournit l'expression de la valeur moyenne de p dans tous les cas, pas seulement pour la particule libre. L'incertitude est maintenant levée et c'est donc la forme $\langle p \rangle_b$, (11.25), qu'il faut retenir ; la valeur moyenne de l'impulsion d'une particule décrite par Ψ est donc :

$$\langle p \rangle = -i\hbar \int_{x \in \mathbb{R}} \Psi^*(x) \frac{\partial \Psi}{\partial x} dx \quad (11.32)$$

ce que l'on peut aussi écrire, compte tenu de (11.23) :

$$\langle p \rangle = \int_{x \in \mathbb{R}} \Psi^*(x) p \Psi(x) dx \quad (11.33)$$

où l'ordre des facteurs dans l'intégrand est évidemment essentiel.

Le calcul allant de (11.31) à (11.32) met en évidence une égalité importante :

$$\int_{x \in \mathbb{R}} \left(-i\hbar \frac{\partial \Psi}{\partial x} \right)^* \Psi(x) dx = \int_{x \in \mathbb{R}} \Psi^*(x) \left(-i\hbar \frac{\partial \Psi}{\partial x} \right) dx \quad (11.34)$$

soit :

$$\int_{x \in \mathbb{R}} (p\Psi)^* \Psi dx = \int_{x \in \mathbb{R}} \Psi^* (p\Psi) dx \quad (11.35)$$

Cette égalité exprime une propriété remarquable de l'opérateur p , appelée *hermiticité*. Tous les opérateurs représentant une grandeur physique doivent être hermitiques : c'est la condition nécessaire et suffisante pour que les valeurs moyennes soient des nombres réels²².

Une fois obtenue la recette pour le calcul de $\langle p \rangle$, l'expression de la valeur moyenne de n'importe quelle fonction de p , $g(p)$, s'en déduit ; en rétablissant le temps partout pour la généralité, on a :

$$\langle g(p) \rangle(t) = \int_{x \in \mathbb{R}} \Psi^*(x, t) g \left(-i\hbar \frac{\partial}{\partial x} \right) \Psi(x, t) dx \quad (11.36)$$

On remplace dans $g(p)$ la "variable" p par $-i\hbar \partial / \partial x$, ce qui donne un nouvel opérateur, que l'on fait agir sur ce qui est à sa droite, et on effectue l'intégration sur x . Par exemple, les fluctuations de p sont mesurées par :

$$(\Delta p)^2 = \langle p^2 \rangle - \langle p \rangle^2 = \langle (p - \langle p \rangle)^2 \rangle \quad (11.37)$$

soit :

$$\Delta p^2 = \int_{x \in \mathbb{R}} \Psi(x)^* p^2 \Psi(x) dx - \left(\int_{x \in \mathbb{R}} \Psi(x)^* p \Psi(x) dx \right)^2 \quad (11.38)$$

²²Visiblement, x satisfait aussi une équation du genre (11.35).

Les écritures avec p – ou $g(p)$ – donnent l'idée d'écrire plus symétriquement la valeur moyenne de la position, ou de n'importe quelle fonction $f(x)$. Comme x (ou $f(x)$) n'est finalement qu'un opérateur *multiplicatif*, on peut écrire tout autant :

$$\langle f(x) \rangle(t) = \int_{x \in \mathbb{R}} \Psi^*(x, t) f(x) \Psi(x, t) dx . \quad (11.39)$$

Il faut enfin savoir écrire la valeur moyenne d'une fonction dépendant à la fois de x et de p . Soit donc une grandeur dynamique représentée classiquement par la fonction $F_{\text{class}}(x, p)$; en général, son représentant quantique n'est pas :

$$F_{\text{class}}(x, p) \rightarrow F\left(x, -i\hbar \frac{\partial}{\partial x}\right) , \quad (11.40)$$

puisque les opérateurs x et p ne commutent pas, alors que dans l'expression de F_{class} , l'ordre des variables classiques est sans importance. Il convient donc de transformer l'expression de la fonction F_{class} avant d'y effectuer la substitution passant des variables classiques aux opérateurs quantiques – en général, une symétrisation suffit. L'exemple suivant montre ce qu'il faut faire. Soit la fonction :

$$F_{\text{class}}(x, p) = xp . \quad (11.41)$$

Comme $xp \neq px$ en termes d'opérateurs, le représentant de (11.41) n'est pas :

$$F(x, p) = x \times \left(-i\hbar \frac{\partial}{\partial x}\right) . \quad (11.42)$$

En revanche, $xp = px$ en termes de variables classiques ; on a donc aussi :

$$F_{\text{class}}(x, p) = \frac{1}{2}(xp + px) , \quad (11.43)$$

dont l'équivalent quantique est l'opérateur :

$$F(x, p) = -\frac{i\hbar}{2} \left(x \frac{\partial}{\partial x} + \frac{\partial}{\partial x} x\right) \quad (11.44)$$

On comprendra par la suite pourquoi cette écriture est de fait acceptable²³. Si l'on omet ce type de précaution, les valeurs moyennes peuvent ne pas être réelles – ce qui constitue un avertissement.

Parmi les fonctions de type $F(x, p)$, l'une joue un rôle de tout premier plan, c'est le Hamiltonien. Pour une particule de masse m et d'énergie potentielle $V(x)$, le Hamiltonien classique est :

$$H_{\text{class}}(x, p) = \frac{1}{2m} p^2 + V(x) . \quad (11.45)$$

Son représentant quantique est l'opérateur :

$$H\left(x, -i\hbar \frac{\partial}{\partial x}\right) = \frac{1}{2m} \left(-i\hbar \frac{\partial}{\partial x}\right)^2 + V(x) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) . \quad (11.46)$$

La généralisation dans \mathbb{R}^3 est :

$$H\left(\vec{r}, -i\hbar \vec{\nabla}\right) = \frac{1}{2m} \left(-i\hbar \vec{\nabla}\right)^2 + V(\vec{r}) = -\frac{\hbar^2}{2m} \Delta + V(\vec{r}) , \quad (11.47)$$

où Δ note le Laplacien. Ceci étant, la valeur moyenne de H – c'est donc la valeur moyenne de l'énergie – dans l'état $\Psi(\vec{r}, t)$ est donnée par²⁴ :

$$\langle H \rangle = \int_{\vec{r} \in \mathbb{R}^3} \Psi^*(\vec{r}, t) \left[-\frac{\hbar^2}{2m} \Delta + V(\vec{r})\right] \Psi(\vec{r}, t) d^3r . \quad (11.48)$$

²³L'opérateur $-i\hbar[x\partial/\partial x + (\partial/\partial x)x]$ est hermitique, alors que (11.42) ne l'est pas. Cette affirmation se vérifie en effectuant des intégrations par parties sur des intégrales où les opérateurs sont associés avec deux fonctions quelconques.

²⁴On remarque que le premier membre de (11.48) est écrit $\langle H \rangle$ et non $\langle H \rangle(t)$. Ce n'est pas une omission : on verra par la suite que la moyenne de l'énergie ne peut dépendre du temps que si H dépend du temps, ce qui n'est pas le cas ici.

Remarque

Les formules précédentes peuvent faire croire que, au contraire de ce qui se passe en Mécanique Classique, la coordonnée q joue un rôle éminent au détriment de son moment conjugué p . En fait il n'en est rien. Il est possible de tout reformuler en échangeant les rôles, semblant cette fois donner la prééminence à p ; il s'agit d'une écriture *différente* de la même théorie, c'est une autre *représentation* formelle.

Pour comprendre ceci, il suffit de revenir à la relation fondatrice :

$$[q, p] = i\hbar \mathbf{1} . \quad (11.49)$$

Il y a plusieurs façons de représenter les opérateurs, qui toutes respectent (11.49). Outre la première substitution exposée ci-dessus, on peut choisir la suivante :

$$q \rightarrow +i\hbar \frac{\partial}{\partial p} , \quad p \rightarrow \times p . \quad (11.50)$$

Il est facile de vérifier que cette association est conforme à (11.49). Maintenant, p une simple variable multiplicative et les opérateurs différentiels agissent sur des fonctions de la variable p ; les fonctions d'onde seront notées Φ , plus précisément $\Phi(p, t)$. On peut travailler avec elles exactement comme avec le premier type rencontré, en suivant cette fois l'association (11.50). Cette façon de représenter la Mécanique Quantique est appelée *représentation- p* , pour la distinguer de la première formulation, dite *représentation- q* . En représentation p , les relations suivantes sont vraies :

$$\langle x \rangle(t) = \int_{p \in \mathbb{R}} \Phi^*(p, t) \left(+i\hbar \frac{\partial \Phi}{\partial p} \right) dp , \quad \langle p \rangle(t) = \int_{p \in \mathbb{R}} p |\Phi(p, t)|^2 dp , \quad etc. \quad (11.51)$$

Il existe un lien étroit entre une fonction d'onde en représentation- q , $\Psi(\vec{r}, t)$ et son équivalent en représentation- p : ces deux fonctions sont transformées de Fourier l'une de l'autre. Très précisément :

$$\Psi(\vec{r}, t) = (2\pi\hbar)^{-3/2} \int_{\mathbb{R}^3} d^3p e^{(i/\hbar) \vec{p} \cdot \vec{r}} \Phi(\vec{p}, t) , \quad (11.52)$$

et inversement :

$$\Phi(\vec{p}, t) = (2\pi\hbar)^{-3/2} \int_{\mathbb{R}^3} d^3r e^{-(i/\hbar) \vec{r} \cdot \vec{p}} \Psi(\vec{r}, t) . \quad (11.53)$$

La fonction d'onde $\Phi(\vec{p}, t)$ a le sens dual de $\Psi(\vec{r}, t)$: la quantité $|\Phi(\vec{p}, t)|^2 d^3p$ est la probabilité élémentaire dP' de trouver la particule avec une impulsion \vec{p} à d^3p près²⁵ ; pour \mathbb{R}^3 , Φ a la dimension [impulsion]^{-3/2}.

11.2.3 Le déterminisme quantique

Il s'agit ici d'abord de préciser les grandes lignes de l'intégration de l'équation de Schrödinger. Par la suite sera donnée une brève discussion des caractères spécifiques du déterminisme comme l'entend la Mécanique Quantique.

L'évolution temporelle de la fonction d'onde est donnée par l'équation de Schrödinger ; outre les conditions aux limites à appliquer – elles sont maintenant précisées ci-dessus – il convient de se donner un état initial : cette équation est une équation aux dérivées partielles du premier ordre en temps ; pour résoudre complètement un problème spécifié par ailleurs, on doit donc disposer, en plus, d'une donnée initiale sous la forme d'une fonction : on déclare que la fonction d'onde à un certain instant, $\Psi(\vec{r}, t = t_0)$, est égale à $\psi(\vec{r})$, fonction connue. Cette information représente une infinité de nombres et, comme dans le cas d'un champ purement classique, on peut dire que le formalisme quantique introduit formellement, pour tout système, une infinité de degrés de liberté ; ici, il s'agit d'une forme bien plus subtile : ce ne sont pas des degrés de liberté liés à la dimension de l'espace (c'est déjà le cas pour une particule dans \mathbb{R}) mais à l'infinité de potentialités qui se présentent à un objet quantique.

²⁵Plus précisément : la probabilité pour qu'une mesure d'impulsion *etc* Il ne s'agit que d'un avatar du postulat fondamental de Born, ce n'est pas un nouveau postulat.

L'information d'un état initial étant fournie, on peut déduire de l'équation de Schrödinger l'état du système à l'instant t ; ceci est l'expression technique du déterminisme quantique. Pour fixer les idées, reprenons le cas d'une particule libre représentée par un paquet d'ondes ; si ce dernier est donné par l'expression suivante à l'instant noté t_0 :

$$\Psi(\vec{r}, t = t_0) \equiv \psi(\vec{r}) = (2\pi)^{-3/2} \int_{\mathbb{R}^3} d^3k A(\vec{k}) e^{i\vec{k}\cdot\vec{r}} , \quad (11.54)$$

alors il est facile d'obtenir le paquet d'ondes qui en est le descendant à l'instant t ; notons que $\psi(\vec{r})$ étant donnée, sa transformée de Fourier $A(\vec{k})$ l'est tout autant :

$$A(\vec{k}) = (2\pi)^{-3/2} \int_{\mathbb{R}^3} d^3r' \Psi(\vec{r}', t_0) e^{-i\vec{k}\cdot\vec{r}'} . \quad (11.55)$$

Pour cette particule libre, l'équation de Schrödinger est :

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \Delta \Psi . \quad (11.56)$$

Visiblement, toute fonction de la forme $e^{i(\vec{k}\cdot\vec{r}-\omega t)}$ est solution (comme on le voit en reportant dans (11.54)), à condition que \vec{k} et ω soient reliés par :

$$\hbar\omega = \frac{\hbar^2 k^2}{2m} \iff \omega \equiv \omega(k) = \frac{\hbar k^2}{2m} , \quad (11.57)$$

ce qui constitue la relation de dispersion. Maintenant, comme (11.56) est une équation linéaire, toute combinaison linéaire de solutions est encore solution. Il en résulte que l'on peut chercher $\Psi(\vec{r}, t)$ sous la forme d'une somme continue (intégrale) avec des coefficients B , pour l'instant quelconques mais qu'il faut précisément trouver :

$$\Psi(\vec{r}, t) = (2\pi)^{-3/2} \int_{\mathbb{R}^3} d^3k B(\vec{k}) e^{i[\vec{k}\cdot\vec{r}-\omega(k)t]} \quad (11.58)$$

En comparant avec la condition initiale (11.54), on voit qu'il faut avoir :

$$B(\vec{k}) e^{-i\omega(k)t_0} = A(\vec{k}) , \quad (11.59)$$

ce qui fournit l'expression intégrale de la solution à l'instant t :

$$\Psi(\vec{r}, t) = (2\pi)^{-3/2} \int_{\mathbb{R}^3} d^3k A(\vec{k}) e^{\frac{1}{i\hbar} \frac{\hbar^2 k^2}{2m} (t-t_0)} e^{i\vec{k}\cdot\vec{r}} ; \quad (11.60)$$

en utilisant (11.55), ceci vaut :

$$\Psi(\vec{r}, t) = (2\pi)^{-3} \int_{\mathbb{R}^3} d^3k \int_{\mathbb{R}^3} d^3r' e^{-i\vec{k}\cdot\vec{r}'} e^{\frac{1}{i\hbar} \frac{\hbar^2 k^2}{2m} (t-t_0)} e^{i\vec{k}\cdot\vec{r}} \Psi(\vec{r}', t_0) . \quad (11.61)$$

En inversant l'ordre des intégrations et en effectuant celle sur le vecteur d'onde, on obtient :

$$\Psi(\vec{r}, t) = \int_{\mathbb{R}^3} d^3r' U(\vec{r} - \vec{r}', t - t_0) \Psi(\vec{r}', t_0) , \quad (11.62)$$

où le noyau U est :

$$U(\vec{\rho}, \tau) = \left(\frac{m}{2i\pi\hbar\tau} \right)^{3/2} e^{\frac{im}{2\hbar\tau} \rho^2} . \quad (11.63)$$

L'expression (11.62) donne la fonction d'onde à l'instant t déduite de la fonction à l'instant initial. $\Psi(\vec{r}, t)$ est une combinaison linéaire de toutes les *valeurs* de la fonction d'onde de départ en tous les points de l'espace. La fonction U pondère le tout : elle incorpore l'importance relative du passage du point \vec{r}' à l'instant t_0 au point \vec{r} à l'instant t et constitue ce que l'on appelle une amplitude de transition. U est appelé propagateur, car c'est lui qui décrit comment la fonction d'onde se déplace (se propage) dans l'espace et dans le temps. Ici, c'est propagateur pour une particule libre, mais la structure fondamentale (11.62) persiste dans les cas plus complexes : pour l'établir, on a seulement utilisé des propriétés générales de l'équation de Schrödinger, en particulier sa linéarité.

Il est utile de remarquer que le propagateur ne dépend ici que de la différence des temps et de la différence des coordonnées ; ceci était prévisible. Le Hamiltonien est indépendant du temps, donc l'origine des temps est sans importance (seul compte le temps écoulé) ; par ailleurs, pour une particule libre, H ne dépend pas de la coordonnée (seules les dérivées spatiales y figurent) ; il n'y a donc pas de point remarquable dans l'espace, au contraire du cas où une origine particulière figure explicitement dans l'expression de l'énergie potentielle. On touche ici du doigt la relation entre symétrie et absence de certaines variables (absentes précisément en raison de cette symétrie). Enfin, l'expression (11.62) montre que l'on doit avoir :

$$\lim_{t \rightarrow t_0} \int_{\mathbb{R}^3} d^3 r' U(\vec{r} - \vec{r}', t - t_0) \Psi(\vec{r}', t_0) = \Psi(\vec{r}, t_0) \quad (11.64)$$

ce qui donne le résultat :

$$\lim_{t \rightarrow t_0} U(\vec{r} - \vec{r}', t - t_0) = \delta(\vec{r} - \vec{r}') \quad (11.65)$$

Il est tout à fait remarquable de voir apparaître l'action classique S dans l'expression du propagateur, ce qui permet d'écrire (au moins pour la particule libre) :

$$\Psi(\vec{r}, t) = \left[\frac{m}{2i\pi\hbar(t - t_0)} \right]^{3/2} \int_{\mathbb{R}^3} d^3 r' e^{\frac{i}{\hbar} S(\vec{r}, \vec{r}'; t, t_0)} \Psi(\vec{r}', t_0) . \quad (11.66)$$

Pour un couple (\vec{r}_0, t_0) donné, on voit que l'amplitude de transition vers le point (\vec{r}, t) ne sera importante (et contribuera donc beaucoup à $\Psi(\vec{r}, t)$) que dans la mesure où la phase est petite devant π , c'est-à-dire si la valeur de l'action pour ces deux couples est au plus comparable à la constante de Planck. A l'opposé, si le chemin de (\vec{r}_0, t_0) à (\vec{r}, t) donne une très grande action (en unité \hbar), la phase sera elle-même très grande. Par un argument de phase stationnaire, on voit que dans une situation classique (où l'action est toujours gigantesque par rapport à \hbar , seul compte (ou à peu près) le chemin pour lequel l'action est stationnaire : c'est la trajectoire classique ! Au contraire, si l'action est comparable à \hbar pour un grand nombre de chemins, les effets quantiques jouent à plein et la notion même de trajectoire, effectivement, disparaît (aucun chemin ne joue de rôle prépondérant). On sent intuitivement qu'il existe une transition continue de la situation classique à la situation quantique : la trajectoire classique devient de plus en plus "floue" et diffuse, jusqu'au point de perdre tout sens par rapport à sa définition ordinaire. Le passage vers le quantique se traduit par une augmentation graduelle de *fluctuations* (la position est incertaine, la tangente devient de plus en plus folle), appelées généralement fluctuations quantiques.

Clairement, comme Ψ s'obtient en effectuant une somme (intégrale) d'amplitudes, la probabilité de présence en un point à l'instant t est le module au carré d'une somme d'amplitudes et non pas la somme de probabilités (cette distinction est le cœur dur de la théorie quantique). L'électron ne passe pas par une fente d'Young ou l'autre²⁶, il y a – en gros – deux chemins possibles, chacun avec son amplitude. La probabilité d'impact en un point de l'écran est le module au carré de cette somme d'amplitudes.

Terminons cette discussion sur le sens de la fonction d'onde par quelques commentaires généraux sur la notion de déterminisme. L'expérience d'Young montre que la dynamique des très petits objets (particules) a un caractère imprévisible, aléatoire : quand on fait passer les électrons un à un, on les voit se matérialiser en des points variables, alors que tous sont issus du même appareil. Ceci est un fait d'expérience, qui semble devoir conduire à l'abandon d'une forme de déterminisme omniprésent et souverain dans la pensée classique. Notons qu'il y a forcément une incertitude sur l'état initial de chaque particule²⁷, qui rend illusoire toute croyance en la reproductibilité de l'expérience portant sur un seul électron : seule la reproductibilité statistique est encore et toujours de règle, et c'est le seul objectif que doit se fixer la théorie²⁸.

L'aspect statistique de la microphysique ne peut toutefois, fondamentalement, se réduire à un manque d'information sur l'état initial. S'impose une révision de ce que l'on appelle l'état : ce n'est plus une position

²⁶ "ou" = somme de probabilités.

²⁷ Des considérations semblables peuvent être discutées dans un cadre classique. Si on considère un ensemble de particules localisées à $t = 0$ mais dont on ne connaît pas avec une précision infiniment grande – c'est toujours le cas ! – les vitesses initiales, la prévision du mouvement ultérieur ne peut qu'être statistique et il est bien évident que le "nuage de points" s'étale au cours du temps (les unes avancent plus vite que les autres). Les astronomes savent bien qu'il faut repointer régulièrement la position des astres et ne pas se contenter d'un mouvement calculé avec des conditions initiales remontant à Mathusalem.

²⁸ La notion de reproductibilité est parfois malmenée en Physique, non que fondamentalement, le déterminisme soit remis en cause ; mais tout simplement parce que certains phénomènes (par exemple l'apparition de la turbulence) semblent dépendre de paramètres non identifiés qui échappent au contrôle de l'expérimentateur, ce qui pose des problèmes de fond sur le choix d'une méthodologie. Plus généralement, il existe des systèmes (non-linéaires) susceptibles de présenter des comportements imprévisibles à court ou moyen terme du fait de leur sensibilité aux conditions initiales (systèmes chaotiques).

et une vitesse, mais c'est une fonction d'onde – qui donne accès au calcul effectif des lois de probabilité. Se placer dans le cadre de la Mécanique Quantique, c'est se mettre dans une vision radicalement différente de – et incompatible avec – la pensée classique ; la Mécanique Quantique ne peut fournir que des informations statistiques, les destins individuels échappent à sa description. En outre, elle refuse l'extrapolation au monde des particules d'images forgées à propos d'objets macroscopiques ; ces dernières avaient finalement acquis le statut de dogmes : force est d'admettre que certains concepts légitimés à une échelle donnée peuvent perdre tout sens sur une autre échelle.

Cette forme d'“indéterminisme” a longtemps agité les esprits²⁹, et a suscité, en particulier, la recherche de ce que l'on a appelé les “variables cachées” : ce seraient des variables non incluses dans la description quantique que l'on connaît aujourd'hui mais qui existeraient bel et bien. Dans cette optique, la Mécanique Quantique est *ipso facto* considérée comme une théorie incomplète, au sens où elle ne donne pas une description exhaustive de la réalité physique. Ces variables cachées seraient, en quelque sorte, tout ce qui est absent dans la formulation actuelle pour supprimer toute indétermination (aléatoire) de principe. Par exemple, l'instant de désintégration d'un noyau instable donné n'est pas calculable dans le formalisme quantique actuel³⁰. Ici, la variable cachée correspondante serait une sorte d'horloge interne déterminant l'instant de désintégration, absente de la description quantique présente. De même, la direction d'émission d'un photon spontané par un atome excité isolé serait inscrite dans l'histoire antérieure de l'atome mais non représentée dans le formalisme connu aujourd'hui. Les expériences d'Alain Aspect ont montré l'impossibilité de certaines variables cachées dites locales. Il existe des théories de variables cachées généralisées, encore compatibles avec les résultats de ces expériences (variables cachées non-locales) ; on peut dire qu'elles sont aussi éloignées de la Physique Classique que la Mécanique Quantique elle-même – ce qui n'est pas une tare – mais leur ôte toutefois une grande part de leur pertinence. Au total, l'intuition n'y gagnerait pas grand'chose.

En pratique, il faut comprendre que le déterminisme est bien présent en Mécanique Quantique ; par rapport à son rôle classique, il a seulement changé de registre. L'équation fondamentale de Schrödinger est une équation du premier ordre en temps : une fois précisée la condition initiale, l'état à l'instant t est unique et parfaitement bien défini. Un état initial donné détermine complètement l'état ultérieur : ce fait est bien la manifestation la plus directe du déterminisme en pleine action – pour autant que c'est bien cette équation qui pilote l'évolution. Ceci revient à se placer dans une situation où l'objet quantique évolue de lui-même, suit son évolution propre, en dehors de toute interaction avec, par exemple, un appareil de mesure. Ce qui est alors prévisible, complètement déterminé, ce ne sont pas des valeurs numériques certaines mais les probabilités de trouver telle ou telle valeur. Le déterminisme quantique porte exclusivement sur l'évolution dans le temps des probabilités : il ne concerne plus l'évolution des nombres mais celle des lois de probabilités de ces nombres.

11.3 Principe d'incertitude de Heisenberg

Ce principe n'en est pas un à proprement parler, dès que l'équation de Schrödinger est admise, mais la terminologie, de nature historique, sera conservée dans la suite. Il se traduit par différentes relations que l'on va essayer de construire intuitivement et en rassemblant des éléments acquis précédemment.

11.3.1 Principe d'incertitude spatial

L'expérience d'Young – et d'autres – impose(nt) de renoncer à la notion même de trajectoire au sens classique. Or une telle trajectoire est bien définie dès que l'on connaît à tout instant la position et la vitesse : la position permet de dessiner un point, la vitesse permet de tracer la tangente ; de proche en proche, on peut dessiner une ligne continue différentiable ; c'est la trajectoire au sens usuel³¹.

²⁹et non des moindres ! Einstein lui-même n'a jamais accepté totalement l'interprétation standard, et a proposé d'ailleurs, en 1935, un fameux paradoxe dit *paradoxe EPR*. Suivant les dernières (et assez récentes) expériences conduites par le groupe d'Alain Aspect dans le milieu des années 80, certaines prévisions de la Mécanique Quantique, parmi les plus “paradoxaux”, sont vérifiées par l'expérience ; elles éliminent la possibilité des variables cachées dites locales.

³⁰On sait seulement trouver la loi de probabilité des instants de mort.

³¹L'équation fondamentale de la Dynamique est du second ordre en temps : elle suppose donc l'existence d'une dérive seconde, ce qui implique que la dérivée première est continue. Une “trajectoire” brownienne – courbe continue mais nulle part dérivable – ne peut être obtenue dans un tel cadre.

Renoncer à pouvoir définir une trajectoire, c'est donc admettre que l'une au moins des deux variables, position ou vitesse, est partiellement indéterminée – en toute généralité, les deux le seront. On pourrait ainsi dans un premier temps imaginer que la position reste parfaitement définie, alors que la vitesse est entachée d'une indétermination fondamentale, comme dans le cas d'une trajectoire brownienne. Rien, *a priori* ne justifie une telle attitude ; au contraire, il semble important intuitivement de continuer à traiter les deux grandeurs position et impulsion de façon symétrique (sans donner à l'une d'entre elles un statut privilégié). En effet, la notion de position est liée à l'aspect localisé d'une particule, que l'on peut mettre en évidence par un point lumineux éphémère sur un écran luminescent. D'un autre côté, l'aspect ondulatoire est directement lié à l'impulsion par la relation de de Broglie $\lambda = h/p$. Or les deux aspects – onde et corpuscule – sont également importants, chacun d'entre eux donnant lieu à des manifestations expérimentales. En outre, l'expérience d'Young montre bien que la coordonnée à elle seule ne peut relever d'une description classique (on ne peut dire par où passe l'électron quand les deux fentes sont ouvertes).

On doit donc essayer de préciser l'incertitude fondamentale par une relation symétrique ; désignons par Δx et Δp les “erreurs” (incertitudes) sur la position et l'impulsion, sans se soucier pour l'instant de donner un sens précis au symbole Δ . La relation la plus simple respectant la symétrie requise est le produit $\Delta x \Delta p$; traduire en équation l'indétermination, c'est finalement écrire que :

$$\Delta x \Delta p \text{ supérieur à quelque chose ,} \quad (11.67)$$

où “quelque chose” est évidemment une grandeur fondamentale, convenablement dimensionnée, indépendante du système considéré. Le produit $\Delta x \Delta p$ a la dimension d'une action ; la constante d'action maintenant incontournable est la constante de Planck. On doit donc s'attendre à une relation du genre :

$$\Delta x \Delta p > \text{nombre} \times \hbar , \quad (11.68)$$

où le nombre est d'ordre 1 et où le choix de \hbar au lieu de h est sans importance et se borne à respecter l'usage (pour l'instant, les incertitudes Δ n'ont même pas été définies précisément). Le point fondamental est que ce *produit* ne peut être rendu *arbitrairement petit*, et c'est bien ce que dit (11.68).

Avant de discuter quelques conséquences de cette relation et de donner des exemples illustrant son importance, il est bon d'en préciser le bien-fondé. Pour ceci, il suffit de revenir aux arguments généraux développés à propos des paquets d'ondes. Pour une fonction $f(x)$ et sa transformée de Fourier $F(k)$, les amplitudes des intervalles où chacune d'entre elles est sensiblement non nulle, Δx et Δk sont reliées par :

$$\Delta x \Delta k \sim 1 . \quad (11.69)$$

Cette relation n'a rien de quantique, elle tient d'ailleurs tout autant si on considère le couple $(f(t), F(\omega))$, quand la transformation de Fourier fait passer de l'espace des temps à celui des pulsations :

$$\Delta \omega \Delta t \sim 1 . \quad (11.70)$$

L'aspect quantique apparaît quand on prend en compte la relation de de Broglie sous la forme $p = \hbar k$. Quand on élimine k au profit de p , (11.69) donne bien ce que l'on attend³² :

$$\Delta x \Delta p \sim \hbar . \quad (11.71)$$

Ainsi se traduit l'indétermination fondamentale rendue logiquement nécessaire par l'ensemble des éléments qui découlent de l'expérience d'Young. C'est la relation (11.71) qui porte le nom de “Principe d'incertitude de Heisenberg”, mais, clairement, il ne s'agit pas à proprement parler d'un principe. On peut démontrer ([7], p. 159) que si deux grandeurs physiques sont représentées par deux opérateurs A et B non-commutatifs tels que :

$$[A, B] = iC \quad (11.72)$$

alors on a strictement :

$$\Delta A \Delta B \geq \frac{1}{2} |\langle C \rangle| . \quad (11.73)$$

³²Il existe une formulation très précise en termes de commutateurs (voir (11.73)), où tous les facteurs numériques seront complètement définis et où de surcroît, \sim devient \geq , ce qui n'est pas sans importance.

Dans ces écritures, les Δ désignent précisément les écarts-types :

$$\Delta A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2} , \quad \Delta B = \sqrt{\langle B^2 \rangle - \langle B \rangle^2} \quad (11.74)$$

et toutes les moyennes sont prises avec le même état Ψ quelconque³³.

La relation d'incertitude (11.71) se généralise comme suit dans \mathbb{R}^3 :

$$\Delta x \Delta p_x \sim \hbar , \quad \Delta y \Delta p_y \sim \hbar , \quad \Delta z \Delta p_z \sim \hbar . \quad (11.76)$$

En revanche, aucune contrainte n'existe sur les incertitudes impliquant deux directions *différentes* de l'espace, sauf évidemment d'être positives :

$$\Delta u \Delta p_v \geq 0 , \quad (u, v = x, y, z; u \neq v) , \quad (11.77)$$

de sorte que toutes les relations d'incertitude position - impulsion peuvent être rassemblées en une seule qui, compte tenu de (11.75), s'écrit :

$$\Delta u \Delta p_v \geq \frac{\hbar}{2} \delta_{uv} , \quad (u, v = x, y, z) . \quad (11.78)$$

L'existence d'une borne inférieure pour le produit $\Delta u \Delta p_u$ a des implications considérables. Par exemple, ceci permet de comprendre que l'atome ne s'effondre pas sur lui-même³⁴ ; l'argument qui suit est essentiellement dimensionnel, ce que ne signifie pas qu'il ne faut pas le prendre au sérieux. L'expression de l'énergie est :

$$E = \frac{\vec{p}^2}{2m} - \frac{e'^2}{r} \quad (11.79)$$

Les valeurs moyennes de r et p , dans un état donné, sont ce qu'elles sont et, en ordre de grandeur (\tilde{r} et \tilde{p}), sont certainement soumises aux relations d'incertitude :

$$\tilde{r} \tilde{p} \sim \hbar . \quad (11.80)$$

L'ordre de grandeur de l'énergie est donc :

$$\tilde{E} \sim \frac{\tilde{p}^2}{2m} - \frac{e'^2}{\tilde{r}} \sim \frac{\tilde{p}^2}{2m} - \frac{e'^2 \tilde{p}}{\hbar} \equiv E(\tilde{p}) . \quad (11.81)$$

Cette expression ne dépend plus que de \tilde{p} , elle a visiblement un minimum quand \tilde{p} varie, atteint pour $\tilde{p} = me'^2/\hbar$. Alors³⁵ :

$$\tilde{E} \geq E(me'^2/\hbar) = -\frac{me'^4}{2\hbar^2} . \quad (11.82)$$

Ainsi, grâce aux relations d'incertitude, l'énergie de l'atome est maintenant bornée inférieurement.

Un autre exemple : la valeur *finie* (non-nulle) de l'énergie fondamentale d'un oscillateur harmonique est une manifestation des relations d'incertitude. En effet, l'énergie est la somme de deux termes positifs : pour qu'elle soit nulle, il faut que chacun d'entre eux le soit ; ici, ceci signifierait que la particule est au repos avec une vitesse nulle : vitesse et position seraient complètement déterminées en même temps, en violation des relations d'incertitude. En d'autres termes, ces dernières interdisent le repos absolu, les fluctuations quantiques ne peuvent être réduites à rien et l'énergie de l'état fondamental est strictement positive (et égale à $\hbar\omega/2$) : c'est ce que l'on appelle l'énergie de point-zéro.

³³Une conséquence de (11.73) est que, strictement (et non plus en ordre de grandeur) :

$$\Delta u \Delta p_u \geq \frac{\hbar}{2} . \quad (11.75)$$

³⁴L'énergie classique de l'atome d'hydrogène n'est pas bornée inférieurement : l'énergie potentielle est d'autant plus petite (négative) que l'électron est proche du noyau ; par le théorème du Viriel, l'énergie totale vaut $+V/2$. L'état le plus stable (de plus basse énergie) correspond donc à $\dots r = 0$.

³⁵On tombe pile sur la valeur *exacte* de l'énergie fondamentale. La coïncidence numérique ne doit pas faire illusion ; l'argument est essentiellement dimensionnel et semi-quantitatif.

Ces fluctuations quantiques donnent à la Mécanique Quantique (qui reste *purement mécanique*, la température est nulle) un caractère essentiellement statistique, tout comme les fluctuations thermiques présentes à toute température finie introduisent un flou dans la description des systèmes thermodynamiques classiques et imposent le recours à une théorie de nature statistique : la vitesse d'un atome de gaz parfait n'est pas calculable en tant que telle³⁶, mais on sait trouver sa moyenne et son écart quadratique. Dans les cas usuels, compte tenu de l'isotropie et de l'homogénéité de l'espace, cet écart quadratique est proportionnel à la température absolue T . L'amplitude des fluctuations thermiques tend donc vers zéro quand $T \rightarrow 0$; sans les effets quantiques, la matière tendrait alors à se figer, étant à température nulle dans un état de repos absolu. La réalité est tout autre : en abaissant la température, il vient un moment où les fluctuations quantiques, toujours présentes mais parfois occultées à haute température par les fluctuations purement thermiques, s'affirment magistralement : tôt ou tard, les effets quantiques finissent par l'emporter au point de devenir prépondérants.

11.3.2 Principe d'incertitude temporel

Le couple (temps, fréquence), ou (temps, pulsation), satisfait des relations de Fourier au même titre que position et impulsion (voir (11.70)). En faisant apparaître l'énergie $E = \hbar\omega$, on a :

$$\Delta E \Delta t \sim \hbar . \quad (11.83)$$

Cette relation est en tout point analogue à celle obtenue avec le couple (x, p) ; toutefois, elle a un statut très différent.

En effet, le paramètre temps est une grandeur qui reste en dehors de toute quantification ; il n'existe pas d'opérateur T représentant le temps, au contraire de l'énergie E , associée à l'opérateur Hamiltonien. Le sens de (11.83) doit donc être précisé. En ce qui concerne ΔE , on peut toujours prendre l'écart-type de H , qui mesure les fluctuations d'énergie ; en revanche, Δt ne peut pas être un écart-type de la variable temps, justement (il serait toujours nul !). En fait, dans (11.83), Δt désigne un temps caractéristique d'évolution, c'est-à-dire un intervalle de temps pendant lequel l'état du système change notablement. Pour éviter les contre-sens ou les confusions, en notant τ une telle échelle de temps, il est préférable de retenir la forme suivante :

$$\Delta E \tau \sim \hbar . \quad (11.84)$$

À titre d'exemple, soit un paquet d'ondes décrit par une certaine fonction $\Psi(x, t)$, permettant de calculer sans ambiguïté des écarts-types Δx et Δp et aussi une impulsion moyenne³⁷ p_0 . À ce paquet est aussi associé un écart-type d'énergie ΔE , que l'on peut calculer exactement, mais qu'il est plus simple d'estimer en écrivant :

$$\Delta E \sim \left(\frac{dE}{dp} \right)_{p_0} \Delta p = \left(\frac{d\omega}{dk} \right)_{k_0} \Delta p \equiv v_g \Delta p . \quad (11.85)$$

La vitesse de groupe v_g est la vitesse de déplacement du paquet d'ondes ; le rapport $\Delta x/v_g$ est le temps nécessaire pour que le paquet se déplace d'une distance égale à sa largeur : c'est bien un temps caractéristique d'évolution, du genre τ dans (11.84). De ce fait :

$$\Delta E \sim v_g \Delta p \sim \frac{\Delta x}{\tau} \Delta p . \quad (11.86)$$

En utilisant maintenant $\Delta x \Delta p \sim \hbar$, la relation (11.84) est reproduite.

³⁶mais, dans la doctrine classique, on admet qu'elle existe.

³⁷En gros, $k_0 = \hbar^{-1}p_0$ donne le maximum de la transformée de Fourier spatiale de Ψ .

Chapitre 12

Magnétisme atomique

Le but de ce chapitre est d'une part de rappeler les origines du magnétisme classique¹, d'autre part de présenter l'expérience fondamentale de Stern et Gerlach, qui apporte la preuve de la quantification du moment cinétique (quantification spatiale). L'analyse de cette expérience confirme par ailleurs la nécessité d'introduire des idées radicalement nouvelles, en particulier la notion de grandeurs incompatibles non mesurables simultanément².

12.1 Magnétisme classique

12.1.1 Moment magnétique

Les sources du magnétisme classique ont été imaginées par Ampère dès 1821 ; pour Ampère, la brique fondamentale de magnétisme est une petite boucle de courant (*ampérien*). Il vaut la peine de remarquer que, dans cette conception, la source du magnétisme est un objet complexe, muni d'une structure interne (représentée essentiellement par un *vecteur* représentant une rotation), au contraire de la source de l'électricité une simple charge ponctuelle, scalaire par nature³. Ultérieurement, avec l'élaboration d'un modèle atomique planétaire, on en est venu tout naturellement à identifier les ampériens avec le mouvement confiné des électrons autour du noyau. Dans la suite, pour simplifier, on considérera un tel atome, muni d'un seul électron. Il convient par ailleurs de garder à l'esprit que, pour des champs usuels, les effets magnétiques restent très petits devant les forces électrostatiques. Enfin, s'il peut paraître obsolète, à ce stade de l'exposé, de revenir à un modèle mécanique d'atome, ceci est pourtant nécessaire pour rappeler les bases classiques du magnétisme – ne serait-ce que pour mieux réaliser en quoi elles sont fondamentalement *illusoires*.

Soit donc un électron (masse m , charge $e < 0$) en rotation autour d'un noyau supposé fixe. Ce système est supposé robuste et "rigide" : le mouvement de l'électron résulte d'un équilibre stable entre l'attraction du noyau et la force centrifuge. Le moment cinétique de l'électron est une constante du mouvement (force centrale) et la trajectoire est plane ; sans nuire à l'essentiel, on la prend circulaire. La fréquence de rotation est très élevée (10^{14} - 10^{15} Hz)⁴.

Ce système possède une propriété vectorielle, un moment magnétique, dont on va voir qu'il s'identifie précisément avec la définition donnée en Électromagnétisme à propos des boucles de circuit macroscopiques.

¹et de rappeler, au passage, que la physique classique est incapable d'expliquer l'existence même du magnétisme.

²notion déjà été rencontrée à propos des fentes d'Young, où on a vu que la mesure de la position, par la diffusion d'un photon, entraîne une perturbation incontrôlable sur l'impulsion.

³Tout porte à croire à l'inexistence du monopôle magnétique.

⁴La considération d'une orbite circulaire simplifie grandement la présentation, qui se veut de toute façon sommaire puisque l'on garde l'esprit l'incompatibilité entre la physique classique et l'existence même du magnétisme. Quoi qu'il en soit, on sait généraliser tous les résultats en invoquant le théorème de Larmor ; selon ce théorème, le mouvement en présence d'un champ magnétique, observé dans le référentiel tournant à la fréquence de Larmor γB autour du champ (qui n'est donc pas un référentiel d'inertie), est exactement le même que dans un référentiel d'inertie en l'absence de champ. L'application répétée de ce théorème reproduit l'essentiel des résultats présentés ici plus simplement.

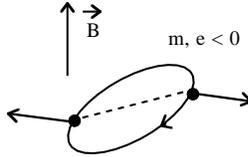


Figure 12.1: Représentation classique de l'ampérien.

Appliquons un (petit) champ magnétique à l'atome ; outre la force centrale due au noyau, l'électron est en plus soumis la force de Lorentz :

$$\vec{F} = e \vec{v} \times \vec{B} = -|e| \vec{v} \times \vec{B} . \quad (12.1)$$

En deux points diamétralement opposés de la trajectoire, les vitesses sont opposées et donc la force de Lorentz a deux valeurs opposées. Le mouvement dynamique interne de l'atome étant supposé rigide, l'application du champ magnétique provoque l'apparition d'un couple qui tend à faire tourner le plan de l'orbite de façon à le rendre perpendiculaire au champ appliqué.

À ce stade, il est utile de fixer les ordres de grandeur. Le rapport entre les intensités des deux forces en présence est (R est le rayon de l'atome) :

$$\phi \equiv \frac{F_{\text{Lorentz}}}{F_{\text{noyau}}} \sim \frac{|e|vB}{e'^2/R^2} = 4\pi\epsilon_0 \frac{vR^2B}{|e|} . \quad (12.2)$$

La vitesse v s'estime en introduisant la fréquence de rotation de l'électron, $\nu = v/(2\pi R)$; avec $\nu \sim 10^{15}$ Hz et $R \sim 1 \text{ \AA}$, on trouve $\phi \sim 4 \times 10^{-6} B$: avec un champ usuel, $\phi \ll 1$: l'application d'un champ magnétique ordinaire constitue bien une très petite perturbation du mouvement de l'électron.

La présence de deux forces d'intensité très différentes introduit dans la dynamique deux échelles de temps nettement différenciées⁵. En pareil cas, la description peut procéder en deux temps. On commence par traiter les variables rapides, en oubliant la (petite) perturbation. Ceci définit le mouvement "interne" et fait apparaître tout naturellement des grandeurs physiques intrinsèques, considérées par la suite comme des attributs donnés une fois pour toutes. Dans le cas présent, il s'agit principalement du moment cinétique \vec{J} résultant de la rotation rapide de l'électron (en l'absence du champ, \vec{J} est une constante du mouvement). Ceci étant fait, on introduit la petite perturbation. En général, une grandeur qui était constante ne l'est plus ; mais parce que la perturbation est *faible*, l'écart à la constance dans le temps sera petit, autrement dit la variation dans le temps sera *lente*. Ici, \vec{J} va se mettre à varier dans le temps à cause de \vec{B} : c'est précisément la précession de Larmor, voir sous-section 12.1.2. En définitive, le champ magnétique provoque un lent mouvement en bloc de l'atome⁶ qui se superpose au mouvement interne de rotation très rapide.

Une telle approche porte généralement le nom d'approximation adiabatique, pour rappeler le fait que les variables rapides s'adaptent à tout instant à un "champ" lentement variable et que, réciproquement, les variables lentes ne voient que le mouvement moyenné dans le temps des variables rapides⁷.

Soit \vec{n} la direction de la normale au plan de la trajectoire circulaire (normale positive) et θ l'angle entre \vec{n} et le champ magnétique appliqué, \vec{B} . Un calcul simple montre que la moyenne dans le temps (sur une période rapide de la rotation électronique) du moment de la force de Lorentz, \vec{M} , est un vecteur perpendiculaire à \vec{n} et à \vec{B} , le trièdre $(\vec{n}, \vec{B}, \vec{M})$ étant direct. La norme de ce vecteur est :

$$|\vec{M}| = \frac{1}{2} eR^2\omega\mathcal{B} \sin \theta . \quad (12.3)$$

⁵Il est facile de comprendre pourquoi l'application d'une petite perturbation se traduit par un effet global à variation lente du mouvement mécanique non perturbé. Il suffit d'imaginer que l'on a décomposé en série de Fourier le mouvement non perturbé, caractérisé par des pulsations propres ω_{n0} et les périodes $T_{n0} = 2\pi/\omega_{n0}$. La petite perturbation modifie (légèrement) ces pulsations (schématiquement, $\omega_n = \omega_{n0} + \delta\omega_{n0}$), d'où l'apparition de modulations temporelles à la période $2\pi/\delta\omega_{n0} \gg T_{n0}$. Un exemple connu est l'avance du périhélie de Mercure, exact analogue des corrections relativistes obtenues par Sommerfeld et expliquant la structure fine de la raie H_α .

⁶C'est précisément le fait de moyenner dans le temps les variables rapides qui permet de définir un système "rigide" bien défini par les grandeurs émergent de cette moyenne. Ces grandeurs constituent l'image effective des variables dynamiques que la moyenne temporelle évacue de fait.

⁷On rencontre ce type d'approximation lors de l'étude des molécules (approximation dite de Born et Oppenheimer).

$\omega = 2\pi\nu$ est la pulsation du mouvement circulaire de l'électron ; il est naturel d'introduire le moment cinétique \vec{J} de module $J = mvR = mR^2\omega$, orienté comme \vec{n} . Finalement, en termes de vecteurs, on a :

$$\vec{M} = \frac{e}{2m} \vec{J} \times \vec{\mathcal{B}} . \quad (12.4)$$

Par identification ($\vec{M} = \vec{\mu} \times \vec{\mathcal{B}}$), on en déduit que le moment magnétique de la "boucle atomique" est :

$$\vec{\mu} = \frac{e}{2m} \vec{J} \equiv \gamma \vec{J} , \quad (12.5)$$

où γ est appelé *facteur gyromagnétique* ; avec la définition (12.5), γ est négatif⁸, le moment cinétique et le moment magnétique étant toujours de sens contraires pour un électron. Un atome planétaire exposé à un champ magnétique est donc soumis à un couple, simplement relié au moment cinétique du mouvement interne et, *via* le facteur gyromagnétique, au moment magnétique de l'atome.

La description précédente en termes de forces (ou de couples) peut être complétée en termes d'énergie, comme toujours. Une fois moyennées les variables rapides, le couple exercé par l'atome sur le champ est donné par \vec{M} , (12.4). Quand le système (atome + champ) est isolé, son énergie est constante, donc l'angle θ ne change pas. Si, de l'extérieur, on applique un couple pour faire varier l'angle de θ à $\theta + d\theta$, le travail *reçu* par le système est $dW = |\mu|B \sin\theta d\theta$ (le signe vient du fait que pour *écarter* le moment magnétique de l'axe, il faut *dépenser* de l'énergie : le travail reçu par le système est bien positif quand $d\theta$ l'est et inversement⁹). La variation correspondante d'énergie est $dE = dW$, travail des forces extérieures. D'où, pour un champ constant et un moment magnétique invariable :

$$dE = |\vec{\mu}| \vec{\mathcal{B}} \sin\theta d\theta \equiv -d(\vec{\mu} \cdot \vec{\mathcal{B}}) \iff E = -\vec{\mu} \cdot \vec{\mathcal{B}} . \quad (12.6)$$

Cette expression montre que la position d'équilibre stable correspond à l'alignement entre $\vec{\mu}$ et $\vec{\mathcal{B}}$, soit $\theta = 0$.

Il est utile de remarquer que la définition microscopique du moment magnétique, (12.5), est identique à celle introduite en Électromagnétisme. On y définit le moment magnétique d'un circuit de surface S parcouru par un courant d'intensité I comme le vecteur normal au circuit orienté, de module égal à IS . I ($= dq/dt$) est la charge traversant la section droite du fil par unité de temps ; dans le cas du "circuit atomique", il passe, en un point donné de la trajectoire, une charge élémentaire toutes les $1/\nu$ secondes : l'intensité "atomique" est donc $|e|\nu$; quant à la surface, elle vaut πR^2 . En ne considérant que les valeurs absolues, il vient $IS = |e|\nu\pi R^2$, soit finalement¹⁰ $IS = |e|/(2m)J$: les deux définitions, macroscopique et microscopique, sont strictement équivalentes.

12.1.2 Précession de Larmor

Les forces (et le couple) étant précisés, on peut décrire la dynamique (lente) de l'atome en présence du champ magnétique, l'existence de μ résultant du mouvement électronique rapide moyenné étant admise une fois pour toutes. $\vec{\mathcal{B}}$ introduit une direction privilégiée, prise comme axe Oz dans la suite.

Quand $\vec{\mathcal{B}} = 0$, le moment cinétique de l'électron est une constante du mouvement et on a :

$$\frac{d\vec{J}}{dt} = 0 \quad (\text{champ nul}) . \quad (12.7)$$

En présence du champ, le second membre n'est plus nul (théorème du moment cinétique), et contient le couple $\vec{M} = \vec{\mu} \times \vec{\mathcal{B}}$ obtenu plus haut :

$$\frac{d\vec{J}}{dt} = \vec{\mu} \times \vec{\mathcal{B}} \iff \frac{d\vec{J}}{dt} = \gamma \vec{J} \times \vec{\mathcal{B}} \quad (\text{champ non-nul}) . \quad (12.8)$$

Cette équation donne un mouvement gyroscopique : le vecteur \vec{J} est constant en norme, fait un angle constant avec le champ et tourne autour de ce dernier à la pulsation $\omega_L = |\gamma|\mathcal{B}$, appelée *pulsation de Larmor*.

⁸Attention à cette convention, qui fluctue d'un auteur à l'autre.

⁹À condition que l'atome et le champ ne constituent pas un système isolé, le couplage $-\vec{\mu} \cdot \vec{\mathcal{B}}$, tend à aligner moment magnétique et champ pour abaisser l'énergie du sous-système (atome+champ).

¹⁰ J est le moment cinétique mvR .

En effet, en multipliant scalairement (12.8) membre à membre par \vec{J} , on obtient :

$$\vec{J} \frac{d\vec{J}}{dt} = \gamma \vec{J} \cdot \vec{J} \times \vec{B} = 0 \iff \vec{J}^2 = C^{\text{ste}} . \quad (12.9)$$

Par ailleurs, (12.8) montre que la dérivée de \vec{J} est perpendiculaire à \vec{B} , donc à Oz : la dérivée de J_z est donc nulle, de sorte que la composante J_z est aussi une constante du mouvement. Enfin, les composantes transverses obéissent à :

$$\frac{dJ_x}{dt} = \gamma \mathcal{B} J_y \equiv \omega_L J_y , \quad \frac{dJ_y}{dt} = -\omega_L J_x . \quad (12.10)$$

En posant $J_+ = J_x + iJ_y$ on trouve $\dot{J}_+ = i\omega_L J_+$, ce qui conduit à :

$$J_x(t) = J_x(0) \cos \omega_L t - J_y(0) \sin \omega_L t , \quad J_y(t) = J_x(0) \sin \omega_L t + J_y(0) \cos \omega_L t . \quad (12.11)$$

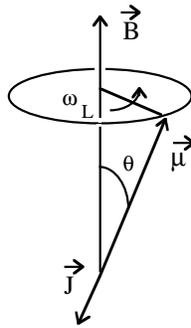


Figure 12.2: Précession de Larmor.

Les composantes transverses ont donc bien un mouvement circulaire uniforme à la pulsation de Larmor ω_L . Evidemment, cette pulsation est très faible devant la pulsation du mouvement interne de l'électron ; avec $\nu = 10^{15}$ Hz, on a :

$$\frac{\omega_L}{\omega} = \frac{\gamma \mathcal{B}}{2\pi \times 10^{15}} = \frac{1.6 \times 10^{-19}}{2\pi \times 2 \times 9 \times 10^{-31} \times 10^{15}} \mathcal{B} \simeq 1.4 \times 10^{-5} \mathcal{B} . \quad (12.12)$$

Avec $\mathcal{B} = 0.1$ T, pendant que le gyroscope fait un tour, l'électron en fait presque un million ! Le mouvement gyroscopique implique donc ici les variables lentes.

12.1.3 Paramagnétisme classique

L'hypothèse du mouvement des électrons au sein de l'atome permet de comprendre l'existence de la brique élémentaire du magnétisme¹¹ (boucle de courant atomique) et donc d'esquisser la théorie du paramagnétisme, due à Langevin. Il s'agit d'expliquer pourquoi certaines substances soumises à un champ extérieur présentent une aimantation induite M proportionnelle au champ appliqué, et dépendant de la température.

Au sein d'une substance dont les molécules possèdent un moment magnétique, chaque moment magnétique est soumis à deux tendances antagonistes : son couplage avec le champ tend à le rendre parallèle à ce dernier, alors que l'agitation thermique tend au contraire à l'en écarter en moyenne¹² ; cette compétition est décrite quantitativement par les probabilités de Boltzmann, $\propto e^{-\beta E}$ avec $\beta = 1/(k_B T)$ et $E = -\vec{\mu} \cdot \vec{B}$. Au total, l'aimantation par unité de volume, \mathcal{M}_V , évidemment dirigée le long du champ appliqué, est donnée par :

$$\mathcal{M}_V = C \int_0^{2\pi} \int_0^\pi \sin \theta d\theta (n\mu \cos \theta) e^{+\beta \mu \mathcal{B} \cos \theta} , \quad (12.13)$$

¹¹en oubliant toutefois le théorème de Miss van Leeuwen ...

¹²Ce sont bien les petits apports d'énergie sous forme de fluctuations thermiques qui permettent à l'angle θ d'un moment donné de varier à tout instant. Cette agitation thermique tend à l'uniformisation, ce qui, comparé à une situation d'alignement, montre que l'agitation thermique tend à éloigner μ de l'axe du champ.

où n est le nombre de moments magnétiques par unité de volume. C est la constante de normalisation des probabilités :

$$C \int_0^{2\pi} 2\pi \sin \theta d\theta e^{x \cos \theta} = 1 . \quad (12.14)$$

Avec (12.13) et (12.14), on peut finalement écrire :

$$\mathcal{M}_V = n\mu \frac{\partial}{\partial x} \int_0^{2\pi} 2\pi \sin \theta d\theta e^{+\beta\mu\mathcal{B} \cos \theta} \quad (x = \beta\mu\mathcal{B}) . \quad (12.15)$$

L'intégrale vaut $(2/x) \sinh x$, de sorte que :

$$\mathcal{M}_V = n\mu \left(\coth \beta\mu\mathcal{B} - \frac{1}{\beta\mu\mathcal{B}} \right) \equiv n\mu L(\beta\mu\mathcal{B}) . \quad (12.16)$$

L s'appelle fonction de Langevin¹³. Les comportements limites de la magnétisation sont :

$$\mathcal{M}_V \simeq \begin{cases} \frac{1}{3} n \frac{\mu^2 \mathcal{B}}{k_B T} & \text{si } k_B T \gg \mu\mathcal{B} \\ n\mu & \text{si } k_B T \ll \mu\mathcal{B} \end{cases} . \quad (12.17)$$

À haute température, l'agitation thermique est violente et la magnétisation est faible ; c'est évidemment le contraire à basse température, où les moments sont presque tous alignés en moyenne le long du champ et ont tendance à se "geler". À toute température, \mathcal{M}_V augmente avec \mathcal{B} . Cette aptitude se mesure par la susceptibilité magnétique (ici, c'est plus précisément la susceptibilité paramagnétique), χ , définie comme la constante de proportionnalité¹⁴ entre magnétisation et champ à petit champ ou, si on préfère :

$$\chi = \left(\frac{\partial \mathcal{M}_V}{\partial \mathcal{B}} \right)_{\mathcal{B}=0} . \quad (12.18)$$

Le calcul à partir de (12.16) donne :

$$\chi = \frac{n\mu^2}{3k_B T} . \quad (12.19)$$

χ est inversement proportionnelle à la température (loi de Curie)

Le paramagnétisme n'existe que pour les substances dont les constituants élémentaires possèdent un moment magnétique μ , ce qui est relativement exceptionnel. Il existe au contraire une forme de magnétisme universel – le diamagnétisme – résultant du fait que la matière est constituée de charges électriques en mouvement confiné : l'application d'un champ induit des moments magnétiques *opposés* au champ (c'est l'expression microscopique de la loi de Lenz). L'effet est petit, de sorte que la susceptibilité magnétique correspondante (qui est négative) est aussi très petite, environ 1000 fois plus faible que son homologue paramagnétique. On retiendra que, si le diamagnétisme est universel, il est complètement occulté par le paramagnétisme quand celui-ci existe.

12.1.4 Expériences de Einstein - de Haas et de Barnett

Soit un solide, de forme cylindrique pour simplifier. Lorsqu'il tourne, sa rotation est caractérisée par un moment cinétique macroscopique \vec{J}_S ; l'énergie cinétique de cette rotation en bloc est $\vec{J}_S^2 / (2I)$ où I est le moment d'inertie par rapport à l'axe de rotation.

¹³La théorie de Langevin donne de bons résultats à haute température (quand des effets quantiques peuvent être ignorés) ; à basse température, la quantification spatiale du moment cinétique finit par apparaître et conduit à des fonctions se substituant à L , appelées fonctions de Brillouin, B_J . En outre, comme on le verra ensuite, la valeur du facteur gyromagnétique classique est incorrecte.

¹⁴Ici, l'aimantation moyenne macroscopique provoquée par l'application d'un petit champ magnétique est de fait *proportionnelle* à ce champ. Ceci est caractéristique d'un système non-critique ; dans le cas contraire, l'aimantation est souvent reliée au champ par un exposant $\beta < 1$, $\mathcal{M}_V \propto \mathcal{B}^\beta$, $\beta < 1$, de sorte que la susceptibilité définie en (12.18) est infinie. Une susceptibilité infinie est la caractéristique d'un système critique.

Supposons de surcroît que les atomes constituant le solide possèdent chacun un moment magnétique $\vec{\mu}_i$ et ont tous le même facteur gyromagnétique γ . A chacun de ces moments correspond un moment cinétique $\vec{J}_{\text{orb}, i}$, puisqu'à tout moment cinétique est associé un moment magnétique et inversement : chaque atome est donc une petite toupie ; la somme de ces moments cinétiques donne un moment total noté \vec{J}_{orb} , relié à la magnétisation macroscopique par :

$$\vec{J}_{\text{orb}} = \gamma^{-1} \vec{M} . \quad (12.20)$$

Le moment cinétique total du barreau paramagnétique est donc :

$$\vec{J} = \vec{J}_S + \vec{J}_{\text{orb}} . \quad (12.21)$$

Lorsque le corps est seulement soumis à un champ magnétique \vec{B} , les forces appliquées sont les forces de Lorentz agissant microscopiquement sur les différents atomes. Ce sont ces forces qui se combinent pour donner lieu à l'apparition d'une aimantation macroscopique \vec{M} et à un moment (couple) égal à $\vec{M} \times \vec{B}$. Comme \vec{M} est parallèle au champ magnétique ($\vec{M} = \chi \vec{B}$), ce moment est nul. Le théorème du moment cinétique donne :

$$\frac{d\vec{J}}{dt} = \vec{M} \times \vec{B} = 0 \iff \frac{d}{dt} (\vec{J}_S + \vec{J}_{\text{orb}}) = 0 . \quad (12.22)$$

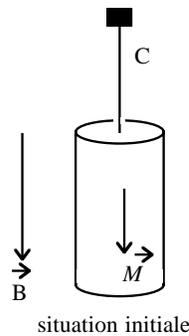


Figure 12.3: Barreau cylindrique paramagnétique au repos.

L'équation (12.22) montre que toute modification de l'aimantation du corps fait varier l'état de rotation macroscopique décrit par \vec{J}_S . En effet, si \vec{M} change de $\Delta \vec{M}$, alors \vec{J}_{orb} varie aussi d'après (12.20), mais comme $\vec{M} = \chi \vec{B}$, le second membre de (12.22) reste nul. Au total, la somme $\vec{J}_S + \vec{J}_{\text{orb}}$ doit rester constante : \vec{J}_S varie donc de $-\gamma^{-1} \Delta \vec{M}$.

Cette idée [14] a été exploitée par Einstein et de Haas en 1915 pour remonter directement aux sources du magnétisme et en particulier au facteur gyromagnétique γ . Un barreau aimanté est suspendu à un fil de torsion de constante C et est soumis à un champ magnétique. Initialement, le barreau est immobile, de sorte que le moment cinétique total se réduit à \vec{J}_{orb} . A un certain instant, on inverse le sens du champ, ce qui inverse la magnétisation et \vec{J}_{orb} ; pour que la somme $\vec{J}_S + \vec{J}_{\text{orb}}$ reste constante, il faut que \vec{J}_S , qui était nul, prenne une valeur finie ; explicitement, la loi de conservation du moment cinétique total (avant=après) s'écrit :

$$\vec{0} + \vec{J}_{\text{orb}} = \vec{J}_S + (-\vec{J}_{\text{orb}}) \iff \vec{J}_S = 2\vec{J}_{\text{orb}} = \frac{2}{\gamma} \vec{M} . \quad (12.23)$$

Ainsi, l'inversion soudaine du champ appliqué met le barreau paramagnétique en rotation. Cette expérience permet :

- de vérifier que γ est bien négatif, en analysant le sens de rotation provoqué par l'inversion du champ.
- de mesurer l'amplitude angulaire A du mouvement de rotation, directement reliée à γ . Un calcul simple montre que :

$$A = \frac{2}{\sqrt{IC}} \frac{\mathcal{M}}{|\gamma|} . \quad (12.24)$$

L'expérience apporte aussi la confirmation du lien indissociable entre le moment magnétique et le moment cinétique. En outre, elle révèle un désaccord systématique sur la valeur de γ : le plus souvent, le $|\gamma|$ mesuré est plus grand d'un facteur 2 par rapport à sa valeur classique (12.5). L'écart est parfois un certain nombre rationnel simple, plus tard appelé facteur de Landé, g_L .

Les expériences de Einstein - de Haas mettent donc en évidence la mise en rotation d'un barreau paramagnétique sous l'effet d'une modification brusque du champ appliqué – et donc de l'aimantation. Elles ont leur symétrique : l'aimantation d'un corps par suite de sa mise en rotation ; ce phénomène est à la base des expériences de Barnett, effectuées en 1914, qui conduisent au même type d'observations.

12.2 Expérience de Stern et Gerlach

Cette expérience (1921¹⁵), est l'une des expériences cruciales de la Physique Atomique. En effet, c'est elle qui démontra l'existence de la quantification du moment cinétique (quantification spatiale) ; en outre, elle sert de modèle à la discussion des fondements de l'interprétation de la Mécanique Quantique.

Dans la discussion précédente, les objets microscopiques étaient supposés munis d'un moment magnétique propre résultant du mouvement orbital des électrons. Les expériences décrites donnent accès au facteur gyromagnétique, mais pas à ce moment magnétique individuel, sauf à faire une théorie *microscopique* de l'aimantation macroscopique \mathcal{M} .

Quelle que soit l'image mécanique que l'on se fait de l'atome, il reste que, situé dans un champ *uniforme* dans l'espace, le petit circuit qu'il constitue est soumis à une force totale nulle. Stern et Gerlach ont eu l'idée d'exploiter le fait que, si au contraire le champ magnétique est spatialement inhomogène, il apparaît une force résultante non nulle : outre la précession (la valeur du champ étant moyennée sur l'espace atomique), il existe une force résultante liée au gradient de champ magnétique. Il est important de noter que ce gradient établit un couplage explicite entre la dynamique *interne* de l'atome (représenté globalement par le moment magnétique de l'atome) et son mouvement *externe* (la position de l'atome dans l'espace – et sa vitesse). C'est l'existence de ce couplage qui permet, observant une variable externe (la position de l'atome) d'en déduire la valeur de la variable interne (le moment magnétique de l'atome) ; en ce sens, une mesure de la position (point d'arrivée sur un écran) permet de remonter à la valeur du moment magnétique de l'atome.

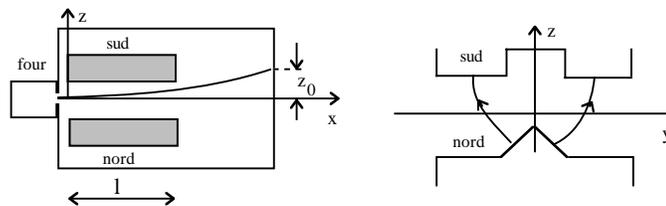


Figure 12.4: Représentation schématique de l'appareil de Stern et Gerlach.

Il n'est évidemment pas possible de fabriquer des gradients intenses à l'échelle de l'atome, dont l'échelle de longueur est l'angström. En revanche, si un atome se "promène" dans l'espace sur une distance macroscopique, il est alors tout à fait envisageable d'obtenir un effet mesurable : c'est en se déplaçant d'un point à l'autre de l'espace que l'atome ressent la variation du champ. C'est l'idée de départ de l'expérience de Stern et Gerlach, que l'on va décrire de façon semi-qualitative et dans une conception classique (prévision de la trajectoire par l'analyse qualitative des forces). Les prévisions classiques sont en totale contradiction avec l'expérience¹⁶.

Des atomes paramagnétiques (d'argent par exemple) sont émis par un four porté à haute température (typiquement 1000 K) et sont injectés à travers une fente verticale de largeur très faible suivant Oy dans

¹⁵100 ans exactement après l'hypothèse d'Ampère.

¹⁶Le traitement quantique de l'expérience montre que tout paquet d'ondes initial (représentant un atome à sa sortie du four) développe deux "bosses" (d'amplitudes de probabilité) lors de son passage à travers l'appareil, nettement différenciées si celui-ci soit assez long. On peut donc prévoir de fait que les impacts se produisent soit en haut de l'écran, soit en bas.

une grande enceinte où règne un très bon vide. Les vitesses initiales sont donc réparties aléatoirement, seuls les atomes dont la vitesse est orientée dans un certain secteur défini par la géométrie précise de l'entrée de l'appareil peuvent effectivement pénétrer dans l'enceinte. Les atomes traversent une région de grande longueur L , où existe un champ magnétique rendu inhomogène par la forme asymétrique des pièces polaires d'un aimant ; la fente d'entrée est située dans le plan de symétrie de l'aimant, de sorte que les atomes sont essentiellement sous l'influence d'un champ inhomogène dirigé le long de Oz et prenant des valeurs différentes suivant l'altitude mesurée par z . On va voir que les atomes sont soumis à une force parallèle à Oz et sont donc déviés hors de l'axe Ox . Les points d'impact des atomes sont observés sur un écran maintenu à basse température pour éviter la diffusion en surface des atomes ainsi recueillis.

Un atome donné possède un moment magnétique intrinsèque¹⁷ $\vec{\mu}$ et est couplé au champ magnétique ; l'interaction est donnée par $-\vec{\mu} \cdot \vec{\mathcal{B}}$. La force totale agissant ce petit gyroscope est donc :

$$\vec{F} = -\vec{\nabla}(-\vec{\mu} \cdot \vec{\mathcal{B}}) = \vec{\nabla}(\vec{\mu} \cdot \vec{\mathcal{B}}) . \quad (12.25)$$

Ainsi, les composantes de la force suivant les axes d'intérêt sont :

$$F_x = \frac{\partial(\vec{\mu} \cdot \vec{\mathcal{B}})}{\partial x} = \mu_x \frac{\partial \mathcal{B}_x}{\partial x} + \mu_y \frac{\partial \mathcal{B}_y}{\partial x} + \mu_z \frac{\partial \mathcal{B}_z}{\partial x} , \quad (12.26)$$

$$F_z = \frac{\partial(\vec{\mu} \cdot \vec{\mathcal{B}})}{\partial z} = \mu_x \frac{\partial \mathcal{B}_x}{\partial z} + \mu_y \frac{\partial \mathcal{B}_y}{\partial z} + \mu_z \frac{\partial \mathcal{B}_z}{\partial z} . \quad (12.27)$$

Chaque atome a un mouvement de précession à la fréquence de Larmor ω_L correspondant à la valeur du champ à l'endroit où il se trouve ; les composantes μ_x et μ_y tournent donc rapidement autour de l'axe Oz à une fréquence ν_L de l'ordre de 10^{10} Hz, cependant que la composante μ_z est une constante du mouvement. Sur une distance macroscopique, le mouvement de précession est ultra-rapide. En effet, si les atomes ont une vitesse de l'ordre de 500 m/s (valeur typique), ils mettent un temps égal à $10^{-2}/500$ s pour parcourir 1 cm, soit 20 μ s. En 20 μ s, il se produit $20 \times 10^4 = 200000$ rotations autour de l'axe du champ. Il en résulte que, pour une observation macroscopique, les composantes de la force \vec{F} peuvent être remplacées par leurs valeurs moyennes dans le temps sur une période de la précession, moyennes qui sont nulles ; seuls les termes proportionnels à μ_z survivent à la moyenne et la force effective sur l'atome a pour composantes :

$$\bar{F}_x = \mu_z \frac{\partial \mathcal{B}_z}{\partial x} , \quad \bar{F}_y = \mu_z \frac{\partial \mathcal{B}_z}{\partial y} , \quad \bar{F}_z = \mu_z \frac{\partial \mathcal{B}_z}{\partial z} . \quad (12.28)$$

Si les atomes sont injectés précisément dans le plan de symétrie vertical de l'appareil, la composante moyenne \bar{F}_y est nulle : par symétrie, la dérivée $\partial \mathcal{B}_z / \partial y$ est nulle dans ce plan puisque \mathcal{B}_z est une fonction paire¹⁸ en y . Par ailleurs, en rendant l'aimant assez long et bien homogène le long de Ox , on peut négliger le terme en $\partial \mathcal{B}_z / \partial x$: il y a une quasi-symétrie de translation le long de Ox de sorte que toute dépendance en x se réduit aux effets de bord ; de toute façon, l'accélération correspondant à cette composante de la force est totalement négligeable, puisque les atomes rentrent dans l'aimant avec une vitesse de l'ordre de 500 m/s ; enfin, ce qui se passe dans l'axe de l'appareil (suivant Ox) est sans grand intérêt. Au total, la force effective subie par chaque atome peut être réduite à sa seule composante \bar{F}_z et, en pratique, *tout se passe comme si* le gradient de champ magnétique était parallèle¹⁹ à Oz .

Chaque atome est donc de fait soumis à une force verticale, que l'on peut d'ailleurs supposer constante si l'inhomogénéité du champ magnétique peut être réduite à un terme linéaire (dans le cas contraire, $\partial \mathcal{B}_z / \partial z$ dépend encore de l'espace – ce détail ne joue que si l'on veut calculer précisément la trajectoire). Le point important est de constater que le signe de la force, donc de la déviation, dépend de la valeur (algébrique) de μ_z . *A priori*, pour chaque moment magnétique individuel, *toutes les orientations sont classiquement possibles*, avec une certaine probabilité²⁰. Pour un grand nombre d'impacts, on devrait donc observer une tache diffuse

¹⁷C'est bien le cas de l'atome d'Ag. Le nombre d'électrons étant *impair*, son moment cinétique total est forcément demi-entier en raison du spin électronique.

¹⁸Ceci suppose que la dépendance de \mathcal{B}_z par rapport à y n'est pas du genre $C|y|$.

¹⁹*Stricto sensu*, le gradient $\vec{\nabla} \mathcal{B}$ ne peut posséder cette propriété, comme le montre la manipulation des relations fondamentales $\vec{\nabla} \times \vec{\mathcal{B}} = 0$ et $\vec{\nabla} \cdot \vec{\mathcal{B}} = 0$.

²⁰Savoir ce que vaut cette probabilité n'est pas si simple. Le faisceau est soumis à un champ variable dans l'espace, de sorte qu'il faut comparer l'échelle de temps de la variation correspondante avec l'échelle de temps de relaxation thermique au sein du faisceau. Si on peut considérer que le faisceau est en équilibre local, alors la probabilité est le poids de Boltzmann calculé avec le champ considéré. Ces incertitudes ne jouent pas un rôle capital dans l'analyse en cours.

d'intensité à peu près constante sur un segment limité essentiellement par l'ouverture de sortie de l'aimant et les valeurs extrêmes de μ_z . Les bords de la tache ne peuvent être abrupts, ne serait-ce que parce qu'il convient de faire une statistique sur les vitesses (thermiques) initiales, dispersées en module et (un peu) en direction. Quoi qu'il en soit dans le détail, la prévision classique est schématisée sur la partie gauche de la figure (12.5).

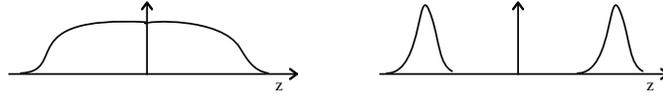


Figure 12.5: Prévision classique (à gauche) et résultat (à droite) de l'expérience de Stern et Gerlach.

Le résultat de l'expérience est en désaccord *total* avec cette prévision : dans le cas de l'atome d'argent, on observe deux taches symétriques et franchement séparées (figure 12.5, à droite), pourvu que l'appareil soit assez long et/ou que la vitesse initiale typique des atomes ne soit pas trop grande (au total, c'est le temps d'interaction entre un atome et l'aimant qui doit être assez long).

Cette expérience fondamentale est confirmée par d'autres, effectuées sur d'autres atomes. Dans tous les cas on obtient des *taches bien séparées*, en nombre pair ou impair suivant l'atome utilisé, mais toujours symétriques par rapport au point milieu. Ces observations sont totalement inexplicables en théorie classique.

L'interprétation détaillée des résultats de l'expérience est sinon complexe, du moins subtile ([15], section 22.5). Les atomes sont des objets quantiques et, *a priori*, la quantification doit porter autant sur les degrés de liberté externes (le mouvement en bloc de l'atome dans l'espace) que sur sa dynamique interne. Il convient donc, en toute rigueur, de définir un paquet d'ondes pour représenter le mouvement externe de l'atome, puis, évidemment, de décrire convenablement la physique interne de celui-ci.

Quoi qu'il en soit, l'utilisation d'atomes assez massifs²¹ permet de se convaincre que la "trajectoire" de l'atome à travers l'appareil va donner lieu à des effets bien localisés, sans pour autant que l'on puisse prédire à l'avance – et en fait décrire – (tout comme pour les fentes d'Young), le destin individuel d'un atome donné. Le paquet d'ondes incontournable représentant un atome en voyage à travers l'appareil est en effet caractérisé par des longueurs très petites par rapport à toutes les autres longueurs pertinentes (macroscopiques) de l'expérience. Pour le mouvement le long de Ox , il n'y a pas de problème de fond, puisque le mouvement n'est finalement intéressant que le long de l'axe perpendiculaire Oz . C'est donc suivant cette direction que la question précédente se pose vraiment. La relation de Heisenberg suivant Oz est :

$$\Delta z \Delta p_z \sim \hbar \iff \Delta z \Delta v_z \sim \frac{\hbar}{M} . \quad (12.29)$$

où M est la masse d'un atome. M valant environ 10^{-25} kg, il faut donc respecter l'ordre de grandeur suivant :

$$\Delta z \Delta v_z \sim 6 \times 10^{-9} \text{ m}^2/\text{s} . \quad (12.30)$$

La fente d'entrée est au mieux une fraction de millimètre, mettons 0,1 mm, ce qui donne la limite-butoir au sens de Heisenberg :

$$\Delta v_z \sim 6 \times 10^{-5} \text{ m/s} . \quad (12.31)$$

Les vitesses thermiques, à $T = 1000$ K, sont de l'ordre de plusieurs centaines de m/s ; on se doute qu'il est possible de réaliser un jet atomique avec une dispersion des vitesses de l'ordre de quelques m/s (ce qui suppose une résolution en vitesse de l'ordre de 1%), mais il est difficile de faire beaucoup mieux de sorte que, en pratique, $(\Delta v_z)_{\text{exp}} \sim 1$ m/s : on est donc très loin, dans les conditions ordinaires de l'expérience, de buter sur la limite de Heisenberg, (12.31) – ce qui ne signifie pas pour autant que la trajectoire de chaque atome est quasi-classique (voir ci-dessous). On note cependant que, pour obtenir une bonne séparation des deux taches, la déviation doit être assez importante : chaque atome doit donc interagir suffisamment *longtemps* avec le gradient de champ magnétique ; toutes choses égales par ailleurs, cette condition se résume en disant que l'appareil (de mesure)

²¹Pour l'argent : $Z = 47$ et la masse molaire est égale à 108 g.

doit être *macroscopique*. En tout état de cause, l'expérience constitue une mesure du moment magnétique de chaque atome pris individuellement : la mesure de la position de l'impact permet de trouver la valeur de μ_z . Le résultat est que, pour l'argent dans son état fondamental, les valeurs trouvées pour μ_z ne peuvent être que ± 1 , dans les bonnes unités : le moment magnétique est *quantifié*, et il en va de même pour la composante J_z du moment cinétique intrinsèque²² de l'atome.

L'existence de deux taches, et de deux seulement, montre donc que toute mesure du moment magnétique de l'atome ne peut produire que deux valeurs, contrairement à la vision classique. Si l'on maintient globalement, et de façon délibérément floue, l'idée d'une relation précise entre moment magnétique et moment cinétique, la conclusion générale qui s'impose est celle d'une quantification du moment cinétique lui-même, l'atome d'argent ne constitue en aucune façon une exception. La mesure de la projection du moment cinétique sur un axe donné ne peut conduire qu'à un nombre *fini* de valeurs discrètes. Dans le cas de l'atome d'argent, les deux seules possibilités sont "en haut ou en bas", que l'on peut noter + et - ; à ces deux possibilités doivent être associés deux états²³, notés u_+ et u_- .

Ceci ne veut pas dire que, avant toute mesure effective (tout impact), l'état de moment cinétique de l'atome est ou + ou -. En réalité, on n'en sait strictement rien²⁴ et, en toute rigueur, cette question n'a pas de réponse en Mécanique Quantique, où seules sont prévisibles les probabilités *d'observer* une valeur ou une autre ; les valeurs "réalisées" préalablement à toute mesure – si tant est qu'elles le sont – échappent à l'observation et sont de ce fait en dehors des prétentions prédictives de la Mécanique Quantique.

Imaginons qu'un atome soit dans l'état + au sortir du four, état représenté par une fonction d'onde $\psi_+(\vec{r}) = f(\vec{r})u_+$; f désigne la partie spatiale : ce peut (doit !) être un paquet d'ondes. Comme la composante suivant Oz ne change pas au cours du mouvement, cet atome ira sûrement se projeter sur la tache de l'écran correspondant à la valeur conventionnellement égale à +1, l'état qui s'est développé au cours du temps à partir de cet état initial ayant une "bosse" de localisation à cet endroit. De même, un atome sortant du four dans l'état - sera représenté au départ par une autre fonction d'onde, $\psi_-(\vec{r}) = f(\vec{r})u_-$; l'impact se produira cette fois au niveau de l'autre tache, -1. Il reste que ces deux états particuliers sont assez exceptionnels et leur considération exclusive serait arbitraire : après tout, juste avant l'entrée de l'aimant, il n'y a même pas de champ magnétique pour définir une direction remarquable et l'atome ne peut savoir qu'il va bientôt rencontrer un champ (inhomogène) dirigé suivant une direction précise (choisie comme axe Oz). Comme la théorie est fondamentalement linéaire, n'importe quelle combinaison linéaire convient tout autant ; l'état le plus général d'un atome à la sortie du four peut ainsi être pris sous la forme :

$$\Psi(\vec{r}) = f(\vec{r}) (c_+u_+ + c_-u_-) . \quad (12.32)$$

Pour un tel état, la valeur *moyenne* du moment cinétique est proportionnelle à :

$$|c_+|^2 - |c_-|^2 . \quad (12.33)$$

Comme $|c_+|^2 + |c_-|^2 = 1$ (normalisation des probabilités), l'expression (12.33) prend ses valeurs sur l'intervalle réel $[-1, +1]$. Une combinaison linéaire comme (12.32) contient potentiellement les deux résultats possibles avec des poids donnés par $|c_{\pm}|^2$. En effet, l'état qui se développe au cours du temps à partir de (12.32) est un état présentant, à la sortie de l'aimant, *deux bosses* de localisation, l'une vers la tache supérieure, l'autre vers la tache inférieure, correspondant respectivement au spin en haut et en bas²⁵. L'atome peut se trouver au voisinage de l'une de ces deux bosses²⁶, et tant qu'on ne l'a pas observé, il est impossible de dire où il est effectivement. Juste avant l'écran, la fonction d'onde présente deux maxima localisés, l'atome peut donc se manifester à un endroit ou un autre un peu plus tard, mais on ne peut dire où l'avance. Seules les *amplitudes* de probabilités ont une évolution totalement déterminée, pas les probabilités.

²²L'adjectif *intrinsèque* fait référence au mouvement des électrons dans l'atome, mais il a aussi une autre signification : Ag possède un nombre impair d'électrons. La somme de leurs moments cinétiques ordinaires (orbitaux) est nulle en moyenne dans l'état fondamental ; mais comme chaque électron possède également un moment cinétique *intrinsèque* (*spin*), valant $\hbar/2$, la somme en nombre impair de ces moments est forcément un demi-entier – égal en fait à $\hbar/2$ pour l'état fondamental. Les deux taches correspondent aux deux valeurs possibles de la projection du moment cinétique total ($J_z = \pm\hbar/2$).

²³On verra par la suite que u_{\pm} sont les états propres de la composante suivant Oz du moment cinétique total de l'atome dans son état fondamental.

²⁴Tout comme Heisenberg ne sait rien – et ne veut rien savoir – au sujet des trajectoires de l'électron au sein de l'atome.

²⁵si $\partial\mathcal{B}_z/\partial z > 0$.

²⁶Les deux bosses représentent une *potentialité de présence* de l'atome et présentent une cohérence de phase : si on retire l'écran et si on utilise par la suite un gradient en sens inverse, les deux bosses se recombinent en une bosse unique. A la question : *où est l'atome juste avant la matérialisation sur l'écran ?* nul ne peut répondre. La question est de même nature que de se demander par où passe l'électron dans l'expérience d'Young.

Quoi qu'il en soit, une chose est sûre : l'expérience de Stern et Gerlach montre que le moment cinétique d'un petit objet comme un atome ne peut avoir, en tant que résultat de sa mesure, qu'un ensemble fini discret de valeurs bien déterminées, contrairement à ce qu'affirme la théorie classique ; ceci est en harmonie avec les recettes de l'Ancienne Théorie des Quanta. Comme on le verra dans la suite, la Mécanique Quantique, elle, fournit un cadre cohérent où cette quantification spatiale est le résultat des équations de départ²⁷.

De surcroît, en anticipant sur la théorie du moment cinétique, cette expérience impose à nouveau la notion de grandeurs non observables simultanément : les composantes J_x , J_y et J_z du moment cinétique satisfont deux à deux des relations de commutation caractéristiques (à l'instar de x et p_x), qui ne sont rien d'autre que les transcriptions quantiques des crochets de Poisson ; pour des raisons dimensionnelles, ces commutateurs sont proportionnels à \hbar . Le produit des incertitudes sur deux composantes est donc borné inférieurement : tout comme avec le couple (position, impulsion), la détermination arbitrairement précise de l'une des composantes conduit à une indétermination complète sur l'autre. On ne peut donc mesurer simultanément deux composantes distinctes du moment cinétique ; on dit souvent que deux telles grandeurs sont incompatibles, une affirmation qui vise la réalité observable et rien de plus : il faut comprendre qu'aucune mesure n'est possible visant à obtenir simultanément la valeur de l'une *et* de l'autre.

²⁷Techniquement, la quantification du moment cinétique (orbital) résulte du fait que l'on impose à la fonction d'onde (spatiale) de reprendre la même valeur lors d'une rotation de 2π .

Chapitre 13

Postulats et structure formelle de la Mécanique Quantique

Le but de ce chapitre est de faire le point sur les idées qui ont été développées jusqu'à présent et, dans un premier temps, d'énoncer progressivement l'ensemble des postulats¹ sur lesquels repose la Mécanique Quantique. Dans un deuxième temps, on présentera une rationalisation et une remise en ordre du formalisme en donnant les bases mathématiques minimales, sur des considérations qui se veulent délibérément plus intuitives que rigoureuses.

13.1 Enoncé des postulats

13.1.1 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 conceptions classiques. Il y a quelque analogie entre la nécessité de renoncer à une vision purement mécaniste des particules et celle, maintenant familière, de décrire les propriétés d'un gaz classique par des théories de nature statistique².

La fonction d'onde de Schrödinger constitue la connaissance ultime³ d'un système relevant de la théorie quantique. Cette fonction d'onde a, fondamentalement, une interprétation statistique. La description de l'état d'un système quantique consiste en l'énoncé de *lois de probabilités* permettant d'en décrire les propriétés et de fournir une explication du monde à l'échelle atomique ou subatomique⁴. Cet objectif doit être bien compris, pour écarter le risque de malentendus ou de paradoxes. Il n'est pas réducteur : renoncer à décrire dans le détail une "réalité" qui échappe l'observation directe est le point vue le plus naturel, l'interrogation sur cette "réalité" étant de nature métaphysique. Il ne doit pas non plus donner à penser ou à croire que la Mécanique Quantique est une théorie incomplète, pour deux raisons au moins.

La première des raisons est que 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

¹Ceux-ci seront donnés dans le même ordre que dans les ouvrages de référence facilement disponibles, comme celui de Cohen-Tannoudji *et al.* [16].

²Les déboires de Boltzmann en raison de l'incompréhension de ses contemporains – qui l'ont conduit au suicide – rappellent que cette révolution statistique, dans le dernier quart du XIX^{ème} siècle, ne s'est pas faite en douceur. Le dogme de Laplace, en tant que principe méthodologique universel, a eu la vie longue ("Donnez-moi tous les x_0 et tous les v_{x0} et je vous prédirai l'avenir du Monde ...").

³à température nulle.

⁴En Physique, expliquer c'est proposer une théorie cohérente qui, rendant compte de phénomènes déjà 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 une simple, et magnifique, rationalisation de l'ensemble des observations.

incomplète une théorie qui rend précisément compte de l'ensemble des expériences effectuées : à ce jour, il n'existe pas d'expérience faisant lever le moindre doute sur 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 a presque érigés en principes. Personne ne songe à décrire un gaz 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 systèmes⁵. Il en va de même aujourd'hui pour la Mécanique Quantique qui, en dépit d'un statut épistémologique propre⁶, peut prétendre au même caractère complet que toute autre théorie rendant compte de l'ensemble des expériences connues⁷.

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. Cette connaissance est multiforme et peut être représentée de plusieurs façons équivalentes ; par exemple, la fonction d'onde au sens de Schrödinger étant donnée, $\Psi(\vec{r}, t)$, il est possible d'introduire sa transformée de Fourier :

$$\Phi(\vec{p}, t) = (2\pi\hbar)^{-3/2} \int_{\mathbb{R}^3} d^3r \Psi(\vec{r}, t) e^{-(i/\hbar)\vec{p}\cdot\vec{r}} \quad (13.1)$$

où \vec{p} est ici un vecteur au sens ordinaire et non l'opérateur $-i\hbar\vec{\nabla}$. Manifestement, la fonction Φ contient exactement la même information que Ψ : connaître l'une de ces fonctions, c'est aussi connaître 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 Ψ et Φ ne sont que deux écritures différentes d'une seule et même abstraction mathématique, représentant l'état intrinsèque du système considéré ; chacune d'entre elles constitue la donnée d'une infinité (non dénombrable) de nombres (les valeurs successives de la fonction considérée). Pour identifier la nature de cet objet, on peut raisonner par analogie avec une autre situation, plus usuelle, et l'idée qui s'impose 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 une suite de trois nombres (les composantes sur la base choisie) ou une autre, mais chacun de ces triplets représente exactement la même réalité – le même vecteur. La notion de vecteur s'impose également au vu de la *linéarité* de l'équation de Schrödinger : toute combinaison linéaire de plusieurs solutions est encore une solution ; la combinaison linéaire est l'opération fondamentale permettant d'introduire la notion d'espace vectoriel. Dans un cadre plus élémentaire, on sait bien que la solution générale d'une équation différentielle linéaire du second ordre est une combinaison linéaire de deux solutions particulières linéairement indépendantes ; dire ceci c'est *ipso facto* déclarer l'existence d'un espace vectoriel de solutions, de dimension égale à 2.

On en vient ainsi à devoir énoncer le premier postulat de la Mécanique Quantique :

Postulat 1 *La connaissance de l'état d'un système à un instant donné est complètement contenue dans un vecteur appelé vecteur d'état. L'espace vectoriel auquel appartient ce vecteur est appelé espace d'états.*

Pour signaler explicitement la nature vectorielle, il est recommandé d'utiliser une notation particulière, afin de distinguer l'objet mathématique vecteur d'une part, ses composantes d'autre part. 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 système à l'instant t : $|\Psi(t)\rangle$. Le corps sur lequel est défini l'espace vectoriel est celui des complexes, \mathbb{C} ; ceci est bien évident puisque i apparaît dans l'équation de Schrödinger et constitue une nécessité fondamentale de la Mécanique Quantique (l'addition des amplitudes complexes de probabilité).

Un vecteur étant donné, le choix d'une base permet de définir les composantes de ce vecteur, lesquelles en constituent la *représentation*. Des exemples de représentations ont déjà été rencontrés. Le tout premier a

⁵Toutefois, l'analogie conceptuelle ne dit pas être poussée trop loin, au risque de contresens. En Théorie Cinétique des Gaz, chaque atome est censé posséder à tout instant une position et une vitesse : on renonce à les calculer, mais on admet qu'elles existent en soi. Pour une particule comme un électron, il n'en va pas ainsi : relire la discussion de l'expérience des trous d'Young (voir aussi [17], § 150 p. 445).

⁶On sait que la formulation précise de la Mécanique Quantique nécessite une référence constante à sa théorie limite, la Mécanique Classique.

⁷et de surcroît avec une précision littéralement diabolique – le calcul théorique du facteur anormal de l'électron donne $g_{\text{th}} = 2.002\,319\,304\,402 \pm 6 \times 10^{-11}$, la valeur expérimentale est $g_{\text{exp}} = 2.002\,319\,304\,376 \pm 8 \times 10^{-12}$!

été celui de la fonction d'onde originelle de Schrödinger, $\Psi(\vec{r}, t)$, dont l'interprétation a été donnée par Born ; elle décrit complètement l'état de la particule et donne la densité de probabilité d'observer cette particule dans un petit volume entourant un point donné de l'espace. Sa transformée de Fourier, $\Phi(\vec{p}, t)$, donne la densité de probabilité d'observer la particule avec une impulsion définie par le vecteur \vec{p} à d^3p près. Dans ces deux cas, l'argument des fonctions d'onde est une variable continue, car rien ne conduit à imaginer que la position et l'impulsion d'une particule 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 mesure 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. Il faut ainsi admettre l'idée générale suivant laquelle il existe des grandeurs qui sont *quantifiées* : les valeurs observées ne peuvent appartenir qu'à un ensemble *discret*, fini ou infini, mais toujours 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 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 en tant que résultat d'une opération de mesure de A . 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 réalise que ceux-ci constituent finalement une “fonction d'onde discrète”, fonction prenant un nombre fini de valeurs sur un ensemble isomorphe aux N premiers entiers et nulle ailleurs⁸. Il conviendra d'établir en temps utile le lien explicite 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 constitue la connaissance complète de l'état du système. Ce postulat nie donc l'existence de variables cachées.

13.1.2 Notion d'observable

En Mécanique Ondulatoire de Schrödinger les grandeurs dynamiques sont représentées par des opérateurs. La correspondance trouvée par Schrödinger est :

$$x \rightarrow \text{multiplication par } x, \quad p_x \rightarrow -i\hbar \times \text{dérivation partielle par rapport à } x. \quad (13.2)$$

C'est la généralisation de cette affirmation qui est érigée en postulat :

Postulat 2 *Toute grandeur physique est représentée par un opérateur agissant dans l'espace des états.*

L'usage a consacré l'adjectif substantivé *observable* pour ces opérateurs ; on dit ainsi que l'opérateur $p_x = -i\hbar\partial/\partial x$ est 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 spécifiques (il faut que les valeurs moyennes soient des quantités réelles) ; ces propriétés seront énoncées par la suite. Un opérateur linéaire quelconque n'a pas le statut d'observable⁹.

En tant qu'opérateurs, les observables obéissent à une algèbre non-commutative ; le produit xp_x est différent du produit $p_x x$: ce fait est traduit par la relation fondamentale :

$$[u, p_v] = i\hbar\delta_{uv}\mathbf{1}. \quad (13.3)$$

Par application systématique et récursive de cette relation, on peut former les relations de commutation pour tous les opérateurs associés à une grandeur classique qui peut s'exprimer comme une fonction des coordonnées et des impulsions ; par exemple, pour le moment cinétique, on calcule facilement le commutateur de J_x avec J_y à partir de la relation fondamentale (13.3) :

$$[J_x, J_y] = [yp_z - zp_y, zp_x - xp_z]. \quad (13.4)$$

⁸En quelque sorte, les c_n constituent les valeurs d'une fonction ponctuelle.

⁹Arrivé à ce stade de l'énoncé des postulats, on mesure la distance qui sépare la description mécanique classique de celle que pourvoit la Mécanique Quantique : l'état d'un système y est décrit par un vecteur et les grandeurs dynamiques par des opérateurs (linéaires) agissant dans l'espace vectoriel des vecteurs d'état.

En distribuant le commutateur, on obtient 4 termes dont deux sont visiblement nuls d'après (13.3) ; il reste :

$$[J_x, J_y] = [yp_z, zp_x] + [zp_y, xp_z] . \quad (13.5)$$

Chacun de ces commutateurs se calcule en remarquant que l'identité suivante est vraie :

$$[A, BC] = [A, B]C + B[A, C] , \quad (13.6)$$

et on trouve :

$$[J_x, J_y] = y(-i\hbar)p_x + p_y(+i\hbar)x = i\hbar(xp_y - yp_x) \equiv i\hbar J_z ; \quad (13.7)$$

les relations de commutation des autres composantes s'obtiennent par des permutations circulaires. Par des moyens analogues, on montre que :

$$[J_u, \vec{J}^2] = 0 \quad (u = x, y, z, \vec{J}^2 = J_x^2 + J_y^2 + J_z^2) . \quad (13.8)$$

Toutes ces relations sont telles que :

$$[A, B] \longleftrightarrow i\hbar \times \{A, B\} , \quad (13.9)$$

où $\{A, B\}$ est le crochet de Poisson des équivalents classiques.

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 , tels que l'action de A sur ceux-ci se traduit par la simple multiplication par un scalaire du corps. Ainsi, en utilisant la notation de Dirac, si Ω est un opérateur quelconque, le couple propre $(\omega, |\omega\rangle)$ satisfait par définition :

$$\Omega|\omega\rangle = \omega|\omega\rangle . \quad (13.10)$$

Il est clair que si $|\omega\rangle$ est vecteur propre, tout vecteur proportionnel $\lambda|\omega\rangle$ est encore vecteur propre et est associé à la même¹⁰ valeur propre. Pour un opérateur quelconque, la valeur propre ω est *a priori* un nombre complexe ; pour les observables, ω sera toujours réel, ce qui traduit évidemment une propriété de l'opérateur lui-même (hermiticité – voir plus loin). On sait par ailleurs que les valeurs propres sont les solutions de l'équation algébrique¹¹ :

$$\text{Det}(\Omega - \omega\mathbf{1}) = 0 . \quad (13.11)$$

Cette équation prend toujours la forme :

$$P(\omega) = 0 , \quad (13.12)$$

où $P(\omega)$ est soit un polynôme de degré N si l'espace vectoriel est de dimension N , soit une série si l'espace est de dimension infinie dénombrable. Il est parfois utile de se souvenir qu'un opérateur donné Ω satisfait sa propre équation caractéristique (théorème de Cayley - Hamilton), c'est-à-dire que l'on a, en tant qu'équation opératorielle :

$$P(\Omega) = \mathbf{0} . \quad (13.13)$$

Ceci étant rappelé, le moment est venu d'établir le théorème suivant, qui jouera un rôle important dans la suite. Pour deux opérateurs dont le commutateur n'est pas nul, l'ensemble des vecteurs propres de l'un est – en tant qu'ensemble – distinct de l'ensemble des vecteurs propres de l'autre¹². En effet, soit A et B deux opérateurs linéaires^{13, 14} tels que $AB \neq BA$ et soit $|a\rangle$ l'un *quelconque* des vecteurs propres de A associé à la valeur propre a :

$$A|a\rangle = a|a\rangle . \quad (13.14)$$

¹⁰Ceci est vrai pour tous les opérateurs *linéaires*.

¹¹appelée équation caractéristique.

¹²Deux tels opérateurs peuvent avoir certains vecteurs propres en commun, mais ceci ne peut être vrai pour *tous* les vecteurs propres.

¹³Pour deux opérateurs antilinéaires, les relations suivantes sont encore vraies si $a \in \mathbb{R}$.

¹⁴Dans la suite, tous les opérateurs sont supposés linéaires, sauf mention contraire.

Par hypothèse, on a :

$$AB|a\rangle \neq BA|a\rangle = Ba|a\rangle = aB|a\rangle \quad \forall |a\rangle . \quad (13.15)$$

Raisonnons maintenant par l'absurde : supposons que le vecteur $|a\rangle$ est aussi vecteur propre de B ; c'est dire qu'il existe un scalaire b tel que $B|a\rangle = b|a\rangle$; dans cette hypothèse, la non-égalité (13.15) devient :

$$ab|a\rangle \neq ab|a\rangle \quad \forall |a\rangle . \quad (13.16)$$

Comme a et b sont des nombres, la dernière non-égalité est absurde. À la réflexion, l'absence de vecteurs propres communs pour deux opérateurs A et B ne commutant pas est un fait bien évident : sur la base propre de l'un d'entre eux, A par exemple, ce dernier 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 évidemment :

$$D_A D_B = D_B D_A \iff [D_A, D_B] = 0 . \quad (13.17)$$

puisque deux matrices diagonales commutent toujours. Comme les relations entre opérateurs se transposent fidèlement au niveau de leurs représentants (leurs matrices), (13.17) serait en violation de $[A, B] \neq 0$. 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 : un vecteur propre de l'un n'est pas en général propre de l'autre. Si ces opérateurs sont des observables, on dit qu'elles sont *incompatibles*¹⁵ ; des exemples de tels couples sont fournis par (u, p_u) , (J_x, J_y) , etc.

A l'inverse, si deux opérateurs commutent, tout vecteur propre de l'un est aussi 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 :

$$A|a\rangle = a|a\rangle , \quad (AB - BA)|a\rangle \equiv AB|a\rangle - BA|a\rangle = 0 . \quad (13.18)$$

La dernière équation s'écrit, en ajoutant des parenthèses redondantes :

$$A(B|a\rangle) = B(A|a\rangle) = B(a|a\rangle) = a(B|a\rangle) , \quad (13.19)$$

ce qui montre que le vecteur $B|a\rangle$ est 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, 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 (qui est la valeur propre de B), on a donc :

$$B|a\rangle = b|a\rangle . \quad (13.20)$$

Pour rappeler ce fait et traiter A et B sur un pied d'égalité, on note $|a, b\rangle$ le vecteur propre commun à A et B . La généralisation au cas où le spectre est dégénéré ne présente aucune difficulté. Si A présente une dégénérescence, il existe alors plusieurs vecteurs propres associés à une même valeur propre : toute combinaison linéaire de ces vecteurs est encore vecteur propre – avec toujours la même valeur propre ; si donc un vecteur propre de A (“tiré au hasard”) n'est pas forcément propre de B , il est toujours possible, en formant précisément les bonnes combinaisons linéaires (qui restent propres de A quoi qu'il arrive), de trouver celles qui sont aussi propres de B .

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 matrice¹⁶ sur la base choisie. Plus particulièrement, les vecteurs propres $\{|a_n\rangle\}_n$ de A peuvent être utilisés pour engendrer effectivement l'espace vectoriel, et en constituent une base¹⁷. Si l'espace des états, \mathcal{E} , est de dimension N , tout vecteur d'état $|\Psi\rangle$ peut s'écrire en combinaison linéaire :

$$|\Psi\rangle = \sum_{m=1}^N c_m |a_m\rangle . \quad (13.21)$$

¹⁵Par le postulat 5 (réduction du paquet d'ondes), il en résultera que l'on ne peut mesurer simultanément deux grandeurs incompatibles.

¹⁶Les matrices des observables (sur la base propre du Hamiltonien) ne sont rien d'autre que les matrices de Heisenberg au sens de sa Mécanique des Matrices.

¹⁷Cette affirmation est évidente pour des raisons physiques. Dans le cas contraire, des valeurs physiques pour A (en tant que résultat d'une mesure de A), seraient inaccessibles lors d'une mesure de A .

Les coefficients $\{c_m\}_m$ jouent le rôle d’une “fonction d’onde discrète”. Si l’état considéré, $|\Psi\rangle$, dépend non trivialement du temps¹⁸ (état non stationnaire), les coefficients $\{c_m\}_m$ (les composantes de $|\Psi\rangle$) dépendent aussi du temps :

$$|\Psi(t)\rangle = \sum_{m=1}^N c_m(t)|a_m\rangle . \quad (13.22)$$

Enfin, il existe des cas plus compliqués où la description complète du système exige simultanément une grandeur continue comme la position, et une grandeur discrète (par exemple, un atome muni d’un moment cinétique et en voyage dans l’espace, sa position étant notée \vec{R}). Dans ce cas, l’état est de la forme :

$$|\Psi(\vec{R}, t)\rangle = \sum_{m=1}^N c_m(\vec{R}, t)|a_m\rangle . \quad (13.23)$$

13.1.3 Résultats possibles de la mesure d’une grandeur physique

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

Postulat 3 *La mesure d’une grandeur physique représentée par l’opérateur (l’observable) A ne peut fournir que l’une des valeurs propres de l’opérateur A .*

De ce postulat, il résulte qu’un opérateur quelconque ne représente pas en général une grandeur physique : un opérateur quelconque a des valeurs propres complexes. Le sens physique du 3^{ème} postulat impose aux opérateurs associés aux grandeurs physiques d’avoir toutes leurs valeurs propres réelles ; de tels opérateurs sont dits hermitiques. Leurs propriétés seront discutées plus loin en détail, mais leur importance extrême justifie que les deux principales d’entre elles soient énoncées dès maintenant :

1. les valeurs propres d’un opérateur hermitique sont toutes réelles
2. deux vecteurs propres d’un opérateur hermitique associés à deux valeurs propres *distinctes* sont orthogonaux, au sens d’un produit scalaire à définir ultérieurement.

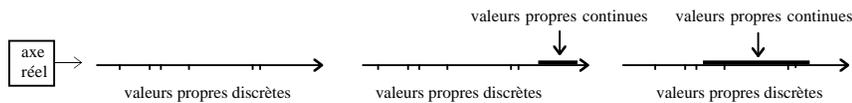


Figure 13.1: Différentes possibilités pour le spectre d’un opérateur hermitique.

On appelle *spectre* d’un opérateur l’ensemble de ses valeurs propres. Comme pour tout opérateur, le spectre d’un opérateur hermitique peut être discret, continu ou les deux la fois ; la figure 13.1 donne quelques dispositions possibles. Les parties discrète et continue, si elles existent simultanément, peuvent être disjointes ou se recouvrir : une valeur propre discrète peut fort bien être “noyée” dans un *continuum* d’autres valeurs propres¹⁹ – cette imbrication étant toutefois assez exceptionnelle. Les valeurs propres discrètes expriment la

¹⁸C’est-à-dire si la dépendance en temps ne se réduit pas à un facteur de phase global, inessentiel de ce fait.

¹⁹Lorsqu’il s’agit du Hamiltonien, dont le spectre constitue l’ensemble des énergies stationnaires possibles du système, 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. Si les fréquences sont commensurables (le rapport de deux quelconques d’entre elles étant un nombre rationnel), le mouvement sera périodique (l’oscillateur harmonique en est un cas extrême : toutes les fréquences sont multiples entiers d’une même fréquence). Si ces rapports sont irrationnels, le mouvement sera pseudo-périodique avec des temps de pseudo-retour (temps de Poincaré) éventuellement très grands. Enfin, si les énergies forment un spectre continu, la dynamique est irréversible. Le couplage à un *continuum* d’énergie est l’ingrédient nécessaire et suffisant de 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 dans la description quantique.

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 valeur des constantes fondamentales à une précision infiniment grande (exemple : l'atome de ^{86}Kr , dont une transition 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 (plus généralement, pour un moment cinétique de valeur J , on observe $2J + 1$ taches).

Une autre illustration est fournie par le processus d'absorption ou d'émission de photon(s) par un atome. Par l'équation de conservation (en émission, pour l'exemple, et dans l'hypothèse du noyau infiniment massif) :

$$E_i = E_f + h\nu \quad , \quad (13.24)$$

l'observation de la fréquence ν constitue bien une mesure de l'énergie de l'atome, à partir d'une origine arbitraire. Une mesure de l'énergie ne peut fournir que l'une des valeurs E_n dont l'ensemble constitue le 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 certain ensemble prédéterminé (le spectre de A). Il faut donc, pour la cohérence de la théorie, énoncer une règle permettant la prédiction effective des résultats d'une mesure ; en vertu du caractère probabiliste, cette règle ne peut être qu'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 :

$$A|a_m\rangle = a_m |a_m\rangle \quad . \quad (13.25)$$

L'absence de dégénérescence signifie qu'à une valeur propre a_{m_0} ne correspond qu'un seul vecteur propre $|a_{m_0}\rangle$; ceci se traduit, dans les notations, par le fait que le ket est spécifié sans ambiguïté par le seul symbole représentant la valeur propre a_m . Les états propres constituent une base sur laquelle on peut décomposer tout état $|\Psi\rangle$:

$$|\Psi\rangle = \sum_m c_m |a_m\rangle \quad . \quad (13.26)$$

Le 4^{ème} Postulat (parfois appelé *principe de décomposition spectrale*) s'énonce alors :

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 $|c_m|^2$.*

On apprécie encore mieux l'importance des états propres d'une observable A en notant que si l'opération de mesure est faite sur un état $|\Psi\rangle$ qui se trouve coïncider 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} (pour un tel état, on trouve a_{m_0} avec probabilité 1). C'est le seul cas où la mesure de A donne un et un seul résultat avec certitude. Par ailleurs, si deux vecteurs d'état ne diffèrent que par une phase globale, ils donnent toujours exactement les mêmes probabilités pour toute opération de mesure et représentent donc le même état physique. 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 état.

Le postulat 4 se généralise au cas où le spectre de A présente une dégénérescence, c'est-à-dire lorsqu'il existe un certain nombre²⁰, $g_m \in \mathbb{N}$, 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 le distinguer de ses semblables. L'ensemble de ces g_m vecteurs engendre le sous-espace propre de dimension g_m associé à la valeur propre a_m . Ainsi, on peut noter :

$$A|a_m, r\rangle = a_m |a_m, r\rangle \quad , \quad r = 1, 2, \dots, g_m \quad . \quad (13.27)$$

²⁰Le cas non-dégénéré correspond à $g_m = 1$.

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

$$|\Psi\rangle = \sum_m \sum_{r=1}^{g_m} c_{m,r} |a_{m,r}\rangle , \quad (13.28)$$

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

Postulat 4 étendu *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 $\sum_{r=1}^{g_m} |c_{m,r}|^2$.*

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 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, comme on le verra par la suite²¹.

Remarque

À plusieurs reprises, l'attention a été attirée sur la distinction essentielle 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 est défini avec arbitraire²² ; plus précisément, ce qui a un sens du point de vue de l'invariance physique – sauf à fournir ou exiger une information supplémentaire –, c'est le sous-espace de tous les vecteurs propres associés à une même valeur propre, 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 direct à chacun de ces vecteurs, il faut et suffit de trouver d'autres observables qui commutent avec A et de déterminer leurs vecteurs propres. Soit par exemple le vecteur $|a_m, b_p\rangle$, propre à la fois à A et à B qui commutent. Si tous les couples de valeurs propres (a_m, b_p) sont distincts (en tant que couples), il n'y a plus de dégénérescence. 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.* Le *ou bien* renvoie à l'addition des probabilités relatives à des événements mutuellement exclusifs, comme l'affirme la théorie des probabilités²³.

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 , $|a_m, b_p, c_q\rangle$, et ainsi de suite. Un ensemble d'observables dont chaque vecteur propre est unique (tous les n -plets (a_m, b_p, c_q, \dots) sont distincts les uns des autres) porte le nom d'ECOC (ensemble complet d'observables qui commutent).

Il reste à dire un mot du cas où l'observable A possède un spectre continu, cas qui englobe en fait le postulat premier de Born à propos 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 suivant :

$$|\Psi\rangle = \int da c(a) |a\rangle . \quad (13.29)$$

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

$$dP(a) = |c(a)|^2 da . \quad (13.30)$$

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 $|\Psi(x)|^2 dx$ désigne la probabilité élémentaire $dP(x)$ d'observer la particule dans l'intervalle $[x, x + dx]$.

²¹Le changement de base est une transformation unitaire conservant les longueurs. $\sum_{r=1}^{g_m} |c_{m,r}|^2$ est le carré de la longueur de la projection sur le sous-espace dégénéré : c'est un invariant dans toute transformation unitaire (rotation généralisée) effectuée à l'intérieur de ce sous-espace.

²²Toute combinaison linéaire de vecteurs propres associés à une même valeur propre est encore vecteur propre.

²³Après tout, on pourrait imaginer des démons (de Schrödinger, de Heisenberg, ...) qui, à l'insu, mesurent simultanément B et trouvent les uns b_1 , les autres b_2 , *etc.*

13.1.4 La réduction du paquet d'ondes

On désigne ainsi un phénomène²⁴ spécifique de la Mécanique Quantique qui apparaît, à ce stade de l'énoncé des principes, comme une nécessité compte tenu des postulats précédents, mais qu'il faut toutefois inclure dans la construction axiomatique. Ce phénomène, assez déconcertant au premier abord, a été l'objet de vifs débats sur la notion de réalité et de la connaissance scientifique que l'on peut souhaiter en avoir. Pour Schrödinger²⁵, la réduction du paquet d'ondes était de la magie et, pour marquer son opposition, formula le fameux paradoxe connu sous le nom du *Paradoxe du Chat de Schrödinger*²⁶.

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_{m_0} , on ne peut que retrouver la même valeur a_{m_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_{m_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 que la même que celle qui a été 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 différents²⁷. 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 cette 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 se soit produit, il n'est plus question de raisonner 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 l'expérience. 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 déjà effectuée. Le processus de réduction du paquet d'ondes lors d'une mesure de A peut s'illustrer comme montré sur la figure 13.2 ; cette évolution – qui est une sorte de projection, au sens géométrique du terme – ne relève pas de l'équation de Schrödinger, en tout cas pas celle où figure le seul Hamiltonien du système mesuré.

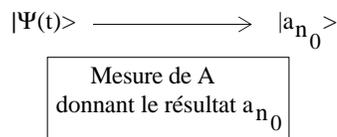


Figure 13.2: Représentation schématique de la réduction du paquet d'ondes.

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

²⁴appelé aussi le *collapse* de la fonction d'onde.

²⁵Selon Etienne Klein, *Conversations avec le Sphinx*, p. 176 (Albin Michel, 1991).

²⁶Les expériences d'Alain Aspect conduisent à affirmer que l'on doit renoncer, soit au postulat de la Relativité affirmant que c est une vitesse limite, soit à l'idée que deux systèmes ayant interagi mais dont l'interaction a cessé sont totalement découplés, indépendants l'un de l'autre. En effet, si on dit que deux tels systèmes sont séparés, la mesure effectuée sur l'un donne instantanément une information sur l'autre système : ceci viole le postulat relativiste. Si on refuse d'en venir ici, alors il faut admettre que les systèmes ne sont pas séparés... même s'il s'agit de deux particules situées à une distance macroscopique l'une de l'autre.

Cette analyse présuppose – ce qui n'est pas le cas en général dans les discussions 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, comme le montre l'expérience suivante. Deux personnes, sur le point de se séparer, tirent chacune (sans la regarder) une boule dans un sac en contenant deux ; les deux personnes savent que le sac contenait une boule blanche et une boule noire. Une fois rendue à destination, chacune d'entre elles, prenant connaissance de la couleur de sa boule, en déduit immédiatement la couleur de la boule emportée par l'autre ; il n'y a là rien de choquant, ni de contradictoire avec quoi que ce soit.

²⁷Ceci impliquerait que la vitesse d'évolution du système dans l'espace des états est infinie, ou, en termes plus imagés, que sa "trajectoire" n'est pas différentiable. Or l'équation de Schrödinger présuppose au contraire l'existence de la première dérivée en temps de la fonction d'onde.

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

Analytiquement, ceci s'écrit comme suit. Supposons que la mesure de A donne la valeur a_m , sachant que l'état juste avant la mesure est $|\Psi\rangle$; cet état peut toujours se décomposer comme suit :

$$|\Psi_{\text{avant}}\rangle = \sum_{m'} \sum_{r=1}^{g_{m'}} c_{m',r} |a_{m'}, r\rangle . \quad (13.31)$$

Alors, juste après la mesure ayant donné la valeur a_m , le système se trouve dans l'état :

$$|\Psi_{\text{après}}\rangle = C \sum_{r=1}^{g_m} c_{m,r} |a_m, r\rangle . \quad (13.32)$$

C est une constante de normalisation assurant que le système repart d'un état normalisé à l'unité ; à une phase globale près sans importance, on a :

$$C = \left[\sum_{r=1}^{g_m} |c_{m,r}|^2 \right]^{-1/2} . \quad (13.33)$$

L'opération de mesure fait donc deux choses : elle projette le vecteur "mesuré" sur le sous-espace propre et elle le dilate²⁸ pour lui donner une longueur unité. Cette projection efface toutes les composantes orthogonales au sous-espace propre pertinent, ne retenant que les composantes dans le sous-espace propre associé à la valeur propre trouvée par la mesure ; les coefficients du développement (13.31) qui survivent à la projection (ceux qui figurent dans (13.32)) conservent *de facto* la relation mutuelle de phase qu'ils avaient juste avant la mesure ; la projection conserve la cohérence de phase, laquelle joue un rôle essentiel dans l'apparition des interférences quantiques.

L'impossibilité de mesurer *simultanément* deux observables incompatibles A et B ($[A, B] \neq 0$) prend maintenant une tout autre dimension ; en effet, à l'issue d'une telle mesure hypothétique, le vecteur d'état devrait être l'un des vecteurs propres de A et de B ; mais on sait que deux opérateurs qui ne commutent pas n'ont pas un système de vecteurs propres en commun. 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 ([15], p. 126) montre que ce type 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 a plus 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 est en réalité âgée de 60 ans, on peut alors affirmer instantanément, que le nombre d'années qui lui restent à vivre (son espérance de vie) est évidemment beaucoup plus faible. Il y a bien, au moment où l'information plus précise sur l'âge 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 a en réalité 60 ans 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 semble toutefois exister une différence essentielle entre l'exemple de Bohm et la situation quantique ; en effet, l'âge 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 ; c'est la fonction d'onde qui sert d'intermédiaire et permet de prévoir et de calculer les probabilités quantiques. Le cadre probabiliste inévitable est donc, en ce sens, "classique". On ne peut toutefois prétendre que Ψ décrit seulement la *connaissance* que l'on a d'un système et non pas l'état du système lui-même :

²⁸Le vecteur projeté a visiblement une norme inférieure ou égale à celle du vecteur juste avant la mesure.

quand un électron se matérialise sur un écran, il s'agit bien d'un processus physique qui se produit, 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 information. La matérialisation d'une particule est toujours 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 réellement²⁹. 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.*

La réduction du paquet d'ondes est discutée 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 n'est pas possible d'affirmer que les problèmes de fond soulevés dans ces débats sont tous complètement résolus³⁰. 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 signifiante 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. L'opération de mesure n'est pas non plus décrite en détail en termes physiques, mais seulement définie par ses *conséquences* (juste après 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 spatialement³¹ et se trouve dans l'état de spin, $|+\rangle$ ou $|-\rangle$, en corrélation parfaite 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 de Cohen-Tannoudji *et al.*, [16], III. E 2. b, sur les appareils de mesure dits insuffisamment sélectifs).

13.1.5 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 dernier³², 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 :*

$$i\hbar \frac{d}{dt}|\Psi(t)\rangle = H(t)|\Psi(t)\rangle , \quad (13.34)$$

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

²⁹Bohm ([15], p. 127) récusé l'idée d'un *collapse* soudain de la fonction d'onde et invoque une perte graduelle de la cohérence de phase lors de l'interaction entre la particule et l'appareil (forcément macroscopique) 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.

³⁰Exemple 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 . . .

³¹Au sens d'une bonne résolution comparée à la distance entre les deux taches.

³²En fait, la disparition des trajectoires provoque en Mécanique Quantique l'*indiscernabilité* des particules identiques. De ce fait, on est conduit à énoncer un postulat supplémentaire (postulat de symétrisation) pour décrire les systèmes contenant plusieurs particules identiques.

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 classique variable dans le temps³³.

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 en est issue à un instant ultérieur est unique. 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 semble qu'il n'y ait 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 au sens propre, cette tâche est hors de portée.

Pour un système isolé, le Hamiltonien ne dépend pas du temps, par définition. Ceci ne veut évidemment pas dire que $|\Psi\rangle$ ne dépend pas du temps, ni, *a fortiori*, que toutes les moyennes de grandeurs physiques calculées avec $|\Psi(t)\rangle$ ne dépendent pas du temps. *Toutes* les valeurs moyennes sont indépendantes du temps seulement dans le cas d'états très particuliers appelés états stationnaires, qui s'écrivent :

$$|\Psi_{\text{st}}(t)\rangle = e^{(i\hbar)^{-1}E_n t} |\psi_n\rangle, \quad (13.35)$$

où E_n est l'une des valeurs propres de H et $|\psi_n\rangle$ le vecteur propre associé. Ces états seront discutés en détail dans le chapitre 15. Un tel état peut être préparé par une opération de mesure de l'énergie, donnant l'énergie E_n ³⁴. 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 l'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 :

$$|\Psi_{\text{st}}(t)\rangle = e^{(i\hbar)^{-1}E_n(t-t_0)} |\psi_n\rangle. \quad (13.36)$$

Si on effectue une nouvelle mesure de l'énergie, on retrouve toujours la même valeur E_n : ceci est l'expression la plus simple de la conservation de l'énergie pour un système isolé (conservatif). La mesure de l'énergie d'un système situé dans un état stationnaire ne modifie donc pas l'énergie de celui-ci et toute autre mesure ultérieure redonnera encore la même énergie³⁵. 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 temps³⁶, que cette observable soit ou non une constante du mouvement. Si A est une constante du mouvement, alors le commutateur $[A, H]$ est nul 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 mesure de A seront toujours des constantes mais la mesure modifiera l'état du système (qui ne sera d'ailleurs plus stationnaire après cette mesure) ; alors les probabilités des résultats d'une nouvelle mesure de A deviendront des fonctions explicites de la durée de l'intervalle de temps séparant les deux mesures de A . 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 . Ceci est à nouveau l'expression de la conservation de l'énergie.

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 – et d'ailleurs, il faut "renormaliser" $|\Psi\rangle$, voir (13.32). Penrose ([8], Ch. 6) discute longuement ces deux types d'évolution, qu'il appelle respectivement **U** et **R**, pour des raisons évidentes. Fondamentalement, on peut dire

³³Bien sûr, si on englobe le champ (les photons) dans la description classique et si l'atome et les photons constituent un système isolé, alors le Hamiltonien global ne dépend pas du temps.

³⁴Ceci 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.

³⁵Ceci 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 . . .

³⁶Le facteur de phase étant un simple scalaire, il commute avec tous les opérateurs (linéaires) et disparaît de toute moyenne, celle-ci contenant par construction la fonction d'onde et sa complexe conjuguée.

que l'évolution \mathbf{U} est complètement déterministe ; c'est au contraire par l'opération \mathbf{R} que surgit 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 essentiel 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.*) – est situé exclusivement dans le processus de réduction du paquet d'ondes.

La distinction entre \mathbf{U} et \mathbf{R} mériterait 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 général³⁷.

13.2 Illustration des postulats : retour sur l'expérience de Stern et Gerlach

L'expérience de Stern et Gerlach se prête remarquablement à la discussion des postulats : elle fait intervenir le cadre quantique *minimum* – il faut au moins deux états³⁸ –, ce qui autorise une discussion dépouillée, réduite à l'essentiel. Le(a) lecteur(rice) est renvoyé(e) à l'ouvrage [16], Ch. IV, pour une illustration des postulats dans le cadre de cette expérience. On y montre notamment comment l'appareil de Stern et Gerlach peut être utilisé comme polariseur ou comme analyseur. La mise en série de deux appareils permet une discussion précise et explicite des postulats.

13.3 Les bases du formalisme de la Mécanique Quantique

Il s'agit ici de donner les fondements du formalisme quantique avec la rigueur du physicien, qui permet de travailler proprement sans sombrer dans la paralysie d'une mathématisation excessive.

13.3.1 Changement de base

Comme l'affirme le Postulat 1, l'une des notions mathématiques fondamentales de la Mécanique Quantique est celle d'espace vectoriel, espace auquel appartient le vecteur d'état $|\Psi(t)\rangle$ d'un système. L'apparition de la structure d'espace vectoriel est clairement une conséquence de la linéarité de l'équation de Schrödinger : si $|\psi_1\rangle$ et $|\psi_2\rangle$ sont deux solutions de cette équation, toute combinaison linéaire $\lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle$ est encore solution. Cette combinaison est l'opération fondamentale de la structure d'espace vectoriel. Le corps est ici celui des nombres complexes : il y a nécessité pour la fonction d'onde de prendre des valeurs complexes et, d'ailleurs, l'équation de Schrödinger contient le nombre imaginaire fondamental, i .

Il est possible formellement d'effectuer toutes les opérations structurelles sur un espace vectoriel sans référence à une base donnée : 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 plus commodément quand une base a été choisie, à condition qu'elle l'ait été judicieusement. Il convient donc ici de discuter quelques aspects mathématiques liés à la notion de base d'un espace vectoriel. Dans l'esprit qui a été précisé au début de ce chapitre, tous les éléments mathématiques sont exposés de façon intuitive.

Une base est un ensemble de vecteurs linéairement indépendants en nombre égal à la dimension de l'espace vectoriel ; en d'autres termes, l'ensemble des N vecteurs $\{|e_n\rangle\}_{1 \leq n \leq N}$ est une base d'un espace vectoriel \mathcal{E} de

³⁷Un 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).

³⁸La plus petite matrice non-triviale est de dimension 2×2 .

dimension N si l'équivalence suivante est vraie :

$$\sum_{n=1}^N \lambda_n |e_n\rangle = 0 \iff \lambda_n = 0 \quad \forall n \in [1, N] \subset \mathbb{N} \quad (13.37)$$

Ceci exprime l'indépendance linéaire des vecteurs $|e_n\rangle$. Il s'agit bien d'une base : il est facile de montrer que tout vecteur $|\psi\rangle$ appartenant à \mathcal{E} peut alors se décomposer sur les $\{|e_n\rangle\}$:

$$\forall |\psi\rangle, \exists \{c_n\}_n : |\psi\rangle = \sum_{n=1}^N c_n |e_n\rangle \quad (13.38)$$

La décomposition est unique ; $\{c_n\}_n$ 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, c'est 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 donc par définition :

$$|\psi\rangle = \sum_{n=1}^N c_n |e_n\rangle, \quad (13.39)$$

et :

$$|\psi\rangle = \sum_{p=1}^N d_p |f_p\rangle ; \quad (13.40)$$

la question est 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 ; il faut donc 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" :

$$|e_n\rangle = \sum_{p=1}^N T_{pn} |f_p\rangle, \quad n = 1, 2, \dots, N. \quad (13.41)$$

Les N^2 nombres T_{np} forment 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 (13.39), il vient :

$$|\psi\rangle = \sum_{n=1}^N \sum_{p=1}^N T_{pn} c_n |f_p\rangle. \quad (13.42)$$

En égalant cette expression à la décomposition (13.40), on trouve :

$$\sum_{n=1}^N \sum_{p=1}^N [T_{pn} c_n - d_p] |f_p\rangle = 0. \quad (13.43)$$

Comme les $\{|f_p\rangle\}$ forment une base, la condition d'indépendance linéaire (13.37) est satisfaite ; en conséquence :

$$d_p = \sum_{n=1}^N T_{pn} c_n, \quad p = 1, 2, \dots, N. \quad (13.44)$$

La démarche peut être faite dans l'autre sens, en se donnant les nombres $\{T'_{np}\}$:

$$|f_p\rangle = \sum_{n=1}^N T'_{np} |e_n\rangle, \quad p = 1, 2, \dots, N, \quad (13.45)$$

et on trouve cette fois :

$$c_n = \sum_{p=1}^N T'_{np} d_p \equiv \sum_{p=1}^N T_{np}^{-1} d_p, \quad n = 1, 2, \dots, N. \quad (13.46)$$

T^{-1} désigne la matrice inverse de T ; d'où :

$$c_n = \sum_{p=1}^N T_{np}^{-1} \sum_{n'=1}^N T_{pn'} c_{n'}, \quad n = 1, 2, \dots, N \quad \iff \quad \sum_{p=1}^N \sum_{n'=1}^N T_{np}^{-1} T_{pn'} = \delta_{nn'}. \quad (13.47)$$

La dernière égalité reproduit la règle de multiplication de deux matrices ; ici il s'agit de la matrice T et de son inverse T^{-1} , dont le produit dans n'importe quel ordre est égal à l'identité $\mathbf{1}$. 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 : en un sens elle l'est car ce qui est derrière, c'est l'invariance de l'objet $|\psi\rangle$ défini intrinsèquement, indépendamment de toute base.

Un vecteur (ket) de l'espace des états est donc 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 :

$$|\psi\rangle, \{|e_n\rangle\}_n \quad \longleftrightarrow \quad \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix}. \quad (13.48)$$

Alors, la relation de changement de base (13.44) s'écrit :

$$\begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} & \dots & T_{1N} \\ T_{21} & T_{22} & \dots & T_{2N} \\ \vdots & \vdots & \dots & \vdots \\ T_{N1} & T_{N2} & \dots & T_{NN} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{bmatrix}. \quad (13.49)$$

La représentation d'un vecteur ket par un vecteur-colonne est évidemment aussi possible pour chacun des vecteurs de la base. On a ainsi :

$$|e_n\rangle, \{|e_{n'}\rangle\}_{n'} \quad \longleftrightarrow \quad \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad (13.50)$$

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

13.3.2 Produit scalaire

Selon le Postulat 4, la probabilité P_m d'obtenir une valeur a_m du spectre de A est le module carré de la composante du vecteur objet de la mesure $|\Psi\rangle$, sur l'état propre $|a_m\rangle$ de l'observable A . Cette opération algébrique – et d'autres d'ailleurs – s'expriment très naturellement à l'aide d'un produit scalaire convenablement défini. La définition ordinaire doit faire l'objet d'une petite généralisation, eu égard au fait que le corps est ici celui des complexes et non pas \mathbb{R} .

Le produit scalaire de deux vecteurs $|\psi\rangle$ et $|\phi\rangle$ est un nombre (scalaire) ; ici c'est un nombre complexe, noté provisoirement $(|\psi\rangle, |\phi\rangle)$. Au contraire de ce qui se passe dans le cas d'un espace vectoriel sur le corps des réels, le produit scalaire est ici *non-commutatif* ; on pose la relation suivante :

$$(|\psi\rangle, |\phi\rangle) = (|\phi\rangle, |\psi\rangle)^*, \quad (13.51)$$

où * note la conjugaison complexe. Cette propriété assure que le carré de la norme d'un vecteur est un réel positif, ce qui permet de définir une distance au sens ordinaire.

Le produit scalaire est linéaire par rapport au *second* partenaire du couple ordonné :

$$(|\psi\rangle, \lambda_1|\phi_1\rangle + \lambda_2|\phi_2\rangle) = \lambda_1(|\psi\rangle, |\phi_1\rangle) + \lambda_2(|\psi\rangle, |\phi_2\rangle) . \quad (13.52)$$

Les équations (13.51) et (13.52) entraînent que le produit scalaire est *antilinéaire* par rapport au *premier* partenaire ; en effet, en conjuguant le premier membre et en renversant l'ordre, on ne fait rien au total, d'où :

$$(\lambda_1|\phi_1\rangle + \lambda_2|\phi_2\rangle, |\psi\rangle)^* = \lambda_1(|\psi\rangle, |\phi_1\rangle) + \lambda_2(|\psi\rangle, |\phi_2\rangle) . \quad (13.53)$$

Reprenant les complexes conjugués :

$$(\lambda_1|\phi_1\rangle + \lambda_2|\phi_2\rangle, |\psi\rangle) = \lambda_1^*(|\phi_1\rangle, |\psi\rangle) + \lambda_2^*(|\phi_2\rangle, |\psi\rangle) , \quad (13.54)$$

qui exprime l'antilinéarité par rapport au premier vecteur du couple.

La norme (longueur) d'un vecteur est la racine carrée du produit scalaire de ce vecteur avec lui-même :

$$\| |\psi\rangle \|^2 = (|\psi\rangle, |\psi\rangle) \iff \| |\psi\rangle \| = \sqrt{(|\psi\rangle, |\psi\rangle)} . \quad (13.55)$$

De (13.54), il résulte notamment :

$$(\lambda|\psi\rangle, \lambda|\psi\rangle) = \lambda^*\lambda(|\psi\rangle, |\psi\rangle) \equiv |\lambda|^2(|\psi\rangle, |\psi\rangle) , \quad (13.56)$$

ce que l'on interprète en disant que si un vecteur est multiplié par un nombre (complexe) quelconque, sa longueur est multipliée par le module de ce nombre, qui joue ainsi le rôle d'un facteur d'échelle.

Deux vecteurs dont le produit scalaire est nul sont dits orthogonaux³⁹. Au total, le produit scalaire ainsi défini est antilinéaire défini positif⁴⁰. Les deux vecteurs étant donnés par leurs composantes sur une base $\{|e_n\rangle\}$:

$$|\psi\rangle = \sum_n x_n |e_n\rangle , \quad |\phi\rangle = \sum_m y_m |e_m\rangle , \quad (13.57)$$

les définitions précédentes donnent immédiatement :

$$(|\psi\rangle, |\phi\rangle) = \sum_n \sum_m x_n^* g_{nm} y_m , \quad g_{nm} = (|e_n\rangle, |e_m\rangle) . \quad (13.58)$$

Des bases remarquables joueront un rôle important dans toute la suite : ce sont les bases orthonormées, dont tous les vecteurs sont de norme unité et deux à deux orthogonaux – elles seront signalées par le symbole \perp_1 ; pour une telle base :

$$g_{nm} = (|e_n\rangle, |e_m\rangle) = \delta_{nm} \quad (\text{base } \perp_1) . \quad (13.59)$$

Alors le produit scalaire de deux vecteurs prend une forme très simple :

$$(|\psi\rangle, |\phi\rangle) = \sum_n \sum_m x_n^* \delta_{nm} y_m = \sum_n x_n^* y_n \quad (\text{base } \perp_1) \quad (13.60)$$

qui généralise la définition ordinaire pour un espace vectoriel sur \mathbb{R} . De même, le carré de la longueur du vecteur $|\psi\rangle$ est⁴¹ :

$$\| |\psi\rangle \|^2 = \sum_n |x_n|^2 \quad (\text{base } \perp_1) \quad (13.61)$$

Dans la suite, sauf mention contraire, toutes les bases seront supposées orthonormées.

³⁹Le vecteur nul est orthogonal à tous les vecteurs.

⁴⁰On dit aussi *sesquilinéaire*.

⁴¹(13.61) montre que seul le vecteur nul est de norme nulle (toutes ses composantes sont nulles) et est bien le seul à être orthogonal à tout autre vecteur.

Toutes ces expressions peuvent recevoir une écriture matricielle. En effet, une base orthonormée étant choisie, l'expression (13.60) du produit scalaire montre qu'il convient d'introduire le vecteur-ligne :

$$\left[x_1^* \quad x_2^* \quad \dots \quad x_N^* \right] . \quad (13.62)$$

Ceci permet d'écrire, partant de (13.57) et pour retrouver (13.60) :

$$(|\psi\rangle, |\phi\rangle) = \left[x_1^* \quad x_2^* \quad \dots \quad x_N^* \right] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} . \quad (13.63)$$

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. Dirac a proposé le symbole $\langle |$ appelé⁴² *bra* :

$$\left[x_1^* \quad x_2^* \quad \dots \quad x_N^* \right] \longleftrightarrow \langle \psi | . \quad (13.64)$$

Alors que le ket est linéaire, le bra est antilinéaire. En effet, soit :

$$\lambda |\psi\rangle \longleftrightarrow \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_N \end{bmatrix} . \quad (13.65)$$

Ce ket est associé à un bra selon :

$$\lambda |\psi\rangle \equiv |\lambda\psi\rangle \longleftrightarrow \langle \lambda\psi | \longleftrightarrow \left[\lambda^* x_1^* \quad \lambda^* x_2^* \quad \dots \quad \lambda^* x_N^* \right] \longleftrightarrow \lambda^* \langle \psi | ; \quad (13.66)$$

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

$$\lambda |\psi\rangle \longleftrightarrow \lambda^* \langle \psi | \equiv \langle \lambda\psi | . \quad (13.67)$$

Ceci posé, le produit scalaire (13.63) devient :

$$(|\psi\rangle, |\phi\rangle) = \langle \psi || \phi \rangle . \quad (13.68)$$

La double barre verticale à droite n'est pas utile et on note plus simplement :

$$(|\psi\rangle, |\phi\rangle) = \langle \psi | \phi \rangle , \quad (13.69)$$

sans perdre de vue que la simple barre est double et que l'on peut détacher les fragments implicites si besoin est :

$$\langle \psi | \phi \rangle \equiv \langle \psi \leftarrow || \rightarrow \phi \rangle \longrightarrow \langle \psi | \phi \rangle . \quad (13.70)$$

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

$$\left[0 \quad 0 \quad \dots \quad 0 \quad 1 \quad 0 \quad \dots \quad 0 \right] \longleftrightarrow \langle e_n | , \quad (13.71)$$

où le 1 est en $n^{\text{ème}}$ position.

Le produit scalaire permet d'exprimer commodément les composantes de tout vecteur. En effet, partant de :

$$|\psi\rangle = \sum_n x_n |e_n\rangle , \quad (13.72)$$

⁴²La terminologie est un néologisme de *bracket*.

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

$$(|e_m\rangle, |\psi\rangle) \equiv \langle e_m|\psi\rangle = \sum_n x_n \langle e_m|e_n\rangle \equiv \sum_n x_n g_{mn} . \quad (13.73)$$

Pour une base orthonormée, ceci devient :

$$\langle e_n|\psi\rangle = x_n , \quad (13.74)$$

qui, lu à l'envers, fournit le résultat annoncé :

$$x_n = \langle e_n|\psi\rangle \quad (\text{base } \perp_1) . \quad (13.75)$$

Alors, (13.72) s'écrit :

$$|\psi\rangle = \sum_n \langle e_n|\psi\rangle |e_n\rangle \equiv \sum_n |e_n\rangle \langle e_n|\psi\rangle . \quad (13.76)$$

Cette relation est vraie $\forall \psi$; d'autre part, on peut dédoubler la barre verticale de droite et "mettre en facteur" $|\psi\rangle$, qui ne dépend pas de l'indice de sommation :

$$|\psi\rangle = \sum_n |e_n\rangle \langle e_n|\psi\rangle \equiv \left[\sum_n |e_n\rangle \langle e_n| \right] |\psi\rangle \quad \forall |\psi\rangle . \quad (13.77)$$

On en déduit :

$$\sum_n |e_n\rangle \langle e_n| = \mathbf{1} , \quad (13.78)$$

où $\mathbf{1}$ désigne l'identité. L'égalité (13.78) s'appelle *relation de fermeture* et exprime le fait que la base $\{|e_n\rangle\}_n$ est *complète*, c'est-à-dire permet d'engendrer tout l'espace vectoriel.

Une base orthonormée étant choisie, on peut en trouver d'autres par des transformations linéaires appropriées⁴³. Soit U une transformation préservant l'orthonormalisation ; le transformé de $|e_p\rangle$ est un certain ket $|f_p\rangle$, qui peut se développer sur la première base :

$$|f_p\rangle \equiv U |e_p\rangle = \sum_n U_{np} |e_n\rangle , \quad (13.79)$$

Par hypothèse sur la nature de U , la base $\{|f_p\rangle\}$ est aussi orthonormée. En explicitant le produit scalaire $\langle f_p|f_{p'}\rangle = \delta_{pp'}$ à l'aide de (13.79), il vient :

$$\sum_n \sum_{n'} U_{np}^* U_{n'p'} \delta_{nn'} = \delta_{pp'} \iff \sum_n U_{np}^* U_{np'} = \delta_{pp'} . \quad (13.80)$$

D'un autre côté, par définition de l'inverse d'une matrice, on a :

$$\sum_n U_{pn}^{-1} U_{np'} = \delta_{pp'} . \quad (13.81)$$

La comparaison de (13.80) et (13.81) montre que pour une transformation préservant les angles et les longueurs, on a :

$$U_{pn}^{-1} = U_{np}^* . \quad (13.82)$$

Ces transformations sont appelées *unitaires*. Sur une base orthonormée, la matrice d'une transformation unitaire est telle que sa transposée conjuguée n'est autre que la matrice inverse – c'est d'ailleurs ce que l'on appelle une matrice unitaire. Pour une transformation unitaire, la relation d'invariance caractéristique du produit scalaire s'écrit :

$$(U|\phi\rangle, U|\psi\rangle) = (|\phi\rangle, |\psi\rangle) \quad \forall |\phi\rangle, |\psi\rangle . \quad (13.83)$$

⁴³Ces transformations généralisent les rotations de l'espace ordinaire.

13.3.3 Généralisations

Les résultats ci-dessus doivent être généralisés de deux façons. 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 établies sur un strict plan mathématique. On se bornera ici, dans l'esprit déjà annoncé, à les donner sans justification réelle, en restant à un niveau intuitif de plausibilité.

Pour un espace de dimension infinie dénombrable, une base contient une infinité dénombrable d'éléments et un vecteur quelconque se décompose suivant :

$$|\psi\rangle = \sum_{n \in \mathbb{N}} c_n |e_n\rangle . \quad (13.84)$$

Il est manifeste qu'une telle écriture, pour recevoir un sens, implique des conditions de convergence topologique. En les supposant toutes réunies, le produit scalaire sur une base orthonormée s'écrit dans les notations précédentes :

$$(|\phi\rangle, |\psi\rangle) = \sum_{n \in \mathbb{N}} x_n^* y_n . \quad (13.85)$$

où les conditions évoquées ci-dessus assurent que la série est convergente. On admet finalement que toutes les propriétés et relations 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. 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 Hilbert⁴⁴.

La deuxième généralisation indispensable est beaucoup moins évidente. Revenons au point de départ, le développement d'un vecteur sur une base, orthonormée ou non :

$$|\psi\rangle = \sum_n c_n |e_n\rangle . \quad (13.86)$$

Les $\{c_n\}$ constituent une suite de nombres que l'on peut considérer comme l'ensemble des valeurs d'une fonction discrète c , définie sur des nombres entiers. Quant aux $\{|e_n\rangle\}_n$, 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, (13.86) devient :

$$|\psi\rangle = \sum_n c_n |n\rangle . \quad (13.87)$$

Généralisons maintenant une telle expression au cas où la variable n , au lieu de prendre des valeurs discrètes, prend des valeurs continues dans⁴⁵ \mathbb{R} , notées ν pour marquer ce fait. La variable c devient alors une fonction au sens ordinaire du terme que, selon l'usage, on notera 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 la plus simple de la sommation discrète est l'intégrale de Riemann ; on écrit ainsi :

$$|\psi\rangle = \sum_n c_n |n\rangle \quad \longrightarrow \quad |\psi\rangle = \int d\nu \psi(\nu) |\nu\rangle . \quad (13.88)$$

Une fois acceptée cette transformation de la règle fondamentale de développement sur une base, on peut effectuer des manipulations similaires à celles portant sur les sommes discrètes, puisque tout repose en définitive

⁴⁴David 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 par exemple Jean Dieudonné, *Abrégé d'histoire des mathématiques*, Hermann, Paris, 1986, p. 193). Le texte complet de la conférence de Hilbert est disponible à l'adresse [18].

⁴⁵C'est juste pour fixer les idées. Plus généralement, ν prend ses valeurs dans $\mathcal{D} \subseteq \mathbb{R}$.

sur le caractère *linéaire* du formalisme et que les opérations de sommation et/ou d'intégration sont elles aussi linéaires. Ainsi, le produit scalaire des deux vecteurs :

$$|\psi\rangle = \int d\nu \psi(\nu)|\nu\rangle, \quad |\phi\rangle = \int d\nu \phi(\nu)|\nu\rangle \quad (13.89)$$

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

$$(|\psi\rangle, |\phi\rangle) \equiv \langle\psi|\phi\rangle = \int d\nu \int d\nu' \psi^*(\nu)\phi(\nu') \langle\nu|\nu'\rangle. \quad (13.90)$$

Les bornes d'intégration 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 (13.90) du produit scalaire doit apparaître comme la généralisation de (13.60), soit :

$$(|\psi\rangle, |\phi\rangle) \equiv \langle\psi|\phi\rangle = \int d\nu \psi^*(\nu)\phi(\nu). \quad (13.91)$$

Ceci n'est possible que si le produit scalaire $\langle\nu|\nu'\rangle$ dans (13.90) est tel que :

$$\langle\nu|\nu'\rangle \begin{cases} = 0 & \text{si } \nu' \neq \nu \\ \neq 0 & \text{si } \nu' = \nu \end{cases}, \quad (13.92)$$

et est l'analogie d'un symbole de Kronecker continu, que l'on peut noter $\delta(\nu, \nu')$ et appeler provisoirement 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 sa valeur, 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 $\nu = \nu'$ du domaine d'intégration, l'intégrale est inchangée. Comme le produit scalaire $\langle\nu|\nu'\rangle$ est nul partout ailleurs, l'une des deux intégrales, celle sur ν' par exemple, est nulle, de sorte que $\langle\nu|\nu'\rangle = 0 \quad \forall \nu, \nu'$: il semble que l'on soit dans une impasse vis-à-vis de (13.92). En réalité, il existe une issue : supprimer un point ne change rien si l'intégrand est *fini* ; tout peut changer si on ôte un point où l'intégrand est infini – 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 limite par laquelle enlever le “bon” point, celui où une “bonne” divergence se produit, n'est pas une opération inoffensive. Dans ce but, on précise (13.92) en posant :

$$\langle\nu|\nu'\rangle = \begin{cases} 0 & \text{si } \nu' \neq \nu \\ +\infty & \text{si } \nu' = \nu \end{cases}, \quad (13.93)$$

ce qui définit une quantité plutôt exotique que l'usage appelle “fonction de Dirac” et note $\delta(\nu - \nu')$:

$$\delta(\nu - \nu') \equiv \langle\nu|\nu'\rangle = \begin{cases} 0 & \text{si } \nu' \neq \nu \\ +\infty & \text{si } \nu' = \nu \end{cases}. \quad (13.94)$$

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 δ ; (13.90) s'écrit maintenant :

$$(|\psi\rangle, |\phi\rangle) \equiv \langle\psi|\phi\rangle = \int d\nu \int d\nu' \psi^*(\nu)\phi(\nu') \delta(\nu - \nu') = \int d\nu \psi^*(\nu)\phi(\nu), \quad (13.95)$$

soit :

$$\int d\nu \psi^*(\nu) \int d\nu' \phi(\nu') \delta(\nu - \nu') = \int d\nu \psi^*(\nu)\phi(\nu). \quad (13.96)$$

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

$$\int d\nu' \phi(\nu') \delta(\nu - \nu') = \phi(\nu) \quad \forall \phi, \quad (13.97)$$

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. Pour

que l'écriture (13.97) – déjà symbolique – ait un sens, il faut évidemment que la valeur de la fonction f au point ν existe. Il est donc utile de préciser :

$$\int_a^b d\nu' \phi(\nu') \delta(\nu - \nu') = \phi(\nu) \quad \forall \text{ fonction } \phi \text{ définie en } \nu \text{ et si } a < \nu < b . \quad (13.98)$$

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

$$\int_a^b d\nu' \delta(\nu - \nu') = 1 , \quad a < \nu < b . \quad (13.99)$$

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, $\delta(\nu)$ est homogène à l'inverse d'une longueur (un nombre d'onde par exemple), si ν est une fréquence, alors $\delta(\nu)$ est homogène à un temps, *etc.* Par ailleurs, on peut remarquer que les relations de fonctionnement de δ généralisent bien celles du symbole de Kronecker, δ_{nm} . En effet :

$$\sum_{n'=1}^N \phi_{n'} \delta_{nn'} = \phi_n \quad \forall \text{ suite } \phi_n \text{ définie en } \nu \text{ et si } 1 \leq n \leq N ; \quad (13.100)$$

ceci est clairement l'analogue de (13.98). En outre, la relation :

$$\sum_{n'=1}^N \delta_{nn'} = 1 \quad (1 \leq n \leq N) ; \quad (13.101)$$

est le pendant de (13.99).

La relation (13.98) doit 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 la surface vaut 1 – conformément à ce qu'exige (13.99). Ce graphe a donc l'allure reportée sur la figure 13.3 ; 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 .

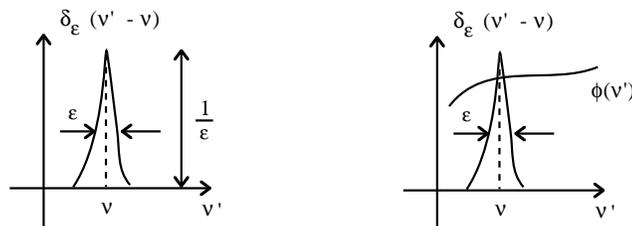


Figure 13.3: Précuseur de la fonction de Dirac.

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 :

$$I = \int_a^b d\nu' \phi(\nu') \delta_\varepsilon(\nu - \nu') . \quad (13.102)$$

Si la fonction ϕ est régulière au voisinage de ν et si elle varie lentement à l'échelle ε , il est évident géométriquement que :

$$I \simeq \phi(\nu) \int_a^b d\nu' \delta_\varepsilon(\nu - \nu') . \quad (13.103)$$

L'intégrale vaut 1 (δ_ε est définie avec une aire égale à 1) ; pour toutes les fonctions régulières, aussi petit que soit ε fini, on a :

$$\int_a^b d\nu' \phi(\nu') \delta_\varepsilon(\nu - \nu') \simeq \phi(\nu) , \quad (13.104)$$

et donc :

$$\lim_{\varepsilon \rightarrow 0} \left[\int_a^b d\nu' \phi(\nu') \delta_\varepsilon(\nu - \nu') \right] = \phi(\nu) , \quad (13.105)$$

Ce qui précède est parfaitement licite ; là où les choses se gâtent et où l'habitude conduit à une certaine désinvolture, c'est lorsque l'on échange l'ordre des opérations limite et intégration. On obtient alors strictement :

$$\int_a^b d\nu' \phi(\nu') \lim_{\varepsilon \rightarrow 0} [\delta_\varepsilon(\nu - \nu')] = \phi(\nu) , \quad (13.106)$$

qui est identique avec (13.97) ; 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'expliquer : il suffit de faire un choix explicite de fonctions reproduisant δ_ε . Quelques choix courants sont les suivants :

$$\delta_\varepsilon(\nu - \nu') = \frac{1}{\sqrt{2\pi\varepsilon}} e^{-(\nu - \nu')^2 / (2\varepsilon^2)} \quad (\text{gaussienne}) , \quad (13.107)$$

$$\delta_\varepsilon(\nu - \nu') = \frac{1}{\pi} \frac{\varepsilon/2}{(\nu - \nu')^2 + \varepsilon^2/4} \quad (\text{lorentzienne}) , \quad (13.108)$$

$$\delta_\varepsilon(\nu - \nu') = \begin{cases} 0 & \text{si } |\nu' - \nu| > \varepsilon/2 \\ \frac{1}{\varepsilon} & \text{si } |\nu' - \nu| < \varepsilon/2 \end{cases} \quad (\text{fonction créneau}) . \quad (13.109)$$

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 celle de Dirac. Avec elles, on peut 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 :

$$\delta(\nu - \nu') = \delta(\nu' - \nu) . \quad (13.110)$$

En outre, raisonnant à nouveau avant le passage à la limite, il est possible d'analyser le rôle opérationnel des fonctions δ_ε et même, *in fine*, de définir les “dérivées” de δ .

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

$$|\psi\rangle = \int d\nu \psi(\nu) |\nu\rangle . \quad (13.111)$$

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

$$\phi(\mu) = \int d\nu T_{\mu\nu} \psi(\nu) . \quad (13.112)$$

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 plutôt appelée noyau⁴⁶ ; on notera ainsi⁴⁷ :

$$\phi(\mu) = \int d\nu T(\mu, \nu) \psi(\nu) . \quad (13.113)$$

⁴⁶ À ne pas confondre avec le noyau d'un opérateur (ensemble des vecteurs dont l'image est le vecteur nul).

⁴⁷ Il y a ici un problème de notation. Comme on va le voir, $\psi(\nu)$ va s'identifier à $\langle \nu | \psi \rangle$, et donc on devrait noter $\langle \mu | \psi \rangle$ les nouvelles composantes ; mais ceci introduirait une confusion : les deux fonctions $\phi(\mu)$ et $\psi(\nu)$ sont deux fonctions différentes.

En exprimant les composantes à l'aide des produits scalaires avec les vecteurs d'une base orthonormée au sens (13.93), on voit que :

$$|\psi\rangle = \int d\nu' \psi(\nu') |\nu'\rangle \iff \langle\nu|\psi\rangle = \int d\nu' \psi(\nu') \langle\nu|\nu'\rangle = \psi(\nu) . \quad (13.114)$$

Il en résulte :

$$|\psi\rangle = \int d\nu \psi(\nu) |\nu\rangle = \int d\nu \langle\nu|\psi\rangle \langle\nu| = \int d\nu |\nu\rangle \langle\nu|\psi\rangle = \left[\int d\nu |\nu\rangle \langle\nu| \right] |\psi\rangle . \quad (13.115)$$

D'où la relation de fermeture :

$$\int d\nu |\nu\rangle \langle\nu| = \mathbf{1} . \quad (13.116)$$

Par ailleurs, le produit scalaire sur une base continue orthonormée s'écrit :

$$\langle\psi|\phi\rangle = \int d\nu \psi^*(\nu) \phi(\nu) , \quad \text{si } \langle\nu|\nu'\rangle = \delta(\nu - \nu') , \quad (13.117)$$

tandis que le carré de la norme est :

$$\langle\psi|\psi\rangle = \int d\nu \psi^*(\nu) \psi(\nu) = \int d\nu |\psi(\nu)|^2 . \quad (13.118)$$

Cette dernière intégrale doit rappeler quelque chose : si la variable ν est la position x d'une particule à une dimension, elle est le carré de la norme la fonction d'onde. 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 (continues) d'un vecteur $|\psi\rangle$:

$$|\psi\rangle = \int dx \psi(x) |x\rangle , \quad \psi(x) = \langle x|\psi\rangle , \quad (13.119)$$

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

$$\int dx |x\rangle \langle x| = \mathbf{1} , \quad \langle x|x'\rangle = \delta(x - x') , \quad (13.120)$$

et on retrouve à ce stade la condition de normalisation sous sa forme familière (transcription de (13.118)) :

$$\| |\psi\rangle \|^2 = \int dx \psi^*(x) \psi(x) = 1 . \quad (13.121)$$

On montrera par la suite que le vecteur $|x_0\rangle$ est propre de l'observable x et associé à la valeur propre x_0 .

Les relations de fermeture peuvent maintenant se combiner les unes avec les autres. Soit $\{a_n(x)\}_n \equiv \{\langle x|a_n\rangle\}_n$ un ensemble de fonctions formant une base complète orthonormalisée⁴⁸ ; en notation de Dirac, on a donc :

$$\sum_n |a_n\rangle \langle a_n| = \mathbf{1} , \quad (13.122)$$

en conséquence de :

$$\int a_n^*(x) a_m(x) dx = \int \langle a_n|x\rangle \langle x|a_m\rangle dx = \delta_{mn} . \quad (13.123)$$

En multipliant membre à membre (13.122) à gauche par $\langle x|$ et à droite par $|x'\rangle$, on fabrique la représentation de l'identité :

$$\sum_n \langle x|a_n\rangle \langle a_n|x'\rangle = \delta(x - x') \quad (13.124)$$

⁴⁸Ce sont par exemple les fonctions propres d'une observable A .

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

$$\sum_n a_n(x) a_n^*(x') = \delta(x - x') \quad (13.125)$$

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

$$\langle x|x'\rangle = \langle x|\mathbf{1}|x'\rangle \iff \langle x|\left[\sum_n |a_n\rangle\langle a_n|\right]|x'\rangle = \delta(x - x') \quad (13.126)$$

En réinsérant $\langle x|$ et $|x'\rangle$ dans la sommation, on retrouve (13.125). Cette dernière forme de la relation de fermeture prendra un sens particulièrement clair lors de l'étude de l'opérateur d'évolution $U(t)$ de la fonction d'onde⁴⁹ ; en effet, anticipant légèrement, celui s'écrit :

$$U(t) = \sum_n |\psi_n\rangle e^{-i\omega_n t} \langle \psi_n| , \quad (13.127)$$

où les $|\psi_n\rangle$ sont les vecteurs propres⁵⁰ de H associés aux énergies propres $E_n \equiv \hbar\omega_n$. En multipliant (13.127) à gauche par $\langle x|$ et à droite par $|x'\rangle$, on trouve le propagateur $U(x, t; x', 0)$:

$$U(x, t; x', 0) = \sum_n \langle x|\psi_n\rangle e^{-i\omega_n t} \langle \psi_n|x'\rangle , \quad (13.128)$$

avec lequel la fonction développée à l'instant t s'écrit :

$$\Psi(x, t) = \int dx' U(x, t; x', 0) \Psi(x', 0) . \quad (13.129)$$

Faisant $t = 0$ dans (13.128) compte tenu de (13.125), on trouve :

$$U(x, t = 0; x', 0) = \delta(x - x') , \quad (13.130)$$

comme l'impose (13.129). (13.130) signifie que $U(x, t = 0; x', 0)$ est le noyau identité : entre $t = 0$ et ... $t = 0$, la fonction d'onde n'évolue pas.

⁴⁹Cet opérateur est unitaire si H est hermitique, d'où la notation $U(t)$.

⁵⁰On pourrait aussi noter $|\psi_n\rangle = |E_n\rangle$.

Chapitre 14

Opérateurs

Ce chapitre est le dernier à être spécialement consacré au formalisme de la Mécanique Quantique et présente quelques propriétés importantes de certains opérateurs linéaires agissant dans l'espace vectoriel des états. Il sera question d'une part des opérateurs représentant les grandeurs physiques (observables), d'autre part des opérateurs unitaires. Pour les premiers, on procédera de façon intuitive, en s'appuyant sur les postulats énoncés antérieurement : ainsi apparaîtra naturellement une classe particulière d'opérateurs (les opérateurs hermitiques), seuls susceptibles de représenter une grandeur physique. Les opérateurs unitaires, eux, ont la propriété de conserver les angles et les longueurs : ils échangent entre elles les bases orthonormées et, par ailleurs, interviennent pour la propagation dans le temps (conservation de la probabilité). Finalement, le point sera fait sur les deux représentations \vec{r} et $-\vec{p}$.

14.1 Propriété fondamentale des observables : hermiticité

Les postulats montrent clairement le rôle central joué par les valeurs propres a_m et vecteurs propres $|a_m\rangle$ d'une observable A . Il convient donc à ce stade d'élucider les conditions qui assurent la réalité du spectre et élèvent ainsi un opérateur au rang d'observable. Les opérateurs représentant les grandeurs physiques sont les opérateurs dits hermitiques, possédant les deux propriétés fondamentales suivantes, déjà énoncées par anticipation :

1. toutes leurs valeurs propres sont réelles
2. deux vecteurs propres associés à deux valeurs propres *distinctes* sont orthogonaux¹.

Ces deux propriétés sont essentielles pour la cohérence de la théorie ; la première est manifestement une nécessité : d'après le postulat 3, la valeur que l'on trouve comme résultat d'une mesure de A appartient au spectre de A et cela n'aurait aucun sens de trouver un nombre complexe. La deuxième propriété assure la cohérence des postulats 4 (probabilité de trouver a_{m_0} avec la probabilité P_{m_0}) et 5 (réduction du paquet d'ondes). En effet, d'après ce dernier, juste après la mesure, le système est dans l'état $|a_{m_0}\rangle$ associé à la valeur propre que l'on vient de trouver ; toute nouvelle mesure immédiate de A doit redonner la même valeur propre a_{m_0} , ce qui impose à $|a_{m_0}\rangle$ d'être orthogonal à tout vecteur propre associé à une valeur propre $a_m \neq a_{m_0}$. L'orthogonalité assure que l'état juste après la mesure a une composante nulle sur tous les états propres ayant une autre valeur propre et la probabilité est bien nulle de trouver une valeur différente de a_{m_0} ; toute nouvelle mesure immédiate donne donc certainement la valeur déjà trouvée.

Les états propres de A sont par définition tels que :

$$A|a_m\rangle = a_m|a_m\rangle . \quad (14.1)$$

¹Deux (ou plus) vecteurs propres associés à la même valeur propre "n'arrivent" pas forcément orthogonaux, mais on peut toujours former les bonnes combinaisons linéaires conduisant à un jeu orthogonal.

L'équation étant homogène, tout vecteur propre d'un opérateur *linéaire* est défini à un facteur près. Ce facteur peut toujours être choisi de sorte que chaque $|a_m\rangle$ soit de norme (longueur) unité, ce que l'on supposera dans la suite :

$$\| |a_m\rangle \|^2 = 1 . \quad (14.2)$$

Si on effectue un grand nombre de mesures de A à partir d'un même état $|\Psi\rangle = \sum_m c_m |a_m\rangle$, on obtient successivement les différentes valeurs constituant le spectre de l'opérateur A ; le résultat de chaque mesure est imprévisible en tant que tel, mais le postulat 4 permet de calculer la probabilité P_m d'obtenir une valeur a_m prescrite à l'avance ; pour un vecteur $|\Psi\rangle$ normalisé à l'unité (toujours supposé dans la suite, sauf mention contraire), on a :

$$P_m = |c_m|^2 . \quad (14.3)$$

Le postulat 4, allié au postulat 5 permet de montrer physiquement que les états propres d'une observable forment une base *complète* de l'espace vectoriel des états \mathcal{E} . Plus précisément, en s'appuyant essentiellement sur ces postulats 4 et 5, on va montrer que ces états propres ont les propriétés suivantes :

1. à eux tous, ils forment un système de générateurs de \mathcal{E}
2. ils sont linéairement indépendants (ils forment donc une base de \mathcal{E})
3. ils sont deux à deux orthogonaux.

Qu'il s'agisse d'un système de générateurs est une évidence physique ; en effet, on doit pouvoir obtenir l'ensemble des valeurs a_m potentielles (une valeur qui ne serait jamais le résultat d'une mesure n'a pas de sens physique), ceci implique qu'un vecteur quelconque $|\Psi\rangle$ a, *a priori*, des composantes non-nulles sur tous les vecteurs propres $|a_m\rangle$, autrement dit ces vecteurs, à eux tous, engendrent tout l'espace vectoriel des états. Pour montrer qu'ils forment une base, il faut et suffit de montrer qu'ils sont linéairement indépendants.

Selon le postulat de réduction du paquet d'ondes, immédiatement après une mesure ayant donné la valeur (non-dégénérée) a_{m_0} à l'instant t_0 , le système est dans l'état propre (normalisé) $|a_{m_0}\rangle$:

$$|\Psi(t_0 + 0)\rangle = |a_{m_0}\rangle . \quad (14.4)$$

Les $\{|a_m\rangle\}$ forment un système de générateurs ; tout vecteur, et en particulier $|\Psi(t_0 + 0)\rangle$ peut donc se développer suivant :

$$|\Psi(t_0 + 0)\rangle = \sum_m c_m |a_m\rangle \iff \sum_m (c_m - \delta_{mm_0}) |a_m\rangle = 0 . \quad (14.5)$$

Pour montrer que les vecteurs propres $\{|a_m\rangle\}$ sont linéairement indépendants, il faut et suffit que cette dernière égalité soit équivalente à :

$$c_m - \delta_{mm_0} = 0 \quad \forall m . \quad (14.6)$$

Effectuons une deuxième mesure de A juste après la première, donc à partir de $|\Psi(t_0 + 0)\rangle$; la probabilité P_m d'obtenir la valeur a_m est égale à $|c_m|^2$ d'après le postulat 4 ; mais, d'après le postulat 5, elle doit être égale à 1 si $m = m_0$ et à zéro si $m \neq m_0$ - il faut donc que c_m soit égal à δ_{mm_0} , ce qui démontre l'indépendance linéaire.

Montrons enfin que les états propres d'une observable sont orthogonaux, en utilisant le postulat 4 ; le carré de la norme d'un vecteur d'état normalisé est :

$$\| |\Psi\rangle \|^2 = (|\Psi\rangle, |\Psi\rangle) = \sum_{n,m} c_n^* c_m (|a_n\rangle, |a_m\rangle) = \sum_n |c_n|^2 + \sum_{n \neq m} c_n^* c_m \langle a_n | a_m \rangle . \quad (14.7)$$

Les $|c_n|^2$ sont les probabilités P_n d'obtenir les différentes valeurs a_n ; la somme de ces probabilités vaut 1. Il reste donc :

$$\sum_{n \neq m} c_n^* c_m \langle a_n | a_m \rangle = 0 \quad \forall |\Psi\rangle . \quad (14.8)$$

Ceci est vrai quel que soit l'état $|\Psi\rangle$, donc quels que soient les coefficients c_n : les produits scalaires sont donc tous nuls. Ainsi, en prenant appui sur les postulats, on peut énoncer la conclusion majeure : les états propres d'une observable forment une base complète orthonormée de l'espace des états. Cette propriété, alliée à la réalité des valeurs propres, contraint les observables à appartenir à une classe d'opérateurs remarquables, les opérateurs hermitiques, aussi qualifiés d'*auto-adjoints*.

Donnons d'abord la définition de l'opérateur adjoint. Soit Ω un opérateur quelconque ; l'opérateur adjoint de Ω , noté Ω^\dagger , est l'opérateur tel que² :

$$(\Omega^\dagger|\psi\rangle, |\phi\rangle) = (|\psi\rangle, \Omega|\phi\rangle) \quad \forall |\psi\rangle, |\phi\rangle . \quad (14.9)$$

En général, $\Omega^\dagger \neq \Omega$, sauf précisément pour les opérateurs hermitiques dont chacun est, par définition, égal à son adjoint. Pour un opérateur hermitique (auto-adjoint) A , on a donc :

$$A^\dagger = A \iff (A|\psi\rangle, |\phi\rangle) = (|\psi\rangle, A|\phi\rangle) \quad \forall |\psi\rangle, |\phi\rangle . \quad (14.10)$$

L'adjoint $(\Omega^\dagger)^\dagger$ de Ω^\dagger est l'opérateur Ω lui-même³. En effet :

$$((\Omega^\dagger)^\dagger|\psi\rangle, |\phi\rangle) = (|\psi\rangle, \Omega^\dagger|\phi\rangle) . \quad (14.11)$$

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

$$(|\phi\rangle, (\Omega^\dagger)^\dagger|\psi\rangle) = (\Omega^\dagger|\phi\rangle, |\psi\rangle) . \quad (14.12)$$

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

$$(|\phi\rangle, (\Omega^\dagger)^\dagger|\psi\rangle) = (|\phi\rangle, \Omega|\psi\rangle) ; \quad (14.13)$$

les vecteurs $|\psi\rangle$ et $|\phi\rangle$ sont quelconques : la relation (14.12) s'élève au niveau des opérateurs, soit $(\Omega^\dagger)^\dagger = \Omega$.

Montrons maintenant que les observables sont de fait des opérateurs hermitiques. Soit une observable A dont on sait que sa base propre $\{|a_n\rangle\}$ est orthonormée. Soit A^\dagger l'adjoint de A ; en écrivant la définition de l'adjoint avec les vecteurs de cette base, on a :

$$(A^\dagger|a_n\rangle, |a_m\rangle) = (|a_n\rangle, A|a_m\rangle) \quad \forall |a_n\rangle, |a_m\rangle . \quad (14.14)$$

Comme $|a_n\rangle$ est propre de A , le second membre de (14.14) est égal à $a_m\delta_{nm}$. Par ailleurs, on peut toujours décomposer sur les $\{|a_m\rangle\}$ – qui forment une base – le vecteur résultant de l'action de A^\dagger sur $|a_n\rangle$; suivant la définition ordinaire de la matrice d'un opérateur sur une base, il vient :

$$A^\dagger|a_n\rangle = \sum_p A_{pn}^\dagger|a_p\rangle , \quad (14.15)$$

et, tenant compte de l'antilinearité du produit scalaire, le premier membre de (14.14) est :

$$\sum_p (A_{pn}^\dagger)^*(|a_p\rangle, |a_m\rangle) = \sum_p (A_{pn}^\dagger)^*\delta_{pm} = (A_{mn}^\dagger)^* ; \quad (14.16)$$

au total, il vient :

$$(A_{mn}^\dagger)^* = a_m\delta_{nm} \iff A_{mn}^\dagger = a_m^*\delta_{nm} . \quad (14.17)$$

Alors (14.15) se récrit :

$$A^\dagger|a_n\rangle = \sum_p a_n^*\delta_{np}|a_p\rangle = a_n^*|a_n\rangle . \quad (14.18)$$

²Cette définition ne fait aucune référence à une base éventuellement définie au préalable.

³L'espace de Hilbert est un espace à métrique symétrique où, par définition, il existe des bases orthonormées dont chacun des éléments a une norme positive (et prise égale à 1) – il n'en va pas toujours ainsi : on connaît des espaces vectoriels (dits symplectiques) dont les éléments sont dits isotropes car leur carré scalaire est nul ([19], § 2. 8).

Ainsi, quand les vecteurs propres de A sont orthogonaux⁴, l'opérateur A^\dagger adjoint de A a les mêmes vecteurs propres (à droite) que A , mais admet ses valeurs propres sont les complexes conjuguées de celles de A . Maintenant, comme A est une observable, ses valeurs propres a_n sont réelles. Il en résulte que les deux relations suivantes sont vraies :

$$A|a_n\rangle = a_n|a_n\rangle, \quad A^\dagger|a_n\rangle = a_n^*|a_n\rangle = a_n|a_n\rangle. \quad (14.19)$$

Elles montrent que A et A^\dagger sont alors un seul et même opérateur : A et A^\dagger , agissant sur le même vecteur d'une base, donnent le même résultat quel que soit ce vecteur de base. Ainsi, tout opérateur auto-adjoint (hermitique) a un spectre de valeurs propres réelles et ceci est, par définition même de l'adjoint, une propriété d'équivalence. Les postulats imposent la réalité des valeurs propres, les observables de la Mécanique Quantique sont donc des opérateurs hermitiques. Un tel opérateur doit être diagonalisable, toujours par référence aux postulats : on peut montrer que c'est bien le cas ; on peut également démontrer que les états propres d'un opérateur hermitique forment une base complète. Au total, la cohérence est complète entre les propriétés requises par les postulats pour tout opérateur représentant une grandeur physique et les propriétés spécifiques des opérateurs hermitiques.

Les opérateurs hermitiques constituent une généralisation des opérateurs symétriques, comme on le verra par la suite. On se souviendra de l'implication :

$$\text{opérateur hermitique} \implies \text{vecteurs propres orthogonaux} . \quad (14.20)$$

dont la réciproque est fautive en général⁵. En revanche, si un opérateur a *toutes* ses valeurs propres réelles et des vecteurs propres orthogonaux, alors il est hermitique.

14.2 Valeur moyenne d'une observable : utilisation de sa base propre

Une observable A étant donnée, ainsi qu'un état $|\Psi\rangle$ et ses répliques sur lesquelles on effectue une série de mesures de A , c'est évidemment la base propre de A qui se prête le mieux au calcul des valeurs moyennes. Il importe de bien comprendre ce que l'on veut calculer. Il est supposé possible de préparer un nombre arbitraire de fois le même état $|\Psi\rangle$ et pour chaque préparation, on effectue une mesure de A . Chaque mesure donne un résultat qui, fondamentalement, est imprévisible, tout comme lorsqu'un dé est lancé. On ne sait pas prévoir le résultat d'une mesure, mais seulement la probabilité d'obtenir une valeur donnée ou une autre du spectre de A . Quand l'opération de mesure est répétée sur le même état $|\Psi\rangle$, on peut dresser une statistique des valeurs obtenues. Si les fluctuations statistiques ont régressé à la précision souhaitable, les fréquences d'observation coïncident pratiquement avec les probabilités théoriques calculées suivant le postulat 4.

Par abus de langage, il est fréquent de parler de "la valeur moyenne de l'observable A dans l'état $|\Psi\rangle$ ". Ceci ne signifie pas que A a, préalablement à toute opération de mesure, une valeur déterminée dont la valeur moyenne est ce que l'on va calculer : en réalité, on n'en sait rien. Tout ce que l'on veut dire est à prendre au sens ci-dessus, c'est-à-dire au sens d'une espérance mathématique en Théorie des Probabilités, comparable au résultat moyen déduit d'une infinité de mesures.

Soit donc un état $|\Psi\rangle$ quelconque, préparable un nombre "infini" de fois, se développant sur les états propres de A suivant :

$$|\Psi\rangle = \sum_m c_m |a_m\rangle. \quad (14.21)$$

La probabilité d'obtenir a_m est $P_m = |c_m|^2$; il en résulte que la valeur moyenne (l'espérance mathématique) de la grandeur A est égale à :

$$\langle A \rangle = \sum_m P_m a_m = \sum_m |c_m|^2 a_m = \sum_m c_m^* c_m a_m. \quad (14.22)$$

⁴L'orthogonalité des vecteurs propres d'un opérateur n'assure pas que cet opérateur est hermitique. Soit A un opérateur hermitique (ses vecteurs propres sont donc orthogonaux). L'opérateur iA admet évidemment les mêmes vecteurs propres que A , ses valeurs propres sont égales à celles de A multipliées par i , et pourtant iA n'est visiblement pas hermitique : $(iA)^\dagger = i^* A = -iA$.

⁵Par exemple : l'opérateur $\Omega = e^{iA}$ est unitaire quand A est hermitique ; en tant que fonction de A , il a les mêmes vecteurs propres, qui sont orthogonaux puisque A est hermitique. En revanche, les valeurs propres de Ω ne sont manifestement pas réelles en général. L'orthogonalité des vecteurs propres d'un opérateur ne suffit donc pas à assurer que cet opérateur est hermitique.

Compte tenu de l'antilinearité du produit scalaire, cette moyenne est égale à :

$$\langle A \rangle = \langle \Psi | A | \Psi \rangle . \quad (14.23)$$

De la même façon, la valeur moyenne du carré de A est :

$$\langle A^2 \rangle = \langle \Psi | A^2 | \Psi \rangle , \quad (14.24)$$

d'où l'écart quadratique de A :

$$\Delta A^2 = \langle \Psi | A^2 | \Psi \rangle - \langle \Psi | A | \Psi \rangle^2 . \quad (14.25)$$

Si $|\Psi\rangle$ est en fait l'un des états propres de A , $|\Psi\rangle = |a_{m_0}\rangle$, tous les coefficients c_m sont nuls sauf c_{m_0} qui est de module 1 ($c_{mm_0} = e^{i\alpha}\delta_{mm_0}$). Il en résulte que :

$$\langle A \rangle = \sum_m |e^{i\alpha}\delta_{mm_0}|^2 a_m = a_{m_0} , \quad \langle A^2 \rangle = \sum_m |e^{i\alpha}\delta_{mm_0}|^2 a_m^2 = a_{m_0}^2 . \quad (14.26)$$

Ceci signifie que les mesures répétées donneront toujours le même résultat a_{m_0} ; évidemment, dans ces conditions, l'écart quadratique est nul, $\Delta A^2 = 0$. Dans cette situation très particulière, l'observable A a donc une valeur certaine et ne présente aucune dispersion. Comme déjà observé, c'est ainsi qu'apparaît la conservation de l'énergie : si $|\Psi\rangle$ est un état propre $|E_{m_0}\rangle$ du Hamiltonien d'un système isolé – donc indépendant du temps – toute mesure de l'énergie donne le même résultat E_{m_0} à tout instant puisque c'est précisément H qui pilote l'évolution de ce système.

Pour trouver toutes les valeurs moyennes des puissances de A , il est utile d'introduire une fonction contenant toutes, souvent appelée fonction caractéristique, ou fonction génératrice. La terminologie est inspirée de la Théorie des Probabilités, mais, évidemment, la “fonction” est ici un opérateur. Considérons en effet par exemple l'opérateur⁶:

$$Z = e^{izA} . \quad (14.27)$$

Il convient tout d'abord de donner un sens à cette expression ; on la comprend comme étant un raccourci d'écriture pour le développement :

$$Z = \sum_{n=0}^{+\infty} \frac{(iz)^n}{n!} A^n \quad (14.28)$$

qui, indépendamment des questions de convergence, implique la puissance $n^{\text{ème}}$ de l'opérateur A , une opération parfaitement définie. La convergence d'une telle série ne pose pas sérieusement de problème : on comprend que si $|a\rangle$ est un vecteur propre de A associé à la valeur propre a , alors :

$$Z|a\rangle = e^{izA}|a\rangle = \sum_{n=0}^{+\infty} \frac{(iz)^n}{n!} A^n |a\rangle = \sum_{n=0}^{+\infty} \frac{(iz)^n}{n!} a^n |a\rangle = e^{iza}|a\rangle . \quad (14.29)$$

La série exponentielle converge partout dans le plan complexe ouvert (z n'est pas forcément réel) ; finalement, la seule condition à respecter pour que cette suite d'opérations ait un sens est que l'opérateur A n'ait aucune valeur propre infinie, ce qui est toujours vrai si A est une observable (aucune grandeur physique ne peut être égale à l'infini et le spectre de A doit contenir toutes les valeurs physiques possibles). On note au passage que pour toute fonction numérique $f(x)$ développable en série (entière par exemple, mais ce n'est pas obligatoire⁷) on pourra définir une fonction $f(A)$ (donc un opérateur) et qu'un calcul semblable à celui effectué ci-dessus montre que si $|a\rangle$ est propre de A avec la valeur propre a alors :

$$f(A)|a\rangle = f(a)|a\rangle , \quad (14.30)$$

où $f(a)$ est maintenant un nombre.

Revenant à la fonction génératrice (14.27), on voit que la valeur moyenne de Z pour l'état $|\Psi\rangle$ est :

$$\langle Z(z) \rangle = \sum_m P_m e^{iza_m} = \sum_m |c_m|^2 e^{iza_m} . \quad (14.31)$$

⁶La notation Z est utilisée en réminiscence d'une fonction de partition.

⁷Avec une fonction multiforme, il suffit de définir convenablement la(es) coupure(s).

Comme :

$$\left(\frac{d^r}{dz^r} \langle Z(z) \rangle \right)_{z=0} = \left(\sum_m (ia_m)^r P_m e^{iza_m} \right)_{z=0} = i^r \sum_m P_m a_m^r, \quad (14.32)$$

on voit à l'envers que :

$$\langle A^r \rangle = i^{-r} \left(\frac{d^r}{dz^r} \langle Z(z) \rangle \right)_{z=0}. \quad (14.33)$$

La fonction $\langle Z(z) \rangle$, si elle est connue, permet de calculer toutes les puissances de A par simple dérivation en $z = 0$.

14.3 Représentation des opérateurs hermitiques et des opérateurs unitaires

La base propre de A est évidemment celle qui se prête le mieux au calcul de l'espérance mathématique de A . Tout opérateur A est intrinsèquement défini par son spectre $\{a_n\}$, indépendamment de toute référence à une base de l'espace vectoriel : par exemple, un opérateur est hermitique quand toutes ses valeurs propres sont réelles et ceci est une propriété en soi de l'opérateur.

D'un autre côté, une base étant donnée, un opérateur Ω est représenté par une matrice, dont les colonnes sont constituées par les composantes des vecteurs de la base transformés par Ω . Par définition, la matrice d'une observable A sur une base $\{|e_n\rangle\}$ permet d'écrire les combinaisons linéaires :

$$A|e_n\rangle = \sum_m A_{mn} |e_m\rangle. \quad (14.34)$$

Supposons toujours par ailleurs que la base est orthonormée et multiplions scalairement par l'un quelconque des vecteurs de base, soit $|a_p\rangle$:

$$\langle |a_p\rangle, A|e_n\rangle = \sum_m A_{mn} \langle |a_p\rangle, |e_m\rangle = \sum_m A_{mn} \delta_{pm} = A_{pn}. \quad (14.35)$$

Lue à l'envers, cette relation montre que l'élément de matrice *sur une base* \perp_1 est donné par :

$$A_{pn} = \langle |a_p\rangle, A|e_n\rangle \equiv \langle e_p | A | e_n \rangle \text{ (base } \perp_1 \text{)}. \quad (14.36)$$

La matrice est la représentation de l'opérateur A sur la base adoptée, tout comme les composantes d'un vecteur sur une base représentent ce vecteur sur cette base. Considérons maintenant une observable A , de spectre $\{a_n\}$ et choisissons comme base précisément les vecteurs propres de A , $|a_n\rangle$; alors, la matrice représentant A est diagonale, ses éléments diagonaux étant les $\{a_n\}$ et les valeurs moyennes de A , ou de $f(A)$, se calculent immédiatement.

Il convient cependant de savoir déterminer ces valeurs moyennes quand la base est quelconque vis-à-vis de A ; dans un tel cas, la matrice de A n'est plus diagonale. Soit donc une autre base orthonormée, $\{|b_p\rangle\}$; on passe d'une base à l'autre par des relations impliquant les éléments de matrice d'un opérateur unitaire l'opérateur⁸ U . Un calcul simple montre que, sur la base $\{|b_p\rangle\}$, l'observable A est représentée par la matrice d'éléments :

$$A_{pr} = \sum_n U_{pn}^{-1} a_n U_{nr} = \sum_n U_{np}^* a_n U_{nr}, \quad (14.37)$$

où l'unitarité de U a été prise en compte pour transformer le second membre. Pour la même raison, (14.37) s'écrit aussi :

$$A_{pr} = \sum_n U_{np}^* a_n U_{rn}^{-1*} = \left(\sum_n U_{np} a_n U_{rn}^{-1} \right)^* \quad (14.38)$$

⁸L'opérateur inverse U^{-1} existe forcément, s'agissant du passage d'une base à une autre.

où on a utilisé le fait que $a_n \in \mathbb{R}$; maintenant, modifiant l'ordre des facteurs dans la somme de droite et en comparant avec (14.37), on voit que :

$$A_{pr} = \left(\sum_n U_{rn}^{-1} a_n U_{np} \right)^* \equiv A_{rp}^* . \quad (14.39)$$

Sur une base \perp_1 , une observable est donc représentée par une matrice qui coïncide avec sa transposée conjuguée – ce que l'on appelle une matrice hermitique :

$$A^\dagger = A \iff A_{pr} = A_{rp}^* \quad (\text{base } \perp_1) . \quad (14.40)$$

Il s'agit bien d'une généralisation des matrices symétriques réelles⁹.

Les mêmes arguments permettent de trouver la matrice de l'adjoint d'un opérateur, Ω^\dagger , connaissant la matrice de Ω . Par la définition de l'adjoint (14.9), on a :

$$(\Omega^\dagger |e_p\rangle, |e_n\rangle) = (|e_p\rangle, \Omega |e_n\rangle) = \langle e_p | \Omega | e_n \rangle = \Omega_{pn} . \quad (14.41)$$

Si on prend les complexes conjugués des deux membres et que l'on renverse l'ordre dans le produit scalaire de gauche :

$$(|e_n\rangle, \Omega^\dagger |e_p\rangle) = \Omega_{pn}^* ; \quad (14.42)$$

mais le premier membre n'est autre que $(\Omega^\dagger)_{np}$. On en déduit :

$$\Omega_{np}^\dagger = \Omega_{pn}^* ; \quad (14.43)$$

La matrice de l'adjoint de Ω s'obtient donc en transposant et en conjuguant la matrice de Ω . Bien sûr, si Ω est hermitique, $\Omega = \Omega^\dagger$, et on trouve $\Omega_{np} = \Omega_{pn}^*$: c'est la relation de symétrie déjà trouvée (voir (14.40)).

Les opérateurs unitaires transforment une base \perp_1 en une autre base \perp_1 (ils conservent les angles et les longueurs). Cette propriété caractéristique se traduit par :

$$(U|\phi\rangle, U|\psi\rangle) = (|\phi\rangle, |\psi\rangle) \quad \forall |\phi\rangle, |\psi\rangle , \quad (14.44)$$

et, en particulier :

$$(U|e_n\rangle, U|e_m\rangle) = (|e_n\rangle, |e_m\rangle) . \quad (14.45)$$

Utilisant la définition première de l'adjoint d'un opérateur¹⁰, il vient :

$$(|e_n\rangle, U^\dagger U |e_m\rangle) = (|e_n\rangle, |e_m\rangle) , \quad (14.46)$$

d'où l'on déduit :

$$U^\dagger U = \mathbf{1} \iff U^\dagger = U^{-1} \iff U U^\dagger = \mathbf{1} . \quad (14.47)$$

De $U^{-1} = U^\dagger$, on déduit que la matrice de l'inverse d'un opérateur unitaire U s'obtient donc en transposant et en conjuguant la matrice de U .

En résumé, on a les propriétés suivantes pour les matrices rapportées à une base orthonormée :

- opérateur adjoint : $\Omega_{np}^\dagger = \Omega_{pn}^*$
- opérateurs hermitiques ($A^\dagger = A$) : $A_{mn} = A_{nm}^*$
- opérateurs unitaires ($U^\dagger = U^{-1}$) : $U_{np}^{-1} = U_{pn}^*$

La propriété d'unitarité ou d'hermiticité d'un opérateur ne se transpose à sa matrice que si la base est orthonormée.

⁹Tout comme les matrices unitaires généralisent les matrices orthogonales.

¹⁰et en utilisant aussi $(U^\dagger)^\dagger = U$.

14.4 Retour sur la notation de Dirac

La notation de Dirac est une façon d'écrire automatiquement toutes les relations traduisant le jeu de l'algèbre vectorielle et se révèle à l'usage d'une extrême commodité. Cette utilité est manifeste quand on revient sur les diverses relations obtenues précédemment en les reprenant dans cette notation.

Par exemple, la relation entre élément de matrice et produit scalaire, (14.36), se retrouve immédiatement comme le montre la séquence suivante :

$$A|e_n\rangle = \mathbf{1}A|e_n\rangle = \left(\sum_m |e_m\rangle\langle e_m| \right) A|e_n\rangle = \sum_m |e_m\rangle\langle e_m|A|e_n\rangle \equiv \sum_m |e_m\rangle A_{mn} . \quad (14.48)$$

La notation de Dirac permet aussi de multiplier automatiquement les opérateurs. On sait que :

$$(\Omega \Xi)_{nm} = \sum_p \Omega_{np} \Xi_{pm} . \quad (14.49)$$

En notation de Dirac, on écrit :

$$\langle e_n|\Omega \Xi|e_m\rangle = \langle e_n|\Omega \mathbf{1} \Xi|e_m\rangle = \sum_p \langle e_n|\Omega|e_p\rangle\langle e_p|\Xi|e_m\rangle \equiv \sum_p \Omega_{np} \Xi_{pm} . \quad (14.50)$$

La commodité de la notation de Dirac est également évidente pour écrire la relation entre les éléments des matrices représentant le même opérateur sur deux bases orthonormées différentes, avec lesquelles on peut écrire de deux façons la relation de fermeture :

$$\sum_n |e_n\rangle\langle e_n| = \sum_p |f_p\rangle\langle f_p| = \mathbf{1} . \quad (14.51)$$

En insérant l'expression de $\mathbf{1}$ convenablement, on a :

$$\begin{aligned} \langle e_n|\Omega|e_m\rangle &= \langle e_n|\mathbf{1}\Omega\mathbf{1}|e_m\rangle \\ &= \sum_p \sum_q \langle e_n|f_p\rangle\langle f_p|\Omega|f_q\rangle\langle f_q|e_m\rangle \equiv \sum_p \sum_q U_{np}\langle f_p|\Omega|f_q\rangle U_{qm}^{-1} . \end{aligned} \quad (14.52)$$

En jouant avec la relation de fermeture, on peut écrire, pour tout opérateur $\Omega = \mathbf{1}\Omega\mathbf{1}$ soit :

$$\Omega = \sum_m \sum_n |e_m\rangle\langle e_m|\Omega|e_n\rangle\langle e_n| = \sum_{mn} \langle e_m|\Omega|e_n\rangle |e_m\rangle\langle e_n| . \quad (14.53)$$

Ceci est, pour un opérateur, l'équivalent du développement d'un vecteur sur une base en termes de ses composantes. Les objets $|e_m\rangle\langle e_n|$, appelés dyades, jouent pour cette décomposition d'un opérateur le même rôle que les éléments d'une base pour la décomposition d'un vecteur¹¹; les Ω_{mn} peuvent être appelés les "composantes" de l'opérateur Ω sur la base $\{|e_m\rangle\}$; le développement précédent, sur la base propre de Ω , prend la forme simple :

$$\Omega = \sum_m |\omega_m\rangle\omega_m\langle\omega_m| \quad (14.54)$$

et, pour des raisons évidentes, s'appelle décomposition spectrale de Ω .

Examinons ensuite la relation d'hermiticité en notation de Dirac. La définition de l'adjoint d'un opérateur Ω est :

$$(\Omega^\dagger|\psi\rangle, |\phi\rangle) = (|\psi\rangle, \Omega|\phi\rangle) . \quad (14.55)$$

¹¹Les opérateurs linéaires agissant dans un espace vectoriel forment une algèbre (deux lois de composition interne, la multiplication et l'addition, et une loi externe, multiplication par un scalaire du corps). Grâce à ces deux lois, l'ensemble de ces opérateurs possède une sous-structure d'espace vectoriel, pour lequel on sait définir des bases. On peut aussi définir le produit scalaire de deux opérateurs (par exemple $(\Omega, \Xi) = \text{Tr}(\Omega^\dagger \Xi)$), ce qui donne à l'espace vectoriel inclus dans l'algèbre une structure de Hilbert.

Transposé dans la notation de Dirac, ceci s'écrit :

$$\langle \Omega^\dagger \psi | \phi \rangle = \langle \psi | \Omega \phi \rangle \equiv \langle \psi | \Omega | \phi \rangle . \quad (14.56)$$

Ainsi, quand on fait passer un opérateur du *ket* au *bra* (ou inversement), *il faut en prendre l'adjoint*. Pour une observable ($A^\dagger = A$), on a :

$$\langle A \psi | \phi \rangle = \langle \psi | A | \phi \rangle \quad (A^\dagger = A) . \quad (14.57)$$

Comme le produit scalaire est changé en son conjugué par transposition des deux vecteurs :

$$\langle \psi | A | \phi \rangle = \langle \phi | A | \psi \rangle^* \quad (A^\dagger = A) ; \quad (14.58)$$

en particulier, si $|\psi\rangle = |\phi\rangle$:

$$\langle \psi | A | \psi \rangle = \langle \psi | A | \psi \rangle^* \quad (A^\dagger = A) \quad (14.59)$$

et ceci n'exprime rien d'autre que la condition de réalité pour l'espérance mathématique d'une observable à la suite de mesures répétées effectuées sur le même état $|\psi\rangle$.

14.5 Combinaisons d'opérateurs

On présente ici quelques opérations fondamentales sur les opérateurs ; les unes sont évidentes et se passent de commentaires ; les autres sont utiles à connaître et il faut savoir qu'elles existent. Comme toujours, la relation entre opérateurs se définit en précisant l'action de ces opérateurs sur un vecteur quelconque de l'espace : le quantificateur (exemple : $\forall |\psi\rangle$) permet d'élever la relation entre vecteurs à une relation entre opérateurs. L'opération la plus simple est la multiplication par un scalaire. Soit Ω un opérateur et λ un nombre du corps ; on définit l'opérateur $\Xi = \lambda\Omega$ par :

$$\Xi |\psi\rangle = \lambda(\Omega |\psi\rangle) \quad \forall |\psi\rangle \quad (14.60)$$

Attention : un scalaire ne commute avec un opérateur que si ce dernier est linéaire. Très précisément :

$$\forall \lambda \in \mathbb{C} , \quad \Omega(\lambda |\psi\rangle) = \begin{cases} \lambda \Omega |\psi\rangle & \text{si } \Omega \text{ est linéaire} \\ \lambda^* \Omega |\psi\rangle & \text{si } \Omega \text{ est antilinéaire} \end{cases} \quad (14.61)$$

Si A est hermitique, λA n'est hermitique que si $\lambda \in \mathbb{R}$ – comme le montre la relation de définition de l'adjoint. On a en effet :

$$(\lambda \Omega)^\dagger = \lambda^* \Omega^\dagger . \quad (14.62)$$

De même, si U est un opérateur unitaire, $T = \lambda U$ n'est pas forcément unitaire. En effet :

$$(\lambda U)^\dagger = \lambda^* U^\dagger = \lambda^* U^{-1} \quad (14.63)$$

Il en résulte que l'on a $T^\dagger = T^{-1}$ si et seulement si $\lambda^* = 1/\lambda$; autrement dit, le scalaire λ doit être de module 1, de la forme $e^{i\alpha}$, avec $\alpha \in \mathbb{R}$.

La somme de deux opérateurs est définie comme :

$$(\Omega + \Xi) |\psi\rangle = \Omega |\psi\rangle + \Xi |\psi\rangle \quad (14.64)$$

En combinant les différentes opérations, on a :

$$(\lambda \Omega + \mu \Xi) |\psi\rangle = \lambda \Omega |\psi\rangle + \mu \Xi |\psi\rangle \quad (14.65)$$

montrant en effet que les opérateurs linéaires définis sur un espace vectoriel forment eux-mêmes un espace vectoriel. Comme la relation de conjugaison hermitique est antilinéaire, on a :

$$(\lambda \Omega + \mu \Xi)^\dagger = \lambda^* \Omega^\dagger + \mu^* \Xi^\dagger . \quad (14.66)$$

Attention, il faut se garder d'écrire :

$$(\Omega + \Xi)^{-1} = \Omega^{-1} + \Xi^{-1} \quad \text{faux !} \quad (14.67)$$

Le produit de deux opérateurs A et B se définit comme on le sait :

$$(AB)|\psi\rangle = A(B|\psi\rangle) \quad (14.68)$$

et est donc le résultat de l'action successive de B , puis de A (attention à l'ordre des opérations : AB est en général différent de BA !).

Examinons maintenant l'opérateur adjoint d'un produit. Soit $C = AB$. Pour trouver C^\dagger en fonction de A^\dagger et de B^\dagger , il suffit de revenir à la définition :

$$((AB)^\dagger|\psi\rangle, |\phi\rangle) = (|\psi\rangle, (AB)|\phi\rangle) . \quad (14.69)$$

A partir du second membre, en écrivant à l'envers la définition de l'adjoint, on a :

$$(|\psi\rangle, AB|\phi\rangle) = (A^\dagger|\psi\rangle, B|\phi\rangle) \quad (14.70)$$

puis :

$$(A^\dagger|\psi\rangle, B|\phi\rangle) = (B^\dagger A^\dagger|\psi\rangle, |\phi\rangle) \quad (14.71)$$

En rapprochant (14.69) et (14.71), on en déduit :

$$(AB)^\dagger = B^\dagger A^\dagger . \quad (14.72)$$

Ainsi, sous l'opération de conjugaison hermitique, l'ordre des opérateurs s'inverse. Plus généralement :

$$(ABC \dots XYZ)^\dagger = Z^\dagger Y^\dagger X^\dagger \dots C^\dagger B^\dagger A^\dagger \quad (14.73)$$

Il en résulte notamment :

$$[\Omega, \Xi]^\dagger = -[\Xi^\dagger, \Omega^\dagger] , \quad (14.74)$$

de sorte que pour deux opérateurs hermitiques A et B :

$$[A, B]^\dagger = -[B, A] \quad (14.75)$$

et c'est pourquoi, en pareille circonstance, il est utile de poser :

$$[A, B] = iD , \quad (14.76)$$

ce qui définit alors un opérateur D lui-même hermitique (exemples : $[x, p_x] = i\hbar\mathbf{1}$, $[J_x, J_y] = i\hbar J_z$, etc.). On voit aussi que le produit de deux opérateurs hermitiques n'est lui-même hermitique que si les deux opérateurs commutent entre eux. Enfin, les résultats ci-dessus conduisent à :

$$e^{itA} \text{ unitaire} \iff t \in \mathbb{R} \text{ et } A = A^\dagger . \quad (14.77)$$

Par exemple : l'opérateur d'évolution $U(t) = e^{-(i/\hbar)Ht}$ d'un système isolé est unitaire puisque le temps t est réel et H est hermitique. Enfin, on vérifie immédiatement que l'inverse d'un produit est le produit renversé des inverses :

$$(UVW \dots Z)^{-1} = Z^{-1} \dots W^{-1}V^{-1}U^{-1} . \quad (14.78)$$

Plus généralement, les différents résultats ci-dessus permettent d'élucider les propriétés d'hermiticité et d'unitarité des fonctions d'opérateurs. En outre, la plupart des opérations faites sur les fonctions numériques sont généralisables aux fonctions d'opérateurs ou aux opérateurs dépendant d'un paramètre scalaire ; on peut ainsi dériver ou intégrer les opérateurs. La vigilance doit cependant toujours être de règle en ce qui concerne la commutativité des produits ; par exemple, pour deux opérateurs quelconques A et B :

$$e^{A+B} \neq e^A e^B . \quad (14.79)$$

En cas de doute sur la légitimité d'une opération algébrique d'apparence banale s'agissant de simples nombres, le mieux est de revenir aux définitions. Par exemple on se convainc de (14.79), en effectuant les développements en série et en remarquant que l'égalité entre *nombres* :

$$e^{x+y} = e^x e^y \quad (x, y \in \mathbb{C}) , \quad (14.80)$$

est une conséquence directe du fait que le produit de deux nombres ordinaires¹² est indifférent à l'ordre des facteurs – ce qui n'est pas le cas pour un produit d'opérateurs.

¹²On connaît des nombres obéissant à une algèbre non-commutative (exemple : les nombres de Grassmann).

14.6 Représentation- \vec{r}

Une représentation, en Mécanique Quantique, résulte finalement du choix d'une base particulière pour rapporter l'espace vectoriel des états. Dans les paragraphes précédents, on a formellement considéré le cas d'une représentation discrète, mais il est également possible d'utiliser des représentations continues. Historiquement, c'est une telle représentation qui est apparue naturellement en premier : la Mécanique Ondulatoire de Schrödinger et son ingrédient primordial, la fonction d'onde ; la fonction d'onde de Schrödinger est la représentation du vecteur d'état où la coordonnée q joue le rôle d'un indice continu pour repérer les vecteurs de base. On sait que l'on peut identifier $\Psi(\nu)$ et $\langle \nu | \Psi \rangle$ dans le cas d'une variable continue. Pour la fonction d'onde de Schrödinger, on en vient ainsi à écrire l'égalité :

$$\langle \vec{r} | \Psi \rangle = \Psi(\vec{r}) , \quad (14.81)$$

Pour donner vraiment un sens à cette relation, dans le contexte désormais familier où $|a_n\rangle$ désigne l'état propre de l'observable A ayant la valeur propre a_n , il convient de définir les vecteurs propres de l'observable position. Afin de faciliter les écritures, on se place dans une seule direction d'espace : la généralisation à \mathbb{R}^3 est immédiate. Soit donc l'observable position, q , pour une particule confinée sur un axe. Si à l'opérateur moment conjugué p correspond l'opérateur différentiel $-i\hbar d/dq$, à la position correspond la simple multiplication par q . Ecrire l'équation aux vecteurs propres pour l'observable q , c'est donc chercher des fonctions $\psi_{q_0}(q)$ telles que :

$$q\psi_{q_0}(q) = q_0\psi_{q_0}(q) , \quad (14.82)$$

q_0 désignant précisément la valeur propre. Une telle équation n'a pas de solution dans l'espace des fonctions ordinaires. En effet, elle s'écrit :

$$(q - q_0)\psi_{q_0}(q) = 0 , \quad (14.83)$$

de sorte que de deux choses l'une :

- ou bien $q \neq q_0$ et alors $\psi_{q_0}(q) = 0$, puisque l'on peut diviser par $(q - q_0)$
- ou bien $q = q_0$ et alors $\psi_{q_0}(q)$ est indéterminée.

Ainsi, la fonction cherchée est nulle partout sauf en un point. On se retrouve en fait ici exactement dans la même situation que celle rencontrée précédemment (Ch. 13), qui a conduit à l'introduction "avec les mains" de la fonction de Dirac ; il est ainsi possible d'identifier suivant :

$$\psi_{q_0}(q) = C \delta(q - q_0) , \quad (14.84)$$

où C est une constante : $C\delta(q - q_0)$ est ainsi une fonction propre de q pour la valeur propre q_0 . Cette fonction est manifestement non-normalisable : on ne sait pas donner de sens au carré d'une "fonction" de Dirac, même par un processus de limite¹³ ; d'ailleurs, la transformée de Fourier de $\delta(x - x_0)$ est $\propto e^{ikx_0}$, qui n'est pas davantage normalisable. En conséquence, la constante C doit être fixée par une convention additionnelle.

Dans le droit fil de la notation utilisée antérieurement, on peut noter $|q_0\rangle$ l'état propre de l'observable q associé à la valeur propre q_0 ; ainsi :

$$q|q_0\rangle = q_0|q_0\rangle , \quad (14.85)$$

et alors (14.84) s'écrit $\langle q|q_0\rangle = C \delta(q - q_0)$. Pour avoir la normalisation à l'unité on prend donc $C = 1$, d'où finalement :

$$\psi_{q_0}(q) = \delta(q - q_0) = \langle q|q_0\rangle , \quad (14.86)$$

Deux états distincts, $|q_0\rangle$ et $|q_1\rangle$ sont orthogonaux puisqu'ils sont propres d'une observable et associés à deux valeurs propres distinctes ; par ailleurs, l'ensemble des $|q\rangle$ doit former une base complète. Il est donc possible de décomposer un vecteur quelconque $|\psi\rangle$ sur une telle base, ce qui revient à écrire :

$$|\psi\rangle = \int dq |q\rangle \langle q|\psi\rangle , \quad (14.87)$$

¹³Par un processus de limite standard, il est facile de montrer qu'un tel carré vaut $+\infty$.

ainsi que toutes les relations donnant le produit scalaire, la norme, *etc.* (voir Ch. 13). En ce qui concerne les éléments de matrice d'une observable A , commençons par le cas le plus simple, celui où A ne dépend que de la coordonnée q .

Sur une base $\{|e_m\rangle\}$ orthonormée donnée, on a :

$$A_{mn} = \langle e_m | A | e_n \rangle . \quad (14.88)$$

Injectons la relation de fermeture continue aux bons endroits; il vient :

$$A_{mn} = \int dq' \int dq'' \langle e_m | q' \rangle \langle q' | A(q) | q'' \rangle \langle q'' | e_n \rangle , \quad (14.89)$$

ce qui s'écrit aussi :

$$A_{mn} = \int dq' \int dq'' e_m^*(q') \langle q' | A(q) | q'' \rangle e_n(q'') . \quad (14.90)$$

Pour calculer $\langle q' | A(q) | q'' \rangle$, il suffit de se souvenir de la règle donnant le résultat de l'action d'une fonction d'opérateur sur un vecteur propre de cet opérateur ; ici il vient :

$$A(q) | q' \rangle = A(q') | q' \rangle \quad (14.91)$$

où $A(q')$ est un simple nombre (c'est la valeur propre de l'opérateur $A(q)$ associée au vecteur propre $|q'\rangle$) ; il en résulte :

$$\langle q'' | A(q) | q' \rangle = A(q') \langle q'' | q' \rangle = A(q') \delta(q' - q'') \quad (14.92)$$

En représentation- q , toute observable ne dépendant que de q est représentée par une "matrice" (noyau) diagonale. On déduit alors de (14.90) :

$$A_{mn} = \int dq' \int dq'' e_m^*(q') A(q') \delta(q' - q'') e_n(q'') = \int dq e_m^*(q) A(q) e_n(q) . \quad (14.93)$$

En particulier, on retrouve bien que la valeur moyenne d'une observable du genre $A(q)$ dans un état $|\Psi\rangle$ est donnée par :

$$\langle A \rangle = \int dq \Psi^*(q) A(q) \Psi(q) = \int dq A(q) |\Psi(q)|^2 ; \quad (14.94)$$

l'expression (14.93) énonce, en représentation- q , la règle d'usage pour calculer l'élément de matrice d'un opérateur ne dépendant que de la coordonnée.

Prenons maintenant le cas d'une observable qui ne dépend que de l'impulsion p , $B(p)$. En injectant deux fois la relation de fermeture, comme précédemment :

$$B_{mn} = \int dq' \int dq'' \langle e_m | q' \rangle \langle q' | B(p) | q'' \rangle \langle q'' | e_n \rangle , \quad (14.95)$$

et la quantité à trouver est $\langle q' | B(p) | q'' \rangle$. Il convient donc en premier lieu d'obtenir le noyau $\langle q' | p | q'' \rangle$ lui-même : celui de toute fonction de p développable en série s'en déduira par la règle habituelle. La question est de donner un sens à l'opération $p|q_0\rangle$, ou, transposé dans le langage des fonctions d'onde, à la quantité $\langle q | p | q_0 \rangle \equiv \langle q | p q_0 \rangle \equiv \langle q | p \psi_{q_0} \rangle \equiv p \delta(q - q_0)$:

$$p \delta(q - q_0) \equiv -i\hbar \frac{\partial}{\partial q} \delta(q - q_0) = ??? \quad (14.96)$$

Ceci revient à définir la "dérivée de la fonction" δ . Pour élucider cette opération, il suffit de réintroduire une fonction du genre δ_ε , dont l'intégrale vaut 1 et dont la dérivée est parfaitement définie. Pour toute fonction $f(q)$ assez régulière, l'intégrale I suivante a un sens :

$$I = \int_{-\infty}^{+\infty} f(q) \delta_\varepsilon(q - q_0) dq . \quad (14.97)$$

L'intervalle d'intégration peut être quelconque, mais doit contenir largement le point intéressant q_0 , faute de quoi $I = 0$. L'intégrale I se transforme par une intégration par parties :

$$I = [f(q)\delta_\varepsilon(q - q_0)]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \frac{\partial f}{\partial q} \delta_\varepsilon(q - q_0) dq . \quad (14.98)$$

Pour une fonction $f(q)$ nulle aux extrémités de l'intervalle d'intégration – et aussi parce que le précurseur δ_ε peut toujours être défini nul au bornes –, le terme tout intégré est nul ; il reste :

$$I = - \int_{-\infty}^{+\infty} \frac{\partial f}{\partial q} \delta_\varepsilon(q - q_0) dq . \quad (14.99)$$

f (et sa dérivée) varie(nt) lentement à l'échelle ε : on peut donc extraire de l'intégrale la dérivée $\partial f/\partial q$ en la prenant en q_0 ; comme la fonction δ_ε a une surface unité, il vient finalement :

$$I \simeq - \left(\frac{\partial f}{\partial q} \right)_{q_0} , \quad (14.100)$$

et, à la limite $\varepsilon = 0_+$, on obtient l'égalité stricte. La dérivée $(\partial/\partial q)\delta(q - q_0)$ agit donc comme suit :

$$\int_{-\infty}^{+\infty} f(q) \frac{\partial}{\partial q} \delta(q - q_0) dq = -f'(q_0) \iff \int_{-\infty}^{+\infty} f(q) \delta'(q - q_0) dq = -f'(q_0) . \quad (14.101)$$

Le résultat (14.101) donne un sens à l'écriture (14.96). La “fonction” δ' est visiblement impaire :

$$\delta'(q - q_0) = -\delta'(q_0 - q) , \quad (14.102)$$

comme on peut s'en convaincre en examinant une fonction précurseur δ_ε ou en se souvenant (sans états d'âme) que la dérivée d'une fonction paire est une fonction impaire. Par ailleurs, ce pas étant franchi, on imagine facilement le mode de définition des dérivées d'ordre supérieur de la fonction δ .

$\delta(q - q_0)$ est une fonction propre de la position ; c'est la fonction d'onde associée au *ket* $|\psi_{q_0}\rangle \equiv |q_0\rangle$, égale à $\psi_{q_0}(q) = \langle q|\psi_{q_0}\rangle \equiv \langle q|q_0\rangle$. Ainsi :

$$p \delta(q - q_0) = -i\hbar \delta'(q - q_0) \iff p \psi_{q_0}(q) = -i\hbar \delta'(q - q_0) . \quad (14.103)$$

soit :

$$p \psi_{q_0}(q) \equiv (p \psi_{q_0})(q) \equiv \langle q|p\psi_{q_0}\rangle = -i\hbar \delta'(q - q_0) , \quad (14.104)$$

mais $p|\psi_{q_0}\rangle \equiv p|q_0\rangle$, d'où en définitive :

$$\langle q|p|q_0\rangle = -i\hbar \delta'(q - q_0) . \quad (14.105)$$

En revenant à des notations plus symétriques, le résultat est donc :

$$\langle q|p|q'\rangle = -i\hbar \delta'(q - q') . \quad (14.106)$$

Comme la fonction δ' est impaire, on a :

$$\langle q|p|q'\rangle^* = +i\hbar \delta'(q - q') = -i\hbar \delta'(q' - q) = \langle q'|p|q\rangle \quad (14.107)$$

qui exprime l'hermiticité de la “matrice” représentant l'observable p sur une base orthonormée au sens généralisé, ici sur la base $\{|q\rangle\}$.

Connaissant le noyau fondamental $\langle q|p|q'\rangle$, donné par (14.106), on peut en déduire d'abord l'élément de matrice p_{nm} sur une base orthonormée $\{|e_n\rangle\}$:

$$\begin{aligned} p_{nm} &= \langle e_n|p|e_m\rangle = \int dq' \int dq'' e_n^*(q') \langle q'|p|q''\rangle e_m(q'') \\ &= -i\hbar \int dq' \int dq'' e_n^*(q') \delta'(q' - q'') e_m(q'') . \end{aligned} \quad (14.108)$$

Compte tenu de (14.101) et (14.102), l'intégrale sur q'' donne $(-1)(-1)e'_m(q')$, d'où :

$$p_{nm} = -i\hbar \int dq e_n^*(q) e'_m(q) \equiv \int dq e_n^*(q) \left(-i\hbar \frac{\partial e_m}{\partial q} \right) , \quad (14.109)$$

comme il se doit. On peut tout autant effectuer d'abord l'intégration sur q' , ce qui permet de retrouver l'hermiticité de p ($p_{nm} = p_{mn}^*$).

L'élément de matrice de toute fonction $B(p)$ développable en polynôme ou en série de p , se trouve en appliquant formellement la règle usuelle de multiplication des matrices. Prenons le cas simple de p^2 ; on a :

$$\langle q' | p^2 | q'' \rangle = \int dq_1 \langle q' | p | q_1 \rangle \langle q_1 | p | q'' \rangle = (-i\hbar)^2 \int dq_1 \delta'(q' - q_1) \delta'(q_1 - q'') \quad (14.110)$$

Il en résulte que l'élément de matrice $p_{nm}^2 = \langle e_n | p^2 | e_m \rangle$:

$$p_{nm}^2 = \int dq' \int dq'' e_n^*(q') \langle q' | p^2 | q'' \rangle e_m(q'') \quad (14.111)$$

s'écrit :

$$p_{nm}^2 = \int dq' \int dq'' e_n^*(q') (-i\hbar)^2 \int dq_1 \delta'(q' - q_1) \delta'(q_1 - q'') e_m(q'') . \quad (14.112)$$

En effectuant l'intégrale sur q'' , puis celle sur q_1 , les règles précédentes donnent :

$$p_{nm}^2 = (-i\hbar)^2 \int dq e_n^*(q) e_m''(q) \equiv \int dq e_n^*(q) \left(-i\hbar \frac{\partial}{\partial q} \right)^2 e_m(q) . \quad (14.113)$$

Au contraire, en faisant l'intégration sur q'' puis celle sur q' , on obtient :

$$p_{nm}^2 = -(-i\hbar)^2 \int dq_1 e_n'^*(q_1) e'_m(q_1) \equiv \int dq \left(+i\hbar \frac{\partial}{\partial q} e_n^*(q) \right) \left(-i\hbar \frac{\partial}{\partial q} e_m(q) \right) . \quad (14.114)$$

La comparaison de (14.113) et de (14.114) montre que :

$$\int dq e_n^*(q) p^2 e_m(q) = \int dq [p e_n(q)]^* [p e_m(q)] , \quad (14.115)$$

où les crochets délimitent la zone d'action des opérateurs p . (14.115) montre que tous les éléments diagonaux ($n = m$) de p^2 sont positifs, ce qui est la moindre des choses. D'une façon générale, l'élément de matrice représentant une observable quelconque $B(p)$ s'obtient par :

$$B(p)_{nm} = \langle e_n | B(p) | e_m \rangle = \int dq e_n^*(q) B \left(-i\hbar \frac{\partial}{\partial q} \right) e_m(q) . \quad (14.116)$$

Pour terminer, examinons le commutateur fondamental $[q, p]$. Pour éviter les confusions entre opérateurs et scalaires, convenons provisoirement de noter les opérateurs en gras ; il vient d'une part (\mathbf{q} est hermitique) :

$$\langle q' | \mathbf{q} \mathbf{p} | q'' \rangle = \langle \mathbf{q} q' | \mathbf{p} | q'' \rangle = q' \langle q' | \mathbf{p} | q'' \rangle = -i\hbar q' \delta'(q' - q'') ; \quad (14.117)$$

d'autre part :

$$\langle q' | \mathbf{p} \mathbf{q} | q'' \rangle = \langle q' | \mathbf{p} | \mathbf{q} q'' \rangle = q'' \langle q' | \mathbf{p} | q'' \rangle = -i\hbar q'' \delta'(q' - q'') ; \quad (14.118)$$

d'où :

$$\langle q' | [\mathbf{q}, \mathbf{p}] | q'' \rangle = -i\hbar (q' - q'') \delta'(q' - q'') . \quad (14.119)$$

Le sens de l'objet au second membre s'élucide comme d'habitude ; pour toute fonction $f(x)$ régulière, on a :

$$\int dx f(x) \delta'(x - x_0) = - \int dx f'(x) \delta(x - x_0) = -f'(x_0) . \quad (14.120)$$

Avec $f(x) = xg(x)$, on en déduit :

$$\int dx xg(x)\delta'(x-x_0) = -[g(x_0) + x_0g'(x_0)] . \quad (14.121)$$

En particulier, si $x_0 = 0$:

$$\int dx xg(x)\delta'(x) = -g(0) \equiv -\int dx g(x)\delta(x) \iff x\delta'(x) = -\delta(x) . \quad (14.122)$$

Ceci donne $-i\hbar(q' - q'')\delta'(q' - q'') = i\hbar\delta(q' - q'')$ d'où, selon (14.119) :

$$\langle q' | [\mathbf{q}, \mathbf{p}] | q'' \rangle = i\hbar\delta(q' - q'') . \quad (14.123)$$

Ceci exprime bien que le commutateur $[\mathbf{q}, \mathbf{p}]$ est l'opérateur identité multiplié par le facteur $i\hbar$.

Avant de poursuivre, la remarque suivante s'impose. Les états propres de la coordonnée apparaissent comme des objets un peu exotiques (ce sont des fonctions de Dirac), mais ceci n'est pas gênant physiquement. En effet, un état propre strict de l'observable q ne saurait avoir un sens physique direct : compte tenu des relations de Heisenberg, une particule qui serait dans cet "état" aurait une impulsion complètement indéterminée entre $\pm\infty$, ce qui n'a pas de sens physiquement (de plus, une impulsion infinie signifie une énergie cinétique infinie¹⁴). En conséquence, tout état physique doit être construit en formant une *combinaison linéaire* des états propres de la coordonnée, en écrivant par exemple :

$$|\psi\rangle = \int dq |q\rangle c(q) . \quad (14.124)$$

La fonction $c(q)$ (qui n'est autre que $\langle q|\psi\rangle = \psi(q)$, c'est la fonction d'onde ordinaire représentant le vecteur $|\psi\rangle$!) peut éventuellement être très bien localisée en q , mais sa largeur est nécessairement finie (elle ne peut être strictement nulle). En définitive, on peut considérer les états propres $|q\rangle$ comme une base commode servant à l'idéalisation algébrique, mais n'ayant, *stricto sensu*, pas de sens physique immédiat. Superposer les états $|q\rangle$ pour construire $|\psi\rangle$ selon (14.124), c'est former un paquet d'ondes.

La coordonnée a joué jusqu'ici un rôle qui peut sembler privilégié ; en réalité, comme déjà vu, il est possible de construire une autre représentation, la représentation- p où l'impulsion prend la place de la coordonnée ; ceci sera fait en détail dans la section suivante. Toujours dans le cadre de la représentation- q , il est loisible de s'intéresser aux états propres de l'impulsion ; comme précédemment, on se borne à examiner le cas de \mathbb{R} , la généralisation à \mathbb{R}^3 étant immédiate. En désignant par ψ_{p_0} un état propre de p , l'équation aux vecteurs propres pour p s'écrit :

$$p\psi_{p_0}(q) = p_0\psi_{p_0}(q) \iff -i\hbar \frac{\partial}{\partial q} \psi_{p_0}(q) = p_0 \psi_{p_0}(q) . \quad (14.125)$$

La solution d'une telle équation est :

$$\psi_{p_0}(q) = \mathcal{N} e^{(i/\hbar)p_0q} . \quad (14.126)$$

où \mathcal{N} est un facteur multiplicatif arbitraire que l'on choisira par commodité. ψ_{p_0} est donc essentiellement une onde plane de vecteur d'onde $k_0 = p_0/\hbar$. Une telle fonction n'est pas normalisable et ne saurait donc représenter un état physique ; ceci est en accord avec les relations d'incertitude de Heisenberg : l'impulsion est parfaitement définie dans un tel état (elle vaut p_0 , sans dispersion), ce qui signifie que la coordonnée est totalement indéterminée ; c'est ceci qui n'a pas de sens physique : ψ_{p_0} décrit une particule qui peut être partout entre $\pm\infty$, or une particule est nécessairement confinée, éventuellement dans un très grand domaine. Pour les mêmes raisons que les $\{|q\rangle\}$ ne forment pas une "base physique", il en va de même pour les états propres de l'impulsion. Le confinement inévitable de la particule a pour conséquence que l'impulsion ne peut être parfaitement définie et, ici encore, les seuls états physiques seront des combinaisons linéaires de la forme :

$$|\psi\rangle = \int dp C(p)|p\rangle , \quad (14.127)$$

où $C(p)$ est une fonction, dont le module est (éventuellement) sensiblement différent de zéro sur un petit intervalle en p entourant p_0 – la largeur de cet intervalle ne peut être nulle si $|\psi\rangle$ représente un état physique.

L'état associé à la fonction d'onde $\psi_{p_0}(q)$ peut être noté $|\psi_{p_0}\rangle$ ou plus simplement $|p_0\rangle$. Il en résulte :

$$\langle q|\psi_{p_0}\rangle \equiv \langle q|p_0\rangle = \mathcal{N} e^{(i/\hbar)p_0q} . \quad (14.128)$$

¹⁴Ici apparaît une idée que l'on retrouvera par la suite : une particule bien localisée est dans un état impliquant des énergies élevées. C'est pourquoi on parle d'*énergie de localisation*.

14.7 Représentation- \vec{p}

Tout ce qui vient d'être fait avec les vecteurs propres de la coordonnée peut l'être avec ceux de l'impulsion, ce qui donne une vision complémentaire du formalisme ; la représentation- \vec{p} est finalement moins souvent utilisée, mais elle est dans certaines situations nettement plus commode et plus transparente physiquement. Par ailleurs, cette dualité illustre bien l'incompatibilité des deux grandeurs coordonnée et impulsion. Il existe également une représentation hybride, due à Wigner, où p et q sont traités sur un pied d'égalité au sein d'une même représentation¹⁵. Les états de base de celle-ci ne sont évidemment propres ni de p , ni de q et l'objet central est une matrice densité et non une fonction d'onde.

Définir la représentation- \vec{p} , c'est finalement montrer comment les observables s'expriment en termes d'opérateurs sur la base propre de l'impulsion. Bien évidemment, les relations fondamentales sont inchangées : elles sont vraies au niveau des opérateurs et ne dépendent donc nullement de l'écriture choisie ; en particulier, la relation fondamentale $[q, p] = i\hbar\mathbf{1}$ est toujours vraie. Par ailleurs, compte tenu de la symétrie entre q et p , et en jouant avec la connaissance acquise suivant laquelle p est représenté par $-i\hbar\partial/\partial q$ en représentation-coordonnée, on se doute que q , en représentation-impulsion, sera donné par quelque chose contenant $\partial/\partial p$, alors que p sera maintenant associé à la simple multiplication par p . Si on reporte la substitution :

$$q \rightarrow C \frac{\partial}{\partial p}, \quad p \rightarrow p \times, \quad (14.129)$$

où C est une constante, dans la relation de commutation fondamentale, on trouve immédiatement que $C = +i\hbar$, ce qui achève de préciser la correspondance :

$$q \rightarrow +i\hbar \frac{\partial}{\partial p}, \quad p \rightarrow p \times, \quad (14.130)$$

Ainsi, dans la représentation- \vec{p} , le Hamiltonien $H = p^2/(2m) + V(q)$ est représenté par l'opérateur :

$$H = \frac{p^2}{2m} + V\left(i\hbar \frac{\partial}{\partial p}\right); \quad (14.131)$$

maintenant, le terme cinétique est un simple opérateur multiplicatif et V donne un opérateur différentiel ; quant à la fonction d'onde, c'est une fonction de p . Par exemple, pour l'oscillateur harmonique, l'équation aux valeurs propres de H s'écrit¹⁶ :

$$\frac{p^2}{2m} \psi(p) + \frac{1}{2}m\omega^2 \left(i\hbar \frac{\partial}{\partial p}\right)^2 \psi(p) = E \psi(p), \quad (14.132)$$

soit :

$$\frac{p^2}{2m} \psi(p) - \frac{\hbar^2}{2}m\omega^2 \psi''(p) = E \psi(p). \quad (14.133)$$

Toutes les relations établies à propos de la représentation-coordonnée se transposent sans aucune difficulté : il y a une symétrie absolue entre ces deux façons d'exprimer la théorie. La relation de fermeture est :

$$\int dp |p\rangle\langle p| = \mathbf{1} \quad (14.134)$$

et ainsi de suite. On sait que les représentations- \vec{r} et $-\vec{p}$ sont reliées par une transformation de Fourier, ce qu'il est facile de retrouver ici. On a :

$$|\psi\rangle = \int dq' |q'\rangle\langle q'|\psi\rangle \equiv \int dq' |q'\rangle\psi_Q(q'). \quad (14.135)$$

¹⁵utilisée parfois pour l'opérateur densité ρ .

¹⁶À des constantes près, l'équation aux valeurs propres de l'oscillateur harmonique est la même dans les deux représentations. Ceci tient au fait que H est quadratique en q et p ; ces deux variables jouent dans ce cas des rôles symétriques.

Il faut ici une notation précise, pour distinguer les différentes fonctions d'onde : ψ_Q désigne la fonction d'onde "ordinaire", au sens de Schrödinger. Injectons maintenant la relation de fermeture (14.134) :

$$|\psi\rangle = \int dq' |q'\rangle \langle q'| \int dp' |p'\rangle \langle p'|\psi\rangle = \int dq' \int dp' |q'\rangle \langle q'|p'\rangle \langle p'|\psi\rangle . \quad (14.136)$$

D'après (14.128), ceci vaut :

$$|\psi\rangle = \int dq' \int dp' |q'\rangle \mathcal{N} e^{(i/\hbar)p'q'} \langle p'|\psi\rangle . \quad (14.137)$$

De la même façon, posons¹⁷ :

$$\langle p|\psi\rangle = \psi_P(p) , \quad (14.138)$$

alors (14.137) devient :

$$|\psi\rangle = \mathcal{N} \int dq' \int dp' |q'\rangle e^{(i/\hbar)p'q'} \psi_P(p') . \quad (14.139)$$

Prenons maintenant le produit scalaire des deux membres avec $|q\rangle$:

$$\langle q|\psi\rangle = \mathcal{N} \int dq' \int dp' \langle q|q'\rangle e^{(i/\hbar)p'q'} \psi_P(p') = \mathcal{N} \int dq' \int dp' \delta(q - q') e^{(i/\hbar)p'q'} \psi_P(p') , \quad (14.140)$$

soit :

$$\psi_Q(q) = \mathcal{N} \int dp e^{(i/\hbar)pq} \psi_P(p) , \quad (14.141)$$

qui exhibe la relation de Fourier entre les deux fonctions ψ_P et ψ_Q . ψ_Q est homogène à $L^{-1/2}$, ψ_P est homogène à $(MLT^{-1})^{-1/2}$. On peut donc choisir \mathcal{N} proportionnel à l'inverse de la racine carrée de la constante de Planck, soit $\mathcal{N} \propto h^{-1/2}$; pour avoir des notations aussi symétriques que possible lors de l'écriture des transformations de Fourier, on pose très précisément (à une dimension d'espace) :

$$\mathcal{N} = \frac{1}{\sqrt{2\pi\hbar}} \equiv (2\pi\hbar)^{-1/2} = \frac{1}{\sqrt{h}} . \quad (14.142)$$

Avec ce choix, (14.128) s'écrit :

$$\langle q|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{(i/\hbar)pq} . \quad (14.143)$$

Ceci permet de retrouver la représentation intégrale connue de la fonction de Dirac ; le produit scalaire $\langle q|q'\rangle$ est égal à $\delta(q - q')$. En injectant (14.134) au milieu de $\langle q|q'\rangle$, compte tenu de (14.143), il vient :

$$\delta(q - q') = \int dp \langle q|p\rangle \langle p|q'\rangle = (2\pi\hbar)^{-1} \int dp e^{(i/\hbar)p(q-q')} = \frac{1}{2\pi} \int dk e^{ik(q-q')} \quad (14.144)$$

soit :

$$\int_{-\infty}^{+\infty} dk e^{ikx} = 2\pi \delta(x) \quad (14.145)$$

La relation entre les deux représentations est précisément, dans \mathbb{R}^d :

$$\psi_Q(\vec{r}) = (2\pi\hbar)^{-d/2} \int e^{(i/\hbar)\vec{p}\cdot\vec{q}} \psi_P(\vec{p}) d^d p \iff \psi_P(\vec{p}) = (2\pi\hbar)^{-d/2} \int e^{(-i/\hbar)\vec{p}\cdot\vec{q}} \psi_Q(\vec{r}) d^d r , \quad (14.146)$$

sachant que :

$$\int_{\mathbb{R}^d} d^d k e^{i\vec{k}\cdot\vec{r}} = (2\pi)^d \delta(\vec{r}) . \quad (14.147)$$

Bien sûr, la relation des représentations par une transformation de Fourier contient en elle-même les relations d'incertitude de Heisenberg pour le couple (coordonnée, impulsion).

¹⁷Antérieurement, on notait $\psi(q)$ la fonction en représentation- q et $\phi(p)$ son équivalent en représentation- p .

Chapitre 15

Evolution temporelle d'un système quantique

15.1 Description de l'évolution dans le temps

L'évolution d'un système quantique, en dehors de toute opération de mesure, est gouvernée par l'équation de Schrödinger :

$$i\hbar \frac{d}{dt} |\Psi(t)\rangle = H |\Psi(t)\rangle , \quad (15.1)$$

où H est le Hamiltonien du système¹, dépendant du temps ou non. H est un opérateur hermitique ; il en résulte immédiatement que la norme de $|\Psi(t)\rangle$ est conservée au cours du temps ; en effet, le vecteur d'état à l'instant $t + dt$ est :

$$|\Psi(t + dt)\rangle = |\Psi(t)\rangle + \frac{dt}{i\hbar} H |\Psi(t)\rangle + \mathcal{O}(dt^2) ; \quad (15.2)$$

le carré de sa norme au même instant est donnée par le produit scalaire :

$$\langle (1 + \frac{dt}{i\hbar} H) \Psi(t) | (1 + \frac{dt}{i\hbar} H) \Psi(t) \rangle ; \quad (15.3)$$

en développant ce produit scalaire suivant les règles établies et en ne retenant que les termes en dt au plus, le second membre est :

$$\langle \Psi(t) | \Psi(t) \rangle + \frac{dt}{i\hbar} \langle \Psi(t) | (H - H^\dagger) | \Psi(t) \rangle + \mathcal{O}(dt^2) ; \quad (15.4)$$

Comme H est hermitique, le terme linéaire en dt est nul de sorte que :

$$\frac{d}{dt} \langle \Psi(t) | \Psi(t) \rangle = 0 \quad \forall t ; \quad (15.5)$$

Vis-à-vis de l'interprétation probabiliste, la constance de la norme est une nécessité : la probabilité de trouver une particule stable dans tout l'espace doit rester constante et égale à 1. A l'inverse, ceci permet d'imaginer comment on doit pouvoir décrire les systèmes instables : d'une façon ou d'une autre, doit apparaître un Hamiltonien effectif

¹Pour un système isolé, le Hamiltonien est indépendant du temps. L'affirmation contenue dans le 6^{ème} postulat consiste à dire deux choses :

1. pour toute partie d'un système isolé, il est encore possible de définir un hamiltonien H (qui alors en général dépend du temps)
2. H est l'opérateur associé à l'énergie du sous-système considéré.

Il n'est pas évident que l'affirmation 2. soit toujours vraie.

non hermitique², dont les valeurs propres seront donc complexes en général. La partie imaginaire sera reliée à la durée de vie de la particule, ou de l'état excité du système considéré.

L'accord est donc parfait entre l'ordre (en temps) de l'équation de Schrödinger et la nécessité pour le Hamiltonien d'être hermitique ; par ailleurs, la présence du nombre imaginaire fondamental i au premier membre est essentielle : à défaut, l'équation du premier ordre serait une simple équation de diffusion dont on sait qu'elle ne donne – en l'absence de sources imposant des valeurs finies des courants – que des états transitoires éphémères.

Ceci étant, il est possible de formaliser de plusieurs façons la “cinématique” d'un système quantique. Il existe essentiellement deux descriptions, celle de Schrödinger et celle de Heisenberg ; une troisième, dite “représentation-interaction”, est aussi très utile³ mais ne sera pas abordée pour l'instant.

Remarque

Une question peut venir à l'esprit : pour quelle raison l'équation de Schrödinger présente-t-elle une dissymétrie entre temps et espace (premier ordre en temps et second ordre en espace) ? Fondamentalement, ceci résulte du couple de relations fondamentales :

$$E = \hbar\omega \quad , \quad p = \hbar k \quad (15.6)$$

jointes à :

$$E = \frac{p^2}{2m} \quad . \quad (15.7)$$

En effet, écrite en termes de grandeurs ondulatoires selon (15.6), cette dernière relation devient :

$$\omega = \frac{\hbar k^2}{2m} \quad (15.8)$$

et présente elle-même cette dissymétrie : elle contient l'inverse d'un temps à gauche et l'inverse du carré d'une longueur à droite. Si on considère que l'onde fondamentale $e^{i(kx-\omega t)}$ doit être solution de l'équation d'onde cherchée, on voit bien qu'il faut la dériver *une fois* en temps et *deux fois* en espace. À partir de l'équation de propagation posée *a priori* :

$$\frac{\partial \Psi}{\partial t} = \gamma \frac{\partial^2 \Psi}{\partial x^2} \quad , \quad (15.9)$$

on trouve le paramètre γ :

$$-i\omega = -k^2\gamma \quad \Longleftrightarrow \quad \gamma = \frac{i\omega}{k^2} \quad ; \quad (15.10)$$

revenant alors aux grandeurs “mécaniques” énergie et impulsion :

$$\gamma = \frac{i\omega}{k^2} = i \frac{E/\hbar}{p^2/\hbar^2} = i \frac{\hbar E}{p^2} = i \frac{\hbar}{2m} \quad . \quad (15.11)$$

γ ne dépend que de la masse de la particule et de la constante fondamentale de Planck, ce qui est satisfaisant. De ce point de vue, une équation de propagation du genre :

$$\frac{\partial^2 \Psi}{\partial t^2} = \gamma \frac{\partial^2 \Psi}{\partial x^2} \quad , \quad (15.12)$$

ne convient pas ; elle donne en effet :

$$\omega^2 = -\gamma k^2 \quad \Longleftrightarrow \quad E^2 = \gamma p^2 \quad \Longleftrightarrow \quad \gamma = \frac{p^2}{4m^2} \quad . \quad (15.13)$$

Avec cette équation, la constante γ est une fonction explicite de l'état dynamique (par p), ce qui n'est pas acceptable. Notons enfin que cette dissymétrie est caractéristique d'une théorie non-relativiste. Au contraire, l'équation de Dirac se construit en posant la combinaison (linéaire) la plus simple de dérivées d'espace et de temps du même ordre.

²Ceci ne peut survenir que dans la mesure où, considérant un système quantifié dans tous ses aspects, on en vient à sommer sur les degrés de liberté d'une partie de ce système ; les degrés de liberté restants sont ceux du complémentaire de ce système, “rhabillé” par les degrés de liberté devenus implicites par l'opération de trace. Un cas standard est celui d'un atome en interaction avec le champ électromagnétique : la sommation sur les degrés du champ peut faire apparaître un Hamiltonien effectif dont les valeurs propres – plus précisément, les résonances de la fonction de Green associée – sont toutes complexes à l'exception de la plus basse ; ceci traduit le fait que l'état fondamental est stable alors que les états excités ont une durée de vie finie et sont le point de départ d'une transition vers les états inférieurs accompagnée de l'émission de photon(s).

³En particulier, la théorie du problème à N corps en fait un usage intensif.

15.1.1 La description de Schrödinger

Cette description⁴ est centrée sur l'évolution du vecteur d'état lui-même (ou de l'une de ses quelconques représentations : fonction d'onde ordinaire, fonction d'onde en représentation- p , développement sur une base discrète, *etc.*). L'équation de Schrödinger (15.1) 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 linéaire⁵ agissant dans l'espace des états ; on pose ainsi :

$$|\Psi(t)\rangle = U(t, t_0)|\Psi(t_0)\rangle . \quad (15.14)$$

En reportant dans (15.1), il vient⁶ :

$$i\hbar \frac{\partial}{\partial t} U(t, t_0)|\Psi(t_0)\rangle = H(t) U(t, t_0)|\Psi(t_0)\rangle . \quad (15.16)$$

Maintenant, l'état initial peut être choisi absolument quelconque ; en particulier, on peut le prendre successive-

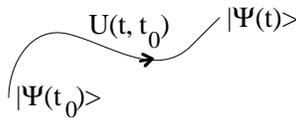


Figure 15.1: $U(t, t_0)$ construit $|\Psi(t)\rangle$ à partir de $|\Psi(t_0)\rangle$.

ment 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 :

$$i\hbar \frac{\partial}{\partial t} U(t, t_0) = H(t) U(t, t_0) . \quad (15.17)$$

C'est évidemment une équation du premier ordre ; par la définition de U , on doit avoir :

$$U(t_0, t_0) = \mathbf{1} \quad \forall t_0 . \quad (15.18)$$

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

$$[U(t, t_0)]^{-1} = [U(t, t_0)]^\dagger . \quad (15.19)$$

L'évolution de t_0 à t peut toujours être décomposée en deux étapes : évolution de t_0 à t_1 puis évolution⁷ de t_1 à t . Par la définition de $U(t, t_0)$, il en résulte :

$$U(t, t_0) = U(t, t_1)U(t_1, t_0) . \quad (15.20)$$

En prenant $t = t_0$ on trouve ainsi :

$$U(t_0, t_1)U(t_1, t_0) = U(t_0, t_0) = \mathbf{1} , \quad (15.21)$$

et en multipliant à gauche par l'inverse, il vient :

$$U(t_1, t_0) = [U(t_0, t_1)]^{-1} . \quad (15.22)$$

⁴aussi appelée "point de vue de Schrödinger" (en anglais : *Schrödinger picture*).

⁵L'opérateur est forcément *linéaire* – il ne saurait être antilinéaire –, puisqu'il existe des transformations infinitésimales infiniment proches de l'identité, celles représentant le passage de t à $t + dt$.

⁶La dérivée de l'opérateur U se définit comme suit :

$$\frac{\partial}{\partial t} U(t, t_0) = \lim_{\delta t \rightarrow 0} \frac{U(t + \delta t, t_0) - U(t, t_0)}{\delta t} . \quad (15.15)$$

⁷Il n'est d'ailleurs pas nécessaire que t_1 se situe entre t_0 et t .

L'inverse de U s'obtient en échangeant 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. En rassemblant (15.19) et (15.22) :

$$U^\dagger(t, t_0) = U^{-1}(t, t_0) = U(t_0, t) . \quad (15.23)$$

Remarque

Une mise en garde est ici nécessaire. L'équation (15.17) porte sur des opérateurs, dont l'algèbre est *a priori* non-commutative. Il en résulte que l'intégration de cette dernière équation ne conduit pas à⁸ :

$$U(t, t_0) = \exp \left(\frac{1}{i\hbar} \int_{t_0}^t H(t') dt' \right) \quad (\text{faux !}) , \quad (15.24)$$

comme ce serait le cas si tous les termes dépendant du temps étaient de simples scalaires. L'expression (15.24) n'est la solution de (15.17) que dans le cas où $H(t)$ commute avec $H(t')$, quels que soient les deux instants t et t' ; ceci n'est réalisé que dans des cas très particuliers – hormis des problèmes triviaux sans intérêt. L'obtention explicite de l'opérateur U dans le cas où le Hamiltonien dépend du temps est en général impossible 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. D'une façon générale, lorsque le problème ne peut être résolu exactement, on recourt aux méthodes de perturbation⁹.

Il est facile de voir où intervient la non-commutativité. Intégrons une fois membre à membre (15.17) ; en utilisant la condition initiale (15.18), il vient :

$$U(t, t_0) = \mathbf{1} + \frac{1}{i\hbar} \int_{t_0}^t H(t') U(t', t_0) dt' . \quad (15.25)$$

Le processus peut s'itérer :

$$U(t, t_0) = \mathbf{1} + \frac{1}{i\hbar} \int_{t_0}^t dt' H(t') \left[\mathbf{1} + \frac{1}{i\hbar} \int_{t_0}^{t'} H(t'') U(t'', t_0) dt'' \right] , \quad (15.26)$$

et ainsi de suite. On met ainsi en évidence une série :

$$U(t, t_0) = \mathbf{1} + \frac{1}{i\hbar} \int_{t_0}^t dt' H(t') + \frac{1}{(i\hbar)^2} \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' H(t') H(t'') + \dots . \quad (15.27)$$

Par inversion de l'ordre des intégrations et échange des variables muettes d'intégration, il est facile de voir que le 3^{ème} terme est aussi :

$$\frac{1}{(i\hbar)^2} \int_{t_0}^t dt' \int_{t'}^t dt'' H(t'') H(t') , \quad (15.28)$$

et on peut l'écrire comme la demi-somme de ses deux expressions :

$$\frac{1}{2} \frac{1}{(i\hbar)^2} \left[\int_{t_0}^t dt' \int_{t'}^t dt'' H(t') H(t'') + \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' H(t'') H(t') \right] . \quad (15.29)$$

Comme $H(t')H(t'') \neq H(t'')H(t')$ en général, on ne peut rassembler les deux intégrales pour former :

$$\frac{1}{2!} \frac{1}{(i\hbar)^2} \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' H(t') H(t'') \quad (15.30)$$

qui est le 3^{ème} terme du développement de $\exp \left[\frac{1}{i\hbar} \int_{t_0}^t H(t') dt' \right]$. Le même argument vaut pour tous les termes du développement en série.

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 ; l'algèbre devient commutative

⁸On omet de toute façon le facteur $U(t_0, t_0) = \mathbf{1}$.

⁹exposées dans le cours de Maîtrise.

et l'équation (15.17) s'intègre formellement comme avec des scalaires. Dès lors, la solution qui vaut $\mathbf{1}$ en $t = t_0$ s'écrit :

$$U(t, t_0) = \exp \left[\frac{1}{i\hbar} (t - t_0)H \right] \quad (15.31)$$

ce que l'on écrit d'habitude :

$$U(t, t_0) = e^{\frac{1}{i\hbar} H (t-t_0)} \quad (15.32)$$

quand on est sûr de voir dans les parenthèses un facteur multiplicatif et non pas une dépendance en temps de H – qui n'existe pas. L'opérateur exponentiel au second membre n'est qu'un raccourci d'écriture pour la série :

$$U(t, t_0) = \sum_{n=0}^{+\infty} \frac{(i\hbar)^{-n}}{n!} H^n (t - t_0)^n . \quad (15.33)$$

Il est manifeste que, en tant que fonction de H , $U(t, t_0)$ commute avec H :

$$[U(t, t_0), H] = 0 . \quad (15.34)$$

Cela vaut la peine de noter que quand H est indépendant du temps, U (donné par (15.32)) ne dépend que de la *différence* des temps t et t_0 ; cette propriété traduit l'invariance par rapport à la translation dans le temps pour un système dont le Hamiltonien est statique : alors seul compte le temps écoulé. En pareil cas, $U(t, t_0)$ est une fonction d'une seule variable que l'on peut simplement noter $U(t - t_0)$ et qui satisfait :

$$U(t, t_0) \equiv U(t - t_0) = U(t)U(-t_0) = U(t)U^\dagger(t_0) , \quad (15.35)$$

où :

$$U(t) = e^{\frac{1}{i\hbar} H t} . \quad (15.36)$$

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 , $\forall t < 0$, à la valeur H_1 , également constante, $\forall t > 0$:

$$H(t) = \begin{cases} H_0 & \text{si } t < 0 \\ H_1 & \text{si } t > 0 \end{cases} \quad (15.37)$$

En décomposant l'évolution de $t_0 < 0$ à $t_1 > 0$ en deux étapes, on peut écrire :

$$U(t_1, t_0) = U(t_1 - 0_+) U(0_+, 0_-) U(0_- - t_0) \quad (15.38)$$

$U(0_+, 0_-)$ est en fait égal à l'identité (voir ci-dessous, (15.45)). Par ailleurs, dans les deux autres intervalles, le Hamiltonien prend une valeur constante, indépendante du temps ; on peut donc utiliser le résultat (15.32) et écrire :

$$U(t_1, t_0) = e^{\frac{1}{i\hbar} H_1 (t_1 - 0_+)} e^{\frac{1}{i\hbar} H_0 (0_- - t_0)} = e^{\frac{1}{i\hbar} H_1 t_1} e^{\frac{-1}{i\hbar} H_0 t_0} . \quad (15.39)$$

Comme H_1 et H_0 ne commutent pas en général entre eux, on n'a pas le droit de regrouper les arguments dans une même exponentielle¹⁰ – qui en effet serait l'intégrale de $H(t)$ dans le cas considéré :

$$U(t_1, t_0) = e^{\frac{1}{i\hbar} H_1 t_1} e^{\frac{-1}{i\hbar} H_0 t_0} \neq e^{\frac{1}{i\hbar} (H_1 t_1 - H_0 t_0)} = e^{\frac{1}{i\hbar} \int_{t_0}^{t_1} H(t') dt'} . \quad (15.40)$$

Cet exemple donne en outre l'occasion de discuter 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. Dans les mêmes notations que précédemment, soit donc une variation instantanée de H passant de H_0 à H_1 à l'instant $t = 0$. Intégrons formellement l'équation de Schrödinger entre deux instants $\pm\delta t$ de part et d'autre de $t = 0$:

$$|\Psi(+\delta t)\rangle - |\Psi(-\delta t)\rangle = \frac{1}{i\hbar} \int_{-\delta t}^{+\delta t} H(t') |\Psi(t')\rangle dt' . \quad (15.41)$$

¹⁰Pour deux opérateurs A et B qui ne commutent pas, $e^A e^B \neq e^{A+B}$.

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 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 :

$$|\Psi(+0)\rangle - |\Psi(-0)\rangle = 0 , \quad (15.42)$$

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, sur un intervalle de durée T ; par exemple, on peut considérer une montée linéaire, ce qui revient à poser explicitement :

$$H(t) = (H_1 - H_0) \frac{t}{T} + H_0 \quad (0 \leq t \leq T) . \quad (15.43)$$

Un calcul simple montre alors que :

$$|\Psi(T)\rangle - |\Psi(0)\rangle = \frac{1}{i\hbar}(H_1 - H_0) \int_0^T \frac{t'}{T} |\Psi(t')\rangle dt' + \frac{1}{i\hbar} H_0 \int_0^T |\Psi(t')\rangle dt' . \quad (15.44)$$

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. En terme de l'opérateur d'évolution, on voit ainsi que, pour un saut fini du Hamiltonien $H(t)$ autour d'un instant quelconque t_0 (modification soudaine) :

$$U(t_0 + 0, t_0 - 0) = \mathbf{1} . \quad (15.45)$$

On retiendra que juste après un tel saut de H , le vecteur d'état est inchangé – il n'a pas encore eu le temps d'évoluer. À l'inverse, on peut montrer que pour un changement infiniment lent (modification adiabatique de $H(t)$), 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 le fondamental de H_1 , mais ce n'est pas obligatoire (théorème de Gell-Mann et Low).

La connaissance du vecteur d'état à l'instant t permet de trouver la moyenne de n'importe quelle observable A à cet instant¹¹, 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 ($|\Psi(t)\rangle$ étant supposé normalisé à l'unité) :

$$\begin{aligned} \langle A(t) \rangle(t) &= \langle \Psi(t) | A(t) | \Psi(t) \rangle = \langle U(t, t_0) \Psi(t_0) | A(t) | U(t, t_0) \Psi(t_0) \rangle \\ &= \langle \Psi(t_0) | U^\dagger(t, t_0) A(t) U(t, t_0) | \Psi(t_0) \rangle ; \end{aligned} \quad (15.46)$$

visiblement, même si A ne dépend pas du temps, sa moyenne en dépend en général par l'intermédiaire de $|\Psi(t)\rangle$:

$$\langle A \rangle(t) = \langle \Psi(t) | A | \Psi(t) \rangle = \text{fonction du temps} ; \quad (15.47)$$

Il n'existe qu'un cas où *toutes* les observables ont une valeur moyenne indépendante du temps c'est celui où $|\Psi(t)\rangle$ est un état stationnaire. Dans ce cas, $|\Psi(t)\rangle$ est de la forme :

$$|\Psi(t)\rangle = \text{état stationnaire} \iff |\Psi(t)\rangle = e^{E_n t / (i\hbar)} |\psi_n\rangle , \quad (15.48)$$

où $|\psi_n\rangle$ est un état propre de H associé à l'énergie E_n ; alors :

$$\langle A \rangle(t) = \langle e^{E_n t / (i\hbar)} \psi_n | A | e^{E_n t / (i\hbar)} \psi_n \rangle = e^{E_n t / (-i\hbar)} e^{E_n t / (i\hbar)} \langle \psi_n | A | \psi_n \rangle = \langle \psi_n | A | \psi_n \rangle = C^{\text{ste}} ; \quad (15.49)$$

Parmi toutes les observables, H joue un rôle éminent, sa moyenne est l'énergie E et s'obtient par :

$$E(t) = \langle H(t) \rangle(t) = \langle \Psi(t) | H(t) | \Psi(t) \rangle . \quad (15.50)$$

La dérivée de E est :

$$\frac{d}{dt} E(t) = \left\langle \frac{d\Psi}{dt} | H(t) | \Psi(t) \right\rangle + \langle \Psi(t) | \frac{dH}{dt} | \Psi(t) \rangle + \left\langle \Psi(t) | H(t) | \frac{d\Psi}{dt} \right\rangle . \quad (15.51)$$

¹¹ au sens d'une espérance mathématique de mesures indépendantes effectuées au même instant.

En utilisant (15.1), il vient :

$$\frac{d}{dt} E(t) = \left\langle \frac{H(t)}{i\hbar} \Psi(t) | H(t) | \Psi(t) \right\rangle + \langle \Psi(t) | \frac{dH}{dt} | \Psi(t) \rangle + \langle \Psi(t) | H(t) | \frac{H(t)}{i\hbar} \Psi(t) \rangle . \quad (15.52)$$

Par l'antilinearité du produit scalaire, le second membre se transforme comme suit :

$$-\frac{1}{i\hbar} \langle \Psi(t) | [H^\dagger(t) - H(t)] H(t) | \Psi(t) \rangle + \langle \Psi(t) | \frac{dH}{dt} | \Psi(t) \rangle . \quad (15.53)$$

d'où, en vertu de l'hermiticité du Hamiltonien¹² :

$$\frac{d}{dt} E(t) = \langle \Psi(t) | \frac{dH}{dt} | \Psi(t) \rangle \iff \frac{dE}{dt} = \left\langle \frac{dH}{dt} \right\rangle . \quad (15.55)$$

Ceci montre 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 une constante. Dans le cas général, la dérivée de $E(t)$ s'obtient simplement en calculant la valeur moyenne de la dérivée de H : comme c'est le Hamiltonien lui-même qui pilote l'état dans le temps, l'évolution de celui-ci s'efface et seule compte la dépendance de H en temps. À titre d'exemple, pour le Hamiltonien variant par saut en $t = 0$, (15.37), on a¹³ :

$$\frac{d}{dt} H(t) = (H_1 - H_0) \delta(t) \iff \frac{dE}{dt} = \langle H_1 - H_0 \rangle(t) \delta(t) , \quad (15.57)$$

d'où :

$$E(0_+) - E(0_-) = \int_{0_-}^{0_+} dt' \langle H_1 - H_0 \rangle(t') \delta(t') = \langle H_1 - H_0 \rangle(0) . \quad (15.58)$$

Ceci donne la variation d'énergie – un saut entre deux valeurs constantes puisque H est constant de chaque côté de $t = 0$ ($E(0_-) = E(t < 0)$, $E(0_+) = E(t > 0)$). En définitive :

$$E(t > 0) = E(t < 0) + \langle \Psi(0) | (H_1 - H_0) | \Psi(0) \rangle . \quad (15.59)$$

Il n'y a variation d'énergie que si $H_1 - H_0$ a une valeur moyenne non-nulle¹⁴ dans l'état $|\Psi(0)\rangle$: une variation de H ne s'accompagne par *forcément* d'une variation de l'énergie $E = \langle H \rangle$.

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 .

15.1.2 La description de Heisenberg

La description de Heisenberg¹⁵ définit une méthode beaucoup plus directe pour calculer l'évolution temporelle des moyennes quantiques. En fait, $|\Psi(t)\rangle$ contient souvent trop d'information pour les questions que l'on se pose et alors, l'intégration explicite de l'évolution de $|\Psi(t)\rangle$ nécessite un effort inutile et excessif¹⁶. 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. Cette dernière situation n'est pas la plus fréquente en pratique : on est le plus souvent intéressé à obtenir les valeurs moyennes à l'instant t de *certaines* observables et non pas de toutes. En pareil

¹²La relation (15.55) n'est qu'un cas particulier d'une propriété tenant au seul fait que le vecteur d'état est normalisé. Plus généralement, quand H dépend d'un paramètre *quelconque* λ , on a :

$$\frac{d}{d\lambda} E(\lambda) = \left\langle \frac{\partial H}{\partial \lambda} \right\rangle . \quad (15.54)$$

Cette relation porte le nom de théorème de Hellmann - Feynman.

¹³Si $Y(t)$ désigne la fonction de Heaviside (échelon-unité) :

$$\frac{dY}{dt} = \delta(t) . \quad (15.56)$$

¹⁴Evidemment, un opérateur non-nul peut fort bien avoir une moyenne nulle dans certains états.

¹⁵On dit aussi "point de vue de Heisenberg" (en anglais : *Heisenberg picture*).

¹⁶Le 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, les équations du mouvement de Heisenberg conduisent au résultat de façon très rapide (voir plus loin).

cas, il est bien plus direct de se poser la question de trouver les équations du mouvement pour ces observables elles-mêmes. Conceptuellement, ces équations présentent d'ailleurs un intérêt majeur : elles sont les pendantes quantiques des équations de la Mécanique Classique. En définitive, alors que la description de Schrödinger porte sur le 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.

Soit un système isolé (son Hamiltonien ne dépend donc pas du temps¹⁷) et considérons d'abord une observable A indépendante du temps. Sa valeur moyenne dans l'état $|\Psi(t)\rangle$ issu de $|\Psi(t_0)\rangle$ est donnée par (15.47), où $U(t, t_0) = U(t - t_0)$ est donné en (15.35). On peut même se débarrasser de t_0 en remarquant que¹⁸ $|\Psi(t_0) = U(t_0)|\Psi(0)\rangle$; en utilisant les relations de groupe (15.35) on trouve :

$$\langle A \rangle(t) = \langle \Psi(0) | U^\dagger(t) A U(t) | \Psi(0) \rangle . \quad (15.60)$$

Cette expression permet de réaliser 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 à $t = 0$. Ce résultat remarquable justifie de donner un nom à cette combinaison ; on note :

$$A_H(t) = U^\dagger(t) A U(t) . \quad (15.61)$$

$A_H(t)$ est le représentant selon Heisenberg de l'opérateur A , ce dernier étant alors appelé, par opposition, représentant de Schrödinger. Comme $U^\dagger(0)$ et $U(0)$ sont chacun égal à l'identité, on a manifestement :

$$A_H(t=0) = A . \quad (15.62)$$

Les représentants de Schrödinger et de Heisenberg coïncident à l'instant origine. Par ailleurs, pour une fonction quelconque d'un opérateur, pourvu qu'elle soit développable en série entière, on a :

$$[f(A)]_H = f(A_H) ; \quad (15.63)$$

en effet, $U^\dagger A^n U = U^\dagger A A A \dots A U = U^\dagger A U U^\dagger A U \dots U^\dagger A U$ puisque $U U^\dagger = U^\dagger U = \mathbf{1}$. Pour la même raison, on vérifie immédiatement que :

$$[A, B]_H = [A_H, B_H] . \quad (15.64)$$

Il existe donc la même relation de commutation entre opérateurs de Schrödinger et opérateurs de Heisenberg :

$$[A, B] = C \iff [A_H, B_H] = C_H . \quad (15.65)$$

Le passage en description de Heisenberg est une transformation *canonique* en ce sens qu'elle conserve la *forme* des équations fondamentales¹⁹.

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. La dérivée de $A_H(t)$ est :

$$\frac{dA_H}{dt} = \frac{dU^\dagger}{dt} A U + U^\dagger A \frac{dU}{dt} . \quad (15.66)$$

Les équations satisfaites par U et son hermitique conjugué sont :

$$\frac{dU}{dt} = \frac{1}{i\hbar} H U(t) , \quad \frac{dU^\dagger}{dt} = -\frac{1}{i\hbar} U^\dagger(t) H^\dagger = -\frac{1}{i\hbar} U^\dagger(t) H . \quad (15.67)$$

où l'hermiticité de H a été utilisée. En reportant dans l'expression de la dérivée de A_H , (15.66), il vient :

$$\frac{dA_H}{dt} = -\frac{1}{i\hbar} U^\dagger(t) H A U(t) + U^\dagger(t) A \frac{1}{i\hbar} H U(t) \iff i\hbar \frac{dA_H}{dt} = U^\dagger(t) [A, H] U \equiv [A, H]_H . \quad (15.68)$$

En vertu aussi de (15.64) :

$$i\hbar \frac{dA_H}{dt} = [A, H]_H \iff i\hbar \frac{dA_H}{dt} = [A_H, H_H] . \quad (15.69)$$

Sous une forme ou une autre, (15.69) constitue l'équation de Heisenberg pour l'observable A . La recette d'écriture de cette équation du mouvement est donc la suivante :

¹⁷Cette hypothèse n'est pas nécessaire pour obtenir (15.69) mais, en pratique, c'est dans ce cas que la description de Heisenberg est utile. Alors, $U(t)$ et H commutent.

¹⁸De toute façon, dans le cas présent (invariance par translation dans le temps), rien n'empêche de choisir l'origine des temps $t = 0$ à l'instant t_0 .

¹⁹Il en va ainsi pour tous les changements de représentation définis par un opérateur unitaire.

1. on calcule le commutateur de A avec H en représentation de Schrödinger
2. on affuble le résultat de l'indice H et on écrit que ceci est $i\hbar dA_H/dt$.

On retrouve une fois de plus que le commutateur de la Mécanique Quantique a la même forme que le crochet de Poisson classique²⁰. Dans la pratique, une fois les idées bien mises en place, on omet le plus souvent l'indice H pour alléger l'écriture.

Dans le cas où l'observable A dépend elle-même du temps avant le passage en Heisenberg, il faut en tenir compte. La valeur moyenne est alors donnée par (15.46) et la dérivation en temps introduit le terme supplémentaire :

$$\langle \Psi(t) | \frac{dA}{dt} | \Psi(t) \rangle = \langle \Psi(0) | U^\dagger(t) \frac{dA}{dt} U(t) | \Psi(0) \rangle . \quad (15.71)$$

L'opérateur en *sandwich* est $(dA/dt)_H$. Ceci étant compris, un calcul analogue à celui fait plus haut montre que l'équation du mouvement de Heisenberg est dans ce cas :

$$i\hbar \frac{dA_H}{dt} = i\hbar \left(\frac{dA}{dt} \right)_H + [A, H]_H . \quad (15.72)$$

Une fois écrites, les équations de Heisenberg donnent par intégration en temps l'expression explicite des observables intéressantes. En prenant 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 C . Pour une telle observable, par définition :

$$[C, H] = 0 . \quad (15.73)$$

Comme $U(t)$ est une fonction de H , (15.73) entraîne que C commute aussi avec U , d'où :

$$[C, U] = 0 \iff C_H = C . \quad (15.74)$$

Une autre façon de voir que C_H est constant (et coïncide avec son représentant de Schrödinger) est d'écrire l'équation de Heisenberg :

$$i\hbar \frac{dC_H}{dt} = [C, H]_H = 0 \iff \frac{dC_H}{dt} = 0 . \quad (15.75)$$

Il en résulte que la valeur moyenne de C dans n'importe quel état ne varie pas au cours du temps : c'est bien une constante du mouvement. Les constantes du mouvement jouent un rôle important : en tant qu'observables commutant avec H , on peut trouver un jeu de vecteurs propres communs à ces constantes et à H . À toute constante du mouvement est associée une propriété de symétrie – d'où la confirmation du rôle majeur de la symétrie pour l'analyse d'un problème donné.

Quelques exemples permettent de montrer explicitement la puissance et l'efficacité des équations de Heisenberg – et aussi leur transparence physique. Même dans le cas simple d'une particule libre à une dimension, on sait que le calcul explicite de $|\Psi(t)\rangle$ est laborieux ; une fois cette étape franchie, il reste encore à calculer les valeurs moyennes avec $|\Psi(t)\rangle$ ainsi obtenu. De plus, le calcul étant fait avec un état initial donné, le choix d'un autre état initial oblige à reprendre le calcul entièrement – sauf si on a l'astuce de penser au propagateur, quand on le connaît. Quoi qu'il en soit, au contraire, le problème se résout facilement et très rapidement à l'aide des équations de Heisenberg. On obtient, avec $H = p^2/(2m)$:

$$i\hbar \frac{dx_H}{dt} = [x, H]_H = i\hbar \frac{p_H}{m} , \quad i\hbar \frac{dp_H}{dt} = [p, H]_H = 0 . \quad (15.76)$$

²⁰Soit $a(q, p)$ une grandeur dynamique classique. Le long de la trajectoire – pour laquelle q et p varient selon les équations canoniques de Hamilton –, on a :

$$\frac{da}{dt} = \frac{\partial a}{\partial q} \dot{q} + \frac{\partial a}{\partial p} \dot{p} \equiv \{a, H\} . \quad (15.70)$$

Pour une particule libre, l'opérateur p est bien une constante du mouvement au sens ci-dessus ; d'où, par intégration :

$$p_H(t) = p, \quad x_H(t) = \frac{p}{m}t + x. \quad (15.77)$$

Disposant d'un état initial $|\Psi\rangle$, les moyennes s'obtiennent immédiatement ($\langle \bullet \rangle = \langle \Psi | \bullet | \Psi \rangle$) :

$$\langle p \rangle(t) = \langle p \rangle, \quad \langle x \rangle(t) = \frac{\langle p \rangle}{m}t + \langle x \rangle, \quad (15.78)$$

Pour trouver l'écart quadratique, il faut d'abord obtenir la moyenne de x^2 ; de (15.77) on déduit (par une élévation au carré) :

$$x_H^2(t) = \frac{p^2}{m^2}t^2 + \frac{1}{m}(xp + px) + x^2, \quad (15.79)$$

d'où :

$$\langle x^2 \rangle(t) = \frac{\langle p^2 \rangle}{m^2}t^2 + \frac{1}{m}\langle xp + px \rangle + \langle x^2 \rangle. \quad (15.80)$$

En définitive, pour la particule libre, l'écart quadratique de la position varie en temps comme suit :

$$\Delta x^2(t) = \frac{\Delta p^2}{m^2}t^2 + \frac{1}{m}(\langle xp + px \rangle - 2\langle x \rangle \langle p \rangle) + \Delta x^2, \quad (15.81)$$

où $\Delta p^2 = \langle p^2 \rangle - \langle p \rangle^2$. Dans le cas où $\langle x | \Psi \rangle \equiv \Psi(x)$ est une fonction réelle et paire en x , ces expressions se simplifient :

$$\langle x \rangle(t) = 0, \quad \langle p \rangle(t) = 0, \quad \langle x^2 \rangle(t) = \frac{\langle p^2 \rangle}{m^2}t^2 + \langle x^2 \rangle, \quad \Delta x^2(t) = \frac{\Delta p^2}{m^2}t^2 + \Delta x^2, \quad (15.82)$$

d'où :

$$\Delta x^2(t) = \frac{\Delta p^2}{m^2}t^2 + \Delta x^2. \quad (15.83)$$

Pour l'oscillateur harmonique ($H = p^2/(2m) + m\omega^2 x^2/2$), on trouve immédiatement :

$$\langle x \rangle(t) = \langle x \rangle \cos \omega t + \frac{\langle p \rangle}{m\omega} \sin \omega t, \quad \langle p \rangle(t) = \langle p \rangle \cos \omega t + m\omega \langle x \rangle \sin \omega t, \quad (15.84)$$

Enfin, pour une particule chargée dans un champ électromagnétique dérivant de (\vec{A}, ϕ) , le Hamiltonien est :

$$H = \frac{1}{2m}(\vec{p} - q\vec{A})^2 + q\phi(\vec{r}). \quad (15.85)$$

Un calcul ([20], p. 178) conduit à :

$$m \frac{d^2 \vec{r}_H}{dt^2} = q\vec{E} + \frac{q}{2} \left(\frac{d\vec{r}_H}{dt} \times \vec{B} - \vec{B} \times \frac{d\vec{r}_H}{dt} \right). \quad (15.86)$$

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

En définitive, on voit bien que les deux points de vue de Schrödinger et de Heisenberg sont équivalents et se distinguent par le choix de l'insertion de la dépendance en temps. En résumé :

- description de Schrödinger : observables fixes, vecteur d'état variable
- description de Heisenberg : observables variables, vecteur d'état fixe.

15.1.3 Le théorème d'Ehrenfest

Ehrenfest [21] fut le premier à montrer les valeurs moyennes obéissent à des équations du mouvement réminiscentes de celles de Newton ; en outre, moyennant une hypothèse supplémentaire (voir ci-dessous) le centre d'un paquet d'ondes suit de fait la trajectoire classique (c'est ce que l'on appelle le théorème d'Ehrenfest). Au lieu de suivre la démarche historique d'Ehrenfest ([15] p. 194, [20] p. 28), il est plus économique d'utiliser les résultats obtenus ci-dessus, et de montrer d'abord que les équations de Heisenberg sont en fait strictement identiques aux équations de Newton²¹. Avec :

$$H = \frac{\vec{p}^2}{2m} + V(\vec{r}) , \quad (15.87)$$

on trouve que les équations de Heisenberg pour \vec{r} et \vec{p} sont :

$$\frac{d\vec{r}_H}{dt} = \frac{\vec{p}_H}{m} , \quad \frac{d\vec{p}_H}{dt} = -\vec{\nabla}V(\vec{r}_H) . \quad (15.88)$$

La première est l'équivalent de $\dot{\vec{r}} = \vec{p}/m$, la seconde de $\dot{\vec{p}} = \vec{F}$. Maintenant, suivant la prescription de Heisenberg, un état initial étant donné, les valeurs moyennes obéissent aux équations différentielles :

$$\frac{d}{dt}\langle\vec{r}\rangle(t) = \frac{\langle p \rangle(t)}{m} , \quad \frac{d}{dt}\langle\vec{p}\rangle(t) = -\langle\vec{\nabla}V(\vec{r})\rangle(t) . \quad (15.89)$$

Au prix d'une relative lourdeur des écritures, il est utile de mentionner la dépendance en temps : ce sont bien les moyennes elles-mêmes qui dépendent du temps. L'intégration explicite de ces équations différentielles fait apparaître des conditions initiales $\langle\vec{r}\rangle(0)$ et $\langle\vec{p}\rangle(0)$: ce sont les moyennes des opérateurs position et impulsion calculées sur l'état initial. Par ces équations du mouvement, le temps est "extrait" de toute l'algèbre quantique et intervient seulement pour construire des équations différentielles portant sur les fonctions (scalaires) que sont ces moyennes. En combinant les deux dernières équations :

$$m \frac{d^2}{dt^2}\langle\vec{r}\rangle(t) = -\langle\vec{\nabla}V(\vec{r})\rangle(t) . \quad (15.90)$$

Ainsi, la valeur moyenne de la position obéit à une équation ressemblant à celle de Newton²² – toutefois, la moyenne entoure V , ce qui est différent de $V(\langle\vec{r}\rangle)$. Si l'on part d'un paquet d'ondes, le *centre*²³ du paquet d'ondes qui s'en déduit à l'instant t (et lui seul²⁴) suit un mouvement obtenu en prenant les valeurs moyennes des équivalents quantiques des grandeurs dynamiques. Ceci constitue le théorème d'Ehrenfest.

On peut d'ailleurs aller un cran plus loin ; en effet, dans le cas où le champ de forces varie très peu sur une longueur de l'ordre de la taille du paquet d'ondes, alors on peut écrire :

$$\langle\vec{\nabla}V(\vec{r})\rangle \simeq \vec{\nabla}V(\langle\vec{r}\rangle) . \quad (15.91)$$

Si l'on accepte cette approximation, l'équation (15.90) pour la moyenne $\langle\vec{r}\rangle$ devient identique à l'équation de Newton : en pareil cas, le centre du paquet d'ondes suit la trajectoire classique. La largeur du paquet d'ondes est liée à l'incertitude en impulsion par la relation de Heisenberg : $\Delta r \sim \hbar/\Delta p$; si on assimile p et Δp , alors $\Delta r \sim \lambda$ au sens de de Broglie : (15.91) dit alors que la limite classique est bien retrouvée si le potentiel varie très lentement à l'échelle de la longueur d'onde.

Bien évidemment, l'approximation (15.91) n'en est pas une dans le cas où $\vec{\nabla}V$ est au plus une fonction linéaire des coordonnées, c'est-à-dire pour un oscillateur harmonique, pour une particule libre ou pour une particule chargée dans un champ électrique constant. Ainsi par exemple, l'oscillateur reproduit exactement pour les valeurs moyennes les équations du mouvement classique – c'est bien ce qu'on a trouvé en (15.84) ; ce système est unique au sens où il est le seul²⁵ à ne pas être le siège du phénomène autrement universel de

²¹du moins tant que des variables spécifiquement quantiques n'interviennent pas.

²²On note que la constante de Planck n'y apparaît plus explicitement.

²³dans la mesure où le paquet est assez symétrique pour que la moyenne de la position coïncide avec le maximum du paquet.

²⁴Bien évidemment, les moyennes des puissances supérieures de la coordonnée obéissent à des équations autrement plus compliquées.

²⁵Cette propriété remarquable est liée au fait que le spectre du Hamiltonien de l'oscillateur harmonique est une suite de niveaux équidistants. Toutes les valeurs moyennes prendront donc tôt ou tard l'allure d'une *série* de Fourier, dont la somme est une fonction périodique du temps. L'oscillateur est le prototype du système *linéaire*.

l'étalement d'un paquet d'ondes²⁶. En ce qui concerne la particule libre, on retrouve bien le mouvement classique en moyenne mais l'étalement est inévitable, en raison notamment de l'absence d'un potentiel de confinement. Dans le cas de la particule accélérée par un champ constant, le paquet d'ondes s'étalement, mais le rapport $\Delta x / \langle x \rangle$ tend vers zéro aux grands temps²⁷ : le paquet s'étalement beaucoup moins vite que son centre ne se déplace. A haute énergie, la localisation reprend le dessus (haute énergie = petite longueur d'onde).

Dans le cas usuel où $\vec{\nabla}V$ est une fonction non-linéaire de la position, la variation à l'instant t de la valeur moyenne de la position est reliée aux moyennes des puissances quelconques de cette même position ; on sent bien se dessiner une hiérarchie d'équations couplées dont la résolution exacte, même dans les cas simples est en général impossible ; on fait alors souvent appel à des méthodes d'approximation.

15.2 Propagateur

On appelle²⁸ propagateur 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 : le traitement du cas contraire relève d'un tout autre formalisme.

La relation donnant la fonction d'onde développée à partir d'un état initial²⁹ a déjà été écrite :

$$\Psi(x, t) = \int_{-\infty}^{+\infty} dx_0 U(x, t; x_0, t_0) \Psi(x_0, t_0) \quad (15.92)$$

(en dimension d , la dimension de U est L^{-d}). En notation de Dirac, on a :

$$|\Psi(t)\rangle = U(t, t_0)|\Psi(t_0)\rangle \iff \langle x|\Psi(t)\rangle = \int_{-\infty}^{+\infty} dx' \langle x|U(t, t_0)|x'\rangle \langle x'|\Psi(t_0)\rangle, \quad (15.93)$$

ce qui permet d'identifier :

$$U(x, t; x', t_0) = \langle x|U(t, t_0)|x'\rangle. \quad (15.94)$$

Si H est indépendant du temps, seule compte la différence des temps ; en pareil cas, on note simplement :

$$U(x, t; x', t_0) \equiv U(x, t - t_0; x'). \quad (15.95)$$

Toutes les fonctions notées U sont des expressions différentes représentant le même objet, à savoir le propagateur du problème considéré. Le noyau $U(x, t; x_0, t_0)$ est en quelque sorte une matrice continue.

L'opérateur $U(t, t_0)$ satisfait l'équation (15.17) avec $U(t_0, t_0) = \mathbf{1}$. En représentation- q , ceci s'écrit :

$$i\hbar \frac{\partial}{\partial t} U(x, t; x_0, t_0) = H \left(x, p = -i\hbar \frac{\partial}{\partial x} \right) U(x, t; x_0, t_0), \quad (15.96)$$

avec :

$$U(x, t_0; x_0, t_0) = \delta(x - x_0). \quad (15.97)$$

U conserve la norme (il est unitaire), ce qui se traduit par :

$$U(x, t; x', t_0)^* = U^{-1}(x, t; x', t_0) \iff U^\dagger(t) = U^{-1}(t) = U(-t). \quad (15.98)$$

²⁶Pour l'oscillateur harmonique, tout paquet d'ondes "respire", en raison du confinement sévère (mais flou) que constitue le potentiel en x^2 .

²⁷ $\langle x \rangle \propto t^2$, $\Delta x \propto t$.

²⁸La terminologie est fluctuante : le propagateur s'appelle aussi fonction de Green ; par ailleurs, le nom propagateur est aussi utilisé dans une acception un peu différente (pour désigner sa transformée de Fourier par exemple). Dans tous les cas, il s'agit d'une représentation ou d'une autre permettant d'écrire explicitement la relation entre un "ancêtre" et sa descendance (filiation). La notion de propagateur n'est pas spécifiquement quantique.

²⁹en raisonnant toujours à une dimension pour simplifier les écritures.

Compte tenu de ceci, il suffit de connaître U pour $t > 0$ par exemple. En définissant un nouvel opérateur $U_+(t)$ suivant³⁰ :

$$U_+(t) = \begin{cases} 0 & \text{si } t < 0 \\ U(t) & \text{si } t > 0 \end{cases}, \quad U_+(0_+) = \mathbf{1}, \quad (15.99)$$

l'objectif suffisant est de trouver $U_+(t)$.

Considérons l'exemple le plus simple, celui d'une particule libre ; dans ce cas, (15.96) est explicitement :

$$i\hbar \frac{\partial}{\partial t} U_+(x, t; x') = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} U_+(x, t; x'). \quad (15.100)$$

La combinaison d'une transformée de Fourier en espace et d'une transformation de Laplace en temps permet de résoudre cette équation ; on trouve ainsi la solution qui satisfait la condition initiale (15.97) :

$$U_+(x, t; x') = \sqrt{\frac{m}{2i\pi\hbar t}} e^{i\frac{m(x-x')^2}{2t}}. \quad (15.101)$$

Ceci donne $U(x, t; x')$ pour $t > 0$. Selon (15.98), l'expression pour $t < 0$ est :

$$U(x, t < 0; x') = U_+^*(x, -t; x') = \left(\sqrt{\frac{m}{2i\pi\hbar t}} \right)^* e^{\frac{-i}{\hbar} \frac{m(x-x')^2}{2(-t)}}; \quad (15.102)$$

si on définit la branche de la racine carrée complexe comme³¹ :

$$(\sqrt{z})^* = \sqrt{z^*}, \quad (15.103)$$

on voit que l'expression (15.102) pour $t < 0$ est la même que (15.101). En définitive, rétablissant l'instant arbitraire t_0 et adoptant des notations plus symétriques, le propagateur pour la particule libre est :

$$U(x, t; x_0, t_0) = \sqrt{\frac{m}{2i\pi\hbar(t-t_0)}} e^{i\frac{m(x-x_0)^2}{2(t-t_0)}} \quad \forall t. \quad (15.104)$$

Il est facile de vérifier que cette fonction satisfait bien l'équation (15.96) et la condition initiale (15.97).

L'expression (15.104) est tout à fait remarquable, en ce sens qu'elle implique l'action *classique* (!?) d'une particule partie de x_0 en t_0 et se trouvant en x à l'instant t . La présence du facteur i/\hbar dans l'exponentielle permet de deviner que si l'action typique du problème est très grande devant \hbar , alors une méthode de phase stationnaire pour évaluer l'intégrale (15.92) donnant la fonction d'onde à l'instant t est efficace. Ce fait permet aussi d'entrevoir la possibilité de définir des "trajectoires" quantiques – dont la superposition linéaire pondérée construit l'amplitude $\Psi(x, t)$ – généralisant la notion de trajectoire au sens classique. Ces idées, développées par Feynman, permettent de reconstruire³² la Mécanique Quantique de façon très physique et nettement plus intuitive³³.

Remarques

1. Parmi toutes les conditions initiales, l'une d'entre elles joue un rôle particulier – à la fois mathématiquement et physiquement – celle où, initialement, la particule est presque parfaitement localisée en un point que l'on peut toujours prendre comme origine dans l'espace ; cette situation est représentée en choisissant une gaussienne très étroite dont l'intégrale du carré vaut 1 :

$$\Psi(x, t = 0) = \frac{1}{\sqrt{\sqrt{2\pi} \delta x}} e^{-\frac{x^2}{4\delta x^2}}. \quad (15.105)$$

³⁰ U_+ est appelé propagateur (ou fonction de Green) avancé(e).

³¹ce qui revient à prendre comme coupure le demi-axe réel négatif.

³²Il faut cependant utiliser un formalisme spécifique (l'intégrale de chemin – en anglais *path integral*) assez lourd ; la lecture du premier chapitre de [22] reste recommandée.

³³En quelque sorte, toutes les "trajectoires" de Feynman (*chemins*) interviennent dans la construction de l'amplitude de probabilité, et sont rassemblées dans une sorte de tube spatio-temporel aux frontières floues ; il n'y a donc ni position, ni vitesse parfaitement définies à un instant donné : l'extension transversale du tube donne une idée de l'incertitude en position, cependant que l'impossibilité de définir une tangente commune aux points rassemblés dans le tube est l'illustration de l'indétermination sur la vitesse. On peut dire que la trajectoire est *floue*. Le passage à la limite classique peut se figurer comme la réduction graduelle de la dimension transversale du tube qui, à la limite $\hbar \rightarrow 0$, redonne une ligne – la trajectoire classique (dans cette limite, on récupère d'un coup d'un seul la ligne et sa tangente)

Le report dans (15.92) (avec $t_0 = 0$) compte tenu de (15.104) donne :

$$\Psi(x, t) = \int_{-\infty}^{+\infty} dx' \sqrt{\frac{m}{2i\pi\hbar t}} e^{\frac{i}{\hbar} \frac{(x-x')^2}{2t}} \frac{1}{\sqrt{\sqrt{2\pi} \delta x}} e^{-\frac{x'^2}{4\delta x^2}} . \quad (15.106)$$

Il s'agit d'une intégrale gaussienne que l'on peut calculer ; on trouve :

$$\Psi(x, t) = \frac{1}{\sqrt{\sqrt{2\pi} \delta x}} \frac{1}{\sqrt{1 + \frac{i\hbar t}{2m\delta x^2}}} \exp\left(-\frac{1}{1 + \frac{i\hbar t}{2m\delta x^2}} \frac{x^2}{4\delta x^2}\right) . \quad (15.107)$$

C'est visiblement la combinaison $\hbar/(2m\delta x^2)$ qui pilote la vitesse d'étalement du paquet d'ondes : plus celui-ci est initialement bien localisé, plus il s'étale vite ; à la limite d'une localisation initiale parfaite, l'aplatissement est quasi instantané. De fait, la bonne échelle de temps est $\tau = 2m\delta x^2/\hbar$ et, pour $t \gg \tau$:

$$|\Psi(x, t)|^2 \simeq \frac{1}{\sqrt{2\pi}} \frac{2m\delta x}{\hbar t} e^{-2\frac{m\delta x^2}{\hbar t} x^2} \quad (t \gg \tau) . \quad (15.108)$$

Cette expression montre $|\Psi(x, t)|^2$ est une gaussienne ultra-plate ; $|\Psi|$ est alors pratiquement constant dans l'espace. On peut vérifier que, si ΔE est l'écart-type de l'énergie dans l'état (15.105), alors on a :

$$\Delta E \tau \sim \hbar . \quad (15.109)$$

2. Lors de la construction de l'équation de Schrödinger – en particulier au moment où a été discutée l'apparition un peu magique du nombre i – il a été noté que cette équation est structurellement identique à l'équation classique de la diffusion déduite de la loi de Fick. Il est facile de voir explicitement que la solution de :

$$\frac{\partial}{\partial T} P(X, T) = D \frac{\partial^2}{\partial X^2} P(X, T) \quad (15.110)$$

qui vaut $\delta(X - X_0)$ en $T = 0$ et le propagateur (15.104) se déduisent en effet l'un de l'autre par un simple changement de variables effectué dans (15.104) :

$$P(X, T) = \xi U(x = \xi X, t = -iT; x_0 = \xi X_0) = \frac{1}{\sqrt{4\pi DT}} e^{-\frac{(X-X_0)^2}{4DT}} , \quad (15.111)$$

où $\xi = \sqrt{\hbar/(2mD)}$.

3. D'une façon générale, le propagateur $U(x, t; x_0, t_0)$ s'exprime aussi sous la forme d'une série impliquant tous les états propres du Hamiltonien, ce qui suppose donc préalablement résolu le problème aux valeurs et vecteurs propres de H . En effet, dans le cas où H ne dépend pas du temps, la relation (15.94) donne :

$$U(x, t; x_0, t_0) = \langle x | e^{\frac{i}{\hbar} H(t-t_0)} | x_0 \rangle . \quad (15.112)$$

Connaissant toutes les solutions de $H|\psi_n\rangle = E_n|\psi_n\rangle$, on en déduit ($E_n = \hbar\omega_n$) :

$$U(t, t_0) = \sum_n |\psi_n\rangle e^{-i\omega_n(t-t_0)} \langle \psi_n | . \quad (15.113)$$

En reportant ce développement dans (15.112) et en utilisant $\langle x | \psi_n \rangle = \psi_n(x)$, il vient :

$$U(x, t; x_0, t_0) = \sum_n \psi_n(x) e^{-i\omega_n(t-t_0)} \psi_n^*(x_0) . \quad (15.114)$$

En faisant $t = t_0$, on retrouve au second membre la relation de fermeture pour une base orthonormée, en conformité avec la condition initiale (15.97).

4. Si on intègre en x (ou en x_0) le propagateur de la particule libre, (15.104), on trouve³⁴ :

$$\int_{\mathbb{R}} dx U(x, t; x_0, t_0) = \sqrt{\frac{m}{2i\pi\hbar(t-t_0)}} \sqrt{\frac{-2\pi\hbar(t-t_0)}{im}} = 1 . \quad (15.116)$$

³⁴sachant que :

$$\int_{\mathbb{R}} dx e^{-i\alpha x^2} = \sqrt{\frac{\pi}{i\alpha}} . \quad (15.115)$$

Cette relation est propre à la particule libre, n'a aucun caractère de généralité, ni aucun sens particulier ; par exemple, dans le cas de l'oscillateur harmonique, on verra que le propagateur est donné par :

$$U(x, t; x_0, 0) = \sqrt{\frac{m\omega}{2i\pi\hbar\sin\omega t}} e^{\frac{im\omega}{2\hbar\sin\omega t} [(x+x_0)^2 \cos\omega t - 2xx_0 \sin\omega t]} . \quad (15.117)$$

Son intégrale vaut $(\cos\omega t)^{-1/2} e^{[m\omega x_0^2/(2i\hbar)] \tan\omega t} \neq 1$, sauf³⁵ si $\omega = 0$ (ou aux instants t_n tels que $\omega t_n = 0$ (2π)). Il n'y a d'ailleurs aucune raison que l'intégrale de U sur l'espace soit égale à 1, sauf précisément pour la particule libre où aucun point de l'espace n'est privilégié.

15.3 Séparation espace – temps et états stationnaires

Considérons le cas où, à $t = 0$, l'état est l'un des états propres du Hamiltonien, soit $|\psi_{m_0}\rangle$ (supposé non dégénéré) d'énergie $E_{m_0} = \hbar\omega_{m_0}$. L'état qui en découle à l'instant t s'obtient par application de l'opérateur d'évolution $U(t)$. Il en résulte, dans les mêmes notations que précédemment :

$$|\Psi(t)\rangle = \sum_n |\psi_n\rangle e^{-i\omega_n t} \langle\psi_n|\psi_{m_0}\rangle = e^{-i\omega_{m_0} t} |\psi_{m_0}\rangle \equiv |\Psi_{\text{st}}(t)\rangle . \quad (15.118)$$

Un tel état, où la dépendance en temps se réduit à un facteur de phase global, est appelé état stationnaire. La fonction d'onde est alors de la forme $\Psi(\vec{r}, t) = e^{-i\omega_{m_0} t} \psi_{m_0}(\vec{r})$ et les variables temps et espace sont séparées. La terminologie vient du fait que la moyenne de toute observable indépendante du temps A dans un tel état est une constante dans le temps, que cette observable soit ou non une constante du mouvement. Il est facile de vérifier ce fait, en raisonnant soit à la Heisenberg, soit à la Schrödinger.

Il en va évidemment tout autrement si l'état initial est une combinaison linéaire de plusieurs états propres d'énergies différentes ; ce cas permet la distinction explicite entre une observable commutant avec H (constante du mouvement) et une observable ne commutant pas avec H . Pour montrer ceci, il suffit de considérer une combinaison à deux termes ; soit donc l'état initial suivant, formé à partir de deux états associés à deux valeurs propres distinctes de H ($E_{m_1} \neq E_{m_2}$) :

$$|\Psi(0)\rangle = c_1 |\psi_{m_1}\rangle + c_2 |\psi_{m_2}\rangle , \quad (15.119)$$

où les deux coefficients c_i sont donnés une fois pour toutes. Le vecteur d'état au temps t est ($E_m = \hbar\omega_m$) :

$$|\Psi(t)\rangle = c_1 e^{-i\omega_{m_1} t} |\psi_{m_1}\rangle + c_2 e^{-i\omega_{m_2} t} |\psi_{m_2}\rangle \quad (15.120)$$

et la valeur moyenne d'une observable quelconque A est :

$$\begin{aligned} \langle A \rangle(t) &\equiv \langle \Psi(t) | A | \Psi(t) \rangle \\ &= |c_1|^2 \langle \psi_{m_1} | A | \psi_{m_1} \rangle + |c_2|^2 \langle \psi_{m_2} | A | \psi_{m_2} \rangle + 2\Re \left(e^{i(\omega_{m_2} - \omega_{m_1})t} c_1 c_2^* \langle \psi_{m_2} | A | \psi_{m_1} \rangle \right) . \end{aligned} \quad (15.121)$$

Maintenant, de deux choses l'une :

- ou bien A est une constante du mouvement ; alors, son commutateur avec H est nul, ainsi que l'élément de matrice $\langle \psi_{m_2} | A | \psi_{m_1} \rangle$, d'où il résulte que $\langle A \rangle$ est indépendant du temps.
- ou bien A n'est pas une constante du mouvement ; alors, en général l'élément de matrice $\langle \psi_{m_2} | A | \psi_{m_1} \rangle$ n'est pas nul. Dans ces conditions, $\langle A \rangle(t)$ oscille dans le temps à la pulsation de Bohr $\omega_{m_2} - \omega_{m_1}$.

La généralisation à une combinaison linéaire quelconque est immédiate. On se convainc ainsi à nouveau que toute observable qui commute avec H a une valeur moyenne constante dans le temps, quel que soit l'état considéré. En revanche, si $[A, H] \neq 0$, les termes oscillants sont bel et bien présents : les probabilités P_m de trouver les différentes valeurs a_m lors d'une mesure de A à l'instant t sont des fonctions explicites du temps écoulé depuis la dernière préparation.

³⁵D'ailleurs, si $\omega = 0$, l'expression (15.117) redonne la particule libre, comme il se doit.

A propos des états stationnaires, on peut démontrer l'analogie quantique du théorème du Viriel ([23], p. 82) bien connu en Mécanique Classique. Sous sa forme générale, ce théorème affirme que pour un mouvement périodique classique et pour un ensemble de N points matériels, l'égalité suivante est vraie :

$$\bar{T} = -\frac{1}{2} \sum_{i=1}^N \overline{\vec{F}_i \cdot \vec{r}_i} , \quad (15.122)$$

où T désigne l'énergie cinétique, \vec{F}_i la force appliquée au point matériel i et où la barre horizontale désigne la moyenne sur une période du mouvement. Le second membre porte le nom de Viriel de Clausius et il existe une démonstration très simple de la loi des gaz parfaits utilisant la relation ci-dessus ([23], p. 84) ; la base de la démonstration de ce théorème est la loi fondamentale de la Dynamique.

En pratique, le théorème du Viriel est très commode lorsque l'énergie potentielle V est une fonction homogène des coordonnées. En considérant un seul point matériel pour simplifier et en prenant par exemple un potentiel à symétrie sphérique de la forme $V(r) = ar^n$ (alors $V'(r) = (n/r)V(r)$), le théorème du Viriel prend immédiatement la forme :

$$2\bar{T} = n\bar{V} . \quad (15.123)$$

Compte tenu de l'identité formelle entre les équations du mouvement classiques et quantiques, on se doute qu'il existe un théorème semblable en Mécanique Quantique, la démonstration s'inspirant de la démonstration classique. Pour une seule particule, considérons la quantité $\vec{r} \cdot \vec{p}$ et son équation de Heisenberg :

$$i\hbar \frac{d}{dt} (\vec{r} \cdot \vec{p})_H = [\vec{r} \cdot \vec{p}, H]_H . \quad (15.124)$$

Un petit calcul montre que :

$$[\vec{r} \cdot \vec{p}, H] = -\vec{r} \cdot \vec{\nabla} V + \frac{\vec{p}^2}{m} . \quad (15.125)$$

Si on prend la moyenne membre à membre de (15.124), tout est nul puisque dans un tel état aucune moyenne ne dépend du temps. Donc, pour un état stationnaire, ou encore plus simplement pour n'importe quel état propre de H :

$$\langle -\vec{r} \cdot \vec{\nabla} V + \frac{\vec{p}^2}{m} \rangle = 0 \iff 2\langle T \rangle = \langle \vec{r} \cdot \vec{\nabla} V \rangle . \quad (15.126)$$

Ceci est l'expression du théorème du Viriel en Mécanique Quantique. Dans le cas où l'énergie potentielle V est à symétrie sphérique, le gradient de V ne produit que la composante radiale dV/dr et (15.126) devient :

$$\langle -\vec{r} \cdot \vec{\nabla} V + \frac{\vec{p}^2}{m} \rangle = 0 \iff 2\langle T \rangle = \langle r \frac{dV}{dr} \rangle . \quad (15.127)$$

Enfin, si V est une fonction homogène de r , $V(r) = ar^n$, alors $V'(r) = nar^{n-1} = nr^{-1}V(r)$ et (15.127) prend la forme simple :

$$2\langle T \rangle = n \langle V(r) \rangle . \quad (15.128)$$

Pour que ces écritures aient un sens, il faut évidemment que les moyennes soient finies, ce qui est le cas pour tout état lié.

Remarque

La forme (15.128) est l'expression traditionnelle du théorème du Viriel, reposant sur l'hypothèse que V est une fonction homogène de degré n . En réalité, la relation importante est (15.127). À une dimension pour simplifier, cette relation est :

$$2\langle T \rangle = \langle x V'(x) \rangle , \quad (15.129)$$

qui découle donc directement des équations de Heisenberg. On peut écrire une relation du genre Viriel chaque fois que l'on sait exprimer simplement la moyenne de xV' en fonction de celle de V , ce qui est possible s'il

existe un lien entre les fonctions V et xV' – par exemple quand V est une fonction homogène. Cette dernière hypothèse, suffisante, n'est toutefois pas nécessaire : il peut arriver que les moyennes elles-mêmes – et non pas $V(r)$ et $V'(r)$ – sont dans en relation simple. À titre d'exemple, soit un potentiel de Dirac $g\delta(x)$ et soit à calculer la moyenne de xV' sur un état normalisable³⁶, $\psi(x)$. Les règles opérationnelles énoncées plus haut à propos de δ (et de ses dérivées) conduisent à :

$$\langle xV'(x) \rangle = g \int_{-\infty}^{+\infty} dx \delta'(x) x |\psi(x)|^2 = -g \int_{-\infty}^{+\infty} dx \delta(x) \left(|\psi(x)|^2 + x \frac{d}{dx} |\psi(x)|^2 \right) \quad (15.130)$$

Le second terme ne contribue pas : la dérivée a bien un saut en $x = 0$ (sa valeur à droite est différente de sa valeur à gauche), mais de toute façon $x\delta(x)$ est nul. Il reste donc :

$$\langle xV'(x) \rangle = -g \int_{-\infty}^{+\infty} dx \delta(x) |\psi(x)|^2 = -g |\psi(0)|^2 \quad (15.131)$$

et ceci n'est autre que $-\langle V \rangle$. Dans le cas du puits de Dirac, on a donc :

$$2\langle T \rangle = -\langle V \rangle \quad (15.132)$$

exactement comme dans le cas du champ coulombien ($n = -1$) – alors qu'ici, $V(x)$ n'est visiblement pas une fonction homogène de degré -1 !

³⁶Pour qu'il y ait des états liés, il faut évidemment que la constante de couplage g soit négative ; ceci étant, on verra qu'un tel puits a toujours un et un seul état lié.

Chapitre 16

Problèmes à une dimension

Ce chapitre présente la méthode de résolution de l'équation aux valeurs propres de Schrödinger dans quelques cas simples (à une dimension d'espace) où les effets quantiques sont le plus souvent spectaculaires. On se souvient que ceux-ci se manifestent en particulier quand le potentiel¹ où se déplace la particule varie vite sur une longueur comparable à la longueur d'onde associée de de Broglie ; c'est évidemment le cas lorsque ce potentiel varie "instantanément" en présentant des discontinuités de première espèce, c'est-à-dire des sauts d'amplitude finie. Plusieurs effets surprenants seront mis en évidence ; on verra ainsi qu'une particule incidente sur un puits d'énergie peut être *réfléchie* (classiquement, elle passe toujours) ; de même, une particule arrivant sur un "mur" avec une petite énergie est capable de "passer à travers le mur" (effet tunnel) alors que, classiquement, elle ne peut que rebondir.

En outre, l'étude de ces situations, en soi physiquement intéressantes, donne l'occasion de se familiariser avec l'équation aux valeurs propres et surtout de réaliser le lien très fort existant entre les conditions imposées à la fonction d'onde, résultant de l'interprétation physique de celle-ci, et l'apparition spontanée de la quantification. Dans la suite on considérera typiquement les trois cas suivants :

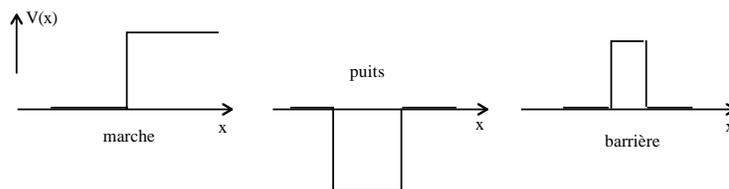


Figure 16.1: Exemples de potentiels constants par morceaux.

L'intérêt de ces potentiels variant par morceaux est de conduire, intervalle par intervalle, à une équation aux valeurs propres très simple à résoudre.

16.1 Propriétés générales des problèmes à une dimension

Comme on l'a vu, l'équation de Schrödinger peut se résoudre en deux étapes, après séparation du temps et de l'espace. En posant :

$$\Psi(\vec{r}, t) = e^{\frac{i}{\hbar} H t} \psi(\vec{r}) , \quad (16.1)$$

l'équation $i\hbar(\partial/\partial t)\Psi = H\Psi$ conduit à l'équation aux fonctions et valeurs propres de H :

$$H\psi(\vec{r}) = E\psi(\vec{r}) , \quad (16.2)$$

¹Conformément à un usage peu heureux, "potentiel" signifie énergie potentielle" de la particule.

En représentation- q , et à une dimension, si $V(x)$ désigne l'énergie potentielle de la particule de masse m , (16.2) s'écrit explicitement :

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V(x)\psi(x) = E\psi(x) . \quad (16.3)$$

Pour la suite de cette discussion générale, pour simplifier l'écriture, on pose :

$$V(x) = \frac{\hbar^2}{2m} U(x) , \quad E = \frac{\hbar^2}{2m} \varepsilon \quad ([U] = [\varepsilon] = \text{L}^{-2}) , \quad (16.4)$$

de sorte que l'équation aux valeurs propres s'écrit :

$$\psi'' + [\varepsilon - U(x)]\psi(x) = 0 . \quad (16.5)$$

Ceci constitue l'équation connue en Mathématiques sous le nom d'équation de Sturm-Liouville ; c'est une équation homogène (si ψ est solution, toute fonction proportionnelle, $\lambda\psi$, est encore solution) du second ordre. Ses solutions, satisfaisant en outre aux conditions requises par les postulats quantiques, seront les fonctions propres associées à l'énergie E (compte tenu du sens que lui attribuent les postulats, toute fonction propre $\psi(x)$ est tenue de posséder des propriétés spécifiques). En outre, le mouvement à $d = 1$ (sur \mathbb{R}) possède des propriétés générales importantes, indépendantes de la forme particulière de $V(x)$, qu'il est bon de connaître.

A priori, rien n'interdit à la fonction d'onde d'avoir des singularités, notamment des divergences ; cependant, l'intégrale de son module carré sur tout intervalle fini doit être finie : dans le cas contraire, la probabilité de trouver la particule dans une région d'extension finie serait infinie, ce qui n'aurait aucun sens². Il en résulte que si, à cause d'une singularité de $V(x)$ en un certain point x_0 , ψ présente une divergence au voisinage de ce point et se comporte comme $(x - x_0)^{-\alpha}$ ($\alpha > 0$) par exemple, l'exposant α ne peut être trop grand : l'intégrale de ψ^2 entre les deux points $x_0 \pm \delta x$ est proportionnelle à $|\delta x|^{-2\alpha+1}$ et doit rester finie dans la limite $\delta x \rightarrow 0$, ce qui impose $2\alpha \leq 1$. C'est bien le sens physique attribué à la fonction d'onde qui proscrit de trop violentes singularités à ψ , quand elles sont rendues possibles mathématiquement en raison d'un accident de l'énergie potentielle.

Par ailleurs, en dehors de toute condition physique résultant des postulats, il faut que l'équation aux valeurs propres, une fois écrite, ait un sens. En tout point x_1 dans le voisinage³ duquel cette équation est écrite, la dérivée seconde doit exister, ce qui entraîne que la fonction ψ et sa dérivée ψ' sont continues en ce point. Cette propriété peut d'ailleurs être établie explicitement. Quand $V(x)$ est une fonction continue, aucun problème ne se présente et il est clair qu'alors ψ'' est également continue⁴. Au contraire, quand le potentiel présente une anomalie, le comportement de ψ et/ou de ses dérivées doit refléter cette singularité d'une façon ou d'une autre.

Examinons ce qui se passe au voisinage d'un point d'abscisse x_0 où la fonction $V(x)$ présente un saut d'amplitude finie⁵. On va voir que si ψ et ψ' restent continues, la dérivée seconde, elle, présente aussi un saut – ce qui est presque évident à la simple vision de (16.5).

Intégrons membre à membre l'équation aux valeurs propres (16.5) entre les deux points $x_0 \pm \delta x$; il vient :

$$\psi'(x_0 + \delta x) - \psi'(x_0 - \delta x) + \int_{x_0 - \delta x}^{x_0 + \delta x} [\varepsilon - U(x)]\psi(x) dx = 0 , \quad (16.6)$$

soit :

$$\psi'(x_0 + \delta x) - \psi'(x_0 - \delta x) + \varepsilon \int_{x_0 - \delta x}^{x_0 + \delta x} \psi(x) dx = \int_{x_0 - \delta x}^{x_0 + \delta x} U(x)\psi(x) dx . \quad (16.7)$$

Le saut de potentiel peut *a priori* induire une singularité de ψ et/ou de ses dérivées ; imaginons qu'il s'agisse d'une divergence comme $(x - x_0)^{-\alpha}$; d'après ce qui précède, il faut que 2α soit inférieur à 1 : l'intégrale du premier membre est dans cette hypothèse proportionnelle à $\delta x^{1-\alpha}$ et donc tend vers zéro à la limite $\delta x \rightarrow 0$.

²Cette affirmation vaut pour tout état, lié ou non-lié.

³intervalle entourant le point considéré.

⁴Il suffit de considérer l'équation (16.3), qui permet d'exprimer ψ'' à l'aide de fonctions toutes continues par hypothèse

⁵On pourrait objecter qu'une telle situation est dénuée de sens physique dans la mesure où une vraie discontinuité n'existe pas en pratique. Il faut comprendre qu'il s'agit d'une modélisation du cas où le potentiel varie très vite sur une distance très petite vis-à-vis de toute échelle de longueur déjà présente dans le problème considéré.

En ce qui concerne l'intégrale du second membre de (16.7), elle donne elle aussi zéro. En effet, en désignant par U_g et U_d les deux limites à gauche et à droite de x_0 , on peut écrire :

$$U(x) = \begin{cases} U_g + \eta_-(x) & \text{si } x \lesssim x_0 \\ U_d + \eta_+(x) & \text{si } x \gtrsim x_0 \end{cases}, \quad (16.8)$$

où les deux fonctions $\eta_{\pm}(x)$ tendent vers zéro si $x \rightarrow x_0$. L'intégration à gauche donne :

$$\int_{x_0-\delta x}^{x_0} [U_g + \eta_-(x)]\psi(x)dx = [U_g + \text{Max}(\eta_-)]\mathcal{O}(\delta x^{1-\alpha}), \quad (16.9)$$

et donne zéro à la limite ; le même argument vaut pour l'intégrale de droite. Au total le second membre est nul à la limite, ce qui établit la continuité de la dérivée ψ' :

$$\lim_{\delta x \rightarrow 0} [\psi'(x_0 + \delta x) - \psi'(x_0 - \delta x)] = 0. \quad (16.10)$$

On retiendra donc :

$$\psi \text{ et } \psi' \text{ sont continues en tout point où } V(x) \text{ a un saut d'amplitude finie.} \quad (16.11)$$

Remarquons par ailleurs que cette propriété assure la continuité locale (en tout point) du courant de probabilité, condition également requise puisqu'un état stationnaire (où, en particulier, la densité de probabilité est constante dans le temps) est construit en multipliant simplement $\psi(x)$ par $e^{\frac{i}{\hbar}Ht}$. Dans un tel état, le courant s'écrit :

$$j(x) = \frac{\hbar}{2im} [\psi^*(x)\psi'(x) - \psi(x)\psi'^*(x)]. \quad (16.12)$$

Comme ψ et ψ' sont continues en x_0 , le courant l'est également⁶. Dans le cas contraire, la singularité de potentiel ferait "disparaître" partiellement la particule (le courant à gauche serait différent du courant à droite, il y aurait une "fuite" de probabilité). Notons enfin que, puisque ψ et ψ' sont continues et que l'équation aux valeurs propres ne contient que ψ et ψ'' , c'est la dérivée seconde à elle seule qui "encaisse" le saut de $V(x)$ – elle a donc une discontinuité (ce qui ne saute pas forcément aux yeux sur le graphe de $\psi(x)$). Ceci se comprend bien par référence à la situation classique : dans ce cas, la discontinuité de l'énergie potentielle produit, par la conservation de l'énergie mécanique, des variations brusques de la vitesse et donc de l'énergie cinétique ; or, en Mécanique Quantique, c'est précisément ψ'' qui est associée à l'énergie cinétique. Notons finalement que le saut de ψ'' est donné par :

$$\psi''(x_0 + 0) - \psi''(x_0 - 0) = (U_d - U_g)\psi(x_0). \quad (16.13)$$

Il est important de se souvenir que la continuité de ψ et de ψ' n'est obtenue que si $V(x)$ a un saut *fini*. Si, au contraire, le saut de $V(x)$ est infini (c'est le cas du puits carré infiniment profond, voir section 16.3.2), la singularité induite est naturellement plus violente : alors la dérivée première ψ' présente également un saut (fini) ; en pareille circonstance, l'équation de Schrödinger ne peut être écrite au point singulier : si ψ' a un saut en un point, ψ'' n'existe pas en ce point ; elle peut au mieux être écrite dans tout demi-voisinage d'un point aussi près que l'on veut mais distinct du point de discontinuité de $V(x)$.

C'est le sens de densité de probabilité postulé et attribué à la fonction d'onde qui limite la sévérité des divergences acceptables ; l'équation de Schrödinger vient alors prendre le relais pour montrer que, même si le potentiel est discontinu, la fonction d'onde et sa première dérivée doivent être continues. Cette propriété sera utilisée à maintes reprises, et fournira des conditions de raccordement de ψ et de ψ' de part et d'autre d'une discontinuité de potentiel ; ces conditions constituent l'origine la quantification.

On doit donc trouver les solutions finies, continues et dérivables de l'équation de Sturm-Liouville (16.5). Comme $U(x)$ est une fonction à valeurs réelles et que, bien sûr l'énergie E est réelle, la partie réelle et la partie imaginaire de $\psi(x)$ satisfont séparément la même équation aux valeurs propres correspondant à la même énergie. En conséquence, lorsque l'état d'énergie E n'est pas dégénéré (on verra que c'est vrai pour tout état lié à une dimension), parties réelle et imaginaire sont simplement proportionnelles et, au total, on peut prendre la fonction d'onde réelle.

⁶On verra en outre que pour tout état lié à une dimension, le courant est forcément nul partout.

La plupart des propriétés générales des solutions de l'équation aux valeurs propres résulte d'un important théorème, dit théorème du Wronskien ([24], p. 83), dont la démonstration est très simple. Soit deux solutions ψ_1 et ψ_2 associées aux deux valeurs propres ε_1 et ε_2 :

$$\psi_1'' + [\varepsilon_1 - U(x)]\psi_1 = 0, \quad \psi_2'' + [\varepsilon_2 - U(x)]\psi_2 = 0. \quad (16.14)$$

Le Wronskien de ces deux solutions est défini comme :

$$W[\psi_1, \psi_2] = \psi_1\psi_2' - \psi_2\psi_1' \equiv \begin{vmatrix} \psi_1 & \psi_2 \\ \psi_1' & \psi_2' \end{vmatrix}. \quad (16.15)$$

Multiplions la deuxième équation (16.14) par ψ_1 et retranchons-lui la première multipliée par ψ_2 ; après simplification il vient :

$$\psi_2''\psi_1 - \psi_1''\psi_2 = [\varepsilon_1 - \varepsilon_2]\psi_1\psi_2. \quad (16.16)$$

Ajoutons et retranchons le produit $\psi_1'\psi_2'$ et intégrons entre deux points d'abscisses a et b ($a < b$) :

$$[\psi_1\psi_2' - \psi_2\psi_1']_a^b = (\varepsilon_1 - \varepsilon_2) \int_a^b \psi_1(x)\psi_2(x) dx. \quad (16.17)$$

Cette relation constitue l'énoncé simplifié (suffisant pour les besoins actuels ou futurs) du théorème du Wronskien.

En particulier, si ψ_1 et ψ_2 sont associées à la même valeur propre ε ($\varepsilon_1 = \varepsilon_2 = \varepsilon$), la variation du Wronskien entre a et b est nulle :

$$[\psi_1\psi_2' - \psi_2\psi_1']_a^b = 0 \iff \psi_1(a)\psi_2'(a) - \psi_2(a)\psi_1'(a) = \psi_1(b)\psi_2'(b) - \psi_2(b)\psi_1'(b) \quad (16.18)$$

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

$$W[\psi_1, \psi_2] = \text{C}^{\text{ste}}. \quad (16.19)$$

Ceci permet d'établir que les états discrets (*i. e.* normalisables) de tout mouvement à une dimension sont non-dégénérés. En effet, pour de tels états, les fonctions ψ_i et ψ_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 :

$$\psi_1(x)\psi_2'(x) - \psi_2(x)\psi_1'(x) = 0 \quad \forall x \iff \psi_2(x) \propto \psi_1(x). \quad (16.20)$$

En définitive, deux fonctions propres associées à la même énergie E sont proportionnelles et donc constituent une seule et même solution. Tout état lié à une dimension est donc non-dégénéré⁷. Il est également possible de démontrer que l'état fondamental ne s'annule pas (n'a pas de noeuds) sauf, bien sûr, aux frontières du domaine accessible : sur cette frontière, tout état propre doit évidemment s'annuler en vertu de la continuité de la fonction d'onde (en tout point) et puisque ψ est nulle là où la particule ne peut se trouver. L'absence de noeuds pour le fondamental peut se percevoir de façon intuitive : plus une fonction a un grand nombre de zéros, plus elle oscille ; plus elle oscille, plus sa dérivée est grande en module. Comme la valeur moyenne de l'énergie cinétique est égale à $(\hbar^2/(2m)) \int |\psi'|^2 dx$, de nombreuses oscillations produisent une grande énergie cinétique moyenne ; au contraire, l'énergie potentielle varie beaucoup moins vite avec le nombre d'oscillations, puisque seule ψ est concernée : des oscillations rapides ont donc tendance à augmenter la valeur totale de l'énergie. La fonction d'onde de l'état fondamental ne peut s'annuler que sur la frontière du domaine accessible à la particule ; quant à celle du $n^{\text{ème}}$ état excité, elle a exactement n zéros⁸.

⁷Voir aussi [7], p. 57 ; en ce qui concerne le spectre continu et sa dégénérescence égale à 2, voir Landau [25] p. 80. Celle-ci résulte du fait que, l'équation étant réelle $-V(x)$ est une fonction à valeurs réelles (condition nécessaire et suffisante pour que la probabilité se conserve) -, ψ et ψ^* satisfont l'une comme l'autre cette équation. C'est lorsque ces deux fonctions sont proportionnelles qu'il n'y a pas en fait de dégénérescence : il en va ainsi pour tous les états liés (mais peut se produire aussi pour le spectre continu). En tout cas, la fonction propre $\psi(x)$ d'un état lié à une dimension peut toujours être supposée réelle.

Prenons par exemple le cas d'une particule confinée sur le demi-axe réel positif (existence d'un mur parfaitement réfléchissant en $x = 0$) mais libre par ailleurs ; alors, $\psi(x \leq 0) = 0$, ce qui contraint la combinaison linéaire $e^{\pm ikx}$ et produit une et une seule solution, à savoir $\sin kx$: pour une valeur de l'énergie, il n'y a plus qu'une seule solution propre et donc plus de dégénérescence ; c'est aussi le cas pour le potentiel de Morse (voir [25] § 23, problème 3 p. 92) et, d'une façon générale, lorsque le potentiel diverge d'un côté et tend vers une valeur finie de l'autre (cette dernière condition assure l'existence d'un spectre continu). La raison profonde de ceci est la suivante : l'existence du mur ou du potentiel qui diverge brise la symétrie gauche-droite, ou la symétrie ($v \rightarrow -v$) ; tout abaissement de symétrie s'accompagne évidemment d'une réduction de la dégénérescence puisque celle-ci traduit l'existence de symétries.

⁸ "théorème d'oscillation", selon Landau ([25], § 21, p. 79).

En outre, compte tenu de l'expression (16.12) du courant de probabilité, on voit que le Wronskien d'une solution et de sa complexe conjuguée et le courant j sont eux-mêmes proportionnels :

$$j(x) = \frac{\hbar}{2im} W[\psi, \psi^*] . \quad (16.21)$$

Comme, d'une façon générale (qu'il s'agisse d'un état lié ou non-lié), le Wronskien est constant, il en résulte que le courant $j(x)$ l'est aussi et prend la même valeur en tout point. Si l'état est lié (donc normalisable), le Wronskien est nul, comme on l'a vu plus haut : le courant est donc nul dans tout état lié à une dimension d'espace, ce qui est évident physiquement. Si l'état est non-lié, le courant est en général⁹ non-nul mais il reste constant dans l'espace.

Tout ceci est bien compatible avec l'équation de conservation, écrite pour une dimension spatiale :

$$\frac{\partial}{\partial t} |\Psi(x, t)|^2 + \frac{\partial}{\partial x} j(x) = 0 ; \quad (16.22)$$

comme un état stationnaire – où $|\Psi|^2$ et j sont constants dans le temps – se fabrique en multipliant simplement un état propre $\psi(x)$ d'énergie E par $e^{\frac{i}{\hbar}Et}$, l'équation de conservation appliquée à un état stationnaire est tout simplement $dj(x)/dx = 0$. Bien évidemment, la dimension 1 joue un rôle crucial : il est bien certain que l'on peut au contraire avoir des courants permanents en dimension $d = 2$ ou plus, courants circulaires par exemple.

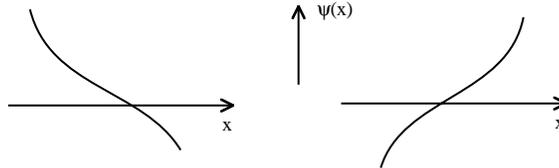


Figure 16.2: Allure de la fonction propre si $E < V_{\min}$.

Une autre propriété importante des états propres est la suivante : toute valeur propre E ne peut être inférieure au minimum de $V(x)$, V_{\min} , supposé exister¹⁰ à distance finie x_{\min} . Notons d'abord que l'on peut se restreindre au cas des états liés : leur énergie est inférieure à celle des états non liés (sauf cas très exotiques) ; il suffit donc ici de montrer que l'énergie de tout état lié est supérieure à V_{\min} . Ceci peut se démontrer graphiquement moyennant un raisonnement par l'absurde. Supposons qu'il existe une valeur propre E inférieure à V_{\min} ; dans ces conditions, l'équation (16.5) montre ψ'' et ψ ont le même signe¹¹ $\forall x$. ψ a donc l'une des deux allures schématisées sur la fig. 16.2.

La concavité est dirigée vers le haut quand la la fonction ψ est positive, vers le bas dans le cas contraire, et la traversée de l'axe horizontal se fait suivant un point d'inflexion¹². Il est bien clair que de telles fonctions ne sauraient être normalisables ; supposer que $E < V(x)$ partout conduit donc à une absurdité et seule reste la possibilité $E \geq V_{\min}$. D'ailleurs, cette inégalité est même stricte : la fonction propre est diffuse dans l'espace, de sorte que la valeur moyenne de $V(x) - V_{\min}$ est strictement positive (intégrale d'une quantité positive) ; à cette moyenne, il convient d'ajouter la moyenne de l'énergie cinétique (qui est évidemment positive), ce qui augmente encore la somme $\langle p^2/(2m) + V \rangle$, laquelle n'est autre que la valeur propre E elle-même. En quelque sorte, l'écart entre E et V_{\min} est une nouvelle manifestation des relations d'incertitude de Heisenberg¹³. On retiendra donc :

$$E > V_{\min} \equiv \text{Min}_x[V(x)] . \quad (16.23)$$

En d'autres termes, il existe toujours quelque part une région accessible classiquement (l'énergie mécanique classique est toujours supérieure ou égale à l'énergie potentielle).

⁹Pour une particule libre confinée sur un demi espace ($x > 0$ pour fixer les idées) grâce à une barrière totalement réfléchissante, mais libre par ailleurs, tous les états sont non-liés et le courant est toujours nul : puisque le mur réfléchit parfaitement, le courant total est nul sur la barrière (courant incident = courant réfléchi) et donc nul partout.

¹⁰Dans toute la discussion, on suppose $V(x)$ borné inférieurement.

¹¹Cela a un sens de parler du signe de ψ puisque, au moins pour les états liés – les seuls qui importent ici –, la fonction d'onde peut toujours être choisie réelle.

¹²On suppose que $V(x)$ n'a pas de saut, donc ψ'' est aussi partout continue.

¹³qui impliquent, en particulier, que le repos absolu n'existe pas pour une particule (voir aussi la discussion dans [16], p. 355).

Le théorème du Wronskien sous sa forme élémentaire (16.17) montre également (ce que l'on sait déjà), que deux états liés sont orthogonaux. En effet, aucun d'entre eux n'est dégénéré (deux d'entre eux correspondent forcément à deux valeurs propres *distinctes*) et peuvent être pris réels. En choisissant pour a et b les limites du domaine de confinement (où nécessairement ψ s'annule s'il s'agit d'un état lié), le Wronskien est visiblement nul, donc il en va de même pour l'intégrale du second membre – laquelle n'est autre que le produit scalaire $\langle \psi_1 | \psi_2 \rangle$ pour deux fonctions réelles à une phase globale près.

Le comportement asymptotique des solutions $\psi(x)$ ($x \rightarrow \pm\infty$) peut maintenant être analysé en détail, indépendamment de la forme précise de la fonction $U(x)$, sans toutefois viser la plus grande généralité mais bien plutôt des situations physiques typiques. Il est clair que le comportement de la fonction d'onde est en réalité conditionné par le signe de la différence $\varepsilon - U$; il est donc utile de discuter les deux cas séparément :

- $\varepsilon - U(x) > 0 \forall x > x_0$ (figure 16.3, à gauche).

Cette condition est en particulier réalisée lorsque le potentiel a une limite U_+ quand $x \rightarrow +\infty$ et si la valeur propre ε est supérieure à U_+ ¹⁴. Elle est également réalisée lorsque $V(x)$, sans avoir de limite, reste néanmoins fini et, par exemple, a une variation périodique. En pareil cas, il est indispensable de faire une exploitation préalable de cette symétrie ; alors, l'analyse à l'infini disparaît et est remplacée par une analyse à distance finie sur une surface caractéristique de la symétrie périodique¹⁵.

Dans le cas où la limite U_+ existe, à condition que $U(x)$ s'en rapproche assez vite¹⁶, il est légitime de raisonner directement sur la forme asymptotique de l'équation (16.5) elle-même :

$$\psi'' + \varepsilon \psi(x) \simeq 0, \quad (16.24)$$

et il en résulte que ψ est (approximativement) une combinaison linéaire d'exponentielles imaginaires, que l'on peut toujours écrire comme :

$$\psi(x) \simeq A_+ \sin(kx - \phi_+) \quad (k = \sqrt{\varepsilon - U_+}). \quad (16.25)$$

Ainsi, la fonction d'onde oscille entre deux valeurs $\pm A_+$ et son module carré est clairement non-normalisable dans tout l'espace.

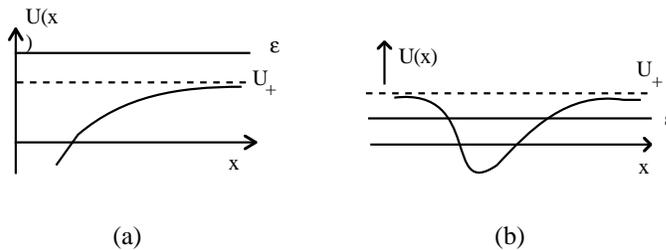


Figure 16.3: Comportements typiques de $V(x)$ à l'infini.

- $\varepsilon - U(x) < 0 \forall x > x_0$. Ce cas est réalisé dans la situation physique illustrée sur la figure 16.3 à droite) et représente un *puits* de potentiel typique, pour les valeurs propres ε inférieures à U_+ ; raisonnant à nouveau avec la forme asymptotique de (16.5), on voit maintenant que ψ se comporte à l'infini comme une combinaison linéaire d'exponentielles¹⁷ *réelles* :

$$\psi(x) \simeq A e^{kx} + B e^{-kx} \quad (k = \sqrt{U_+ - \varepsilon}), \quad (16.26)$$

L'inégalité $\varepsilon < U(x) \forall x > x_0$ est aussi satisfaite si le potentiel diverge à l'infini ; comme $U(x)$ ne peut plus être pris à peu près constant mais dépend alors de x jusqu'au bout, l'analyse est nettement plus complexe : elle sera entreprise dans le détail à propos de l'oscillateur harmonique (Ch. 17) présentant d'ailleurs un cas exemplaire. D'une façon générale, en pareil cas, toutes les énergies propres sont quantifiées (il n'existe

¹⁴en pratique, on prend souvent l'origine des potentiels à l'infini, auquel cas cette limite est nulle.

¹⁵La périodicité du potentiel conduit au théorème de Bloch qui est à la base des propriétés des électrons dans les solides (théorie des bandes).

¹⁶Il faut que $U(x)$ tende vers U_+ plus vite que $1/x$ ([24], p. 86).

¹⁷Il est clair qu'il faudra en temps utile invoquer une bonne raison pour éliminer les fonctions divergentes.

pas d'états non-liés), au contraire de ce qui va être montré maintenant¹⁸ pour des potentiels constants par morceaux.

16.2 La quantification comme conséquence des conditions imposées à Ψ

L'équation de Sturm - Liouville (16.5) admet une infinité non dénombrable de solutions et n'impose en aucune façon, par elle-même, la moindre condition à la valeur propre E : pourvu que l'on donne deux conditions – par exemple les valeurs de ψ et de ψ' en un point donné – il est en principe possible de trouver la solution $\psi(x)$ satisfaisant ces deux conditions “initiales” ; en pratique, la résolution peut être difficile, mais ceci est une autre affaire. En définitive, ce n'est pas l'équation aux valeurs propres en elle-même qui est de nature à restreindre les valeurs possibles de l'énergie E : c'est le contenu physique attribué à la fonction d'onde et exprimé par les postulats qui va tout naturellement conduire à la quantification de l'énergie des états liés. Ce contenu physique, allié à l'équation fondamentale, engendre la quantification, le fait majeur incontournable, imposé par l'expérience, et qui a disqualifié la Mécanique Classique. Au total, c'est bien l'apparition spontanée de la quantification de l'énergie dans la formulation de Schrödinger “habillée” avec les prescriptions de Born et Jordan, qui constitue le premier succès spectaculaire du nouvel édifice théorique.

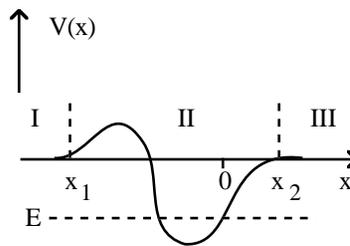


Figure 16.4: Exemple de potentiel non-nul sur un intervalle fini $[x_1, x_2]$; ici, $V_{\min} < E < 0$

Soit V_{\min} ($\leftrightarrow U_{\min}$) la plus petite valeur de l'énergie potentielle et considérons le cas (très usuel) où $V(x)$ a la même limite en $x = -\infty$ et en $x = +\infty$, valeur prise comme origine de l'énergie potentielle ; dans ces conditions, les limites U_{\pm} coïncident et sont égales à zéro. Pour simplifier encore la discussion, supposons que $V(x)$ s'annule *strictement* au-delà d'une certaine valeur x_2 et en-deçà d'une valeur x_1 (fig. 16.4). L'analyse ci-dessus conduit à envisager successivement deux cas¹⁹, $E < 0$, puis $E > 0$.

16.2.1 $\varepsilon < 0 \iff E < 0$

Comme on cherche une solution à énergie négative, il s'agit bien d'un état lié normalisable au sens usuel. Dès que x est inférieur à x_1 (région I), $V'(x)$ est nul par hypothèse de sorte que la fonction d'onde est strictement égale à :

$$x < x_1 < 0 : \quad \psi(x) = A e^{kx} + B e^{-kx} \quad , \quad k = \sqrt{-\varepsilon} \quad . \quad (16.27)$$

Il est clair que cette fonction ne peut représenter un état lié : un tel état correspond physiquement au cas où la particule reste localisée essentiellement autour d'un centre fixe, ce qui signifie que le module carré de la fonction d'onde décroît à grande distance et tend d'ailleurs vers zéro assez vite pour que son intégrale soit finie ; ψ donnée

¹⁸Quoi qu'il en soit, tout ce qui est discuté et/ou affirmé ici ne concerne que le mouvement à une dimension. Dans \mathbb{R}^3 , il peut exister des états liés “noyés” dans le *continuum* des états de diffusion ; à ce sujet, voir par exemple [26]. De tels états n'ont manifestement pas d'équivalents classiques.

¹⁹Comme la limite de $V(x)$ à l'infini est nulle, c'est précisément la valeur zéro pour l'énergie qui sépare les deux types d'états (liés : $E < 0$, non-liés, $E > 0$).

par (16.27), non seulement ne décroît pas à l'infini, mais diverge exponentiellement ! Dans la région I, le seul choix possible physiquement est $B = 0$, donnant dans cet intervalle la seule classe de solutions acceptables :

$$x < x_1 < 0 : \quad \psi(x) = A e^{kx} \quad (16.28)$$

Il convient maintenant de prolonger sur tout l'axe réel la fonction précisément égale à $A e^{kx}$ dans la région I. Dans la région II, ψ est de la forme :

$$x_1 < x < x_2 : \quad \psi(x) = \alpha e_1(x) + \beta e_2(x) . \quad (16.29)$$

En effet, l'équation est une équation différentielle homogène du second ordre, dont la solution générale s'exprime en combinaison linéaire de deux solutions particulières indépendantes, ici notées $e_1(x)$ et $e_2(x)$; celles-ci dépendent bien évidemment de la forme particulière de la fonction $V(x)$. α et β sont deux constantes fixées de façon unique par l'écriture en x_1 de la continuité de ψ et de sa dérivée²⁰. Les équations de continuité sont :

$$A e^{kx_1} = \alpha e_1(x_1) + \beta e_2(x_1) , \quad k A e^{kx_1} = \alpha e_1'(x_1) + \beta e_2'(x_1) , \quad (16.30)$$

et constituent un système homogène permettant de trouver les deux constantes α et β . Manifestement, chacune d'entre elles est proportionnelle à A , celle-ci étant une fonction de x_1 et surtout de l'énergie, par l'intermédiaire de k . Ceci étant fait, la fonction d'onde est complètement déterminée (à un facteur global près) $\forall x \leq x_2 - y$ compris en x_2 (par sa limite à gauche), ce qui est important pour la suite, et les deux constantes α et β ressortent comme des fonctions de l'énergie E .

Enfin, dans la région III – où $V(x)$ est à nouveau nul –, la solution la plus générale est de la forme :

$$x > x_2 > 0 : \quad \psi(x) = C e^{-kx} + D e^{+kx} , \quad k = \sqrt{-\varepsilon} . \quad (16.31)$$

et, pour la même raison que B est nul dans la région I, il faut ici $D = 0$. Maintenant, l'idée est précisément de montrer que cette dernière condition n'est satisfaite que si E – c'est-à-dire ε – prend certaines valeurs bien particulières, à l'exclusion de toute autre. Pour cela, posons le problème à l'envers et cherchons à exprimer la constante D en fonction de ce que l'on sait déjà, compte tenu de la nécessité de satisfaire les équations de raccordement en $x = x_2$. Les conditions de continuité de ψ et ψ' s'écrivent respectivement :

$$\alpha e_1(x_2) + \beta e_2(x_2) = C e^{-kx_2} + D e^{+kx_2} , \quad (16.32)$$

$$\alpha e_1'(x_2) + \beta e_2'(x_2) = -k C e^{-kx_2} + k D e^{+kx_2} . \quad (16.33)$$

Ce système linéaire permet de trouver D sous la forme d'une combinaison linéaire de α et β ; comme ces deux constantes sont des fonctions maintenant connues de l'énergie E , D ressort *in fine* sous la forme d'une certaine fonction $\Phi(E)$, caractéristique du potentiel particulier considéré :

$$D = \Phi(E) . \quad (16.34)$$

D'après la première constatation, il faut que D soit nul ; ceci ne peut se produire que pour des énergies bien particulières, qui ne sont rien d'autre que les zéros de la fonction $\Phi(E)$ (voir fig. 16.5). Ainsi, une condition nécessaire pour que ψ ait les qualités requises (sommabilité du module carré) est de choisir l'énergie E dans un ensemble discret $\{E_n\}_n$.

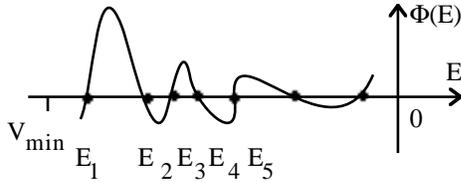
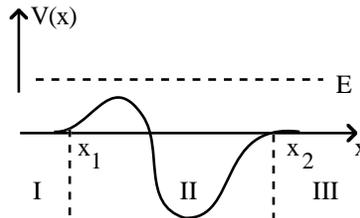
L'argument qui précède est tout à fait général : la quantification de l'énergie des états liés est une conséquence directe des postulats précisant notamment le sens physique de la fonction ψ et la dotant *ipso facto* de propriétés spécifiques.

16.2.2 $\varepsilon > 0 \iff E > 0$

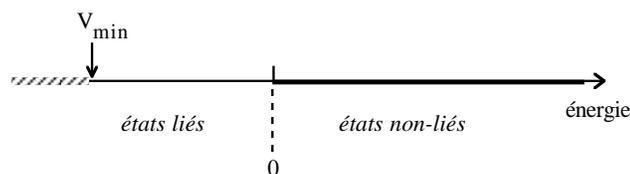
La disposition est maintenant celle illustrée sur la fig. 16.6 ; ε est positif, c'est-à-dire, essentiellement, supérieur à la limite de $U(x)$ en $\pm\infty$. Dans ce cas, dans les régions I et III, toutes les solutions sont de la forme :

$$x < x_1 \text{ ou } x > x_2 : \quad \psi(x) = C_{\pm} \sin(kx - \phi_{\pm}) , \quad k = \sqrt{+\varepsilon} . \quad (16.35)$$

²⁰Il faut bien voir que les deux expressions (16.27) et (16.29) sont les deux expressions dans deux régions adjacentes d'une seule et même fonction.

Figure 16.5: Illustration des zéros de la fonction $\Phi(E)$.Figure 16.6: Exemple de potentiel non-nul sur un intervalle fini $[x_1, x_2]$; maintenant $E > \text{Max}_x[V]$.

Aucune de ces fonctions n'est normalisable – mais toutes peuvent être combinées pour former des paquets d'onde, d'extension spatiale aussi grande que l'on veut, ce qui signifie que l'énergie (associée à k) est définie à une précision aussi grande que l'on veut. Il en résulte qu'aucune condition supplémentaire n'est à prendre en compte (étant entendu qu'avec la convention de potentiel nul à l'infini, E doit être positif) et toutes les fonctions de cette famille sont acceptables ; en conséquence, aucune condition ne vient sélectionner les valeurs possibles de l'énergie E : il n'y a pas de quantification, le spectre d'énergie est continu. Ces solutions donnent une densité de probabilité de présence non nulle à une distance arbitrairement grande de la région où le potentiel varie et représentent des états non-liés ; elles servent de base à la description des phénomènes de collision – ou, plus exactement, de diffusion – par le potentiel $V(x)$, qui modélise une cible désincarnée physiquement.

Figure 16.7: Allure typique du spectre de H pour un potentiel $V(x)$ borné de limite nulle à l'infini ($d = 1$).

On retiendra donc les résultats essentiels suivants, sans oublier qu'ils sont relatifs à un puits de potentiel présentant les caractéristiques énoncées ci-dessus (dimension $d = 1$, existence d'une limite finie à l'infini – prise d'ailleurs comme origine de l'énergie potentielle) :

- l'énergie des états liés ne peut prendre que certaines valeurs discrètes et aucune valeur propre discrète n'est dégénérée. Toutes ces valeurs sont supérieures au minimum de potentiel V_{\min}
- l'énergie des états non-liés forme un *continuum*.

Ces résultats sont illustrés sur la fig. 16.7. Dans le cas où $\lim_{|x| \rightarrow \infty} V(x) = +\infty$ (exemple : l'oscillateur

harmonique), il n'existe évidemment que des états liés (à l'infini, on a alors $\psi'' \simeq U(x)\psi$) ; comme U diverge, il faut que ψ tende vers zéro très vite, de sorte que, au total, ψ'' tend également vers zéro.

Finalement, on voit qu'il existe *a priori* deux types d'états :

1. les états décrits par les solutions qui décroissent assez vite à l'infini de sorte que l'intégrale dans tout l'espace du carré de leur module soit finie. Cette condition de normalisabilité a une conséquence extraordinaire : pour qu'il en soit ainsi, il faut et il suffit que la valeur propre E (l'énergie) prenne une suite de valeurs discrètes bien définies, en nombre fini ou infini, typiques du potentiel et dépendant des autres grandeurs physiques en jeu (la masse de la particule, par exemple). Ainsi, au prix de l'introduction de la fonction d'onde, elle-même dotée de propriétés résultant du sens physique qui lui est attribué, se trouve justifiée et comprise l'existence d'une quantification de l'énergie. Ces états, où la fonction d'onde s'annule rapidement à l'infini représentent clairement des situations où la particule reste confinée autour d'une origine typique du potentiel V . On parle naturellement d'états liés, tout comme l'électron de l'atome de Bohr est considéré comme lié s'il décrit une orbite circulaire²¹. Il convient de remarquer que de telles solutions n'existent pas toujours (un potentiel arbitraire n'a pas forcément des états liés). Physiquement, cela se comprend : le potentiel doit être suffisamment attractif pour que la particule puisse s'y accrocher ; d'ailleurs, pour un potentiel donné l'existence de l'état lié dépend également de la masse de la particule : d'une façon générale, plus la masse est faible, plus le nombre d'états liés est petit – toutes choses égales par ailleurs.
2. les solutions non normalisables, soit parce que la fonction d'onde ne décroît pas assez vite, soit tout simplement parce qu'elle n'a pas de limite à l'infini (par exemple, quand ψ oscille). Clairement, ce type de fonction représente un état où la particule a une probabilité finie de se trouver dans un intervalle fini situé à une distance arbitrairement grande : il s'agit donc d'un état non-lié. Dans un tel cas l'énergie n'est pas quantifiée.

La conclusion majeure est la quantification de l'énergie des états liés en conséquence de l'équation fondamentale de la théorie. On pourrait objecter bien sûr que le pourquoi de la quantification n'est toujours pas compris, puisque, à proprement parler, l'équation de Schrödinger ne se démontre pas, mais se pose à la suite d'un développement d'arguments heuristiques. Il y a toutefois une différence majeure avec le statut de la physique classique complétée par les conditions énigmatiques de Bohr, qui, en tant que telle, constituait une théorie hybride – pour tout dire bâtarde – remplie de contradictions (l'atome reste électrodynamiquement instable) : la nouvelle mécanique est, elle, cohérente et le problème de la stabilité de l'atome est parfaitement compris – pourvu qu'on l'aborde dans un cadre complètement quantique.

16.3 Le puits carré

L'énergie potentielle V figurant dans l'équation aux valeurs propres décrit le système physique considéré. Ainsi, dans le cas d'un atome hydrogénoïde, on a :

$$V(r) = -\frac{Ze'^2}{r}, \quad e'^2 = \frac{e^2}{4\pi\epsilon_0}. \quad (16.36)$$

où r est la distance entre l'électron de masse m et le noyau (supposé infiniment massif) situé à l'origine des coordonnées. On sait résoudre l'équation aux valeurs propres correspondante (voir cours de Maîtrise) et obtenir à la fois les fonctions d'onde et les valeurs de l'énergie. Pour les états liés, on trouve des valeurs propres quantifiées et données par :

$$E_n = -\frac{mZ^2e'^4}{2n^2\hbar^2}, \quad n \in \mathbb{N}^*. \quad (16.37)$$

Ce résultat coïncide d'ailleurs très exactement avec celui issu de l'Ancienne Théorie des Quanta de Bohr - Wilson - Sommerfeld.

La caractéristique essentielle du potentiel créé par le noyau de l'atome d'hydrogène est d'être attractif, ce qui signifie, l'origine des potentiels étant prise à l'infini, que l'énergie potentielle est négative $\forall r$ fini. Une telle

²¹Au contraire d'une particule α venant de l'infini et y retournant après déviation par un noyau.

situation peut être modélisée de façon beaucoup plus simple : d'une part on se met à une dimension d'espace, d'autre part on prend un potentiel ayant fondamentalement la même allure ; bien sûr, le spectre dépendra dans ses détails de la forme précise du potentiel étudié, mais on peut espérer que l'essentiel des phénomènes observés dans un cas simple sera cependant exemplaire, et pourra même servir de guide pour l'analyse ultérieure de cas plus complexes.

16.3.1 Le puits fini

Soit un puits de potentiel symétrique, pour lequel $V(x)$ varie comme indiqué sur la fig. 16.8. Ceci modélise un centre attractif où l'énergie potentielle de la particule est soit nulle si la distance est supérieure à $a/2$, soit constante et négative si la particule est très près du centre attractif. C'est clairement un modèle d'atome, qui, pour être rustique, n'en est pas moins exemplaire. Les deux caractéristiques importantes du puits sont sa largeur a et sa profondeur V_0 . Ces deux grandeurs définissent une échelle de longueur et une échelle d'énergie.

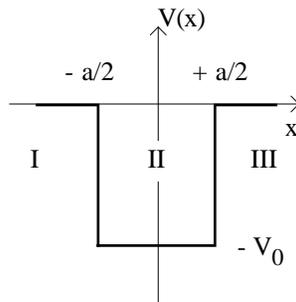


Figure 16.8: Puits carré à une dimension ($V_0 > 0$).

Une autre énergie est disponible : c'est celle que l'on peut fabriquer avec la constante de Planck, la distance a et la masse m de la particule :

$$E_{\text{loc}} = \frac{\hbar^2}{ma^2} . \quad (16.38)$$

On va voir que le nombre d'états liés dépend fortement du rapport E_{loc}/V_0 . Quand ce rapport est petit devant 1, le nombre d'états liés est élevé ; ceci correspond au cas où la masse m est grande et, bien sûr, ce rapport est nul à la limite classique ($\hbar = 0$) où il existe une infinité (non dénombrable) d'états liés. A l'inverse, si ce rapport est grand, le nombre d'états liés est faible²². Ainsi, plus la masse diminue plus il est "difficile" de localiser la particule quantique dans le puits²³, en la plaçant dans un état lié offert par ce dernier. En définitive, comme la relation $E_{\text{loc}} \sim V_0$ exprime la condition d'existence d'un unique état lié (voir détails plus loin), E_{loc} est appelée "énergie de localisation" de la particule.

Il n'est pas inutile de rappeler ce que donnerait la Mécanique Classique dans ce cas. Il convient de distinguer deux cas, selon que l'énergie de la particule, E , qui est une constante du mouvement, est positive ou négative :

- $E > 0$

En posant $E = mv_0^2/2$ ($v_0 > 0$), la conservation de l'énergie s'écrit :

$$\frac{1}{2}mv_0^2 = \frac{1}{2}mv^2 + V(x) \quad (16.39)$$

et les deux intervalles sont à considérer :

$$- x < -a/2 \text{ ou } x > +a/2$$

Dans ces intervalles, $V(x)$ étant nul, on a simplement $v = v_0$.

²²Il peut même ne pas exister du tout d'état lié si E_{loc}/V_0 est trop grand et s'il existe une barrière impénétrable en-dehors du puits, qui confine le mouvement quantique sur une demi-droite ; en l'absence d'une telle barrière, un puits de largeur finie a toujours, à une dimension, au moins un état lié (voir ci-dessous).

²³Comme d'habitude, les effets quantiques sont d'autant plus spectaculaires que la masse est petite (le mouvement d'une particule classique d'énergie totale négative sera toujours, au contraire, limité strictement au puits).

– $-a/2 < x < +a/2$ Alors :

$$\frac{1}{2}mv_0^2 = \frac{1}{2}mv^2 - V_0 \iff v = \pm\sqrt{v_0^2 + \frac{2}{m}V_0} \quad (16.40)$$

Ainsi, une particule de masse m venant de la gauche avec une vitesse constante v_0 , est instantanément accélérée par la chute de potentiel en $x = -a/2$ (ceci, par $F = -V'(x)$, correspond à une force impulsionnelle tirant vers la droite), retrouve une vitesse constante v donnée par (16.40), avant d'être instantanément freinée en $x = +a/2$ pour retrouver sa vitesse initiale. La force est nulle partout, sauf en $x = -a/2$ où elle est dirigée vers la droite (accélération instantanée) et dirigée vers la gauche en $x = +a/2$ (freinage instantané), pour un mouvement de gauche à droite (fig. 16.9, (a)).

• $E < 0$

L'énergie totale E reste évidemment bornée vers le bas par $-V_0$. Maintenant, on pose :

$$E = -\frac{1}{2}mv_0^2 . \quad (16.41)$$

et la conservation de l'énergie s'écrit :

$$-\frac{1}{2}mv_0^2 = \frac{1}{2}mv^2 + V(x) . \quad (16.42)$$

Ceci montre clairement que x ne peut pas prendre de valeurs en dehors de l'intervalle $[-a/2, +a/2]$, là où $V(x)$ est nul. Pour $-a/2 < x < +a/2$:

$$-\frac{1}{2}mv_0^2 = \frac{1}{2}mv^2 - V_0 \iff v = \pm\sqrt{-v_0^2 + \frac{2}{m}V_0} . \quad (16.43)$$

Maintenant, le mouvement est confiné entre $\pm a/2$ à vitesse constante en module (fig. 16.9, (b)). La particule rebondit élastiquement sur les murs, l'énergie totale étant forcément comprise entre $-V_0$ (particule immobile) et zéro (lorsque $v_0 = \sqrt{2V_0/m}$).

Cette distinction entre les états $E > 0$ (non-liés) et $E < 0$ (liés) va évidemment se transposer en Mécanique Quantique, avec toutefois des différences importantes. Cette classification en tout cas se retrouve dans l'analyse de la solution de l'équation aux valeurs propres $H\psi = E\psi$. Dans un but de clarté, on examine successivement ces deux cas.

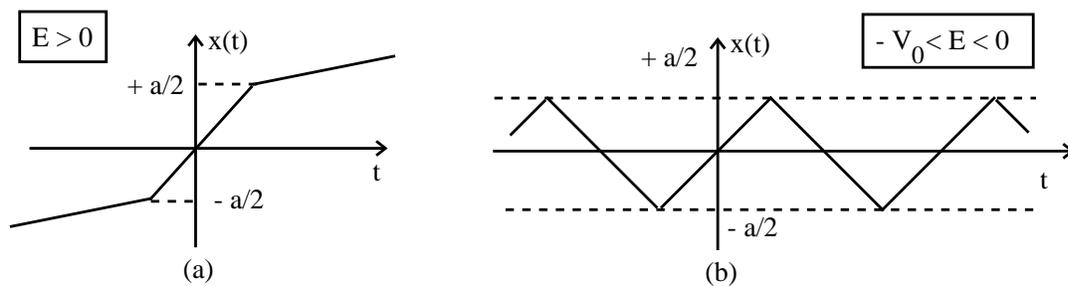


Figure 16.9: Mouvement d'une particule classique en présence d'un puits carré.

Etats liés ($E < 0$)

Dans la région I ($x < -a/2$, $V(x) = 0$), l'équation aux valeurs propres s'écrit simplement :

$$-\frac{\hbar^2}{2m}\psi''(x) = E\psi(x) . \quad (16.44)$$

Comme l'énergie E est négative, il existe une quantité réelle k , homogène à l'inverse d'une longueur et telle que :

$$E = -\frac{\hbar^2 k^2}{2m} , \quad (16.45)$$

ce qui permet de réécrire l'équation (16.44) sous la forme :

$$\psi''(x) - k^2 \psi(x) = 0 . \quad (16.46)$$

La solution de cette équation différentielle du second ordre est :

$$\psi(x) = A e^{kx} + B e^{-kx} . \quad (16.47)$$

où A et B sont deux constantes d'intégration. En fait, la constante B doit être nulle : compte tenu du sens physique attribué à la fonction d'onde, $\psi(x)$ ne peut diverger quand x tend vers $-\infty$. Il vient donc, pour $x < -a/2$:

$$\psi(x) = A e^{kx} . \quad (16.48)$$

De la même façon, quand $x > +a/2$ (région III), la seule solution physiquement acceptable de (16.46) est de la forme :

$$\psi(x) = C e^{-kx} . \quad (16.49)$$

où C est à nouveau une constante, pour l'instant arbitraire.

Quand x est compris entre $\pm a/2$ (région II), l'équation aux valeurs propres est :

$$-\frac{\hbar^2}{2m} \psi''(x) + (-V_0) \psi(x) = E \psi(x) . \quad (16.50)$$

La quantité $E + V_0$ ne peut être négative d'après (16.23). On peut donc introduire un nombre d'onde *réel* K tel que :

$$E + V_0 = \frac{\hbar^2 K^2}{2m} \iff k^2 + K^2 = \frac{2m}{\hbar^2} V_0 \equiv k_0^2 \iff E = \frac{\hbar^2}{2m} (K^2 - k_0^2) . \quad (16.51)$$

Les seules solutions acceptables seront paramétrées par K , que l'on peut d'ailleurs toujours prendre positif, par définition²⁴. De surcroît, avec $E < 0$:

$$\frac{\hbar^2 K^2}{2m} = E + V_0 < V_0 \iff K^2 < k_0^2 . \quad (16.52)$$

Dans ces notations, l'équation aux valeurs propres dans la région centrale (II) est :

$$\psi''(x) + K^2 \psi(x) = 0 . \quad (16.53)$$

Sa solution générale est :

$$\psi(x) = \alpha e^{iKx} + \beta e^{-iKx} . \quad (16.54)$$

A ce stade, on dispose donc des trois expressions suivantes pour la fonction propre générique, chaque expression étant associée à un intervalle donné pour x :

$$\psi(x) = \begin{cases} A e^{kx} & \text{si } x < -a/2 \\ \alpha e^{iKx} + \beta e^{-iKx} & \text{si } -a/2 < x < +a/2 \\ C e^{-kx} & \text{si } x > +a/2 \end{cases} . \quad (16.55)$$

Les quatre constantes A , C , α et β ne sont pas indépendantes les unes des autres : elles doivent être choisies de sorte que les trois expressions ci-dessus représentent une seule et même fonction dans différents intervalles ;

²⁴Indépendamment du contexte, K est de toute façon soit réel, soit imaginaire pur, puisque E et V_0 sont réels.

cette fonction est la fonction propre associée à une valeur propre²⁵ E . En conformité avec la discussion générale, $V(x)$ ayant un saut fini, ψ et ψ' sont partout continues²⁶ ; les coefficients α et β vont s'obtenir à partir de A et C en écrivant la continuité en $x = \pm a/2$ de la fonction propre et de sa dérivée²⁷. Au total, quatre conditions doivent être écrites, qui vont visiblement produire un système *homogène* pour les quatre constantes inconnues. En conséquence, il n'existera de solution non-triviale que si un certain déterminant est nul²⁸, soit si une certaine fonction de k et K (c'est-à-dire de l'énergie E) est nulle. Le problème en cours d'analyse fournit ainsi la première illustration explicite de la relation intime existant entre les conditions imposées à ψ et la quantification de l'énergie.

Écrivons maintenant explicitement les conditions de raccordement ; en $x = -a/2$:

$$\begin{cases} \text{continuité de } \psi \text{ en } -a/2 : & Ae^{-ka/2} = \alpha e^{-iKa/2} + \beta e^{iKa/2} \\ \text{continuité de } \psi' \text{ en } -a/2 : & kAe^{-ka/2} = iK\alpha e^{-iKa/2} - iK\beta e^{iKa/2} \end{cases} . \quad (16.56)$$

Considérant ceci comme un système pour α et β , sa solution est :

$$\alpha = \frac{1}{2} e^{iKa/2} e^{-ka/2} \left(1 - i \frac{k}{K} \right) A , \quad \beta = \frac{1}{2} e^{-iKa/2} e^{-ka/2} \left(1 + i \frac{k}{K} \right) A . \quad (16.57)$$

Le raccord en $x = +a/2$ s'écrit :

$$\begin{cases} \text{continuité de } \psi \text{ en } +a/2 : & Ce^{-ka/2} = \alpha e^{+iKa/2} + \beta e^{-iKa/2} \\ \text{continuité de } \psi' \text{ en } +a/2 : & -kCe^{-ka/2} = iK\alpha e^{+iKa/2} - iK\beta e^{-iKa/2} \end{cases} , \quad (16.58)$$

d'où, de la même façon :

$$\alpha = \frac{1}{2} e^{-iKa/2} e^{-ka/2} \left(1 + i \frac{k}{K} \right) C , \quad \beta = \frac{1}{2} e^{+iKa/2} e^{-ka/2} \left(1 - i \frac{k}{K} \right) C . \quad (16.59)$$

De (16.57) on déduit le rapport α/β , où A n'apparaît plus, et de même avec (16.59), C disparaissant à son tour. En égalant les deux expressions du rapport α/β ainsi obtenues, on trouve :

$$e^{+iKa} \frac{1 - i(k/K)}{1 + i(k/K)} = e^{-iKa} \frac{1 + i(k/K)}{1 - i(k/K)} \iff \left[\frac{(K/k) + i}{(K/k) - i} \right]^2 = e^{2iKa} . \quad (16.60)$$

Il est facile de voir que cette équation n'admet que des solutions réelles en K , conformément à l'affirmation ci-dessus : compte tenu de sa définition, et en-dehors de toute autre considération, K est soit réel, soit imaginaire pur ; dans ce dernier cas, le premier membre de (16.60) serait négatif, alors que le second serait positif. Par ailleurs, K et k dépendent de la valeur propre E : au total, (16.60) constitue une équation pour E . On va voir que cette équation a une suite *discrète* de solutions : l'énergie des états liés du puits carré est donc quantifiée.

En tout état de cause, il y a deux familles de solutions :

1. on extrait la racine carrée de (16.60) avec le signe + :

$$\frac{K + ik}{K - ik} = +e^{iKa} . \quad (16.61)$$

Cette équation se résout en K/k :

$$\frac{k}{K} = \cot \frac{Ka}{2} \iff \cos^2 \frac{Ka}{2} = \left(\frac{K}{k_0} \right)^2 . \quad (16.62)$$

²⁵L'opérateur parité Π commute avec H : il en résulte que si ψ est propre de H avec l'énergie E , la fonction transformée par Π , $\Pi\psi(x) = \psi(-x)$, est encore fonction propre avec la *même* énergie. Comme le spectre d'énergie (pour les états liés) n'est pas dégénéré, les deux fonctions ψ et $\Pi\psi$ sont proportionnelles ; en d'autres termes toute fonction propre de H est aussi fonction propre de Π et, de ce fait, est soit paire, soit impaire.

²⁶Les conditions qui en résultent sont maintenant les seules à prendre en compte puisque les constantes d'intégration correspondant aux termes divergents (B et D) ont été éliminées d'emblée afin de simplifier les calculs ultérieurs. On pourrait d'ailleurs ici adopter une autre présentation, suivant dans le détail la discussion générale donnée en 16.2 : il suffirait de garder D jusqu'au bout et de l'exprimer en fonction de A par raccord de proche en proche en $x = \pm a/2$, pour finalement obtenir explicitement la fonction notée $\Phi(E)$ dans la discussion générale ; l'équation donnant les valeurs propres prendrait alors la forme $\Phi(E) = 0$, comme indiqué alors.

²⁷On pourrait aussi écrire la continuité de la dérivée logarithmique ψ'/ψ .

²⁸Dans la suite, on n'introduira pas explicitement ce déterminant – qui joue un rôle analogue à $\Phi(E)$ –, en éliminant les constantes de proche en proche.

Comme $\tan(Ka/2)$ vaut ici k/K qui est positif, ceci est équivalent à :

$$\left| \cos \frac{Ka}{2} \right| = \frac{K}{k_0} \quad \text{et} \quad \tan \frac{Ka}{2} > 0 . \quad (16.63)$$

Ces solutions s'obtiennent donc en coupant la cosinusoïde rectifiée avec la droite K/k_0 et en ne conservant que les intersections pour lesquelles $Ka/2$ est compris entre 0 et $\pi/2$ (modulo π), c'est-à-dire Ka compris entre $2n\pi$ et $(2n+1)\pi$.

2. On extrait la racine carrée de (16.60) avec le signe $-$:

$$\frac{K+ik}{K-ik} = -e^{iKa} . \quad (16.64)$$

Un calcul analogue conduit maintenant à $K/k = -\tan(Ka/2)$ puis à :

$$\left| \sin \frac{Ka}{2} \right| = \frac{K}{k_0} \quad \text{et} \quad \tan \frac{Ka}{2} < 0 . \quad (16.65)$$

Il faut donc maintenant couper la sinussoïde rectifiée par la même droite que précédemment, mais en ne gardant ici que les intersections telles que Ka soit compris entre $(2n+1)\pi$ et $(2n+2)\pi$.

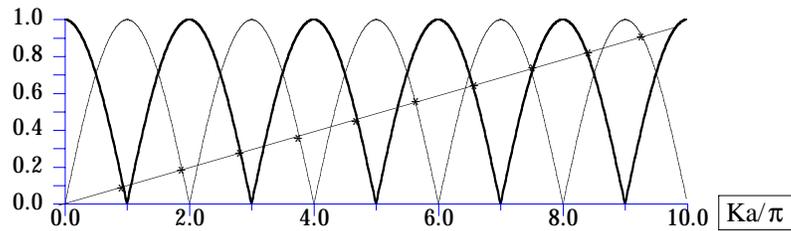


Figure 16.10: Solution graphique donnant les valeurs de Ka/π pour le puits carré.

En définitive, une arche étant donnée (sur une sinussoïde ou une cosinussoïde), l'intersection à retenir doit se trouver sur l'arc *descendant*, situé à droite du maximum. On obtient ainsi une suite de points d'abscisses K_n , fixant, par (16.51), les valeurs possibles de l'énergie, E_n . Au total, imposer la continuité de ψ et de ψ' produit la quantification de l'énergie :

$$E_n = -V_0 + \frac{\hbar^2 K_n^2}{2m} . \quad (16.66)$$

K (et k) ne peut donc prendre ses valeurs que dans une suite discrète bien définie. K décrit les oscillations de la fonction propre à l'intérieur du puits. La longueur d'onde correspondante, $\lambda_n = 2\pi/K_n$, est d'autant plus petite que l'énergie de l'état lié est grande (moins l'état est lié, plus la fonction propre oscille – un état non-lié oscille partout). Par ailleurs, k^{-1} définit l'échelle de longueur caractéristique de la décroissance de la fonction propre à l'extérieur du puits (voir (16.48) et (16.49)) ; cette longueur est d'autant plus grande que l'état est excité. En d'autres termes, la fonction propre de l'état fondamental “oscille peu” (en fait, pas du tout, au sens où elle ne s'annule pas à distance finie) et décroît vite à l'extérieur du puits, un état faiblement lié (E un peu en-dessous de V_0) oscille beaucoup et décroît lentement pour $|x| > a/2$. Par ailleurs, le nombre de solutions est piloté par la valeur de k_0 , défini en (16.51), de sorte que l'équation de la droite coupant les co- et les sinussoïdes rectifiées est :

$$\frac{\hbar}{\sqrt{2mV_0}} K \equiv \frac{\pi\hbar}{\sqrt{2ma^2V_0}} \frac{Ka}{\pi} . \quad (16.67)$$

Plus V_0 est grand, plus la pente est faible et plus grand est le nombre d'états liés ; il en va de même, à V_0 donné, si la masse de la particule augmente : le puits offre bien plus d'états liés à une particule massive qu'à une particule légère. Ceci est relié au fait que la quantité \hbar^2/ma^2 mesure l'énergie de localisation de la particule : pour une particule légère, cette énergie est grande et la particule “a du mal” à se localiser dans un puits peu profond. Alors que le piégeage dans un puits est toujours possible classiquement, il devient difficile pour une particule de petite masse : ici encore, les effets quantiques sont essentiels lorsque la masse est faible. Pour une

particule très massive, les niveaux sont très serrés (il y en a beaucoup à l'échelle V_0), c'est le contraire si la masse est faible. L'idée de quantification est bien d'autant plus manifeste que la masse est petite.

Pour une masse donnée, la droite se redresse quand V_0 décroît (puits de moins en moins profond) ; la première intersection avec la sinusöïde se rapproche par en dessus de π/a . En dessous de cette valeur-seuil pour V_0 , V_{0s} , il ne subsiste plus qu'un seul et unique état, résultant de la seule intersection acceptable entre la droite et la première arche de la fonction $\cos(Ka/2)$. La valeur-seuil est telle que $(k_0^{-1})(\pi/a) = 1$, soit $ak_0 = \pi$, soit :

$$V_{0s} = \frac{\pi^2 \hbar^2}{2ma^2} \equiv \frac{\pi^2}{2} E_{\text{loc}} \sim E_{\text{loc}} . \quad (16.68)$$

Un puits carré à une dimension²⁹ de largeur finie a donc toujours au moins un état lié. Visiblement, les autres solutions qui apparaissent les unes après les autres quand V_0 augmente au-delà de V_{0s} restent en nombre fini tant que V_0 est fini. Il est en outre évident géométriquement que si $k_0 a$ est très grand devant 1, le nombre d'états liés est lui-même très élevé. Ceci correspond aux valeurs de $V_0 \gg E_{\text{loc}}$:

$$V_0 \gg E_{\text{loc}} \iff \text{beaucoup d'états liés} . \quad (16.69)$$

L'ordre de grandeur du nombre d'états liés, N_b , est donné par $(V_0/E_{\text{loc}})^{1/2} = k_0 a$. Si $N_b \gg 1$, les premiers états ($1 \leq n \ll N_b \sim k_0 a$) sont presque équidistants en k , puisque la droite est presque horizontale ; les valeurs correspondantes de K , K_n , sont très voisines de $n\pi/a$ où n est un entier strictement positif :

$$E_n \simeq -V_0 + \frac{n^2 \pi^2 \hbar^2}{2ma^2} \quad (1 \leq n \ll N_b, N_b \sim k_0 a \gg 1) . \quad (16.70)$$

Les valeurs possibles de l'énergie étant maintenant trouvées, il est possible d'écrire explicitement la fonction d'onde pour un état n donné. D'après (16.57) (ou (16.59)), le rapport α/β est donné par :

$$\frac{\alpha}{\beta} = \frac{K - ik}{K + ik} e^{iKa} . \quad (16.71)$$

Pour l'état fondamental (celui de plus basse énergie, $n = 1$) comme pour tous ceux qui résultent d'une intersection avec la cosinusöïde rectifiée (n impair), la valeur propre s'obtient par (16.61) ; il en résulte que le rapport α/β vaut simplement 1. Pour cette famille de solutions, $\alpha = \beta$ c'est-à-dire que la fonction d'onde, dans la région centrale, est paire et est donnée par :

$$\psi_n(x) = 2\alpha \cos K_n x \quad (-a/2 \leq x \leq a/2, n \text{ impair}) . \quad (16.72)$$

Par le jeu des conditions de raccordement, cette symétrie de parité s'étend en-dehors de l'intervalle $\pm a/2$ et la fonction d'onde est globalement symétrique. En effet, on a :

$$A = 2\alpha e^{k_n a/2} \cos \frac{K_n a}{2} , \quad C = 2\alpha e^{-k_n a/2} \cos \frac{K_n a}{2} . \quad (16.73)$$

Au total :

$$\psi_n(x) = \begin{cases} 2\alpha \cos \frac{K_n a}{2} e^{k_n [x+(a/2)]} & \text{si } x < -a/2 \\ 2\alpha \cos K_n x & \text{si } -a/2 < x < +a/2 \\ 2\alpha \cos \frac{K_n a}{2} e^{-k_n [x-(a/2)]} & \text{si } x > +a/2 \end{cases} . \quad (16.74)$$

Par construction, ψ_n et sa dérivée sont continues partout : les différents arcs se raboutent donc avec une tangente commune. En revanche, la dérivée seconde a un saut – ce qui ne se voit pas très facilement à l'œil (variation du rayon de courbure et/ou changement de signe de ψ'' sans passer par zéro).

Pour l'autre famille de solutions (n pair), K est donné par (16.64), de sorte que le rapport α/β vaut maintenant -1 et la fonction est impaire dans la région II ; à nouveau, les conditions de raccordement exportent cette symétrie dans les régions I et III. Dans la région centrale :

$$\psi_n(x) = 2\alpha \sin K_n x \quad (-a/2 \leq x \leq a/2, n \text{ pair}) . \quad (16.75)$$

²⁹Ceci n'est plus vrai dès \mathbb{R}^2 . Physiquement : la particule dans le plan (et *a fortiori* en dimension supérieure) peut "tourner" autour du puits sans y tomber. Alors, pour qu'il existe des états liés, il faut que le puits soit assez profond et/ou assez étendu.

Ainsi, est obtenue une succession alternative d'états de parité bien déterminée, l'état fondamental ($n = 1$) étant symétrique (pair), le premier excité ($n = 2$) étant antisymétrique (impair), le second excité ($n = 3$) pair, etc. (on peut d'ailleurs démontrer en toute généralité que l'état fondamental est caractérisé par une fonction d'onde ne s'annulant jamais, sauf aux points formant la frontière avec les régions de l'espace où la particule ne peut se trouver³⁰). Le nombre de noeuds (zéros) des états propres augmente d'une unité quand on passe d'une valeur propre à la valeur propre immédiatement supérieure.

Il n'est pas surprenant que l'on puisse caractériser les états propres par leur symétrie : le Hamiltonien est invariant dans la symétrie par rapport à l'origine et par conséquent commute avec l'opérateur correspondant (appelé parité, noté Π) ; il est donc possible de trouver un ensemble de vecteurs propres communs à H et Π . Toute opération de symétrie est par nature incapable de changer l'énergie ; comme il s'agit ici d'un mouvement lié à une dimension (associé en tant que tel à un spectre non dégénéré), toute fonction propre de H est *ipso facto* fonction propre de Π .

Reste enfin à trouver la constante α ; elle va être fixée en écrivant la condition de normalisation. En définitive, une fois l'énergie fixée (et cette normalisation achevée), il existe bien une et une seule fonction propre (en conformité avec la discussion générale, le spectre des états liés ne présente pas de dégénérescence).

Effectuons explicitement la normalisation des fonctions propres. Comme le module au carré de ψ est pair $\forall n$, il suffit d'écrire :

$$2 \int_0^{+\infty} dx |\psi_n(x)|^2 = 1 . \quad (16.76)$$

Un calcul sans difficulté à partir de (16.74) donne, pour les états pairs (n impair) :

$$2\alpha^2 \left(a + \frac{2}{k_n} \right) = 1 \iff 2\alpha = \sqrt{\frac{2}{a + 2(k_0^2 - K_n^2)^{-1/2}}} . \quad (16.77)$$

Ainsi, la valeur de la fonction propre ψ_n d'un état pair (n impair) au point $x = \pm a/2$ est égale à :

$$|\psi_n(\pm a/2)| = \sqrt{\frac{2}{a + 2(k_0^2 - K_n^2)^{-1/2}}} \cos \frac{K_n a}{2} = \sqrt{\frac{2}{a + 2(k_0^2 - K_n^2)^{-1/2}}} \frac{K_n}{k_0} . \quad (16.78)$$

Le signe à choisir provient de l'égalité $|\cos K_n a/2| = K_n/k_0$. Pour un état profond, K_n est très petit devant k_0 et donc :

$$|\psi_n(\pm a/2)| \simeq \sqrt{\frac{k_0}{1 + (ak_0/2)}} \frac{K_n}{k_0} \ll 1 . \quad (16.79)$$

Pour un état profond, la fonction propre est donc très petite sur les limites du puits. Pour l'état le moins lié, la différence $k_0 - K_{N_b}$ est très petite devant $1/a$ (elle peut même être nulle si V_0 est bien ajusté) ; le terme $\cos Ka/2$ est très voisin de 1 en module et il reste :

$$|\psi_n(\pm a/2)| \simeq \sqrt{\sqrt{k_0^2 - K_n^2} \frac{K_n}{k_0}} \simeq (k_0^2 - K_n^2)^{1/4} \quad (K_n \lesssim k_0) . \quad (16.80)$$

Au total, la valeur aux bornes du module de la fonction varie comme schématisé sur la fig. 16.11. Pour les états très profonds en énergie, $|\psi|$ est petit en $x = \pm a/2$; on verra par la suite que, dans le cas du puits infiniment profond, ψ s'annule strictement aux bornes, quel que soit l'état propre considéré.

Pour l'état le moins lié ($n = N_b$, K_n juste en dessous de k_0), ka est très petit devant 1 : à l'extérieur du puits, ψ est donc lentement décroissante à l'échelle a . Comme la norme au total vaut l'unité et que sa plus grande fraction vient de l'extérieur du puits, ψ prend des valeurs faibles à l'intérieur de celui-ci – en particulier aux bornes. On sent bien alors que l'état très peu lié est très diffus et que la fonction propre correspondante est très large : c'est bien l'image que l'on peut se faire d'un état à faible énergie de liaison – tout comme, dans

³⁰Ici (parce que le puits est de profondeur finie), ces points sont à l'infini : la fonction propre fondamentale ne s'annule qu'à l'infini.



Figure 16.11: Variation de $|\psi(\pm a/2)|$ en fonction du nombre quantique n .

l'atome d'hydrogène, les états très excités correspondent à des électrons très délocalisés³¹. Un tel état est en quelque sorte la métamorphose (V_0 augmentant, toutes choses égales par ailleurs) d'un état non-lié de petite énergie (positive), devenu faiblement lié parce que k_0 vient de franchir une valeur-seuil $2n\pi/a$, donnant ainsi naissance à un nouvel état très peu lié, nécessairement voisin de ce qu'il était avant de se lier, au moins sur une échelle de longueur d'ordre a .

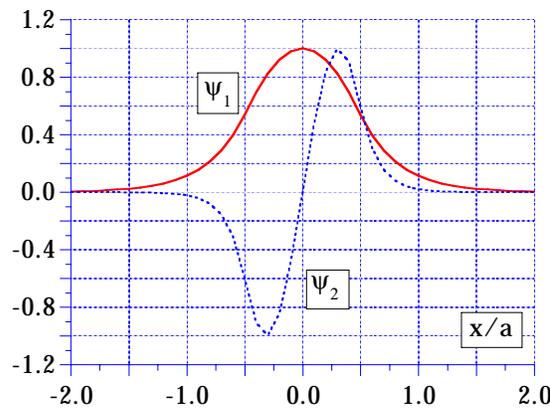


Figure 16.12: Fonctions d'onde (non normalisées) de l'état fondamental et du premier état excité du puits carré.

Les deux fonctions d'onde associées à l'état fondamental et au premier état excité sont tracées sur la fig. 16.12 (les fonctions représentées ne sont pas normalisées). En regardant bien, on voit le changement de signe de ψ_1'' (et de ψ_2'') en $x = \pm a/2$ (la phase est choisie de sorte que ψ_1 soit partout positive) ; à l'intérieur du puits, la dérivée seconde est négative puisque $E > V(x) = -V_0$, elle est positive à l'extérieur puisque $E < 0 = V(x)$. Il y a donc une inflexion de $\psi_1(x)$ aux frontières du puits mais ce n'est pas une inflexion ordinaire : la courbure change de signe de façon discontinue, puisque c'est la dérivée seconde à elle seule qui encaisse de saut de potentiel.

À ce stade, une remarque s'impose : la fonction d'onde est non-nulle partout et, en particulier à l'extérieur du puits. Physiquement, ceci signifie que, bien que la particule ait une énergie négative (*i. e.* inférieure au bord supérieur du puits), il existe une probabilité finie, non-nulle, de la trouver en dehors du puits. La particule classique (d'énergie totale négative) rebondit élastiquement sur les murs du potentiel de confinement (elle ne peut en sortir, l'énergie cinétique serait négative !) ; la particule quantique, elle, peut "s'égarer" à l'extérieur du puits, en dépit d'une énergie apparemment insuffisante, sans pouvoir toutefois aller très loin puisque la fonction d'onde décroît exponentiellement³². Cet effet, spécifiquement quantique et d'autant plus marqué que la masse est petite, s'appelle effet-tunnel et fera l'objet d'une discussion approfondie dans un cas où il apparaît encore plus spectaculaire (franchissement d'une barrière, donnant lieu à un courant non nul).

La possibilité de ces excursions interdites classiquement se mesure par la probabilité P de trouver la

³¹Les atomes dits de Rydberg (dont le prototype est l'atome d'hydrogène dans un état très excité) ont une dimension gigantesque, qui peut atteindre 1000 Å (on se souvient que le rayon des orbites de Bohr croît comme n^2).

³²Toutefois, le courant de probabilité est nul partout, en dedans comme en dehors du puits (simplement parce que la fonction d'onde est réelle, ou plutôt, peut toujours être prise réelle, comme toujours s'agissant d'un état lié à une dimension). Bien que la particule ait la possibilité d'explorer l'extérieur du puits, ces excursions doivent être vues comme des allers-retours donnant finalement un courant nul. C'est la généralisation au domaine quantique de la notion de rebond élastique sur un mur : la particule quantique "s'enfonce" un peu dans le mur avant de repartir en arrière.

particule hors du puits ; pour un état normalisé à l'unité, elle est donnée par :

$$P = 2 \int_{a/2}^{+\infty} |\psi_n(x)|^2 dx . \quad (16.81)$$

Pour un état pair, en revenant au calcul de la constante de normalisation α , on voit que :

$$P = \frac{2}{2 + ak_n} \cos^2 \frac{K_n a}{2} = \frac{2}{2 + a\sqrt{k_0^2 - K_n^2}} \frac{K_n^2}{k_0^2} . \quad (16.82)$$

P est visiblement une fonction croissante de K_n . Pour un état très profond ($K_n \ll k_0$) :

$$P \simeq \frac{2}{2 + ak_0} \frac{K_n^2}{k_0^2} \ll 1 \quad (\text{état profond}) . \quad (16.83)$$

Au contraire, pour l'état le moins lié (K_n juste en dessous de k_0) :

$$P \simeq \frac{2}{2 + ak_0} \quad (\text{état très peu lié}) . \quad (16.84)$$

qui est un nombre voisin de 1, sauf si $ak_0 \gg 1$: si donc cette inégalité est respectée, la particule a pratiquement autant de chance d'être à l'extérieur ou à l'intérieur du puits. Entre cet état et l'état non-lié de plus basse énergie (E juste positive), il n'y a en réalité pas grande différence, en dépit des apparences formelles. Le premier décroît certes exponentiellement, mais l'échelle de longueur k_n^{-1} est très grande puisque la différence $k_0^2 - K_n^2$ est petite ; le second, comme on va le voir, a une variation oscillatoire mais les oscillations ont une très grande longueur d'onde puisque ce qui va maintenant jouer le rôle symétrique de k est très petit. Notons enfin que, toutes choses égales par ailleurs, P est une fonction décroissante de k_0 , donc une fonction décroissante de la masse : plus la masse est faible, plus l'effet quantique est important. Ceci peut d'ailleurs se voir explicitement en considérant l'état fondamental, le seul qui existe quelle que soit la masse. Le bon paramètre sans dimension est ici $k_0 a$ avec :

$$k_0 a \ll 1 \iff \text{masse faible} , \quad k_0 a \gg 1 \iff \text{masse élevée} . \quad (16.85)$$

Lorsque $k_0 a \ll 1$, la droite coupant la cosinusoïde est presque verticale de sorte que $K_1 a \ll 1$ et l'équation aux valeurs propres pour K_1 est approximativement :

$$1 - \frac{1}{8} K_1^2 a^2 \simeq \frac{K_1}{k_0} , \quad (16.86)$$

et comme l'abscisse de l'intersection est très voisine de $k_0 a$, on a :

$$K_1 a \simeq 1 - \frac{1}{8} k_0^2 a^2 \iff k_1 \simeq \frac{1}{2} k_0^2 a . \quad (16.87)$$

Dans ces conditions :

$$P \simeq 1 - \frac{1}{2} k_0^2 a^2 ; \quad (16.88)$$

la particule légère ($k_0 a \ll 1$), bien que liée dans le puits, est majoritairement à l'extérieur du puits. Au contraire, si $k_0 a \gg 1$, alors $K_1 \simeq \pi/a \ll k_0$ et $P \ll 1$.

Remarquons enfin que, comme les fonctions propres peuvent toujours être prises réelles, le courant est identiquement nul dans tout état lié. Ceci est spécifique de la dimension 1 d'espace³³.

Etats non-liés ($E > 0$)

Dans la région I ($x \leq -a/2$), $V(x) = 0$ et l'équation aux valeurs propres est encore (16.44), avec maintenant $E > 0$. En conséquence, et pour garder k réel (positif), on pose à partir d'ici :

$$E = \frac{\hbar^2 k^2}{2m} \quad (k > 0) , \quad (16.89)$$

³³Pour l'atome d'hydrogène, dans un état propre de la composante J_z , il existe un courant circulaire non-nul autour de l'axe Oz .

ce qui permet de réécrire (16.44) sous la forme :

$$\psi''(x) + k^2\psi(x) = 0 . \quad (16.90)$$

La solution de cette équation différentielle du second ordre est évidemment :

$$\psi(x) = A e^{ikx} + B e^{-ikx} , \quad (16.91)$$

où A et B sont deux constantes d'intégration. Ces constantes seront déterminées ultérieurement quand on aura spécifié complètement la situation physique dont on désire rendre compte.

La fonction étant ainsi précisée dans la région I, il convient de poursuivre la résolution dans les deux régions II et III, et utiliser explicitement des conditions de raccordement assurant la continuité aux points de discontinuité du potentiel. Dans la région II ($-a/2 \leq x \leq +a/2$), l'équation est (16.50) ; en posant comme auparavant :

$$E + V_0 = \frac{\hbar^2 K^2}{2m} , \quad k_0^2 = \frac{2mV_0}{\hbar^2} , \quad (16.92)$$

on a maintenant (compte tenu de (16.89)) :

$$k^2 + k_0^2 = K^2 \quad (K \in \mathbb{R}_+) , \quad (16.93)$$

et l'équation aux valeurs propres est encore de la forme (16.53). Ses solutions sont du type :

$$\psi(x) = \alpha e^{iKx} + \beta e^{-iKx} . \quad (16.94)$$

Les coefficients α et β sont reliés aux constantes A et B ; en écrivant la continuité en $x = -a/2$, il vient :

$$\begin{cases} \text{continuité de } \psi \text{ en } -a/2 : & A e^{-ika/2} + B e^{ika/2} = \alpha e^{-iKa/2} + \beta e^{iKa/2} \\ \text{continuité de } \psi' \text{ en } -a/2 : & ikA e^{-ika/2} - ikB e^{ika/2} = iK\alpha e^{-iKa/2} - iK\beta e^{iKa/2} \end{cases} . \quad (16.95)$$

Ces deux équations fournissent un système linéaire pour les deux inconnues α et β que l'on peut donc exprimer en fonction de A et B . On trouve ainsi :

$$\alpha = \frac{1}{2} e^{iKa/2} \left[e^{-ika/2} \left(1 + \frac{k}{K} \right) A + e^{ika/2} \left(1 - \frac{k}{K} \right) B \right] , \quad (16.96)$$

$$\beta = \frac{1}{2} e^{-iKa/2} \left[e^{-ika/2} \left(1 - \frac{k}{K} \right) A + e^{ika/2} \left(1 + \frac{k}{K} \right) B \right] . \quad (16.97)$$

Enfin, dans la région III ($x \geq +a/2$), où $V(x)$ est à nouveau identiquement nul, la solution générale de l'équation aux valeurs propres est de la forme :

$$\psi(x) = C e^{ikx} + D e^{-ikx} . \quad (16.98)$$

De la même façon, les coefficients C et D se trouvent en écrivant les conditions de continuité en $x = +a/2$:

$$\begin{cases} \text{continuité de } \psi \text{ en } +a/2 : & \alpha e^{+iKa/2} + \beta e^{-iKa/2} = C e^{ika/2} + D e^{-ika/2} \\ \text{continuité de } \psi' \text{ en } +a/2 : & iK\alpha e^{iKa/2} - iK\beta e^{-iKa/2} = ikC e^{ika/2} - ikD e^{-ika/2} \end{cases} . \quad (16.99)$$

Ceci fournit un système linéaire pour C et D que l'on peut exprimer en fonction de α et β (maintenant connus en fonction de A et B) :

$$C = \frac{1}{2} e^{-ika/2} \left[e^{iKa/2} \left(1 + \frac{K}{k} \right) \alpha + e^{-iKa/2} \left(1 - \frac{K}{k} \right) \beta \right] , \quad (16.100)$$

$$D = \frac{1}{2} e^{ika/2} \left[e^{iKa/2} \left(1 - \frac{K}{k} \right) \alpha + e^{-iKa/2} \left(1 + \frac{K}{k} \right) \beta \right] . \quad (16.101)$$

Au total, il est possible d'exprimer C et D en fonction de A et B . En écrivant (16.96), (16.97), (16.100) et (16.101) sous forme matricielle et en effectuant le produit des deux matrices, il vient :

$$\begin{bmatrix} C \\ D \end{bmatrix} = \begin{bmatrix} e^{-ika} \left(\cos Ka + i \frac{K^2+k^2}{2kK} \sin Ka \right) & i \frac{K^2-k^2}{2kK} \sin Ka \\ -i \frac{K^2-k^2}{2kK} \sin Ka & e^{ika} \left(\cos Ka - i \frac{K^2+k^2}{2kK} \sin Ka \right) \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix}. \quad (16.102)$$

Les constantes A et B , ou C et D , dépendent du problème choisi. Le vecteur densité de courant est évidemment dirigé suivant Ox . Compte tenu de (16.91), de (16.98) et de l'expression générale du courant (16.12), on trouve :

$$j(x < -a/2) = \frac{\hbar k}{m} [|A|^2 - |B|^2], \quad j(x > a/2) = \frac{\hbar k}{m} [|C|^2 - |D|^2]. \quad (16.103)$$

$\hbar k/m$ est une vitesse ; il en résulte que chaque module carré représente une densité. La conservation de la probabilité totale dans un état stationnaire impose au courant d'être constant dans l'espace ; on doit donc avoir en particulier³⁴ :

$$j(x < -a/2) = j(x > a/2) \iff \frac{\hbar k}{m} [|A|^2 - |B|^2] = \frac{\hbar k}{m} [|C|^2 - |D|^2], \quad (16.104)$$

ce qui se vérifie facilement. Ici, le vecteur d'onde est le même à gauche et à droite (puisque V prend la même valeur pour $x < -a/2$ et pour $x > +a/2$) ; après simplification, la conservation du courant donne ici la relation suivante entre les différentes constantes :

$$|A|^2 - |B|^2 = |C|^2 - |D|^2. \quad (16.105)$$

Pour décrire la situation où aucune particule n'est émise en $x = +\infty$ en direction du potentiel, il faut prendre $D = 0$, ce qui assure que, du côté $x > a/2$, le vecteur courant ne contient qu'une seule composante, proportionnelle à $+\hbar k/m$. Dans ces conditions, la source de particules est en $x = -\infty$ et $|A|^2(\hbar k/m)$ est le vecteur courant injecté par la source, j_{source} - qui est bien de la forme densité \times vitesse ; pour $x < -a/2$, j est la somme algébrique du courant vers la droite $j_{\text{source}} = |A|^2(\hbar k/m)$ et du courant vers la gauche $j_{\text{réfléchi}} = |B|^2(\hbar(-k)/m)$; pour $x > -a/2$, j se réduit au courant transmis $j_{\text{transmis}} = |C|^2(\hbar k/m)$. Ainsi, quand la source est en $-\infty$, la conservation du courant s'écrit, après simplification par $\hbar k/m$, la vitesse étant la même pour $x < -a/2$ et pour $x > a/2$:

$$j_{\text{source}} + j_{\text{réfléchi}} = j_{\text{transmis}} \iff |A|^2 - |B|^2 = |C|^2, \quad (16.106)$$

cependant que la fonction d'onde est alors plus précisément :

$$\psi(x < -a/2) = A e^{ikx} + B e^{-ikx}, \quad \psi(x > a/2) = C e^{ikx}. \quad (16.107)$$

On doit d'ores et déjà remarquer que, contrairement à ce qui se passerait en Mécanique Classique, il existe toujours une onde *réfléchie* (voir fig. 16.13).

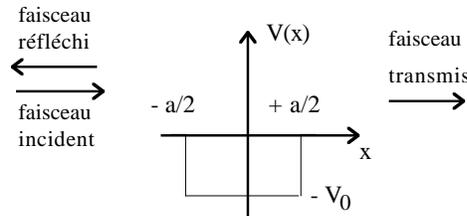


Figure 16.13: Représentation schématique des faisceaux incident, réfléchi et transmis par le puits carré lorsque la source est à l'infini à gauche.

Avec le choix $D = 0$, l'équation matricielle donne immédiatement le rapport B/A , puis le rapport C/A . Ces quantités permettent de définir un coefficient de réflexion R et un coefficient de transmission T , obtenus en formant les rapports appropriés des diverses composantes du courant total :

$$R = \frac{j_{\text{réfléchi}}}{j_{\text{source}}} = \frac{|B|^2}{|A|^2}, \quad T = \frac{j_{\text{transmis}}}{j_{\text{source}}} = \frac{|C|^2}{|A|^2}. \quad (16.108)$$

³⁴ $j(|x| < a/2)$ prend également la même valeur !

La somme $R + T$ vaut bien 1, eu égard à (16.106). En posant :

$$X = \left(\frac{K^2 - k^2}{2kK} \right)^2 \sin^2 Ka , \quad (16.109)$$

on a :

$$R = \frac{X}{1+X} , \quad T = \frac{1}{1+X} . \quad (16.110)$$

Dans les variables k et K , R et T s'écrivent :

$$R = \frac{(K^2 - k^2)^2 \sin^2 Ka}{(2kK)^2 + (K^2 - k^2)^2 \sin^2 Ka} , \quad T = \frac{(2kK)^2}{(2kK)^2 + (K^2 - k^2)^2 \sin^2 Ka} . \quad (16.111)$$

Ainsi, la présence du potentiel donne lieu, à gauche, à deux ondes $e^{\pm ikx}$ incidente et réfléchie, caractérisées par les amplitudes A et B , ainsi que par un certain déphasage δ . Au total, on peut écrire :

$$B = i\sqrt{R} e^{-ika} e^{i\delta} A , \quad (16.112)$$

auquel cas, selon (16.102) avec $D = 0$:

$$\cos \delta = \frac{\cos Ka}{\sqrt{\cos^2 Ka + \left(\frac{K^2 + k^2}{2kK} \right)^2 \sin^2 Ka}} , \quad \sin \delta = \frac{\frac{K^2 + k^2}{2kK} \sin Ka}{\sqrt{\cos^2 Ka + \left(\frac{K^2 + k^2}{2kK} \right)^2 \sin^2 Ka}} . \quad (16.113)$$

Dans ces notations, on a aussi :

$$C = \sqrt{T} e^{-ika} e^{i\delta} A . \quad (16.114)$$

R , T et δ dépendent de k (et sont paramétrés³⁵ par V_0), c'est-à-dire finalement de l'énergie de la particule incidente. En posant :

$$\varepsilon = \frac{E}{V_0} , \quad \mu = \frac{2ma^2 V_0}{\hbar^2} \equiv 2 \frac{V_0}{E_{\text{loc}}} \sim N_b^2 , \quad (16.115)$$

il vient :

$$X = \frac{\sin^2 \sqrt{\mu(1+\varepsilon)}}{4\varepsilon(1+\varepsilon)} , \quad \tan \delta = \frac{2\varepsilon + 1}{2\sqrt{\varepsilon(1+\varepsilon)}} \tan \sqrt{\mu(1+\varepsilon)} . \quad (16.116)$$

Naturellement, le coefficient de transmission tend vers 1 aux hautes énergies : quand la particule a une énergie très grande par rapport à la profondeur du puits, ce dernier est quasiment inopérant. De même, si $\mu \gg 1$ – ce qui correspond à la limite presque classique – T oscille extrêmement vite puisque X varie très vite (en gros une oscillation pour $\Delta(\mu\varepsilon)$ d'ordre π^2). Si T est “moyenné” sur une telle oscillation³⁶, on obtient une variation très rapide mais monotone, le T moyen passant de zéro à un sur un intervalle en énergie très petit ; corrélativement, le coefficient de réflexion est presque toujours très petit, conformément à ce que l'on attend de la limite presque classique. Il reste toutefois que les effets quantiques, même “faibles” dans cette limite, sont singuliers³⁷ : des oscillations brutales apparaissent, donnant à T une variation presque partout très différente de la fonction échelon prévue par la Mécanique Classique (ou encore : R , identiquement nul si $h = 0$, présente des accidents spectaculaires près de $\varepsilon = 0$ même dans la situation presque classique).

De fait, T (et R) ont une variation non-triviale par rapport à l'énergie. A cause du sinus au dénominateur par X , T présente des oscillations en énergie et résonne à la valeur 1 à chaque fois que X est nul, soit pour les valeurs de l'énergie telles que :

$$\sqrt{\mu(1+\varepsilon)} = \nu\pi \iff E_\nu = \nu^2 \frac{\pi^2 \hbar^2}{2ma^2} - V_0 \iff K_\nu = \nu \frac{\pi}{a} \quad (\nu > \pi^{-1} \sqrt{\mu(1+\varepsilon)}) . \quad (16.117)$$

³⁵si $V_0 = 0$ (pas de puits), alors $K = k$ et donc $R = 0$, $T = 1$ et $\delta = 0$, comme il se doit.

³⁶Ce qui sera de fait observé si la résolution en énergie est médiocre comparée à la largeur d'une oscillation.

³⁷L'extrême singularité des corrections quantiques a déjà été mentionnée : il suffit de se souvenir de l'expression du propagateur, qui contient toujours un terme du genre $e^{iS/\hbar}$ (la fonction $e^{1/z}$ a une singularité essentielle en $z = 0$).

Ces valeurs de résonance sont données exactement par la même formule que les énergies des premiers états liés quand le puits est très profond (voir (16.70)), mais l'énergie E est au total positive. En quelque sorte, les énergies donnant lieu à résonance sont le "prolongement analytique" (du côté $E > 0$) de l'expression de l'énergie obtenue pour les états liés. En outre, il s'agit visiblement d'un effet d'interférences, comme en atteste la relation $K_\nu = \nu\pi/a$, entre la face avant et la face arrière du puits. Ces résonances sont dites de basse énergie puisqu'elles sont manifestes seulement dans le cas où l'énergie de la particule n'est pas trop grande par rapport à V_0 (si l'énergie E est trop grande devant V_0 , T vaut pratiquement 1). On se doute – même si le modèle utilisé est un peu rustique – qu'il y a là le principe d'une investigation des propriétés énergétiques d'une cible donnée : en faisant diffuser des projectiles et en analysant ces résonances de basse énergie, on peut obtenir des renseignements très précis sur la nature de la cible (et de ses états liés). Par ailleurs, à cause du facteur μ , ces oscillations sont rapides et fines si μ est grand³⁸, c'est-à-dire si V_0 est grand par rapport à l'énergie de localisation E_{loc} ; c'est le cas pour un V_0 donné lorsque la masse de la particule devient assez grande. D'un autre côté, si μ est d'ordre 1, la première oscillation se produit lorsque ε est d'ordre π^2 , mais elle est invisible car alors le facteur $1/[\varepsilon(1+\varepsilon)]$ est d'ordre 10^{-2} et T a pratiquement atteint sa valeur finale égale à 1 – à $(10^{-2})^2$ près.

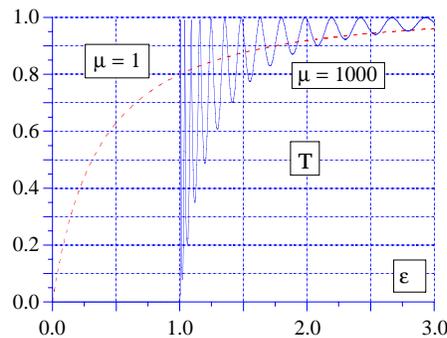


Figure 16.14: Variation du coefficient de transmission T d'un puits carré en fonction de l'énergie et pour deux valeurs de μ (voir (16.115)).

En définitive, la visibilité des résonances de basse énergie reflète l'existence d'un grand nombre d'états liés pour le potentiel dont on analyse, ici, les états stationnaires de diffusion (masse élevée, énergie de localisation faible, grand nombre d'états liés). Ces effets sont importants à basse énergie (ε nettement plus petit que 1, résonances de basse énergie). Enfin, on note que, toujours à basse énergie, T est très petit si μ est petit (en particulier : masse faible) : une particule très légère présente un comportement "hyper-quantique", retournant majoritairement en arrière.

La fonction d'onde est toujours définie à un facteur près (l'équation aux valeurs propres est homogène). À gauche de $x = -a/2$, elle peut s'écrire :

$$\psi(x < -a/2) = A \left[e^{ikx} + i\sqrt{R} e^{i\delta(k)} e^{-ik(x+a)} \right] , \quad (16.118)$$

où le déphasage δ est une fonction de k (voir (16.113)), c'est-à-dire de l'énergie incidente de la particule, et des paramètres a et V_0 définissant ici complètement le potentiel (voir fig. 16.15). De même :

$$\psi(x > a/2) = A\sqrt{T} e^{i\delta(k)} e^{ik(x-a)} . \quad (16.119)$$

On conçoit aisément que l'analyse expérimentale des faisceaux transmis ou réfléchis (autant par leurs intensités que par le déphasage) puisse fournir des renseignements sur le potentiel. C'est en fait le principe même de toute expérience de diffusion, expérience qui permet, moyennant l'étude des particules diffusées, de remonter jusqu'aux caractéristiques du potentiel diffuseur.

Notons qu'ici il existe une dégénérescence d'ordre deux ; en effet, si la source est située en $x = +\infty$, alors la constante D est différente de zéro et c'est A qui est nul. Au total, on obtient exactement les mêmes résultats où k (qui est positif par définition) est changé en $-k$ dans toutes les formules. Dans une telle opération, l'énergie ne change pas. Ce que l'on observe ici n'est rien d'autre que l'invariance par renversement du temps : il est en effet clair que changer k en $-k$ revient à prendre les complexes conjuguées des fonctions propres ainsi trouvées

³⁸Noter cependant la dépendance en $\mu^{1/2}$: il faut choisir des μ très grands pour que les effets décrits soient bien visibles.

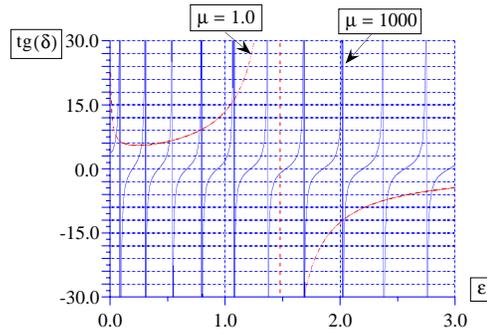


Figure 16.15: Variation du déphasage δ en fonction de k pour deux valeurs du paramètre μ (voir (16.115)).

(le déphasage δ est bien, tout comme k , une fonction impaire de k). Dans le cas des états liés, les fonctions d'onde peuvent être prises réelles et cette dégénérescence n'apparaît pas.

Remarque

La matrice \mathbf{M} reliant A et B à C et D (ou inversement) possède des propriétés spécifiques liées à la conservation du courant de part et d'autre du puits potentiel, conservation qui s'exprime par la relation (16.104), laquelle se lit aussi :

$$|A|^2 + |D|^2 = |B|^2 + |C|^2 . \quad (16.120)$$

La vitesse de la particule étant la même à gauche et à droite du puits, les coefficients $|A|^2$ et $|C|^2$ (resp. $|D|^2$ et $|B|^2$) mesurent à eux seuls le courant de probabilité vers la droite (resp. vers la gauche). À partir de cette relation, on peut montrer que la matrice \mathbf{M} s'exprime commodément à l'aide d'une matrice S unitaire (voir [7], p. 96 et sq.).

16.3.2 Le puits infini

Un cas particulier intéressant est celui où le puits de potentiel a une profondeur infinie ($V_0 = +\infty$, voir fig. 16.16) ; dans ce cas, les résultats peuvent être très simplement formulés. En redéfinissant l'origine des potentiels, on peut prendre :

$$V(x) = \begin{cases} 0 & \text{si } |x| < a/2 \\ +\infty & \text{si } |x| > a/2 \end{cases} . \quad (16.121)$$

Les résultats de ce puits peuvent être obtenus en prenant la limite $V_0 \rightarrow +\infty$ dans les formules ci-dessus : il suffit de raisonner avec la quantité $E + V_0$ puisque l'origine des énergies est maintenant choisie au fond du puits.

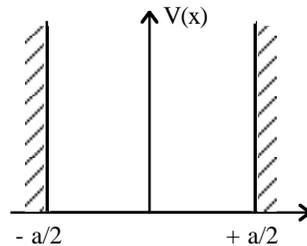


Figure 16.16: Puits de potentiel infini

Comme le potentiel présente un saut infiniment grand, les conclusions de la discussion générale concernant la continuité de ψ et de ψ' ne peuvent plus être invoquées ; à l'inverse, rien n'empêche de les utiliser tant que la limite V_0 infini n'est pas prise. Toutefois, une chose est claire : comme il faudrait fournir une énergie infinie pour extraire la particule du puits, celle-ci est strictement localisée entre les deux points d'abscisses $\pm a/2$; en d'autres termes, la probabilité de trouver la particule en dehors du puits est maintenant *nulle* ; cette probabilité, donnée par le module carré de $\psi(x)$, ne peut être nulle que si la fonction d'onde est nulle : quand le puits est infiniment profond, $\psi(x) = 0 \forall |x| > a/2$. De surcroît, le puits infini ne peut évidemment pas avoir d'états

non-liés : *toutes* les énergies possibles seront donc quantifiées. C'est bien ce que va produire le processus de limite.

Après le changement d'origine des énergies pour le puits fini de profondeur V_0 (on prend maintenant l'origine au fond du puits, soit $(E_n + V_0)_{(16.66)} \rightarrow E_n$), les énergies propres sont données par (voir (16.66)) :

$$E_n = \frac{\hbar^2 K_n^2}{2m} , \quad (16.122)$$

où les K_n sont fixés par les mêmes équations que ci-dessus et sont donc obtenus en intersectant la droite K/k_0 avec les sinusoides et cosinusoides rectifiés ; il s'agit toujours des états liés : après changement d'origine des énergies, ceci suppose que les E_n sont strictement inférieures à $+V_0$, qui est pour l'instant fini.

Quand $V_0 \rightarrow +\infty$, il en va de même pour k_0 et la pente de la droite tend vers zéro. Il en résulte que, à la limite, les K_n sont simplement donnés par :

$$K_n = n \frac{\pi}{a} \iff a = n \frac{\lambda_n}{2} . \quad (16.123)$$

Les modes stationnaires correspondent donc aux cas où la dimension du puits est égale à un nombre entier de fois la demi-longueur d'onde (comme une corde vibrante dont les extrémités sont fixées). Les énergies propres du puits infiniment profond sont donc égales à :

$$E_n = \frac{n^2 \pi^2 \hbar^2}{2ma^2} \quad (n \in \mathbb{N}^*) . \quad (16.124)$$

Maintenant que V_0 est infini, l'inégalité $E < V_0$ n'est plus restrictive : l'entier n est donc quelconque (mais strictement positif).

Il est facile de voir que toutes les fonctions propres sont nulles aux bornes ; comme toute fonction d'onde est nulle en dehors du puits, on trouve que malgré le saut infini de potentiel, ψ reste continue sur la frontière (mais sa dérivée a évidemment un saut). Pour le puits de profondeur finie, la valeur d'un état pair en $x = \pm a/2$ est donnée par l'expression (16.78). Quand k_0 devient infiniment grand, la racine carrée tend vers $\sqrt{2/a}$ et le facteur K_n/k_0 tend vers zéro, ce qui prouve l'affirmation que ψ s'annule aux bornes ; il en va évidemment de même pour les états impairs.

Ainsi, quand le potentiel présente un saut infini, la fonction d'onde reste continue au point de discontinuité ; en revanche, bien sûr, la dérivée de ψ présente un saut : elle est non-nulle sur la "surface" intérieure du puits, et nulle en dehors ; une violente singularité du potentiel affecte maintenant la dérivée mais la fonction propre, elle, reste continue partout. Ce résultat est général et doit être retenu.

Une fois que l'on est persuadé que ψ reste continue malgré le saut infini de $V(x)$, il est facile de retomber sur ses pieds. Dans le puits, l'équation aux fonctions propres admet des solutions qu'il est ici commode d'écrire d'emblée sous la forme :

$$\psi(x) = A \sin Kx + B \cos Kx , \quad (16.125)$$

où, comme avant dans la région centrale, $E = \hbar^2 K^2/(2m)$. Ecrivons maintenant que toutes ces fonctions s'annulent en $x = \pm a/2$:

$$-A \sin \frac{Ka}{2} + B \cos \frac{Ka}{2} = 0 , \quad A \sin \frac{Ka}{2} + B \cos \frac{Ka}{2} = 0 . \quad (16.126)$$

En faisant la somme et la différence de ces deux équations, on obtient :

$$A \sin \frac{Ka}{2} = 0 , \quad B \cos \frac{Ka}{2} = 0 . \quad (16.127)$$

d'où la première famille de solutions :

$$B \cos \frac{Ka}{2} = 0 \iff \frac{Ka}{2} = \frac{\pi}{2} + n\pi \quad \text{et} \quad A = 0 , \quad (16.128)$$

et :

$$\psi_n = B \cos(2n+1) \frac{\pi}{a} \quad (n \in \mathbb{N}) . \quad (16.129)$$

L'autre famille correspond à :

$$A \sin \frac{Ka}{2} = 0 \iff \frac{Ka}{2} = n'\pi \quad \text{et} \quad B = 0 . \quad (16.130)$$

et :

$$\psi_n = A \sin 2n' \frac{\pi}{a} \quad (n' \in \mathbb{N}^*) . \quad (16.131)$$

Les valeurs de K sont donc $(2n+1)\pi/a$ pour les solutions paires, $(2n')\pi/a$ pour les solutions impaires ; la suite des K_n et des $K_{n'}$ est isomorphe à celle des nombres entiers impairs et pairs, les énergies sont finalement, quel que soit $n \geq 1$, par l'expression (16.124). Après ordonnancement suivant les états d'énergie croissante et comme pour le puits fini, on obtient à nouveau une alternance d'états pairs (fondamental, deuxième excité, ...) et d'états impairs (premier excité, troisième excité, etc.). Ainsi, la condition aux limites spécifique du puits infini (ψ nulle aux bornes) reproduit très exactement ce qui est obtenu en prenant la limite infinie du puits fini. Notons enfin que l'écart entre deux niveaux consécutifs diminue si la masse devient très grande et tend vers zéro si $\hbar \rightarrow 0$. Une grande masse donne donc un quasi-*continuum* d'énergie : une fois encore, les effets quantiques tendent à s'effacer quand, toutes choses égales par ailleurs, on considère des particules de plus en plus massives.

Remarques

1. Dans le cas du puits infini, il est possible – et finalement plus judicieux après coup, à condition de savoir au préalable que ψ reste une fonction continue – de choisir autrement l'origine sur l'axe Ox , en la prenant au bord de gauche du puits – alors, le bord droit est à l'abscisse a . Dans ces conditions, les fonctions propres sont toutes des sinus :

$$\psi_n = \sqrt{\frac{2}{a}} \sin n\pi \frac{x}{a} \quad (n \in \mathbb{N}^*) , \quad (16.132)$$

les énergies étant évidemment inchangées, toujours données par (16.124)).

2. Il est intéressant d'examiner la limite des très grands nombres quantiques ; d'une façon générale, de (16.124)) on déduit :

$$E_{n+1} - E_n = (2n+1)\pi^2 \frac{\hbar^2}{2ma^2} , \quad (16.133)$$

et, si $n \gg 1$:

$$E_{n+1} - E_n \simeq n\pi^2 \frac{\hbar^2}{ma^2} . \quad (16.134)$$

Par ailleurs, la moyenne du carré de la vitesse dans l'état n est telle que :

$$\frac{1}{2} m \langle v^2 \rangle_n = E_n \iff \langle v^2 \rangle_n = n^2 \pi^2 \frac{\hbar^2}{m^2 a^2} . \quad (16.135)$$

Il en résulte que pour $n \gg 1$:

$$E_{n+1} - E_n \simeq \frac{\pi \hbar}{a} \sqrt{\langle v^2 \rangle_n} \iff \omega_{n+1} - \omega_n \simeq \frac{\pi}{a} \sqrt{\langle v^2 \rangle_n} . \quad (16.136)$$

D'un autre côté, pour une particule classique confinée entre deux murs distants de a et ayant une vitesse v (celle-ci est constante en module), la période du mouvement est $T = 2a/v$, correspondant à une pulsation $\omega = \pi v/a$. (16.136) montre que la fréquence de Bohr associée à deux états adjacents (en tout cas très voisins, $|n-m| \ll n, m$) tend donc vers la (une) pulsation fondamentale du mouvement classique dans la limite des très grands nombres quantiques. Une fois encore, on retrouve bien l'idée maîtresse de Bohr (Principe de Correspondance) généralisée et exploitée à fond par Heisenberg pour construire sa Mécanique des Matrices.

16.4 La marche de potentiel

Une marche de potentiel (voir fig. 16.17) modélise l'interface entre deux milieux où l'énergie potentielle prend deux valeurs constantes. Très grossièrement, la région (I) est par exemple l'intérieur d'un métal et la région (II) représente le vide ; en pareil cas, V_0 est le travail de sortie W_S .

Comme dans le cas du puits carré, il convient de distinguer les deux cas, selon que l'énergie est plus petite ou plus grande que V_0 . Dans le cas où E est supérieure à V_0 et où il existe une source de particules en $x = -\infty$, on trouvera dans la région de gauche à la fois un courant transmis et un courant *réfléchi* ; l'existence de ce dernier est manifestement un effet purement quantique : une particule classique incidente de la gauche sur la marche et ayant une énergie (cinétique) supérieure à V_0 est instantanément freinée en $x = 0$ mais continue sa route dans la région (II) et se retrouve donc à droite avec certitude ; à l'opposé, si l'énergie est inférieure à V_0 , la particule rebondit élastiquement sur le mur et repart en arrière : la probabilité est nulle de la trouver à droite. Au contraire, comme on va le voir, une particule quantique d'énergie inférieure à V_0 peut *pénétrer* dans la région (II), inaccessible classiquement. Il reste que le coefficient de réflexion R est égal à 1, comme dans le cas classique, le courant étant nul dans la région (II) ; l'image³⁹ que l'on peut se forger est celle d'une particule pénétrant plus ou moins profondément dans la barrière avant de rebrousser chemin et revenir dans la région $x < 0$.

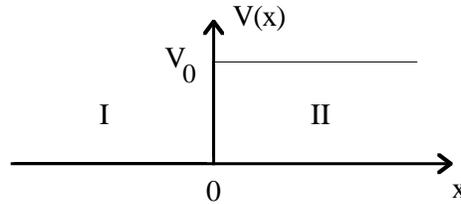


Figure 16.17: Marche de potentiel de hauteur V_0

Ici, l'énergie est forcément positive, puisque la plus petite valeur de $V(x)$ est nulle. On pose donc :

$$E = \frac{\hbar^2 k^2}{2m} \quad (k \in \mathbb{R}_+) . \quad (16.137)$$

Dans la région (I), l'équation aux valeurs propres prend la forme :

$$\psi'' + k^2 \psi = 0 . \quad (16.138)$$

Dans la région (II), cette équation est :

$$\psi'' + \left(\frac{2m}{\hbar^2} k^2 - V_0 \right) \psi = 0 . \quad (16.139)$$

En posant comme avant $V_0 = \hbar^2 k_0^2 / (2m)$ et :

$$K^2 = k_0^2 - k^2 , \quad (16.140)$$

(16.139) devient :

$$\psi'' - K^2 \psi = 0 . \quad (16.141)$$

Traisons d'abord en détail ce dernier cas ($E < V_0$, $K \in \mathbb{R}_+$). Les solutions s'écrivent :

$$\psi(x) = \begin{cases} A \cos kx + B \sin kx & \text{si } x < 0 \\ C e^{-Kx} & \text{si } x > 0 \end{cases} , \quad (16.142)$$

où on a d'emblée éliminé l'exponentielle divergente du côté $x > 0$. Les équations de continuité de ψ et ψ' donnent :

$$A = C , \quad kB = -KC , \quad (16.143)$$

³⁹ confortée par l'existence d'un retard à la réflexion, voir ci-dessous.

d'où :

$$\psi(x) = \begin{cases} A \left[\cos kx - \frac{K}{k} \sin kx \right] & \text{si } x < 0 \\ A e^{-Kx} & \text{si } x > 0 \end{cases} . \quad (16.144)$$

Pour des raisons qui doivent maintenant être claires, aucune autre condition n'est à prendre en compte et la valeur propre E n'est donc pas quantifiée. Dans la région $x < 0$, la fonction d'onde s'écrit aussi :

$$\psi(x) = \frac{A}{2} \left[\left(1 + i \frac{K}{k} \right) e^{ikx} + \left(1 - i \frac{K}{k} \right) e^{-ikx} \right] \equiv \tilde{A} (e^{ikx} + e^{-2i\delta} e^{-ikx}) . \quad (16.145)$$

Le poids de l'onde réfléchie est donc égal à celui de l'onde incidente (le coefficient de réflexion R est égal à 1, le courant $j(x)$ est nul dans la région $x > 0$ – où la fonction propre est réelle), mais il existe un *retard* à la réflexion mesuré par le déphasage⁴⁰ 2δ . Ce dernier s'exprime aussi comme :

$$\tan \delta = \sqrt{\frac{k_0^2}{k^2} - 1} . \quad (16.146)$$

δ est donc maximum pour l'énergie minimum ($k = 0$) et vaut alors $\pi/2$; sa plus petite valeur est nulle, atteinte au moment où l'énergie E devient juste égale à V_0 . Le courant étant nul à droite, il l'est aussi à gauche : le courant réfléchi est en module égal au courant incident fixé par la source. En ce qui concerne ce dernier, j_{inc} , il est le produit de $\hbar k/m$ et du module carré du coefficient de e^{ikx} de l'expression valide pour $x < 0$, soit (16.144) :

$$j_{\text{inc}} = \frac{\hbar k |A|^2}{4m} \left(1 + \frac{K^2}{k^2} \right) = \frac{\hbar k |A|^2}{4m} \frac{V_0}{E} . \quad (16.147)$$

Pour $x > 0$, la fonction d'onde décroît exponentiellement : même si la particule ne peut aller très loin dans cette région, il n'en reste pas moins qu'elle peut effectivement s'y trouver. La probabilité totale de trouver la particule à droite est :

$$P_{\text{droite}} = \int_0^{+\infty} dx |\psi(x)|^2 = \frac{|A|^2}{2K} . \quad (16.148)$$

En dimension un, le courant j est homogène à l'inverse d'un temps. Le rapport $P_{\text{droite}}/j_{\text{inc}}$ est donc un temps, τ , que l'on peut interpréter comme le temps de séjour de la particule dans la région interdite classiquement ; avec les résultats ci-dessus, on trouve :

$$\tau = \frac{P_{\text{droite}}}{j_{\text{inc}}} = \sqrt{\frac{E}{V_0 - E}} \frac{\hbar}{V_0} . \quad (16.149)$$

L'énergie étant fixée, τ est une fonction décroissante de la hauteur de la barrière. Pour une vitesse donnée, une particule massive passera moins de temps dans la région de droite qu'une particule légère. Au total, la proportion de particules qui pénètrent dans la région interdite est d'autant plus faible que la barrière est haute.

Dans l'autre cas, $E > V_0$, avec les mêmes notations que précédemment, K est imaginaire pur ($K = i|K|$) ; toute fonction propre est de la forme :

$$\psi(x) = \begin{cases} A e^{ikx} + B e^{-ikx} & \text{si } x < 0 \\ C e^{i|K|x} + D e^{-i|K|x} & \text{si } x > 0 \end{cases} , \quad (16.150)$$

où A , B , C et D sont de nouvelles constantes à trouver. Les conditions de raccordement en $x = 0$ s'écrivent :

$$A + B = C + D , \quad k(A - B) = |K|(C - D) , \quad (16.151)$$

de sorte que :

$$C = \frac{1}{2} \left[\left(1 + \frac{k}{|K|} \right) A + \left(1 - \frac{k}{|K|} \right) B \right] , \quad D = \frac{1}{2} \left[\left(1 - \frac{k}{|K|} \right) A + \left(1 + \frac{k}{|K|} \right) B \right] . \quad (16.152)$$

Si la source est en $x = -\infty$, la constante D doit être nulle ; dans ces conditions, les rapports B/A et C/A sont :

$$\frac{B}{A} = \frac{k - |K|}{k + |K|} \equiv \frac{k + iK}{k - iK} , \quad \frac{C}{A} = \frac{2k}{k + |K|} \equiv \frac{2k}{k - iK} . \quad (16.153)$$

Comme le nombre d'onde n'est pas le même à gauche et à droite (respectivement : k et $|K|$) ; la valeur constante du potentiel n'est pas la même à gauche et à droite, donc, pour une énergie E donnée, la vitesse n'est pas la même des deux côtés), il convient de prendre garde en définissant les coefficients de réflexion et de transmission. Le courant incident est $j_{\text{inc}} = (\hbar k/m)|A|^2$, le courant réfléchi est $j_r = (\hbar k/m)|B|^2$ et le courant transmis est⁴¹

⁴⁰Dans le cas $E > V_0$, il n'existe pas de déphasage (voir [16], p. 70).

⁴¹Attention ! C'est bien K (et non k) qui apparaît dans le courant pour $x > 0$.

$j_{\text{tr}} = (\hbar|K|/m)|C|^2$. On en déduit $T = j_{\text{tr}}/j_{\text{inc}}$ et $R = j_{\text{r}}/j_{\text{inc}}$:

$$R = \frac{|B|^2}{|A|^2} = \left(\frac{k+iK}{k-iK} \right)^2, \quad T = \frac{|C|^2}{|A|^2} = \frac{|K|}{k} \left(\frac{2k}{k-iK} \right)^2 = \frac{-4ikK}{(k-iK)^2}. \quad (16.154)$$

et on vérifie que $R + T = 1$ comme il se doit.

En rassemblant les résultats pour toutes les valeurs possibles de l'énergie, on voit que le coefficient de réflexion R vaut 1 tant que $E \leq V_0$, puis décroît et tend vers zéro à haute énergie ; sa décroissance pour $E \gg V_0$ est en $(V_0/E)^2$.

Notons que le courant total a en général les expressions :

$$j(x) = \begin{cases} (\hbar k/m)(|A|^2 - |B|^2) & \text{si } x < 0 \\ (|K|/m)(|C|^2 - |D|^2) & \text{si } x > 0 \end{cases}. \quad (16.155)$$

de sorte que la conservation du courant s'écrit⁴² *ici* :

$$k(|A|^2 - |B|^2) = |K|(|C|^2 - |D|^2). \quad (16.156)$$

16.5 La barrière de potentiel

L'étude des états liés du puits carré, tout autant que celle des états d'énergie inférieure à la hauteur de la marche, a en particulier révélé une propriété assez extraordinaire : quand l'énergie de la particule est *inférieure* à celle fixant le haut du puits (état lié, $-V_0 < E_n < 0$), ou le haut de la marche ($0 < E < V_0$), la probabilité n'est pas nulle de trouver la particule en dehors du puits, ou au-delà de la marche. Dans ces régions, interdites classiquement, la fonction d'onde décroît, certes exponentiellement, mais n'est pas identiquement nulle (on a vu toutefois que le courant y est identiquement nul). La longueur typique de décroissance est $1/k_n$; dans le cas du puits carré, cette quantité est reliée au vecteur d'onde k_n de l'état lié considéré par :

$$k_n = \sqrt{k_0^2 - K_n^2} = \frac{1}{2m} \sqrt{2m(-E_n)} = \hbar^{-1} \sqrt{(p^2)_n}. \quad (16.157)$$

Cette échelle de longueur est donc d'autant plus grande que l'état est moins lié ($|E|$ petit), traduisant d'ailleurs le fait que les états sont naturellement d'autant plus diffus (moins localisés) qu'ils correspondent à une grande énergie.

Ce phénomène de pénétration dans les régions inaccessibles classiquement est encore plus spectaculaire si on considère non pas un puits mais une barrière⁴³ de potentiel, dans le cas où l'énergie est inférieure à la hauteur V_0 de la barrière. Une particule classique incidente de la gauche avec une énergie (cinétique) E *inférieure* à V_0 rebondit simplement sur le mur. On va voir que la particule quantique, dans les mêmes conditions ($E < V_0$!), peut *passer* à travers la barrière (le courant n'est pas nul à droite si la source est en $x = -\infty$) : c'est ce que l'on appelle l'effet tunnel (passer de l'autre côté de la montagne sans l'escalader).

Pour mettre en évidence l'effet en question, seuls importent les états d'énergie inférieure à V_0 ; pour les autres, on trouve qu'il existe à la fois une onde transmise et une onde réfléchie, phénomène surprenant au premier abord mais auquel on doit maintenant être habitué ; examinons donc exclusivement les états $E < V_0$, la méthode étant toujours la même : résolution de l'équation aux valeurs propres dans chaque région et raccordement aux points où $V(x)$ a un saut. Il est inutile de faire les calculs : tout ce qui a été fait dans le cas du puits carré est utilisable directement, à condition simplement de changer V_0 en $-V_0$. On pose donc :

$$E = \frac{\hbar^2 k^2}{2m}, \quad E - V_0 = \frac{\hbar^2 K^2}{2m}. \quad (16.158)$$

Comme $E < V_0$, K est maintenant imaginaire pur ; les relations de continuité écrites à propos des états non-liés du puits fini restent valables et il suffit, si on souhaite les expliciter, de prendre en compte le fait que maintenant $K = i|K|$. Les relations de continuité se traduisent toujours par (16.102) (mais comme K est maintenant

⁴²Noter que la présence de k et de $|K|$ aux premier et second membre respectivement.

⁴³Une barrière est la juxtaposition d'une marche et d'une "antimarche" de potentiel.

imaginaire pur $\cos ix = \cosh x$, $\sin x = i \sinh x$). Les coefficients R et T sont toujours donnés par (16.111). En particulier, on trouve T sous la forme :

$$T = \frac{4E(V_0 - E)}{4E(V_0 - E) + V_0^2 \sinh^2 \sqrt{2ma^2(V_0 - E)}/\hbar^2} \equiv \frac{4\varepsilon(1 - \varepsilon)}{4\varepsilon(1 - \varepsilon) + \sinh^2 \sqrt{\mu(1 - \varepsilon)}} , \quad (16.159)$$

où μ et ε sont définis en (16.115). Cette expression se simplifie pour une barrière haute ($E \ll V_0$ soit $\varepsilon \ll 1$) et large ($k_0 a \gg 1$) ; dans ces conditions :

$$T \simeq 16 \frac{E}{V_0} \exp \left[-2\hbar^{-1} \sqrt{2ma^2(V_0 - E)} \right] , \quad (16.160)$$

Dans ces conditions extrêmes, le coefficient de transmission est exponentiellement petit et, tant que ε n'est pas trop proche de 1, croît en gros comme ε . Quand $\varepsilon \rightarrow 1$ (à gauche), T tend vers $4/(4 + \mu)$.

Dans l'autre cas, $E > V_0$, le coefficient de transmission est donné par :

$$T = \frac{4E(E - V_0)}{4E(E - V_0) + V_0^2 \sin^2 \sqrt{2ma^2(E - V_0)}/\hbar^2} \equiv \frac{4\varepsilon(\varepsilon - 1)}{4\varepsilon(\varepsilon - 1) + \sin^2 \sqrt{\mu(\varepsilon - 1)}} , \quad (16.161)$$

Ceci achève de déterminer T quelle que soit la valeur de l'énergie. Bien évidemment, les deux expressions de T , (16.159) et (16.161), donnent la même valeur pour $E = V_0$ à savoir $T = 4/(4 + \mu)$. A nouveau, des résonances très fines se produisent à chaque fois que le sinus s'annule et alors T vaut exactement 1. La plupart des remarques faites à propos des états non-liés du puits fini, peuvent être reprises ici ; en particulier, si $\mu \gg 1$ (vrai si la particule est très massive), T vaut à peu près soit 0 soit 1 (suivant que l'énergie E est plus petite ou plus grande que V_0) ; la montée entre 0 et 1 (T est une fonction continue !) se fait sur un intervalle d'ordre $1/\mu$.

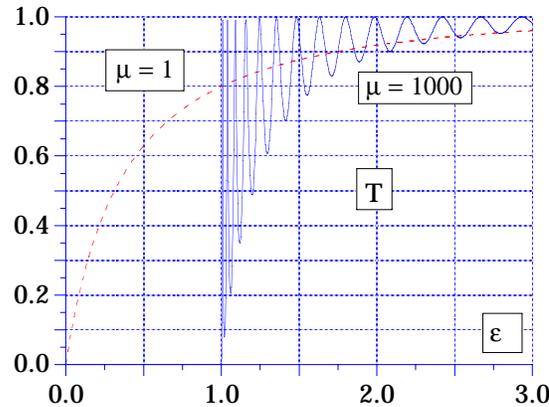


Figure 16.18: Variation de la transmission de la barrière en fonction de E/V_0 , pour deux valeurs de $\mu = 2ma^2V_0/\hbar^2$.

Le coefficient de transmission dépend de façon spectaculaire de la masse de la particule. Pour une particule de masse élevée ($\mu \gg 1$), le coefficient est pratiquement nul tant que l'énergie est inférieure à V_0 . Au contraire, pour une particule légère, T prend des valeurs importantes, même à basse énergie. Ceci est confirmé par une application numérique ; pour une barrière atomique :

$$a = 1 \text{ \AA} , \quad V_0 = 2 \text{ eV} . \quad (16.162)$$

Si la particule incidente est un électron ($m = 9 \times 10^{-31} \text{ kg}$) d'énergie 1 eV⁴⁴, on trouve :

$$T_{\text{électron}} \simeq 0.78 . \quad (16.163)$$

En revanche, dans les mêmes conditions, pour un proton, on trouve :

$$T_{\text{proton}} \simeq 4 \times 10^{-19} . \quad (16.164)$$

Ainsi, pour cette barrière de dimensions atomiques, l'électron se comporte ultra-quantiquement, alors que le proton, en raison de sa masse élevée, est quasi-classique : la probabilité de trouver ce dernier au-delà de la barrière est très petite.

⁴⁴Une telle énergie est faible, mais pour un électron, elle correspond à une vitesse de l'ordre de 600 km/s.

L'effet tunnel joue un rôle fondamental en microphysique et possède bien des applications. C'est lui qui est responsable de la désintégration α de certains noyaux. On fabrique des diodes utilisant cet effet (diode tunnel). Le microscope à effet tunnel permet d'obtenir des images extraordinairement précises de la surface des métaux et constitue un appareil d'une sensibilité extraordinaire⁴⁵. Le maser à ammoniac utilise une transition tunnel (l'atome d'azote passe d'un côté à l'autre du plan des trois atomes d'hydrogène à une fréquence relativement faible (correspondant à une longueur d'onde centimétrique : $\nu \sim 10^{11} \text{ s}^{-1}$, $\lambda \sim 1 \text{ cm}$). Un mélange isotopique peut être enrichi dans l'isotope le moins massif par passage à travers une (ou plusieurs) barrière(s), *etc.*

⁴⁵La sensibilité résulte essentiellement de la dépendance *exponentielle* du courant tunnel par rapport à la distance entre la pointe et la surface.

Chapitre 17

L'oscillateur harmonique

Le but de ce chapitre est de donner un traitement élémentaire et aussi complet que possible de l'oscillateur harmonique linéaire (au sens : une dimension d'espace). Comme on le sait, ce système a été le terrain d'essai de la Mécanique des Matrices de Heisenberg et une partie des résultats établis dans la suite sont déjà connus (par exemple : la quantification de l'énergie, $E_n = (n + \frac{1}{2})\hbar\omega$), de sorte que ce dernier chapitre, d'une certaine façon, referme la boucle. Le point de départ sera ici l'équation aux valeurs propres de Schrödinger, dont on s'attachera à déterminer les solutions (modes propres), rencontrant à nouveau le rôle crucial des conditions requises pour toute fonction acceptable quant à l'apparition de la quantification¹ – dans une situation toutefois moins simple qu'au chapitre précédent puisqu'ici le potentiel varie continûment. En outre, l'introduction des opérateurs de création et d'annihilation donnera l'occasion de toucher du doigt une étape supplémentaire de l'élaboration de la Théorie Quantique², où on laisse “flotter” le nombre de particules, bosons ou fermions.

17.1 L'importance de l'oscillateur harmonique

L'oscillateur harmonique joue en Physique un rôle de premier plan. Le plus souvent, ce système apparaît pour la première fois à propos des oscillations d'une masse (ponctuelle) m fixée à l'extrémité d'un ressort parfait³ de raideur k . En l'absence de frottements, on sait que, à partir d'un état initial hors d'équilibre (vitesse initiale non nulle et/ou position initiale distincte de la position d'équilibre), le point matériel effectue des oscillations à une fréquence ν_0 bien déterminée, reliée à la pulsation $\omega_0 = (k/m)^{1/2}$ par $\nu_0 = \omega_0/2\pi$. En présence de frottements, les oscillations sont amorties, le caractère éphémère du mouvement étant une simple conséquence de la non-conservation de l'énergie mécanique du système constitué par la masse et le ressort, le milieu extérieur absorbant peu à peu l'énergie mécanique initialement disponible sous forme potentielle et/ou cinétique ; l'amortissement est caractérisé par une échelle de temps τ que l'on peut appeler durée de vie. On distingue usuellement, deux types de régimes (facilement mis en évidence théoriquement lorsque la force de freinage est proportionnelle à la vitesse⁴) : le régime sous-amorti, pour lequel la petite masse effectue un grand nombre d'oscillations durant le temps τ ($\omega_0^{-1} \ll \tau$) et le régime sur-amorti où, le frottement étant très fort, la masse n'a guère le temps d'osciller et perd très vite toute son énergie mécanique ($\omega_0^{-1} \gg \tau$). On caractérise souvent la nature du régime par la valeur du facteur de qualité $Q = \omega_0\tau/2\pi = \nu_0\tau$; Q est très grand pour le régime sous-amorti, très petit dans le cas contraire. Lorsque τ est très grand par rapport à toute autre échelle de temps déjà présente par ailleurs, le système physique considéré peut, de fait, être assimilé à un oscillateur idéal, non-dissipatif.

Dans le cas le plus simple, l'oscillateur n'implique qu'un seul degré de liberté (coordonnée pour la masse fixée au ressort, angle par rapport à la verticale pour le pendule dans la limite des petites oscillations, *etc.*) ; il existe évidemment des oscillateurs à plusieurs degrés de liberté, qui peuvent d'ailleurs être isotropes ou anisotropes. Tous ont la même propriété remarquable : ce sont des systèmes linéaires. En effet, comme l'énergie potentielle est – par définition – quadratique par rapport aux degrés de liberté, la force est une fonction du

¹Comme on le verra, l'oscillateur harmonique ne possède que des états liés.

²appelée “Seconde Quantification”.

³On parle alors d'oscillateur *matériel*.

⁴Ce que l'on appelle le frottement *fluide*.

premier degré (linéaire) de ces mêmes variables ; il en résulte que l'équation fondamentale de la dynamique est une équation (ou un système d'équations) linéaire. C'est la raison profonde pour laquelle l'oscillateur harmonique a une dynamique très simple, aussi bien d'ailleurs en Mécanique Classique qu'en Mécanique Quantique.

Les aspects dynamiques se retrouvent dans la susceptibilité⁵ de l'oscillateur, χ , fonction qui traduit la réponse du système à une excitation extérieure ; comme l'oscillateur est un système linéaire, l'étude de la réponse peut être abordée en décomposant la perturbation externe en ses composantes de Fourier et l'analyse de la variation en fréquence engendre spontanément une fonction $\chi(\omega)$. Dans le cas du frottement faible, cette fonction présente une variation résonnante à une valeur très proche de ω_0 , si proche d'ailleurs que le plus souvent on admet que la résonance se produit à ω_0 strictement (le déplacement de la résonance est du second ordre par rapport à $1/(\omega_0\tau)$ qui, dans le cas sous-amorti, est très petit devant 1) ; le maximum est d'autant plus aigu et la largeur d'autant plus faible que le frottement est faible et, en pratique, c'est la dissipation d'énergie induite par le couplage avec l'extérieur sous forme de frottements qui empêche le système d'exploser⁶. En présence de frottement fort, $\chi(\omega)$ est une fonction variant lentement, très large et passant par un maximum pour une fréquence, très différente de ω_0 , dont la dépendance par rapport à τ ne peut plus être ignorée.

On sait qu'il existe également des oscillateurs électriques, le plus simple étant constitué par un circuit LC parallèle ; la fréquence propre est alors $\omega_0 = \sqrt{LC}$ et la composante dissipative est constituée par les inévitables résistances présentes dans le circuit (fils de connexion, résistance de fuite du condensateur, résistance de la bobine). Un intérêt majeur d'un tel dispositif est son aptitude à filtrer les différentes composantes de Fourier d'une source et constitue la base du circuit d'accord d'un appareil de réception. Par des associations diverses, on peut aussi fabriquer des filtres passe-haut, passe-bas ou passe-bande, utilisés par exemple dans les enceintes acoustiques.

L'apparition de l'oscillateur harmonique ne se limite pas, loin de là, aux exemples qui viennent d'être cités. D'ailleurs, l'atome de Thomson classique est un oscillateur harmonique dans le régime sous-amorti⁷. D'une façon tout à fait générale, l'oscillateur surgit dès qu'un système caractérisé par un degré de liberté q possède une "position" d'équilibre stable q_0 résultant de l'existence d'un minimum pour une fonction $V(q)$ qui joue le rôle d'une énergie potentielle dans une équation du mouvement, pourvu que l'on se borne à examiner les petites oscillations autour de la position d'équilibre q_0 ($V'(q_0) = 0$). Par hypothèse, dans ce contexte, on introduit le développement :

$$V(q) = V(q_0) + \frac{1}{2}(q - q_0)^2 V''(q_0) + \dots , \quad (17.1)$$

et, par la suite, on ne considère que les termes explicitement écrits. Par identification avec l'énergie potentielle d'un oscillateur matériel parfait :

$$V(x) = \frac{1}{2}kx^2 \equiv \frac{1}{2}m\omega_0^2 x^2 , \quad (17.2)$$

on voit que :

$$V''(q_0) = m\omega_0^2 . \quad (17.3)$$

L'hypothèse de l'équilibre *stable* est contenue dans le fait que la dérivée seconde $V''(q_0)$ est positive. Dans le cas des états métastables (quand q_0 est un minimum relatif assez profond de $V(q)$), ω_0 représente la fréquence d'oscillation sur une échelle de temps petite par rapport au temps d'échappement (par fluctuations thermiques, par exemple).

Ce point étant réalisé, on imagine facilement l'immense variété des situations où l'oscillateur harmonique apparaît spontanément. Par exemple, une molécule diatomique AB est caractérisée notamment par la longueur

⁵À prendre au même sens que la *susceptibilité* de caractère, traduisant l'aptitude pour une personne à partir "au quart de tour".

⁶Les frottements, souvent considérés comme nuisibles, sont aussi malgré tout d'une utilité extrême. Un galvanomètre non-amorti est totalement inutilisable (l'aiguille ne cesse d'osciller !). Un régulateur à boules trop bien "huilé" se met à osciller et ne remplit plus son rôle (d'une façon générale, un amortissement optimisé est essentiel dans tous les dispositifs de régulation). Pour un système à très grand nombre de degrés de liberté, l'existence d'un équilibre macroscopique résulte d'une compétition subtile entre les fluctuations (thermiques par exemple) et la dissipation d'énergie ; ceci se traduit par les théorèmes fondamentaux dits "théorèmes de fluctuation-dissipation". Enfin, dans le cas de systèmes complexes, les frottements sont déterminants pour limiter et contenir les explosions chaotiques : on s'accorde parfois à penser que les *krachs* boursiers peuvent résulter d'une mise en communication trop fiable (il manque un peu de "bruit") et trop rapide des différents agents prenant des décisions ultra-rapides inconsidérées, déterminées dans la hâte par le comportement erratique des uns et des autres.

⁷L'électron effectue environ un million de vibrations avant de chuter au centre.

d'équilibre de la liaison chimique A-B ; mais les noyaux ne sont pas immobiles et, dans les cas ordinaires, effectuent des *petites*⁸ vibrations autour de leurs positions d'équilibre⁹. D'une façon générale, la théorie des vibrations moléculaires, et donc celle des réactions chimiques où intervient une étape dissociative¹⁰, fait un usage intensif de la théorie de l'oscillateur harmonique. Il en va de même dans un solide : les ions vibrent autour de leurs positions stables et les mouvements de petite amplitude sont décrits par des excitations (par rapport à un état d'immobilité complète), appelées généralement *phonons* car certaines d'entre elles sont responsables de la propagation du son dans le cristal. Ces phonons sont des "particules", au sens où il existe une relation de dispersion reliant l'énergie à l'impulsion.

Il existe bien d'autres cas, en Physique de la matière condensée, où l'oscillateur harmonique joue un rôle majeur, et où interviennent d'ailleurs des degrés de liberté qui ne sont pas "matériels" au sens usuel. Il en va ainsi lorsqu'une structure magnétique peut exister parce que, pour des raisons subtiles¹¹, des moments permanents préalablement existants, peuvent s'organiser pour donner lieu à l'apparition d'un ordre à longue distance (ferromagnétisme, antiferromagnétisme). Dans les systèmes possédant cette inclination, on définit des excitations de type magnétique (par exemple les *magnons*) qui sont des excitations élémentaires au-dessus de l'état fondamental ordonné dans l'espace (tous les moments pointent dans la même direction, ou sont alternativement en haut et en bas ; tout écart par rapport à cet ordre ne peut apparaître que par un apport d'énergie, par fluctuations thermiques par exemple). Un magnon est l'excitation, compatible avec la symétrie de translation discrète, construite sur tous les états possibles où une seul spin est retourné (a "flippé"). À basse densité, les magnons ont les propriétés d'oscillateurs harmoniques.

L'oscillateur harmonique s'introduit aussi lors de l'étude des propriétés magnétiques de l'atome. En effet, si un champ magnétique statique \vec{B} parallèle à Oz est appliqué à un atome, il apparaît d'une part un couplage du genre Zeeman donnant lieu au paramagnétisme ordinaire, d'autre part un terme diamagnétique proportionnel à $(x^2 + y^2)$ constituant un oscillateur harmonique à deux dimensions et produisant les niveaux quantiques dits niveaux de Landau. Ces niveaux jouent un rôle important en Astrophysique : en champ intense, la distribution électronique d'un atome a tendance à "s'aplatir" puisque le champ intense fait tourner rapidement les électrons dans un plan qui lui est perpendiculaire ; dans des conditions physiques exotiques extrêmes apparaissant dans certains milieux astrophysiques, l'atome a tendance à devenir quasi-bidimensionnel ; il en résulte une spectroscopie bien différente de celle connue sur la Terre.

Enfin, il existe un autre cas très important où l'oscillateur harmonique joue un rôle fondamental : c'est lors de l'étude du champ électromagnétique. L'analyse du rayonnement thermique a donné l'occasion de voir que le champ électromagnétique (classique) dans une cavité possède un ensemble infini de modes propres, chacun d'entre eux étant strictement identique à un oscillateur harmonique (immatériel) de fréquence bien déterminée ; cette identification est possible par la forme du Hamiltonien du champ (voir Chapitre 4). Comme toute quantification peut s'effectuer à partir de l'expression du Hamiltonien (classique), la Théorie Quantique du rayonnement repose, elle aussi, sur celle de l'oscillateur harmonique.

Ces différents exemples montrent l'omniprésence de l'oscillateur harmonique en Physique et justifient une présentation détaillée de la quantification de ce système très simple. Celle-ci ne concernera qu'un oscillateur non-dissipatif : la théorie quantique de la dissipation est un monde en soi et sort résolument du cadre de ce cours.

⁸Cette hypothèse est essentielle pour que l'on puisse s'en tenir au développement (17.1).

⁹Cette vision résulte de l'approximation dite de Born - Oppenheimer (aussi dite adiabatique), justifiée par le fait que les noyaux ont un mouvement lent (compte tenu de leur grande inertie) par rapport à celui des électrons, très rapides en raison de leur faible masse ; bien évidemment, la distance internucléaire d'équilibre dépend à son tour du mouvement électronique, dont les noyaux ne voient, en quelque sorte, que la moyenne temporelle.

¹⁰Si l'amplitude des vibrations devient grande, la molécule peut se dissocier. Pour un solide, c'est la fusion qui commence.

¹¹L'origine de ces structures n'est pas l'existence d'interactions du genre dipôle magnétique entre les moments permanents, même si ces interactions jouent finalement un rôle important pour la structure macroscopique (domaines de Bloch) ; en réalité, ce sont les électrons, leurs interactions électrostatiques et le principe de Pauli qui sont la base de tout. D'une part, c'est la structure électronique des ions du solide qui produit ou ne produit pas les moments permanents individuels ; d'autre part, c'est le mouvement électronique dans le système étendu dans l'espace qui, par des mécanismes subtils, provoque l'apparition d'une interaction effective entre ces moments magnétiques.

17.2 Résolution de l'équation aux valeurs propres

Il s'agit ici de résoudre l'équation aux valeurs et vecteurs propres survenant après la séparation espace-temps de l'équation de Schrödinger. Pour fixer les idées, on raisonne dans le cas d'un oscillateur linéaire matériel de masse m , de pulsation notée maintenant ω , vibrant le long de l'axe Ox . Le Hamiltonien est :

$$H(x, p) = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 . \quad (17.4)$$

Suivant la prescription de Schrödinger, l'équation aux valeurs et vecteurs (fonctions) propres pour $\psi(x)$ prend la forme :

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + \frac{1}{2}m\omega^2 x^2 \psi(x) = E\psi(x) . \quad (17.5)$$

Comme $V(x)$ diverge aux grands x , toutes les solutions physiquement acceptables tendent vers zéro à l'infini¹², suffisamment vite pour être de module carré sommable : l'oscillateur harmonique ne possède donc que des états liés. Par ailleurs, l'énergie potentielle est une fonction paire de la coordonnée ; le Hamiltonien est donc invariant quand on change x en $-x$. Cette propriété de symétrie doit se retrouver, physiquement, dans les propriétés des valeurs et fonctions propres. Le système se moque du sens positif conventionnel de l'axe Ox et, par conséquent, toute fonction propre $\psi(x)$ (associée à l'énergie E) se transforme, par l'inversion d'espace Π , en une fonction $\Pi\psi(x) = \psi(-x)$ qui correspond sûrement à la même énergie E ; il serait en effet absurde d'imaginer qu'une opération de symétrie – qui par construction ne change rien à l'état du système – peut changer l'énergie de ce dernier.

Par ailleurs, on sait que les états discrets (liés) de tout mouvement à une dimension sont non-dégénérés : c'est donc aussi le cas pour l'oscillateur harmonique linéaire. Comme les deux fonctions $\psi(x)$ et $\Pi\psi(x)$ correspondent la même énergie, ces deux fonctions sont en fait proportionnelles : c'est dire que $\psi(x)$ est aussi fonction propre de Π . L'opérateur Π est manifestement unitaire ($\Pi^\dagger = \Pi^{-1}$) puisque changer le sens positif sur l'axe Ox ne change évidemment pas l'intégrale définissant le produit scalaire ; par ailleurs, le carré de Π , Π^2 , est égal à l'identité car effectuer deux fois l'inversion d'espace revient à ne rien faire. Il en résulte :

$$\Pi^2 = \mathbf{1} \iff \Pi = \Pi^{-1} , \quad (17.6)$$

et, au total :

$$\Pi = \Pi^{-1} = \Pi^\dagger . \quad (17.7)$$

Π est hermitique et possède donc des valeurs propres réelles ; comme $\Pi^2 = \mathbf{1}$, ces valeurs propres sont les racines carrées de 1 soit ± 1 . Toutes les fonctions propres de (17.5) sont donc telles que :

$$\Pi\psi(x) = \pm \psi(x) \iff \psi(-x) = \pm \psi(x) . \quad (17.8)$$

En d'autres termes, toute fonction propre d'état lié de l'oscillateur harmonique à une dimension est soit paire, soit impaire.

L'exposé qui suit reprend la méthode dite polynômiale due à Sommerfeld (1929). Pour des raisons de commodité, il est utile de récrire l'équation aux valeurs et vecteurs propres en termes de grandeurs sans dimension. Posons :

$$\xi = \kappa x , \quad u(\xi) = \psi(x) \iff \psi(x) = u(\xi = \kappa x) . \quad (17.9)$$

où κ est l'inverse d'une longueur ; on a :

$$\frac{d^2\psi}{dx^2} = \kappa^2 \frac{d^2u}{d\xi^2} . \quad (17.10)$$

de sorte que l'équation (17.5) devient :

$$-\frac{d^2u}{d\xi^2} + \frac{m^2}{\hbar^2\kappa^4}\omega^2\xi^2 u(\xi) = E \frac{2m}{\hbar^2\kappa^2} u(x) . \quad (17.11)$$

¹²La divergence de V écarte la possibilité de solutions oscillantes à l'infini.

En posant maintenant :

$$\kappa^4 = \frac{m^2 \omega^2}{\hbar^2}, \quad \lambda = E \frac{2m}{\hbar^2 \kappa^2} = \frac{2E}{\hbar \omega} \iff E = \frac{1}{2} \lambda \hbar \omega, \quad x = \sqrt{\frac{\hbar}{m \omega}} \xi, \quad (17.12)$$

il vient :

$$\frac{d^2 u}{d\xi^2} + (\lambda - \xi^2) u(\xi) = 0. \quad (17.13)$$

λ est la valeur propre adimensionnée. Une fois mise sous cette forme, cette équation est dite de Weber, et on peut en trouver une discussion approfondie dans les ouvrages de Mathématiques¹³. Avant d'en démarrer la résolution effective, et pour circonscrire d'emblée le domaine d'intérêt, il est utile d'obtenir le comportement asymptotique des solutions $u(\xi)$, au seul vu de l'équation (17.13).

À condition que ξ soit assez grand, on peut négliger la valeur propre λ devant ξ et écrire :

$$\frac{d^2 u}{d\xi^2} - \xi^2 u(\xi) \simeq 0 \quad (|\xi| \rightarrow +\infty). \quad (17.14)$$

L'inspection de cette équation simplifiée montre que toutes les fonctions de la forme $\xi^\mu e^{\pm \frac{\xi^2}{2}}$ conviennent, μ étant un exposant pour l'instant quelconque¹⁴ ; en effet, on a :

$$\frac{d^2}{d\xi^2} [\xi^\mu e^{\pm \frac{\xi^2}{2}}] = [\mu(\mu-1)\xi^{\mu-2} \pm (\mu+1)\xi^\mu \pm (\mu\xi^{\mu-1} \pm \xi^{\mu+1})\xi] e^{\pm \frac{\xi^2}{2}}, \quad (17.15)$$

où les signes \pm sont solidaires. En ne retenant que les termes dominants (de plus haute puissance en ξ), le premier membre de l'équation prise dans sa forme asymptotique se réduit à :

$$\pm (\pm \xi^{\mu+1}) \xi e^{\pm \frac{\xi^2}{2}} - \xi^2 \xi^\mu e^{\pm \frac{\xi^2}{2}}, \quad (17.16)$$

et est bien nul. De surcroît, l'exponentielle divergente est évidemment à rejeter : elle ne saurait conduire à une fonction satisfaisant la condition de normalisabilité, requise pour les états liés recherchés. Globalement, l'analyse ci-dessus suggère de faire un changement de fonction inconnue et de poser :

$$u(\xi) = H(\xi) e^{-\frac{\xi^2}{2}}, \quad (17.17)$$

où H est une fonction que l'on peut chercher *a priori* sous la forme d'une série (pas forcément entière) – dont on va voir qu'elle doit de réduire à un polynôme afin que la fonction $u(\xi)$ soit normalisable¹⁵. Il reste maintenant à déterminer H , dont on trouve facilement l'équation différentielle par substitution de (17.17) dans (17.13) :

$$H''(\xi) - 2\xi H'(\xi) + (\lambda - 1) H(\xi) = 0. \quad (17.18)$$

Cherchons maintenant les solutions de (17.18) sous la forme :

$$H(\xi) = \xi^s \sum_{\nu \in \mathbb{N}} a_\nu \xi^\nu, \quad (17.19)$$

où a_0 est supposé non-nul (dans le cas contraire, il suffirait de redéfinir l'exposant s du préfacteur). Par ailleurs, comme $|\psi|^2$ doit être intégrable sur tout intervalle fini, en particulier autour de l'origine, il faut $2s > -1$; en réalité, comme l'énergie potentielle est continue en $x = 0$, la dérivée ψ' elle-même est continue, ce qui impose $s \geq 0$. Reportant un tel développement dans (17.18) et identifiant les coefficients d'une même puissance de ξ , on trouve la récurrence¹⁶ :

$$s(s-1)a_0 = 0, \quad (s+1)s a_1 = 0. \quad (17.20)$$

¹³On peut par exemple se reporter à Morse et Feshbach [27]. Il convient toutefois de rappeler que la discussion purement mathématique d'une telle équation est insuffisante pour la Physique, qui impose de surcroît certaines propriétés bien définies aux solutions acceptables.

¹⁴L'exposant μ est *a priori* quelconque mais positif, car rien dans l'équation ne saurait faire diverger u en un point à distance finie ; si μ n'est pas entier, ξ^μ désigne une fonction multiforme convenablement définie sur sa surface de Riemann.

¹⁵Comme rappelé ci-dessus, l'oscillateur ne possède que des états liés, donc toutes les fonctions propres doivent être normalisables.

¹⁶la relation (17.21) est encore valide pour $\nu = -2, -1$ à condition de convenir que $a_{-2} = a_{-1} = 0$. À *contrario*, on devine qu'il existe des solutions ayant $a_{-1} \neq 0$ et/ou $a_{-2} \neq 0$: ces solutions se trouvent écartées de fait par la nécessité pour les solutions physiques de posséder les bonnes propriétés.

$$(s + \nu + 2)(s + \nu + 1) a_{\nu+2} - (2s + 2\nu + 1 - \lambda) a_{\nu} = 0 \quad (\nu \in \mathbb{N}) . \quad (17.21)$$

La récurrence couple entre eux, de deux en deux, les coefficients de la série (17.19). a_0 étant différent de zéro, la première équation donne $s = 0$ ou $s = 1$; la seconde donne $s = 0$ ou $a_1 = 0$ ou les deux à la fois. Au total, la série (17.19) est finalement entière.

- Quand s est choisi nul, a_1 est indéterminé ; si a_1 est aussi choisi nul, alors les coefficients $a_3, a_5, \text{etc.}$ sont également nuls et la fonction ainsi trouvée est paire. Si a_1 est pris non-nul, la fonction $H(\xi)$ contient des termes pairs (puisque $a_0 \neq 0$) et impairs (puisque $a_1 \neq 0$), or on sait qu'elle doit être de parité déterminée. Donc, si s est pris nul, le seul choix possible est corrélativement $a_1 = 0$, ce qui donne une fonction $H(\xi)$ paire
- Dans l'autre cas possible, $s = 1$, alors nécessairement $a_1 = 0$ et de proche en proche, tous les a_{2p+1} le sont aussi : la série ne contient que des termes en ξ^{2p} , elle est paire mais au total la solution $H(\xi)$ est impaire, puisque, avec $s = 1$, la série est multipliée par ξ .

En définitive, on voit que $H(\xi)$ est une fonction de la forme :

$$H(\xi) = \xi^s \sum_{p \in \mathbb{N}} a_{2p} \xi^{2p} , \quad (17.22)$$

et qu'elle est donc paire si $s = 0$, impaire si $s = 1$.

On va maintenant montrer – c'est l'étape cruciale introduisant la quantification de l'énergie E – que cette série doit avoir *tous ses coefficients nuls à partir d'un certain rang*, afin d'obtenir une fonction $u(\xi)$ ayant les propriétés requises. En effet, en partant de la relation de récurrence (17.21) réécrite avec $\nu = 2p$:

$$(s + 2p + 2)(s + 2p + 1) a_{2p+2} - (2s + 4p + 1 - \lambda) a_{2p} = 0 \quad (p \in \mathbb{N}) , \quad (17.23)$$

on voit que, pour p grand, le comportement asymptotique du rapport a_{2p+2}/a_{2p} est :

$$\frac{a_{2p+2}}{a_{2p}} \simeq \frac{4p}{(2p)(2p)} = \frac{1}{p} \simeq \frac{1}{p+1} \quad (p \gg 1) . \quad (17.24)$$

Or ceci est très exactement aussi le rapport des coefficients de deux termes consécutifs du développement en série entière de e^{ξ^2} :

$$e^{\xi^2} = \sum_{p \in \mathbb{N}} \frac{1}{p!} \xi^{2p} \equiv \sum_{p \in \mathbb{N}} c_{2p} \xi^{2p} \quad \text{avec} \quad \frac{c_{2p+2}}{c_{2p}} = \frac{1}{p+1} . \quad (17.25)$$

Au total, si la série (17.22) est une vraie série (*i. e.* si la somme contient une infinité de termes), le comportement dominant de $H(\xi)$ aux grands ξ (piloté par les a_{2p} , $p \gg 1$) est $\xi^s e^{\xi^2}$; suivant (17.17), celui de la fonction propre $u(\xi)$ est donc $\xi^s e^{\xi^2} e^{-\frac{\xi^2}{2}} = \xi^s e^{+\frac{\xi^2}{2}}$: malgré le facteur exponentiel $e^{-\frac{\xi^2}{2}}$ introduit d'emblée en (17.17), la fonction u diverge exponentiellement quand $\xi \rightarrow +\infty$. À ce stade, la seule issue possible est que la fonction $H(\xi)$ se réduise à un simple *polynôme*, de degré éventuellement aussi grand que l'on veut, mais fini. Alors, aussi grand que soit ce degré, en allant assez loin (en ξ) l'exponentielle $e^{-\frac{\xi^2}{2}}$ finira par l'emporter sur le polynôme et la fonction u tendra vers zéro (très vite d'ailleurs en raison du facteur gaussien) ; manifestement u est alors de module carré sommable. Revenant à la relation de récurrence (17.21), on voit que ceci est vrai si la valeur propre réduite λ n'est pas quelconque mais satisfait :

$$2s + 4p_0 + 1 - \lambda = 0 \quad \Longleftrightarrow \quad \lambda = 4p_0 + 2s + 1 , \quad (17.26)$$

où p_0 est un certain entier positif ou nul mais à part cela quelconque. En effet, reportant l'expression de λ qui résulte de (17.26), la récurrence (17.23) devient :

$$(s + 2p + 2)(s + 2p + 1) a_{2p+2} - (4p - 4p_0) a_{2p} = 0 . \quad (17.27)$$

Tôt ou tard, p augmentant à partir de 0, le facteur de a_{2p} devient nul pour $p = p_0$; alors a_{2p+2} est aussi nul et, de proche, il en va de même pour tous les a_{2p+2r} . Comme s peut prendre les deux valeurs 0 ou 1, les valeurs

possibles de λ sont de la forme $4p_0 + 1$ ou $4p_0 + 3$, résultats que l'on peut condenser en une seule et même écriture :

$$\lambda = 2n + 1 \quad (n \in \mathbb{N}) . \quad (17.28)$$

Revenant à la valeur propre E , reliée à λ par (17.12), on obtient le résultat fondamental¹⁷ :

$$E \in \{E_n\}_n , \quad E_n = \left(n + \frac{1}{2}\right) \hbar\omega . \quad (17.29)$$

Une fois encore, ce sont les propriétés requises pour les états propres, elles-mêmes découlant des postulats, qui engendrent spontanément la quantification de l'énergie. Conformément à l'habitude, le polynôme associé à $\lambda = 2n + 1$ est noté H_n ; certaines de ses propriétés seront données plus loin.

Les $\{E_n\}_n$ constituent le spectre du Hamiltonien de l'oscillateur harmonique. La propriété la plus remarquable de ce spectre est d'être constitué de niveaux *équidistants*, deux niveaux consécutifs étant séparés par $\hbar\omega$; cette propriété est spécifique de la dépendance parabolique du potentiel. Selon la règle de Bohr, toutes les transitions en émission comme en absorption seront caractérisées par une pulsation multiple entier de ω , spécifique de l'oscillateur considéré. Toutes les observables dynamiques auront de ce fait un comportement strictement périodique : toutes, en effet, pourront s'exprimer sous la forme d'une certaine *série* de Fourier.

On sait que les énergies des niveaux du puits infini augmentent comme n^2 ; pour l'oscillateur harmonique, elles croissent seulement comme n . On retiendra l'image suivant laquelle plus le potentiel diverge rapidement, plus les niveaux d'énergie s'écartent vite les uns des autres ; en quelque sorte, resserrer le potentiel expulse les niveaux vers le haut.

Un aspect non prévu par la première forme de l'Ancienne Théorie des Quanta¹⁸ est la valeur non-nulle de l'énergie de l'oscillateur dans son état fondamental :

$$E_{n=0} = \frac{1}{2} \hbar\omega . \quad (17.30)$$

Cette énergie, dite de point zéro, est proportionnelle à la constante de Planck. Cette énergie résiduelle, essentielle, est une manifestation des relations d'incertitude de Heisenberg et représente l'irréductibilité des fluctuations quantiques.

L'analyse du rôle des différentes conditions aux limites montre bien que tous les états acceptables sont associés à une valeur quantifiée de l'énergie : un état non-lié peut ne pas être sommable, mais il doit rester borné ; or on a vu que si la série a une infinité de termes non nuls (condition *sine qua non* pour que l'énergie soit quelconque, non quantifiée), la fonction propre correspondante diverge comme $e^{\frac{\xi^2}{2}}$ à l'infini. Autrement dit, comme prévu, il n'existe pas, pour l'oscillateur harmonique, d'états non-liés ; ceci résulte évidemment du fait que le potentiel devient arbitrairement grand à grande distance : aucun état ne saurait exister, où la particule liée harmoniquement aurait la possibilité de s'échapper¹⁹.

Il reste à dire quelques mots sur les différentes fonctions propres elles-mêmes. Compte tenu des définitions de ξ , u , \dots , toute fonction propre – indicée par le seul entier n puisqu'il n'y a pas de dégénérescence – est de la forme :

$$\psi_n(x) = C_n H_n \left(\sqrt{\frac{m\omega}{\hbar}} x \right) e^{-\frac{m\omega x^2}{2\hbar}} . \quad (17.31)$$

C_n est une constante de normalisation²⁰, assurant que l'intégrale de $|\psi_n|^2$ est égale à 1 ; C_n dépend aussi de la façon précise dont sont définis les polynômes H_n (voir plus loin). ψ_n est ainsi le produit d'un polynôme de degré n – qui oscille –, et d'une gaussienne s'éteignant sur la distance fixe $\sqrt{\hbar/(m\omega)}$, indépendante de l'entier n .

¹⁷déjà obtenu par Heisenberg dans sa Mécanique des Matrices.

¹⁸Toutefois, le calcul précis dans l'approximation semi-classique donne effectivement le terme additif $1/2$.

¹⁹Cette conclusion est tirée en dimension 1 : en dimension supérieure (3 par exemple) il conviendrait d'examiner le rôle du terme centrifuge (énergie cinétique de rotation) ; en fait, la conclusion reste la même : un oscillateur à trois dimensions peut toujours se décomposer en trois oscillateurs linéaires indépendants et pour chacun d'entre eux, il n'existe que des états liés.

²⁰L'expression normalisée de ψ_n est donnée en (17.52).

Les polynômes H_n sont bien connus et portent le nom de polynômes de Hermite ; ils satisfont l'équation différentielle homogène :

$$H_n''(\xi) - 2\xi H_n'(\xi) + 2n H_n(\xi) = 0 . \quad (17.32)$$

déduite de (17.18) avec $\lambda = 2n + 1$, $n \geq 0$. Ils sont de la forme :

$$H_n(\xi) = \xi^s \sum_{p=0}^{p_0(n)} a_{2p} \xi^{2p} , \quad (17.33)$$

avec les correspondances²¹ :

$$s = \begin{cases} 0 , & \lambda = 4p_0 + 1 = 2n + 1 \iff n \text{ pair et } p_0 = \frac{n}{2} \\ 1 , & \lambda = 4p_0 + 3 = 2n + 1 \iff n \text{ impair et } p_0 = \frac{n-1}{2} \end{cases} \quad (17.34)$$

de sorte que finalement :

$$H_n(\xi) = \begin{cases} \sum_{p=0}^{n/2} a_{2p}^{(n)} \xi^{2p} & (n \text{ pair}) \\ \xi \sum_{p=0}^{(n-1)/2} a_{2p}^{(n)} \xi^{2p} & (n \text{ impair}) \end{cases} \quad (17.35)$$

Les H_n sont donc alternativement pairs ($s = 0$) et impairs ($s = 1$) en ξ selon que l'entier n (positif ou nul) est pair ou impair, leur degré étant dans tous les cas égal à n . Comme ils satisfont une équation différentielle homogène, chacun d'entre eux est défini à un facteur près ; le facteur global arbitraire d'un H_n donné n'est, pour l'instant, en aucune façon relié à celui des autres (l'équation différentielle caractéristique ne dépend que du n considéré). Par la suite, c'est une convention *supplémentaire* (très précisément le choix du coefficient du terme de plus haut degré, à savoir $a_n^{(n)} = 2^n$) qui permettra de relier un H_n à ses "voisins" $H_{n\pm 1}$; ceci étant fait, on peut établir la récurrence importante caractéristique :

$$H_n' = 2n H_{n-1} . \quad (17.36)$$

En quelque sorte, l'équation différentielle fondamentale définit une large famille de polynômes dont la parenté (indirecte) se limite à satisfaire cette équation ; la convention supplémentaire permet de relier entre eux les polynômes, suivant une filiation directe.

Un outil extrêmement commode en pratique²² est ce que l'on appelle la fonction génératrice des $H_n(\xi)$, notée $F(t, \xi)$ et définie comme :

$$F(t, \xi) = \sum_{n \in \mathbb{N}} \frac{H_n(\xi)}{n!} t^n . \quad (17.37)$$

Il est clair que F contient l'information globale sur les H_n et que ceux-ci peuvent s'en déduire – on verra plus loin en détail comment obtenir les H_n à partir de F . La question importante est donc de trouver une équation pour F . Cette dernière s'établit en utilisant précisément la récurrence (17.36). Un calcul simple montre que :

$$\frac{\partial}{\partial \xi} F(t, \xi) = 2t F(t, \xi) . \quad (17.38)$$

Cette équation différentielle s'intègre immédiatement et donne :

$$F(t, \xi) = e^{2t\xi} F(t, 0) . \quad (17.39)$$

²¹Pour une valeur de n donnée, l'équation différentielle (17.13) (du second ordre) a deux solutions linéairement indépendantes, correspondant aux deux valeurs possibles de s (0 ou 1) ; mais seule l'une d'entre elles se réduit comme il le faut à un polynôme : c'est celle qui correspond à $s = 0$ (resp. $s = 1$) suivant que n est pair (resp. impair). C'est en effet cette association qui assure que a_{2p} , par la relation de récurrence, sera bien nul à partir d'un certain rang. En quelque sorte, pour un n donné, seule la "moitié" des solutions de l'équation différentielle est acceptable ; ceci est finalement assez usuel : lors de l'étude des puits et des barrières, on a éliminé l'exponentielle divergente, ce qui revient à ne considérer que l'une des deux familles de solutions de l'équation du second ordre.

²²Notamment pour vérifier sans peine que deux fonctions ψ_n et $\psi_{n'}$, $n \neq n'$, sont bien orthogonales et pour trouver la constante de normalisation C_n dans (17.31).

La valeur $F(t, 0)$ s'obtient en sachant que²³ :

$$H_n(0) = \begin{cases} (-1)^{n/2} \frac{n!}{(n/2)!} & \text{si } n \text{ est pair} \\ 0 & \text{si } n \text{ est impair} \end{cases} \quad (17.40)$$

d'où l'on déduit facilement :

$$F(t, 0) \equiv \sum_{n \in \mathbb{N}} \frac{H_n(0)}{n!} t^n = e^{-t^2} . \quad (17.41)$$

Il en résulte :

$$F(t, \xi) = e^{2t\xi} e^{-t^2} = e^{+\xi^2} e^{-(t-\xi)^2} . \quad (17.42)$$

Montrons maintenant comment obtenir effectivement H_n à partir de $F(t, \xi)$. Par la définition même de F , (17.37), on a :

$$H_n(\xi) = \left[\frac{\partial^n}{\partial t^n} F(t, \xi) \right]_{t=0} . \quad (17.43)$$

de sorte que, suivant (17.42) :

$$H_n(\xi) = e^{+\xi^2} \left[\frac{\partial^n}{\partial t^n} e^{-(t-\xi)^2} \right]_{t=0} = e^{+\xi^2} (-1)^n \left[\frac{\partial^n}{\partial \xi^n} e^{-(t-\xi)^2} \right]_{t=0} = (-1)^n e^{+\xi^2} \frac{\partial^n}{\partial \xi^n} e^{-\xi^2} . \quad (17.44)$$

Cette dernière relation est d'ailleurs souvent prise comme définition des H_n . Elle montre de surcroît que les zéros de H_n sont tous réels. En effet, $e^{-\xi^2}$ tend vers zéro à l'infini, ainsi que toutes ses dérivées ; supposons maintenant que la $(n-1)$ ^{ème} dérivée, $d^{n-1}e^{-\xi^2}/d\xi^{n-1}$, ait exactement $n-1$ zéros réels ; la dérivée de cette fonction a au moins un zéro réel de plus ; mais comme cette dérivée $n^{\text{ème}}$ est un polynôme de degré n simplement multiplié par $e^{-\xi^2}$, elle a exactement n zéros réels ou complexes : tous les zéros de H_n sont donc réels. Chaque H_n ayant tous ses zéros réels, la fonction d'onde ψ_n représentant le $n^{\text{ème}}$ état excité a donc exactement n noeuds à distance finie.

Quelques relations se révéleront utiles dans la suite :

1. En utilisant la relation de Fourier :

$$\int_0^{2\pi} \frac{d\phi}{2\pi} e^{i(n-n')\phi} = \delta_{nn'} , \quad (n, n') \in \mathbb{Z}^2 , \quad (17.45)$$

on obtient une autre expression de H_n à l'aide de la fonction génératrice :

$$H_n(\xi) = n! \int_0^{2\pi} \frac{d\phi}{2\pi} e^{-in\phi} F(t = e^{i\phi}, \xi) . \quad (17.46)$$

2. En utilisant l'équation différentielle (17.32) et en dérivant par rapport à ξ la relation de récurrence fondamentale (17.36), on peut facilement établir d'autres équations de récurrence telles que :

$$H'_n(\xi) = 2\xi H_n(\xi) - H_{n-1}(\xi) , \quad 2\xi H_n(\xi) = H_{n+1}(\xi) + 2n H_{n-1}(\xi) . \quad (17.47)$$

3. Les polynômes de Hermite possèdent bien d'autres propriétés (voir par exemple [24]) ; il est notamment possible d'en donner différentes représentations, sous forme d'intégrales par exemple. Une expression de cette sorte est la suivante²⁴ :

$$H_n(\xi) = \pi^{-1/2} e^{\xi^2} \int_{-\infty}^{+\infty} du (-2iu)^n e^{-u^2+2iu\xi} . \quad (17.48)$$

²³ $H_{n \text{ impair}}(0)$ est évidemment nul ; quant à $H_{n \text{ pair}}(0)$, la valeur donnée en (17.40) est celle usuellement choisie en théorie de ces polynômes. Elle assure que la fonction génératrice $F(t, \xi)$ a une forme simple, débarrassée de tout facteur inessentiel.

²⁴Avec cette expression intégrale de ψ_n , il est facile d'établir explicitement la relation de fermeture des fonctions propres $\sum_n \psi_n^*(x) \psi_n(x') = \delta(x-x')$ - ce qui est une autre façon de se convaincre que tous les états acceptables physiquement sont bien les seuls états discrets déjà obtenus.

4. Les premiers polynômes de Hermite sont²⁵ :

$$H_0(\xi) = 1, \quad H_1 = 2\xi, \quad H_2 = 4\xi^2 - 2, \quad H_3 = 8\xi^3 - 12\xi, \quad H_4(\xi) = 16\xi^4 - 48\xi^2 + 12. \quad (17.49)$$

$$H_5 = 32\xi^5 - 160\xi^3 + 120\xi, \quad H_6(\xi) = 64\xi^6 - 480\xi^4 + 720\xi^2 - 120. \quad (17.50)$$

Les coefficients deviennent très vite très grands. Par exemple, H_{10} (intervenant sur la fig. 17.1) est :

$$H_{10} = 1024\xi^{10} - 23040\xi^8 + 161280\xi^6 - 403200\xi^4 + 302400\xi^2 - 30240. \quad (17.51)$$

Muni de cet arsenal, il n'est pas difficile de montrer que les fonctions propres normalisées²⁶ ψ_n sont très précisément données par :

$$\psi_n(x) = (2^n n!)^{-1/2} \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) e^{-\frac{m\omega x^2}{2\hbar}}. \quad (17.52)$$

Comme le polynôme H_n a exactement n zéros réels, les fonctions propres s'annulent n fois. La densité de probabilité $|\psi_n(x)|^2$ présente donc des oscillations de plus en plus nombreuses au fur et à mesure que l'on monte dans les états excités, sur une même longueur (l'échelle de longueur est fixée une fois pour toutes, c'est $\sqrt{\hbar/(m\omega)}$). Le facteur exponentiel $e^{-(\frac{\xi^2}{2})}$ est indépendant de n , de sorte que, globalement, $|\psi_n(x)|^2$ sera sensiblement non-nul sur un intervalle en ξ indépendant de n et valant quelques unités (en fait, dès que ξ vaut environ 6, $|\psi_n(x)|^2$ est pratiquement nul). Les oscillations seront donc à la fois plus nombreuses et plus resserrées quand n devient grand.

Cette image de la densité de probabilité de présence est très différente de celle que l'on peut construire avec la théorie classique. Dans un tel cadre, on peut définir $\rho_{cl}(x)$, probabilité de présence d'une particule oscillant suivant²⁷ $x_0 \cos \omega t$, en considérant la fraction de temps (relativement à une période T) passée au point d'abscisse x , à dx près ; sur de simples considérations dimensionnelles, cette fraction est proportionnelle à $1/(|v|T)$, où x est la vitesse au point d'abscisse x ; il faut ajouter un facteur 2 pour tenir compte des deux signes possibles de la vitesse (quand la particule est près de l'abscisse x , elle peut avoir une vitesse dirigée vers la droite ou vers la gauche) :

$$\rho_{cl}(x) = \frac{2}{|v|T}. \quad (17.53)$$

La conservation de l'énergie s'écrit :

$$\frac{1}{2}mv^2 + \frac{1}{2}m\omega^2 x^2 = \frac{1}{2}m\omega^2 x_0^2 \iff |v| = \omega \sqrt{x_0^2 - x^2}. \quad (17.54)$$

Comme $\omega T = \pi$, il vient :

$$\rho_{cl}(x) = \frac{1}{\pi} \frac{1}{\sqrt{x_0^2 - x^2}}. \quad (17.55)$$

Cette densité²⁸ diverge aux points où la particule classique fait demi-tour, exprimant le fait qu'elle passe relativement beaucoup de temps en ces points où la vitesse s'annule. $\rho_{cl}(x)$ est évidemment identiquement nulle pour $|x|$ supérieur à x_0 (région interdite classiquement).

Pour effectuer la comparaison entre cette densité classique et la densité quantique $|\psi_n(x)|^2$, il convient de se fixer un critère rendant significative ce rapprochement. Le plus naturel est de comparer deux situations de même énergie (totale), l'une classique, l'autre quantique. Suivant (17.54) et (17.29), on pose donc :

$$\frac{1}{2}m\omega^2 x_0^2 = \left(n + \frac{1}{2}\right) \hbar\omega, \quad (17.56)$$

²⁵ Attention aux coquilles dans [24].

²⁶ Conformément aux résultats généraux, toutes les fonctions propres sont réelles.

²⁷ correspondant au cas où la particule est lâchée en $x = x_0$ avec une vitesse nulle. L'énergie est donc $(1/2)m\omega^2 x_0^2$.

²⁸ On vérifie facilement que la densité $\rho_{cl}(x)$ est normalisée à 1 pour $x \in [-x_0, +x_0]$.

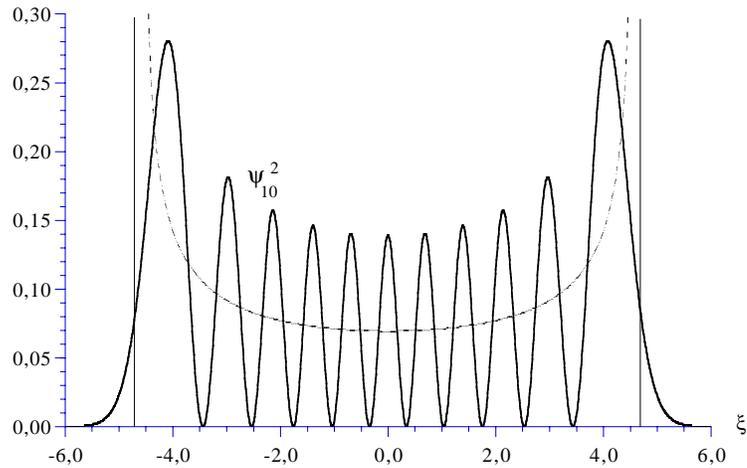


Figure 17.1: Comparaison de la densité de probabilité de présence pour l'état $n = 10$ ($E = (21/2)\hbar\omega$) (trait continu) et de la densité classique d'une particule de même énergie (pointillés).

Ceci ce qui permet de relier biunivoquement x_0 et E_n , donc de comparer précisément $\rho_{cl}(x)$ et $|\psi_n(x)|^2$. On constate d'abord que la densité $|\psi_n(x)|^2$ est très petite là où la densité $\rho_{cl}(x)$ est strictement nulle. Par ailleurs, les différences sont très marquées pour les petites valeurs de n : ρ_{cl} est une fonction relativement régulière (à ses divergences près aux bornes), tandis que $|\psi_n(x)|^2$ oscille et s'annule n fois, toujours sur le même intervalle. Ces différences tendent à s'estomper quand n devient grand : quand n augmente, si $|\psi_n|^2$ se met à osciller de plus en plus vite, il apparaît toutefois que ρ_{cl} passe à peu près à mi-hauteur par rapport aux maxima et aux zéros de $|\psi_n|^2$. En d'autres termes, si on lisse les "fluctuations" de $|\psi_n|^2$ par une intégration sur un intervalle d'ordre la distance entre deux zéros (ou maxima) consécutifs, la moyenne ainsi obtenue et ρ_{cl} ont tendance à se rapprocher l'une de l'autre quand n augmente : on retrouve l'habituelle fusion entre les résultats classiques et quantiques à la limite des très grands nombres quantiques, avec toutefois la persistance, même dans cette limite, d'effets hautement singuliers²⁹ (voir fig.17.1).

Il est facile de comprendre pourquoi, quand $n \gg 1$, $|\psi_n|^2$ oscille beaucoup au voisinage du centre, avec une amplitude lentement variable. En effet, la région centrale définie comme :

$$\frac{1}{2}m\omega^2 x^2 \ll E_n \quad (17.57)$$

est relativement étendue si $n \gg 1$ et, pour ces valeurs de x , l'énergie cinétique est dominante, de sorte que E_n peut être remplacé par $p_0^2/(2m)$ où $p_0 = m\omega x_0$ représente l'impulsion maximum de la particule classique. Dans cette région, l'équation aux valeurs propres est approximativement :

$$-\frac{\hbar^2}{2m} \psi'' \simeq E_n \psi(x) \simeq \frac{p_0^2}{2m} \psi(x) \iff \psi'' + \frac{p_0^2}{\hbar^2} \psi(x) \simeq 0 \quad (17.58)$$

et est formellement la même que celle d'une particule libre³⁰ ; dans cette zone centrale :

$$\psi(x) \simeq A e^{\frac{i}{\hbar} p_0 x} + B e^{-\frac{i}{\hbar} p_0 x} \quad (17.59)$$

Une telle fonction a un comportement oscillant ; par ailleurs, pour ces états fortement excités, l'impulsion sera relativement peu dispersée autour des deux valeurs $\pm p_0$.

Il est également instructif d'examiner les valeurs moyennes de l'impulsion et de la coordonnée dans les différents états propres. On note tout d'abord que, compte tenu de la symétrie (paire ou impaire) des fonctions propres, toutes les valeurs moyennes $\langle x^{2r+1} \rangle$ sont nulles ($r \in \mathbb{N}$) ; de surcroît, toutes les fonctions propres peuvent être choisies réelles, et, par conséquent, toutes les moyennes $\langle p^{2r+1} \rangle$ sont également nulles, d'où :

$$\Delta x^2 = \langle x^2 \rangle, \quad \Delta p^2 = \langle p^2 \rangle. \quad (17.60)$$

²⁹L'extrême singularités des corrections quantiques se mesure au fait que les oscillations sont de plus en plus nombreuses quand n augmente, mais se produisent sur une échelle de longueur fixe, indépendante de n . Pour une résolution spatiale donnée δx , les oscillations deviennent de fait invisibles si $n \gtrsim \sqrt{\hbar/(m\omega)}/\delta x$. Ce type de phénomène – comportement quasi-classique avec persistance d'oscillations rapides à petite échelle, ici en position, là en énergie –, a déjà été observé lors de l'étude des coefficients de réflexion et de transmission pour le puits carré et la barrière de potentiel.

³⁰Dans la région centrale, le potentiel est relativement plat ; si en plus l'énergie est très élevée, tout se passe à peu près (dans la zone centrale) comme si la particule était libre.

D'un autre côté, par le théorème du Viriel, les valeurs moyennes de l'énergie cinétique et de l'énergie potentielle sont égales entre elles et valent donc chacune la moitié de l'énergie totale :

$$\left\langle \frac{p^2}{2m} \right\rangle = \left\langle \frac{1}{2} m \omega^2 x^2 \right\rangle = \frac{1}{2} \left(n + \frac{1}{2} \right) \hbar \omega , \quad (17.61)$$

d'où :

$$\langle p^2 \rangle = \left(n + \frac{1}{2} \right) m \hbar \omega \quad \langle x^2 \rangle = \left(n + \frac{1}{2} \right) \frac{\hbar}{m \omega} , \quad (17.62)$$

et par conséquent :

$$\Delta x \Delta p = \left(n + \frac{1}{2} \right) \hbar . \quad (17.63)$$

En particulier, pour l'état fondamental ($n = 0$) :

$$\Delta x \Delta p = \frac{\hbar}{2} . \quad (17.64)$$

On sait d'un théorème général démontré antérieurement³¹ que $\hbar/2$ est la plus petite valeur possible dans l'absolu du produit des incertitudes sur la position et l'impulsion. L'état fondamental de l'oscillateur harmonique (de forme strictement gaussienne) représente la distribution "minimale" compatible avec les relations d'incertitude.

Remarque

Le Hamiltonien (17.4) est remarquable au sens où x et p y jouent des rôles symétriques³² : chacune de ces grandeurs fondamentales apparaît de la même façon (au carré). Il en résulte que, à des recalages près des constantes physiques, l'équation aux fonctions propres a la même forme dans les représentations- q et $-p$. L'une est (17.5), l'autre s'écrit :

$$\frac{p^2}{2m} \phi(p) - \frac{1}{2} m \hbar^2 \omega^2 \frac{d^2 \phi}{dp^2} = E \phi(p) . \quad (17.65)$$

Dans le même esprit que pour la représentation- q , introduisons des grandeurs sans dimension :

$$\varpi = \frac{p}{\sqrt{m \hbar \omega}} , \quad \phi(p) = v(\varpi) , \quad E = \frac{1}{2} \lambda \hbar \omega ; \quad (17.66)$$

alors, (17.65) prend la forme :

$$\frac{d^2 v}{d\varpi^2} + (\lambda - \varpi^2) v(\varpi) = 0 , \quad (17.67)$$

strictement identique à (17.13). Bien évidemment, les conditions requises pour $\phi(p)$ sont les mêmes que celles pour $\phi(x)$. Les conditions aux limites et l'équation étant identiques, on obtient $\phi(p)$ par simple substitution dans (17.52) des grandeurs appropriées. Ainsi³³, à une phase globale près :

$$\phi_n(p) = (2^n n!)^{-1/2} \left(\frac{1}{\pi m \hbar \omega} \right)^{1/4} H_n \left(\frac{p}{\sqrt{m \hbar \omega}} \right) e^{-p^2/(2m \hbar \omega)} . \quad (17.68)$$

Il est facile de vérifier que (à une phase éventuelle près), l'expression (17.68) est bien la transformée de Fourier de $\psi_n(x)$ (17.52), conformément à la relation générale :

$$\phi_n(p) = \frac{1}{\sqrt{2\pi \hbar}} \int_{-\infty}^{+\infty} dx e^{-\frac{i}{\hbar} p x} \psi_n(x) . \quad (17.69)$$

³¹Si $[A, B] = iC$, alors $\Delta A \Delta B \geq |\langle C \rangle|/2$ (voir [7]).

³²Cette propriété sera exploitée à fond dans la section 17.3.

³³Pour plus de détails sur la représentation- p de l'oscillateur harmonique, voir [16], complément D_V.

17.3 Quantification canonique, opérateurs de création et d'annihilation

La procédure qui va être décrite est due à Dirac³⁴ et utilise systématiquement le fait que le Hamiltonien de l'oscillateur harmonique (17.4) est quadratique en x et p et peut même recevoir une forme plus symétrique en définissant de nouveaux opérateurs hermitiques X et P sans dimension :

$$X = \sqrt{\frac{m\omega}{\hbar}} x , \quad P = \frac{1}{\sqrt{m\hbar\omega}} p . \quad (17.70)$$

La relation de commutation fondamentale $[x, p] = i\hbar\mathbf{1}$ revêt maintenant la forme :

$$[X, P] = i\mathbf{1} . \quad (17.71)$$

et, avec ces variables, (17.4) s'écrit simplement :

$$H = \frac{\hbar\omega}{2}(X^2 + P^2) . \quad (17.72)$$

Introduisons maintenant deux opérateurs notés a et a^\dagger , hermitiques conjugués l'un de l'autre :

$$a = \frac{1}{\sqrt{2}}(X + iP) , \quad a^\dagger = \frac{1}{\sqrt{2}}(X - iP) . \quad (17.73)$$

Les relations inverses s'écrivent :

$$X = \frac{1}{\sqrt{2}}(a + a^\dagger) , \quad P = \frac{1}{i\sqrt{2}}(a - a^\dagger) , \quad (17.74)$$

et aussi :

$$x = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger) , \quad p = \frac{1}{i}\sqrt{\frac{m\hbar\omega}{2}}(a - a^\dagger) . \quad (17.75)$$

Les deux opérateurs a et a^\dagger ne sont pas hermitiques, à cause du facteur i devant P (qui est hermitique). Leur relation de commutation se déduit de celle de X et P ; on trouve facilement :

$$[a, a^\dagger] = \mathbf{1} . \quad (17.76)$$

Le Hamiltonien s'exprime très simplement à l'aide des a et a^\dagger . En effet, partant de (17.72) et en utilisant (17.74) :

$$H = \frac{\hbar\omega}{2} \left[\left(\frac{1}{\sqrt{2}}(a + a^\dagger) \right)^2 + \left(\frac{1}{i\sqrt{2}}(a - a^\dagger) \right)^2 \right] = \frac{\hbar\omega}{2}(aa^\dagger + a^\dagger a) . \quad (17.77)$$

En remplaçant aa^\dagger par $aa^\dagger + 1$ en vertu de (17.76), il vient finalement l'expression remarquable :

$$H = \hbar\omega \left(a^\dagger a + \frac{1}{2} \right) \equiv \hbar\omega \left(N + \frac{1}{2} \right) . \quad (17.78)$$

L'opérateur $N = a^\dagger a$ est manifestement hermitique et sans dimension ; par comparaison avec les résultats obtenus dans la section 17.2 (et notamment l'expression des énergies possibles (17.29)), on voit que le spectre de N est constitué de tous les entiers positifs ou nuls et que ses vecteurs propres sont les mêmes que ceux de H . Toutefois, il est très instructif de retrouver ceci en raisonnant directement sur N en procédant comme suit ; cette démarche à l'élégance extrême est due à Dirac.

En utilisant l'identité $[AB, C] = A[B, C] + [A, C]B$ on obtient :

$$[N, a] = [a^\dagger a, a] = a^\dagger[a, a] + [a^\dagger, a]a = -a . \quad (17.79)$$

³⁴C'est le point de départ de sa théorie quantique du champ électromagnétique.

et, comme $[N, a^\dagger] = -[N^\dagger, a]^\dagger = -[N, a]^\dagger$, on a aussi :

$$[N, a^\dagger] = a^\dagger . \quad (17.80)$$

Soit $|\nu\rangle$ un vecteur propre de N , *supposé de norme finie* :

$$N|\nu\rangle = \nu|\nu\rangle . \quad (17.81)$$

En utilisant (17.79), il vient :

$$Na|\nu\rangle \equiv ([N, a] + aN)|\nu\rangle = -a|\nu\rangle + a\nu|\nu\rangle = (\nu - 1)a|\nu\rangle . \quad (17.82)$$

ce qui se lit :

$$N(a|\nu\rangle) = (\nu - 1)(a|\nu\rangle) , \quad (17.83)$$

et montre que le vecteur $a|\nu\rangle$ est aussi vecteur propre de N mais cette fois avec la valeur propre $\nu - 1$ – à condition bien sûr que ce vecteur ne soit pas le vecteur nul. La norme carrée de ce vecteur est :

$$\langle a|\nu\rangle, a|\nu\rangle = \langle \nu|a^\dagger a|\nu\rangle \equiv \langle \nu|N|\nu\rangle = \nu \langle \nu|\nu\rangle . \quad (17.84)$$

Comme la norme carrée de tout vecteur est positive ou nulle, on en déduit $\nu \geq 0$: toutes les valeurs propres de l'opérateur N sont donc des nombres réels positifs ou nuls. Maintenant, on peut recommencer ces opérations et considérer :

$$\begin{aligned} Na^2|\nu\rangle &= ([N, a^2] + a^2N)|\nu\rangle = ([N, a]a + a[N, a])|\nu\rangle + a^2N|\nu\rangle \\ &= (-a^2 - a^2)|\nu\rangle + a^2N|\nu\rangle = (\nu - 2)a^2|\nu\rangle . \end{aligned} \quad (17.85)$$

Ceci qui montre que le vecteur $a^2|\nu\rangle$ est propre de N avec la valeur propre $\nu - 2$. Par application répétée de a , on peut ainsi mettre en évidence une suite de vecteurs propres de N ayant les valeurs propres $\nu, \nu - 1, \nu - 2, \dots$. Cette suite ne peut être infinie, puisque l'on vient de montrer que les valeurs propres de N sont toutes positives ou nulles ; si ν n'était pas un entier, on finirait par construire un état propre de N avec une valeur propre strictement négative. Il faut donc que, tôt ou tard, la différence ($\nu -$ un entier) soit nulle, ce qui montre que ν est lui-même un entier, positif ou nul³⁵. Alors, la séquence de vecteurs se termine avec :

$$Na^\nu|\nu\rangle = (\nu - \nu)a^\nu|\nu\rangle = 0 . \quad (17.86)$$

Revenant à (17.83) pour la plus petite valeur propre $\nu = 0$ (état fondamental), on a :

$$Na|0\rangle = (0 - 1)a|0\rangle \iff N(a|0\rangle) = -(a|0\rangle) ; \quad (17.87)$$

mais comme N ne peut avoir de valeur propre négative, ceci n'est possible que si $a|0\rangle$ est le vecteur nul :

$$a|0\rangle = 0 . \quad (17.88)$$

On voit ainsi que l'opérateur a *détruit* l'état fondamental (souvent appelé dans ce contexte : état vide, ou plus simplement, |vide>). D'une façon générale, l'opérateur a engendre, à partir de $|\nu\rangle$, le vecteur propre de N ayant la valeur propre $\nu - 1$, réduite d'une unité par rapport à la première : en quelque sorte, a *annihile* une excitation et fait passer d'un état à l'état immédiatement inférieur en énergie. Comme on ne peut pas détruire ce qui n'existe pas, a agissant sur l'état |vide> (le fondamental) donne zéro.

La même gymnastique peut s'effectuer à partir de l'autre relation, (17.80). On a ainsi :

$$Na^\dagger|\nu\rangle \equiv ([N, a^\dagger] + a^\dagger N)|\nu\rangle = a^\dagger|\nu\rangle + a^\dagger\nu|\nu\rangle = (\nu + 1)a^\dagger|\nu\rangle . \quad (17.89)$$

Le vecteur $a^\dagger|\nu\rangle$ est donc propre de N avec la valeur propre $\nu + 1$; sa norme est :

$$\langle a^\dagger|\nu\rangle, a^\dagger|\nu\rangle = \langle \nu|aa^\dagger|\nu\rangle = \langle \nu|a^\dagger a + 1|\nu\rangle \equiv \langle \nu|N + 1|\nu\rangle = (\nu + 1)\langle \nu|\nu\rangle . \quad (17.90)$$

³⁵Ce résultat s'obtient aussi en se souvenant que l'énergie est supérieure au minimum de $V(x)$, ici égal à zéro. En vertu de (17.78), il faut donc que $\nu + 1/2$ soit toujours positif. Si ν n'était pas entier, tôt ou tard la quantité $\nu + 1/2 -$ entier deviendrait négative.

Cette norme n'est jamais nulle puisque $\nu \geq 0$; par application répétée de a^\dagger , on peut ainsi engendrer une suite infinie de vecteurs propres de N correspondant aux valeurs propres $\nu + 1, \nu + 2, \text{etc.}$ L'opérateur a^\dagger fait ainsi passer d'un état $|\nu\rangle$ à l'état immédiatement supérieur $|\nu + 1\rangle$: il crée une excitation supplémentaire. Par ailleurs, il est clair que le spectre de N n'est pas limité vers le haut.

Au total, le spectre de N est constitué de tous les entiers positifs ou nuls. On notera désormais $\nu = n$ pour rappeler ce fait et on désignera par $|n\rangle$ un vecteur propre normalisé de N :

$$N|n\rangle = n|n\rangle, \quad \langle n|n\rangle = 1. \quad (17.91)$$

Les deux opérateurs a et a^\dagger respectivement annihile et crée une excitation ; pour cette raison, ils sont respectivement appelés opérateur d'annihilation et opérateur de destruction. L'un crée, l'autre annihile un quantum d'excitation : on peut imaginer que, dans un contexte à préciser, ces opérateurs créent ou annihilent des *particules*, la notion de particule étant alors identifiée avec celle de quantum élémentaire (d'un champ). L'opérateur N , lui, "compte" le nombre d'excitations (= nombre de quanta, *i.e.* nombre de particules) : il est pour cette raison naturellement appelé opérateur "nombre de particules". Compte tenu de l'expression de H (17.78), il résulte que H a bien pour valeurs propres les énergies $\hbar\omega(n + 1/2)$.

Remarque

Dans la première présentation (méthode polynômiale de Sommerfeld), le lien entre la nécessité d'avoir des états propres normalisables et la quantification des valeurs propres de H a été clairement mis en évidence. Ce lien est peu visible dans la discussion ci-dessus mais il est évidemment toujours présent. En réalité, on a supposé implicitement dès le début que tous les états propres de N étaient de *norme finie*.

Précisons maintenant l'action des opérateurs a et a^\dagger sur les états propres $|n\rangle$ normalisés, que l'on suppose désormais normalisés à l'unité. Comme $a^\dagger|n\rangle$ est propre de N avec la valeur propre $n + 1$, ce vecteur est proportionnel à $|n + 1\rangle$ et on peut écrire :

$$a^\dagger|n\rangle = C_n|n + 1\rangle \quad (\langle n|n\rangle = 1 \quad \forall n), \quad (17.92)$$

où C_n est une constante à trouver. L'équation (17.90), avec $\nu = n$, montre que le carré de la norme de $a^\dagger|n\rangle$ vaut $n + 1$; d'où, par (17.92) :

$$C_n^2 = n + 1. \quad (17.93)$$

À une phase près, on peut toujours choisir C_n réel positif. Avec cette convention, (17.92) devient :

$$a^\dagger|n\rangle = \sqrt{n + 1}|n + 1\rangle. \quad (17.94)$$

Les éléments de matrice de a^\dagger sur la base³⁶ $\{|n\rangle\}_n$ s'en déduisent :

$$\langle n'|a^\dagger|n\rangle = \sqrt{n + 1}\langle n'|n + 1\rangle = \sqrt{n + 1}\delta_{n', n+1}. \quad (17.95)$$

La matrice de a^\dagger sur la base $\{|n\rangle\}_n$, ordonnée naturellement, est donc la matrice infinie :

$$a^\dagger = \begin{bmatrix} 0 & 0 & 0 & 0 & \dots \\ \sqrt{1} & 0 & 0 & 0 & \dots \\ 0 & \sqrt{2} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}. \quad (17.96)$$

Par application répétée de a^\dagger à l'état fondamental $|n = 0\rangle$, il vient :

$$a^\dagger|0\rangle = \sqrt{1}|1\rangle, \quad (a^\dagger)^2|0\rangle = a^\dagger\sqrt{1}|1\rangle = \sqrt{1.2}|2\rangle, \quad (a^\dagger)^3|0\rangle = \sqrt{1.2.3}|3\rangle \quad \text{etc} \quad (17.97)$$

et, en toute généralité :

$$(a^\dagger)^n|0\rangle = \sqrt{n!}|n\rangle \iff |n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle. \quad (17.98)$$

³⁶C'est bien une base orthonormée : tous les $|n\rangle$ sont propres d'un opérateur hermitique (N, H) et il n'y a pas de dégénérescence.

Par ailleurs, comme a est le hermitique conjugué de a^\dagger et que la base $\{|n\rangle\}_n$ est orthonormée, on a :

$$\langle n|a|n'\rangle = \langle n'|a^\dagger|n\rangle^* = \sqrt{n+1} \delta_{n',n+1} . \quad (17.99)$$

d'où :

$$a|n\rangle = \sqrt{n}|n-1\rangle \quad (17.100)$$

et la matrice de a sur la base $\{|n\rangle\}_n$:

$$a = \begin{bmatrix} 0 & \sqrt{1} & 0 & 0 & \dots \\ 0 & 0 & \sqrt{2} & 0 & \dots \\ 0 & 0 & 0 & \sqrt{3} & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} . \quad (17.101)$$

On retrouve bien que a , agissant sur l'état fondamental (vide) donne zéro : si on veut détruire une excitation (particule) qui n'existe pas, on obtient rien du tout.

Ayant trouvé les matrices de a et a^\dagger , il est facile d'en déduire celles³⁷ de la position x et de l'impulsion p , en utilisant (17.75) :

$$x = \sqrt{\frac{\hbar}{2m\omega}} \begin{bmatrix} 0 & \sqrt{1} & 0 & 0 & \dots \\ \sqrt{1} & 0 & \sqrt{2} & 0 & \dots \\ 0 & \sqrt{2} & 0 & \sqrt{3} & \dots \\ 0 & 0 & \sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} , \quad (17.102)$$

$$p = \sqrt{\frac{m\hbar\omega}{2}} \begin{bmatrix} 0 & -i\sqrt{1} & 0 & 0 & \dots \\ i\sqrt{1} & 0 & -i\sqrt{2} & 0 & \dots \\ 0 & i\sqrt{2} & 0 & -i\sqrt{3} & \dots \\ 0 & 0 & i\sqrt{3} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} . \quad (17.103)$$

Il est facile de vérifier avec ces matrices que $[x, p] = i\hbar\mathbf{1}$.

Etablissons maintenant le lien explicite entre l'approche de Dirac et la formulation polynômiale de Sommerfeld opérant directement sur l'équation de Schrödinger. Comme on l'a vu, l'état $|n\rangle$ peut s'obtenir par application répétée de a^\dagger sur l'état fondamental (voir (17.98)). Par ailleurs, compte tenu de (17.70) et de (17.73), l'expression de a^\dagger en représentation coordonnée est ($p \rightarrow -i\hbar(d/dx)$) :

$$a^\dagger = \frac{1}{\sqrt{2}} \left[\sqrt{\frac{m\omega}{\hbar}} x - i \frac{1}{\sqrt{m\hbar\omega}} (-i\hbar) \frac{d}{dx} \right] . \quad (17.104)$$

Il vient donc, en utilisant (17.98) :

$$\psi_n(x) \equiv \langle x|\psi_n\rangle = \frac{1}{\sqrt{n!}} \langle x|(a^\dagger)^n|0\rangle = \frac{1}{\sqrt{n!}} \langle x| \left[\sqrt{\frac{m\omega}{2\hbar}} x - \sqrt{\frac{\hbar}{2m\omega}} \frac{d}{dx} \right]^n |0\rangle , \quad (17.105)$$

soit, en introduisant à nouveau $\xi = \sqrt{m\omega/\hbar} x$:

$$\psi_n(x) = \frac{1}{\sqrt{n!}} \langle x| \left[\frac{1}{\sqrt{2}} \xi - \frac{1}{\sqrt{2}} \frac{d}{d\xi} \right]^n |0\rangle = \frac{1}{\sqrt{2^n n!}} \langle x| \left[\xi - \frac{d}{d\xi} \right]^n |0\rangle . \quad (17.106)$$

³⁷Les matrices de x et de p peuvent aussi s'obtenir directement en suivant l'approche algébrique de Heisenberg (Mécanique des Matrices).

$|x\rangle$ et $|\xi\rangle$ sont deux vecteurs proportionnels, $|x\rangle = C|\xi\rangle$ ($[C] = L^{-1/2}$). La constante de proportionnalité s'obtient en écrivant le produit scalaire³⁸ :

$$\delta(x - x') \equiv \langle x|x'\rangle = C^2 \langle \xi|\xi'\rangle = C^2 \delta(\xi - \xi') = C^2 \sqrt{\frac{\hbar}{m\omega}} \delta(x - x') , \quad (17.107)$$

d'où :

$$C = \left(\frac{m\omega}{\hbar}\right)^{1/4} \iff |x\rangle = \left(\frac{m\omega}{\hbar}\right)^{1/4} |\xi\rangle . \quad (17.108)$$

L'état fondamental $\psi_0(x)$ est associé à la fonction u correspondant à la valeur propre réduite $\lambda = 1$ et est donc solution de (voir (17.13)) :

$$\frac{d^2u}{d\xi^2} + (1 - \xi^2)u = 0 , \quad (17.109)$$

On vérifie facilement que la seule solution acceptable est une gaussienne, précisément $e^{-\frac{\xi^2}{2}}$, correspondant évidemment à $H_0(\xi) \equiv 1$. La normalisation de cette gaussienne ajoute un facteur $\pi^{-1/4}$, de sorte que l'expression de la fonction d'onde représentant l'état $|n\rangle$ est, selon (17.106) et compte tenu de (17.108) :

$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}} \left(\frac{m\omega}{\hbar}\right)^{1/4} \langle \xi | \left[\xi - \frac{d}{d\xi}\right]^n |0\rangle = \frac{1}{\sqrt{2^n n!}} \left(\frac{m\omega}{\hbar}\right)^{1/4} \left[\xi - \frac{d}{d\xi}\right]^n \pi^{-1/4} e^{-\frac{\xi^2}{2}} . \quad (17.110)$$

Un calcul un peu long mais sans difficulté, montrant notamment (par récurrence) que :

$$\left[\xi - \frac{d}{d\xi}\right]^n e^{-\frac{\xi^2}{2}} = (-1)^n e^{\frac{\xi^2}{2}} \frac{d^n}{d\xi^n} e^{-\xi^2} . \quad (17.111)$$

permet de vérifier que l'expression (17.110) coïncide très exactement avec (17.52).

17.4 Le propagateur de l'oscillateur harmonique

Le propagateur $U(x, t ; x_0, t_0)$ a été défini au chapitre 16 et permet d'écrire la fonction d'onde à l'instant comme une somme (ici une intégrale) d'amplitudes de transition à partir de l'état initial :

$$\Psi(x, t) = \int_{\mathbb{R}} dx_0 U(x, t ; x_0, t_0) \Psi(x_0, t_0) . \quad (17.112)$$

Il est la matrice continue de l'opérateur d'évolution $U(t, t_0)$:

$$U(x, t ; x_0, t_0) = \langle x|U(t, t_0)|x_0\rangle , \quad U(t, t_0) = e^{\frac{1}{i\hbar}H(t-t_0)} . \quad (17.113)$$

En tant que tel, $U(x, t ; x_0, t_0)$ peut se développer sur les états propres de H . Ici, on a (en prenant $t_0 = 0$) et avec $E_n = (n + \frac{1}{2})\hbar\omega$:

$$U(x, t ; x_0, 0) = \sum_{n \in \mathbb{N}} \langle x|\psi_n\rangle e^{-i(n+\frac{1}{2})\omega t} \langle \psi_n|x_0\rangle . \quad (17.114)$$

Compte tenu de l'expression des fonctions propres (17.52), il vient :

$$U(x, t ; x_0, 0) = \frac{1}{2^n n!} \sqrt{\frac{m\omega}{\pi\hbar}} e^{-i\omega t/2} \sum_{n \in \mathbb{N}} e^{-\frac{\xi^2}{2}} e^{-\frac{\xi_0^2}{2}} H_n(\xi) H_n(\xi_0) e^{-in\omega t} . \quad (17.115)$$

Pour calculer la série, on remplace chaque polynôme de Hermite par son expression intégrale (17.48), ce qui fait apparaître une série que l'on peut resommer :

$$\sum_{n=0}^{+\infty} \frac{1}{n!} (-2uu' e^{-i\omega t})^n = e^{-2uu' e^{-i\omega t}} . \quad (17.116)$$

³⁸On utilise au passage $\delta(ax) = |a|^{-1}\delta(x)$.

Ceci étant fait, il reste à calculer des intégrales du genre :

$$\int_{-\infty}^{+\infty} du \int_{-\infty}^{+\infty} du' e^{-2uu'} e^{-i\omega t} e^{-u^2+2iu\xi} e^{-u'^2+2iu'\xi_0} . \quad (17.117)$$

Il s'agit d'intégrales gaussiennes, que l'on sait donc effectuer, par exemple en diagonalisant la forme quadratique en (u, u') apparaissant dans les exponentielles. Après un calcul un peu long mais sans difficulté, on obtient (rétablissant t_0) :

$$U(x, t ; x_0, t_0) = \sqrt{\frac{m\omega}{2i\pi\hbar \sin \omega(t-t_0)}} e^{\frac{i}{\hbar} \frac{m\omega}{2 \sin \omega(t-t_0)} [(x^2+x_0^2) \cos \omega(t-t_0) - 2xx_0]} . \quad (17.118)$$

La quantité figurant dans l'exponentielle, en facteur de i/\hbar , n'est autre que ... l'action classique d'une particule liée harmoniquement, partie de x_0 à $t = t_0$, et parvenue en x à l'instant t ! Ainsi, tout comme dans le cas de la particule libre, le propagateur fait apparaître la quantité caractéristique $S_{cl}(x, t ; x_0, t_0)$. Bien que cette resurgence soit fréquente, il ne s'agit toutefois pas d'une propriété générale et l'on connaît des exemples où le propagateur fait intervenir des "chemins" sans équivalents classiques ([28], p. 39). Par ailleurs, dans la limite $\omega \rightarrow 0$, (17.118) redonne bien le propagateur de la particule libre obtenu au chapitre précédent.

Donnons maintenant un exemple de propagation ; soit à $t = 0$ l'état initial (normalisé, a est l'inverse du carré d'une longueur) :

$$\Psi_a(x, t = 0) = \left(\frac{a}{\pi}\right)^{1/4} e^{-\frac{ax^2}{2}} . \quad (17.119)$$

L'état à l'instant t est alors donné par :

$$\Psi_a(x, t) = \left(\frac{a}{\pi}\right)^{1/4} \sqrt{\frac{m\omega}{2i\pi\hbar \sin \omega t}} e^{\frac{i}{\hbar} \frac{m\omega}{2} x^2 \cot \omega t} \int_{\mathbb{R}} dx' e^{\frac{i}{\hbar} \frac{m\omega}{2 \sin \omega t} (x'^2 \cos \omega t - 2xx')} e^{-\frac{ax'^2}{2}} . \quad (17.120)$$

L'intégrale se calcule sachant que :

$$\int_{\mathbb{R}} dx' e^{-Ax'^2+Bx'} = \sqrt{\frac{\pi}{A}} e^{\frac{B^2}{4A}} \quad (-\pi < \arg A < +\pi) . \quad (17.121)$$

Après calcul explicite, on en déduit l'expression de $\Psi(x, t)$:

$$\Psi_a(x, t) = \left(\frac{a}{\pi}\right)^{1/4} \frac{1}{\sqrt{\cos \omega t + \frac{i\hbar a}{m\omega} \sin \omega t}} \exp\left(-\frac{a}{2} \frac{1 + \frac{im\omega}{a\hbar} \tan \omega t}{1 + \frac{ia\hbar}{m\omega} \tan \omega t} x^2\right) . \quad (17.122)$$

La fonction d'onde est périodique, de période $T = 2\pi/\omega$: toutes les valeurs moyennes sont donc des fonctions périodiques de même période. Ceci est vrai, pour l'oscillateur harmonique, quel que soit l'état initial³⁹.

L'expression (17.122) devient très simple dans le cas où $a = m\omega/\hbar$, car alors l'état de départ n'est autre que l'état propre fondamental $\psi_{n=0}(x)$; pour cette valeur de a , (17.122) devient :

$$\Psi_{a=m\omega/\hbar}(x, t) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-i\frac{\omega t}{2}} \exp\left(-\frac{a}{2} x^2\right) \equiv e^{-i\frac{\omega t}{2}} \Psi(x, 0) , \quad (17.123)$$

comme il se doit puisqu'il s'agit alors d'un état stationnaire.

Dans le cas général, le module carré de $\Psi_a(x, t)$ donne un sentiment global sur l'évolution au cours du temps de l'état préparé. C'est une gaussienne normalisée :

$$|\Psi_a(x, t)|^2 = \frac{1}{\sqrt{2\pi} \Delta x(t)} e^{-\frac{x^2}{2\Delta x^2(t)}} , \quad (17.124)$$

dont l'écart quadratique *oscille* dans le temps :

$$\Delta x^2(t) = \frac{1}{a^2} \cos^2 \omega t + \left(\frac{\hbar}{m\omega}\right)^2 \sin^2 \omega t . \quad (17.125)$$

³⁹On peut s'en convaincre en utilisant les équations du mouvement de Heisenberg, ou en se référant au spectre de H (échelle à barreaux équidistants).

Conformément à la remarque générale faite plus haut, $\Delta x^2(t)$ a une variation périodique. Par ailleurs, comme on l'a vu dans le chapitre 16, de tels résultats peuvent s'obtenir beaucoup plus vite par l'intégration des équations du mouvement de Heisenberg.

Remarque

On verra que l'objet central de la Mécanique Statistique Quantique est un opérateur, l'opérateur densité ρ , qui généralise la fonction de Liouville à la situation quantique. Pour un système canonique, l'opérateur ρ est de la forme $Z^{-1} e^{-\beta H}$, où Z est une facteur de normalisation⁴⁰ et où $\beta = (k_B T)^{-1}$; ρ satisfait donc l'équation (dite de Bloch) :

$$\frac{\partial \rho}{\partial \beta} = -H \rho, \quad (17.126)$$

identique à celle satisfaite par l'opérateur d'évolution $U(t)$ où l'on a fait la substitution⁴¹ $t \rightarrow -i\hbar\beta$. De l'expression du propagateur (17.118), on déduit celle de ρ , à un facteur près :

$$\rho(x, \beta; x') \propto U(x, t = -i\hbar\beta; x', 0) = \sqrt{\frac{m\omega}{2\pi\hbar \sinh \beta\hbar\omega}} e^{-\frac{1}{\hbar} \frac{m\omega}{2 \sinh \beta\hbar\omega} [(x^2 + x'^2) \cosh \beta\hbar\omega - 2xx']} \quad (17.127)$$

L'opérateur convenablement normalisé ($\text{Tr}\rho = 1 \iff \int_{\mathbb{R}} dx \rho(x, \beta; x' = x) = 1$) est :

$$\rho(x, \beta; x') = \sqrt{\frac{m\omega}{\pi\hbar} \tanh \frac{\beta\hbar\omega}{2}} e^{-\frac{1}{\hbar} \frac{m\omega}{2 \sinh \beta\hbar\omega} [(x^2 + x'^2) \cosh \beta\hbar\omega - 2xx']} . \quad (17.128)$$

⁴⁰ Z est la fonction de partition canonique ; le facteur Z^{-1} assure que la trace de ρ est égale à l'unité (c'est une somme de probabilités).

⁴¹On dit que l'on passe en temps imaginaire.

Bibliographie

- [1] Werner HEISENBERG, *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen*, Zeitschrift für Physik, **33**, 879 (1925)
- [2] Max JAMMER, *The Conceptual Development of Quantum Mechanics* (Mc Graw – Hill, New York, 1966)
- [3] Albert EINSTEIN - Max BORN, *Correspondance* (Éditions du Seuil, Paris, 1972)
- [4] Erwin SCHRÖDINGER, *Quantisierung als Eigenwertproblem*, Annalen der Physik, **79**, 489 (1926)
- [5] Erwin SCHRÖDINGER, *Mémoires sur la Mécanique Ondulatoire* (J. Gabay, Paris, 1988)
- [6] D. I. BLOKHINTSEV, *Principes essentiels de la Mécanique Quantique*, (Dunod, Paris, 1968)
- [7] Eugen MERZBACHER *Quantum Mechanics*, (John Wiley, New York, 1970)
- [8] Roger PENROSE, *L'esprit, l'ordinateur et les lois de la Physique*, p. 242 (InterÉditions, Paris, 1992)
- [9] Max BORN, Zeitsch. f. Phys., **38**, 803 (1926)
- [10] Werner HEISENBERG, Zeitsch. f. Phys., **43**, 172 (1927)
- [11] Niels BOHR, Naturwiss., **16**, 145 (1928) ; *ibid.* **17**, 483 (1929) ; *ibid.* **18**, 73 (1930)
- [12] Paul Adrien Marie DIRAC, Proc. Roy. Soc., **A117**, 610 (1928) ; *ibid.* **A118**, 351 (1928)
- [13] Erwin SCHRÖDINGER, *Physique Quantique et représentation du monde* (collection Points-Sciences, Paris, 1992)
- [14] Bernard CAGNAC et Jean-Claude PEBAY-PEYROULA, *Physique Atomique*, tome 1 (Dunod, Paris, 1971)
- [15] David BOHM, *Quantum Theory* (Prentice Hall, Englewood, 1951)
- [16] Claude COHEN-TANNOUDJI, Bernard DIU et Franck LALOË, *Mécanique Quantique* (Hermann, Paris, 1977)
- [17] Edouard CHPOLSKI, *Physique Atomique*, tome 1 (Mir, Moscou, 1977)
- [18] <http://aleph0.clarku.edu/~djoyce/hilbert/problems.html>
- [19] Henri BACRY, *Leçons sur la théorie des groupes et les symétries des particules élémentaires* (Gordon et Breach, New York, 1967)
- [20] Leonard I. SCHIFF, *Quantum Mechanics* (McGraw-Hill, New York, 1968)
- [21] Paul EHRENFEST, Zeit. für Physik., **45**, 455 (1927)
- [22] Richard P. FEYNMAN et A. R. HIBBS *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965)
- [23] Hubert GOLDSTEIN *Classical Mechanics* (Addison-Wesley, New York, 1980)
- [24] Albert MESSIAH, *Mécanique Quantique tome I*, (Dunod, Paris, 1965)
- [25] Lev LANDAU et Evguéni LIFSHITZ, *Mécanique Quantique*, (Mir, Moscou, 1967)

- [26] F. H. SILLINGER et D. R. HERRICKS, *Phys. Rev.*, **11**, 446 (1975).
- [27] Philip M. MORSE et Herman FESHBACH, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953)
- [28] Lawrence S. SCHULMAN, *Techniques and Applications of Path Integration* (John Wiley, New York, 1981)