

Introduction à la mécanique quantique

Cours d'ouverture, EPF 3eme année

Fabien Besnard

6 février 2013

Table des matières

1	Avant-Propos	3
2	La physique classique ou le triomphe du mécanisme	5
3	La période de fermentation : 1900–1923	7
3.1	Les quanta	7
3.1.1	Le problème du corps noir et l’hypothèse de Planck	7
3.1.2	L’effet photoélectrique	10
3.2	Les atomes	11
3.3	Autres arguments	14
4	La transition vers une nouvelle mécanique : 1923–1926	15
4.1	Les ondes de matière	15
4.2	L’équation de Schrödinger	17
4.3	Le hasard et l’incertitude	18
4.3.1	L’expérience des fentes d’Young	19
4.4	La mécanique des matrices	23
4.5	La fusion des deux nouvelles mécaniques	25
5	La formulation mathématique de la mécanique quantique	29
5.1	Introduction	29
5.2	Rappels de mathématiques	29
5.2.1	Bras et Kets	29
5.2.2	Opérateurs sur un espace de Hilbert. Théorème spectral.	31
5.2.3	Codiagonalisation	32
5.2.4	Norme d’opérateur. Exponentielle d’opérateur.	33
5.2.5	Avertissement sur la dimension infinie	33
5.2.6	c -nombres et q -nombres	34
5.3	Les postulats de la mécanique quantique	34
5.3.1	Vecteurs d’états, Espace des états	34
5.3.2	Observables	35
5.3.3	Interprétation probabiliste	36
5.3.4	L’équation d’évolution	37
5.3.5	Réduction du paquet d’onde	39
6	Premières applications des postulats	42
6.1	Espérance et écart-type d’une variable dynamique	42
6.2	Exemple : l’ion H_2^+	43
6.3	Relations d’incertitude et conséquences	44
6.3.1	Énoncé des relations	44
6.3.2	Conséquences	46

6.4	Le théorème d'Ehrenfest	47
6.5	L'oscillateur harmonique	49
6.6	Effet tunnel	53
6.7	Moment cinétique et spin	57
6.7.1	Moment cinétique	57
6.7.2	Moment cinétique orbital. Application à l'atome d'hydrogène.	59
6.7.3	Existence du spin	60
6.7.4	Algèbre du spin 1/2	62
6.7.5	Spin et statistique	64
7	Paradoxes et Interprétations	68
7.1	Chat de Schrödinger	68
7.2	Intrication. Paradoxe EPR.	71
8	Conclusion	76
A	Solutions des exercices	78
B	Condensé de mécanique analytique	83
C	Produit tensoriel	85
D	Constantes	87
	Bibliographie	88

Chapitre 1

Avant-Propos

« I can safely say that nobody understands quantum mechanics. » Richard Feynman

Cette déclaration de l'un des plus grands physiciens quantiques du XX^e siècle a de quoi surprendre. Si Feynman ne comprenait pas la mécanique quantique, qui le pourra ? Si « comprendre » signifie interpréter les résultats de la théorie en termes simples utilisant la logique et le bon sens de la vie de tous les jours, alors Feynman avait sûrement raison, tant la mécanique quantique défie l'intuition. Pourtant, avec un peu d'habileté et beaucoup de courage, on peut en maîtriser le formalisme mathématique et en tirer toutes sortes de prédictions, dont aucune n'a jamais été prise en défaut.

Dans ce cours d'ouverture, nous retracerons les grandes lignes de la découverte de cette théorie stupéfiante. Nous réfléchirons aux célèbres paradoxes liés à la « réduction du paquet d'onde », ou au phénomène d'intrication quantique.

Le niveau sera intermédiaire entre un ouvrage de vulgarisation et un véritable cours d'introduction à la théorie. Notre objectif sera double : d'une part guider les élèves et les accompagner à la découverte d'une théorie réputée difficile, mais dont la connaissance est pourtant indispensable à qui veut comprendre le monde qui l'entoure et la technologie moderne, et d'autre part les confronter aux difficultés conceptuelles liées à son interprétation, car, comme disait Niels Bohr : « Quiconque n'a pas été choqué par la mécanique quantique ne l'a pas encore comprise. »

La confection d'un cours d'introduction à la mécanique quantique est un véritable casse-tête. Le sujet est si vaste qu'il faut nécessairement faire des choix drastiques, en particulier lorsque le temps d'exposition est limité.

Nous avons par exemple choisi de ne pas ou peu parler des applications pratiques, qui sont pourtant légion. Mais d'une part, nous ne disposons pas toujours des compétences requises pour parler de ces questions et nous n'aurions pu que recopier des ouvrages existants, ce qui ne présente guère d'intérêt, et d'autre part nous imaginons que les futurs ingénieurs de l'École y seront confrontés par la suite, et disposeront alors des capacités pour faire le lien nécessaire avec la théorie et les concepts, sur lesquels nous nous sommes focalisés.

Un autre écueil est la place dévolue aux mathématiques. Nous avons choisi de faire peu de rappels et de nous reposer essentiellement sur les connaissances en algèbre linéaire et multilinéaire qui sont forcément excellentes chez tous les élèves de l'EPF !

Enfin, l'exposé de la mécanique quantique elle-même pose problème : doit-on suivre le cheminement historique des idées, de nature inductive, ou poser d'emblée les postulats afin d'en déduire les conséquences par la voie déductive ? Si la seconde option paraît séduisante par sa clarté et son efficacité conceptuelle, elle nous semble trop violente pour l'esprit : les concepts quantiques sont si étranges qu'il est impensable de ne pas donner des motivations très fortes pour les adopter. Sans cela, on court le risque du rejet. Nous avons donc suivi pendant la plus grande partie de ce cours une approche historique, qui n'est cependant pas celle d'un historien des sciences ! En effet, il s'agit plutôt de raconter une histoire, dans laquelle on a décidé a posteriori, et pour la bonne cause, de mettre en avant certains aspects et d'en occulter d'autres, que de raconter l'Histoire, avec tous ses détours, ses fausses pistes et ses balbutiements. En effet, les

arguments qui ont convaincu sur le moment les physiciens des années 1900–1935, période qui recouvre à peu près ce que nous allons raconter, ne sont pas forcément ceux que nous trouvons les plus éclairants aujourd’hui. Or ce sont bien les esprits d’aujourd’hui auxquels sont destinés ces pages. Et nous espérons qu’elles toucheront leurs cibles.

Nous remercions d’avance les lecteurs qui voudront bien nous faire part des erreurs, coquilles, et autres maladresses que nous avons inévitablement commises.

Chapitre 2

La physique classique ou le triomphe du mécanisme

En 1900, la physique que l'on appelle aujourd'hui « classique », née des méditations de Galilée au tout début du XVII^e siècle, est à son apogée. Trois siècles de progrès lui ont permis de rendre compte de la plupart des phénomènes observés, avec une précision parfois extraordinaire. En témoigne par exemple la prédiction de l'existence, et la détermination de la position, de la planète Neptune par Le Verrier, en 1846, uniquement à l'aide des lois de Newton et de l'observation de l'orbite d'Uranus. Mais il y a plus remarquable encore, peut-être, que la précision des prédictions : c'est la marche alors ininterrompue vers l'unification des phénomènes au sein d'un même schéma explicatif. Les phénomènes thermiques, par exemple, ont longtemps été décrits à l'aide d'un mystérieux « fluide calorique ». Mais au milieu du XIX^e siècle, la physique statistique, c'est-à-dire les lois de la mécanique alliées à l'hypothèse des atomes, a permis de comprendre que la vraie nature de la chaleur était l'agitation moléculaire. À peu près au même moment, deux autres phénomènes sont unifiés par Maxwell au sein d'une même théorie : l'électricité et le magnétisme. La théorie de Maxwell permet d'identifier la lumière à une onde électromagnétique. Il suffit alors d'imaginer un support matériel à cette onde, l'éther, pour espérer unifier toute la physique. Ainsi, à l'aube du XX^e siècle, il est possible d'espérer décrire la totalité des phénomènes physiques par les interactions de particules de matière via les forces gravitationnelles et électriques.

Ce tableau d'une physique unifiée et triomphante doit bien sûr être quelque peu tempéré. Il existe alors encore de nombreuses questions sans réponse. Le tableau périodique des éléments, publié en 1869 par Mendeleïev, et la formule de Balmer (1885), tous deux fruits de l'expérience et de l'observation, n'ont pas encore d'explications théoriques. Dans le ciel serein de l'astronomie, tout semble aller pour le mieux, mais certains spécialistes sont pourtant préoccupés par une très légère déviation entre l'orbite théorique et l'orbite observée de la planète Mercure. Avec le recul, nous savons que ces défis n'auraient jamais pu être relevés à l'aide de la physique classique, mais à l'époque, l'espoir était encore permis. Il existait cependant deux autres problèmes où la théorie entraînait cette fois en conflit violent avec l'observation. Pour les présenter, laissons la parole à un éminent représentant de la science de son temps : Lord Kelvin.

*« The beauty and clearness of the dynamical theory, which asserts heat and light to be modes of motion, is at present obscured by two clouds. The first came into existence with the undulatory theory of light [...] It involved the question 'How could Earth move through an elastic solid, such as essentially is the luminiferous ether?' The second is the Maxwell-Boltzmann current doctrine regarding the partition of energy. »*¹ William Thomson (Lord Kelvin) « Nineteenth-century clouds over the dynamical theory of heat and light », 1900

1. « La beauté et la clarté de la théorie dynamique, qui affirme que la chaleur et la lumière sont des modes du mouvement, est à présent obscurcie par deux nuages. Le premier vint à l'existence avec la théorie ondulatoire de la lumière [...] Il met en jeu la question : 'Comment la Terre pourrait-elle se déplacer dans un solide élastique, telle qu'est essentiellement l'éther lumineux ?' Le second est la doctrine actuelle de Maxwell-Boltzmann sur la répartition de l'énergie. »

Le premier nuage identifié par Thomson a finalement conduit Albert Einstein à formuler la théorie de la relativité restreinte (1905). Cette histoire a été contée ailleurs (voir [6]). Le second nuage auquel se réfère Thomson est le problème du corps noir, qui ne sera résolu qu'au prix d'une hypothèse très curieuse, l'hypothèse des « quanta », qui donnera son nom à toute une nouvelle physique.

Nous devons avertir le lecteur que certains physiciens travaillant dans le domaine quantique ont l'habitude de nommer « physique classique » toute la physique basée sur des quantités continues et dont les prédictions sont déterministes. Ceci englobe non seulement toute la physique du XIX^e siècle, mais également la théorie de la relativité restreinte et même la théorie de la relativité générale. Nous pensons qu'il s'agit d'une tendance à minimiser le caractère fondamental et révolutionnaire des deux théories de la relativité. Il nous semble préférable de distinguer quatre physiques : la physique classique, la physique relativiste (relativité restreinte), la relativité générale, et la physique quantique. Dans les pages qui suivent, nous parlerons de la physique quantique en partant de ce que nous supposons connu du lecteur, à savoir les bases de la physique classique. Des considérations relativistes ont parfois joué un rôle important pendant la période de tumulte qui a vu l'émergence de la physique quantique. Nous les passerons cependant volontairement sous silence. Nous reviendrons brièvement en conclusion aux rapports qu'entretiennent physique quantique et relativité.

Chapitre 3

La période de fermentation : 1900–1923

3.1 Les quanta

3.1.1 Le problème du corps noir et l'hypothèse de Planck

Chacun sait que les corps émettent des rayonnements électromagnétiques en fonction de leur température. Un morceau de métal, par exemple, émet dans les infrarouges à température ambiante, puis se met à briller dans le rouge si on le chauffe suffisamment. Pour comprendre ce phénomène, il est utile de définir *les corps noirs*. Par définition, un corps noir est un corps qui absorbe tous les rayonnements électromagnétiques, sans en réfléchir ou en transmettre aucun. Il s'agit bien sûr d'une situation idéalisée, mais un morceau de graphite en est une bonne approximation. On peut également simuler un corps noir à l'aide d'un four (dont les parois internes sont supposées parfaitement réfléchissantes) percé d'un petit trou. Le rayonnement qui entre dans le trou est « piégé » à l'intérieur de la boîte. Le rayonnement piégé entre en équilibre thermique avec le four, et ce qui sort du trou est une très bonne approximation du rayonnement d'un corps noir, et peut être analysé.

De par sa définition même, la seule source de rayonnement d'un corps noir est l'agitation thermique de ses propres molécules (les molécules sont électriquement neutres, mais c'est le mouvement des particules chargées à l'intérieur de celles-ci qui induit le rayonnement¹). Du fait de cette agitation, un corps noir... n'est pas noir ! S'il est assez froid, il apparaît noir à un œil humain parce qu'il n'émet pas dans le spectre visible. La figure 3.1 montre la puissance émise par unité de surface d'un four chauffé à 1000 K dans l'intervalle de fréquence $[\nu; \nu + d\nu]$. La forme « en cloche » de cette courbe se comprend aisément. En effet, l'énergie cinétique moyenne² des particules du four est directement proportionnelle à la température de celui-ci. Ainsi, la distribution des énergies cinétiques ressemble à une courbe en cloche autour de cette moyenne. Comme les particules d'énergie plus élevées émettent dans les fréquences plus hautes, la distribution des énergies émises en fonction de la fréquence a ce même profil.

Le spectre visible est dans la gamme de fréquences 400–790 THz. On voit que dans le cas de la figure 3.1, l'énergie est presque entièrement rayonnée dans les infrarouges. Si on augmente la température du four, le maximum de la courbe va se déplacer vers les hautes fréquences³. La forme de la courbe est toujours à peu près celle d'une cloche. Quand la température est assez grande, l'énergie émise dans les fréquences rouges sera suffisante pour être perceptible, et le four commencera à rougir faiblement, mais l'essentiel du rayonnement émis à ce moment sera toujours dans les infrarouges. Quand le four est suffisamment chaud,

1. Pour décrire le phénomène plus précisément, il nous faudrait anticiper et parler des électrons qui sautent d'une couche à l'autre à l'intérieur des atomes.

2. Calculée dans un référentiel lié au four.

3. On peut montrer que ce déplacement est simplement proportionnel à la température : c'est la loi de Wien.

le pic d'émission se trouve environ au milieu de la gamme visible, ce qui signifie qu'une énergie notable est émise dans toutes les fréquences du spectre visible, et la lumière émise par le four apparaît blanche.

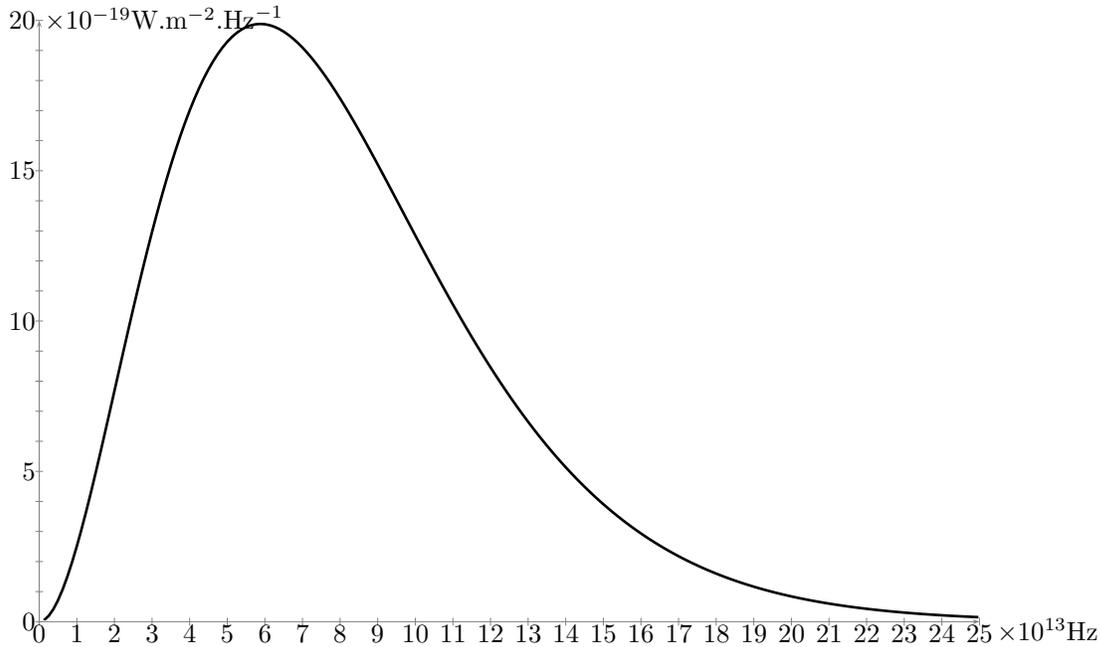


FIGURE 3.1 – Densité de puissance émise par unité de surface d'un corps noir en fonction de la fréquence

Tout ceci est assez intuitif et correspond bien aux observations empiriques que l'on peut faire. Mais un problème survient quand on essaie de retrouver le profil d'émission par la théorie. En effet, on peut penser à utiliser le théorème d'équirépartition de l'énergie. Ce théorème de mécanique statistique nous dit que, pour un système à l'équilibre thermique, chaque degré de liberté contribue de façon égale à l'énergie totale du système. De plus, cette contribution est égale à $\frac{1}{2}kT$, où k est la constante de Boltzmann et T la température absolue. Par exemple, pour un système constitué de molécules, l'énergie cinétique totale provient pour parts égales de leurs mouvement de translation et de rotation respectifs⁴.

Maintenant il y a un gros problème, car la théorie de Maxwell nous enseigne que le nombre de modes d'oscillations du champ électromagnétique⁵ dans un intervalle $d\nu$ est proportionnel à $\nu^2 d\nu$. Quand la fréquence tend vers l'infini (ou la longueur d'onde vers 0), le nombre de degrés de liberté explose, et l'énergie totale émise tend vers l'infini ! Comme le problème surgit pour les courtes longueurs d'onde, ou les hautes fréquences, on a appelé ça la *catastrophe ultraviolette*. Cette catastrophe met en évidence un conflit entre la physique statistique, à travers le théorème d'équirépartition de l'énergie, et la théorie du champ électromagnétique.

En 1900 la plupart des physiciens ne sont toutefois pas prêts à reconnaître l'existence d'une profonde incohérence au sein de la théorie classique. Rappelons-nous qu'il s'agit seulement d'un « petit nuage ». Selon Max Planck, le problème vient d'une application sauvage du théorème d'équirépartition. Planck pense que ce théorème n'est pas universellement valide. Il cherche donc un autre moyen de calculer le profil d'émission du corps noir. Après bien des déboires (cf [11], [16]), Planck finit par se résoudre à adopter une hypothèse *ad hoc* : les échanges d'énergie entre la matière et le rayonnement se font par paquets discrets de valeur $h\nu$, où h est une constante universelle, qui sera plus tard nommée en son honneur. À l'aide de cette hypothèse, le physicien allemand démontre une loi⁶, qui porte aujourd'hui son nom, en excellent accord avec les données expérimentales :

4. Il y a trois degrés de liberté dans chaque cas !

5. Ce nombre s'obtient en déterminant toutes les ondes stationnaires de fréquence donnée vérifiant les conditions aux bords à l'intérieur de la cavité.

6. Pour une démonstration de la loi de Planck on pourra consulter [1], p. 97 et suivantes.

$$M_{\nu}^0(\nu, T) = \frac{2\pi h\nu^3}{c^3} \frac{1}{e^{\frac{h\nu}{kT}} - 1} \quad (3.1)$$

La quantité M_{ν}^0 porte le nom étrange « d'exittance » et s'exprime $\text{W.m}^{-2}.\text{Hz}^{-1}$. Il s'agit de la puissance rayonnée dans tout un demi-espace par un élément de surface de corps noir, porté à la température T , dans l'intervalle de fréquence $[\nu; \nu + d\nu]$. La température T est bien sûr la température absolue, k est la constante de Boltzmann et c est la vitesse de la lumière. Planck est loin de considérer « l'hypothèse des quanta » comme la porte vers une nouvelle physique. Bien au contraire, il estime qu'il s'agit simplement d'un artifice qu'une compréhension plus profonde des échanges entre matière et rayonnement doit pouvoir expliquer. C'est Albert Einstein qui ira plus loin, en considérant que ce ne sont pas seulement les échanges d'énergie qui sont quantifiés, mais l'énergie elle-même, et qu'il s'agit d'un phénomène physique fondamental. Cela lui permettra d'expliquer l'effet photoélectrique, nous y reviendrons.

Mais j'ai gardé le plus beau pour la fin ! La plus belle confirmation de la loi de Planck est venue de la cosmologie. En effet, le fond diffus cosmologique est un rayonnement électromagnétique qui baigne tout l'univers. Il a été émis, non pas au moment du big-bang, comme on le lit trop souvent, mais quelques milliers d'années après, lorsque l'univers était juste assez refroidi pour permettre aux atomes d'hydrogène de se former. Du fait de l'expansion de l'univers, ce rayonnement est aujourd'hui très dilué. On observe qu'il s'agit d'un rayonnement thermique de corps noir à la température de 2,725 K. En fait, c'est le corps noir le plus parfait qu'on connaisse !

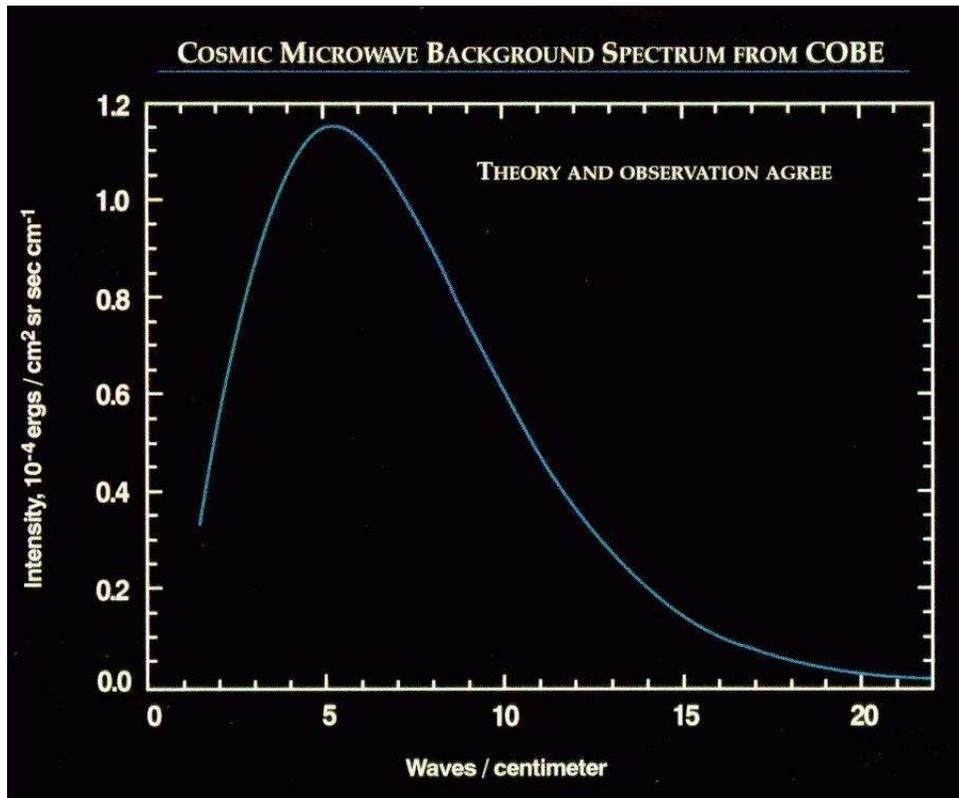


FIGURE 3.2 – Le spectre du fond diffus cosmologique, observé par le satellite COBE.

3.1.2 L'effet photoélectrique

Hertz a observé en 1887 que certains métaux produisent du courant électrique lorsqu'ils sont exposés à de la lumière ultraviolette. Dans les années qui suivirent sa découverte, plusieurs caractéristiques de cet *effet photoélectrique* furent mises en évidence.

1. Plus la fréquence de la lumière est élevée, plus l'énergie cinétique des électrons libérés par le métal est grande.
2. Il existe un seuil de fréquence, dépendant du métal, en dessous duquel aucun effet n'est observé.
3. L'intensité du courant électrique dépend de l'intensité de la lumière, mais...
4. l'énergie cinétique des électrons libérés est indépendante de l'intensité lumineuse.

La théorie ondulatoire de la lumière ne permettait pas de comprendre ces observations. En effet, si on peut facilement imaginer qu'une onde électromagnétique, ce qu'est la lumière selon la théorie ondulatoire, arrache des électrons aux atomes du métal, on peut difficilement comprendre qu'en augmentant l'intensité de cette onde on augmente simplement le nombre d'électrons arrachés, mais pas leur vitesse ! Par ailleurs, il est possible d'envisager un effet de seuil en physique classique : pour s'échapper du métal, les électrons ont besoin d'une certaine énergie qui dépend de la nature de ce métal. La lumière transportant une énergie proportionnelle à son intensité, on s'attendrait classiquement à ce qu'une partie de cette énergie soit transférée aux électrons du métal. Même si l'on peut admettre que l'efficacité de ce transfert dépende de la fréquence, certains électrons devraient quand même réussir à s'échapper avant que le seuil ne soit atteint. On s'attendrait également à l'existence d'un certain délai, correspondant au temps nécessaire aux électrons pour accumuler assez d'énergie. Or le phénomène est discontinu (on n'observe absolument rien en dessous du seuil de fréquence) et immédiat.

En 1905, l'année même où il publie sa théorie de la relativité restreinte et un article sur le mouvement brownien qui prouve l'existence des atomes, Albert Einstein donne une explication très simple de l'effet photoélectrique. Inspiré par Planck, mais plus radical que lui, Einstein fait l'hypothèse que la lumière est constituée de particules transportant chacune une énergie

$$E = h\nu \tag{3.2}$$

où h est la constante de Planck. Einstein appelle « quanta de lumière » ces particules qui seront plus tard baptisées des *photons*. Lors d'un choc, un photon confère toute son énergie à l'électron. Pour s'échapper du métal, celui-ci doit effectuer un travail W . Si $h\nu > W$, l'électron s'échappe avec une énergie cinétique

$$E_c = h\nu - W \tag{3.3}$$

Si on augmente l'intensité lumineuse, on augmente le nombre de photons et donc le nombre d'électrons émis, mais ils ont tous l'énergie cinétique déterminée par (3.3).

L'explication d'Einstein est remarquable de simplicité, mais il est tout simplement impossible d'abandonner la théorie ondulatoire pour revenir à la théorie corpusculaire, car la première explique de nombreux phénomènes dont on ne peut pas rendre compte avec la seconde, notamment les phénomènes d'interférence et de diffraction (pour un exemple original, voir [12] p. 18). C'est pourquoi l'hypothèse des photons a été initialement accueillie avec le plus grand scepticisme. Planck, par exemple, bien que très impressionné par le génie d'Einstein, écrit en 1913 une lettre en faveur de sa candidature à l'Académie prussienne, dans laquelle il s'exprime en ces termes :

« Bien qu'il ait parfois manqué sa cible dans ses spéculations, par exemple dans son hypothèse des quanta de lumière, cela ne peut pas vraiment être retenu contre lui. Car il n'est pas possible d'introduire des idées fondamentalement nouvelles dans les sciences les plus exactes sans prendre parfois un risque. » (cité par [14], p. 382)

Encore plus ironique, c'est dans le but de démolir l'hypothèse du photon que Robert Millikan entreprit, entre 1912 et 1917, de vérifier expérimentalement la relation affine (3.3) prédite par Einstein. À sa grande déception, et malgré sa volonté de mettre en défaut cette relation en variant les conditions expérimentales, Millikan ne fit que la confirmer. Ceci contribua grandement à l'attribution du prix Nobel à Einstein en 1921, et... à Millikan en 1923.

3.2 Les atomes

La découverte de la structure interne de l'atome va poser un autre grave problème à la physique classique. En effet, J.J. Thomson a montré dès 1897 qu'on pouvait arracher des particules électriquement chargées aux atomes : les électrons. Étant donnée la neutralité des atomes, il fallait bien qu'il existât une, ou des, charges positives à l'intérieur de ceux-ci. C'est en 1907 que l'expérience de Rutherford va montrer que la charge positive est concentrée dans un tout petit noyau. Le modèle de l'atome qui émerge alors est celui d'un système solaire en miniature, où la force de Coulomb remplace la gravitation. Dans le cas de l'atome le plus simple, celui d'hydrogène, on a donc un électron en orbite autour d'un noyau positif. Cette image, qui semble émerger d'elle-même de l'expérience, est cependant en contradiction complète avec les lois de l'électromagnétisme. En effet, une charge en mouvement accéléré émet des ondes électromagnétiques et perd en conséquence de l'énergie. L'électron doit donc spiraler vers le noyau en rayonnant, et on peut calculer que ce processus ne prend qu'une fraction de seconde. En d'autres termes, la physique classique est violemment incompatible avec la stabilité de l'atome !

D'autres énigmes provenaient de la spectroscopie. Au cours du XIX^e siècle, des données de plus en plus nombreuses avaient été accumulées sur le spectre d'émission (ou d'absorption, car ce sont les mêmes) des éléments chimiques. Tout d'abord, les raies ont des valeurs bien précises, caractéristiques des éléments étudiés. Après la découverte de la structure interne de l'atome, on comprendra que les raies spectrales sont émises par les électrons, mais d'après la théorie classique du rayonnement on se serait attendu à un spectre continu et pas à un spectre discret. C'est déjà une première énigme. La structure du spectre de l'hydrogène va en fournir une autre, qui est en même temps un indice précieux. En 1885, Johann Jakob Balmer, un professeur de mathématiques de Bâle, découvre une formule empirique qui donne les valeurs des longueurs d'onde des 4 raies visibles du spectre de l'hydrogène :

$$\nu = R_H \left(\frac{1}{4} - \frac{1}{n^2} \right) \quad (3.4)$$

où R_H est une constante (la constante de Rydberg de l'hydrogène) valant approximativement $3,29163 \times 10^{15}$ Hz, et n un entier prenant successivement les valeurs 3, 4, 5 et 6. Un peu plus tard, Rydberg généralisa la formule de Balmer sous la forme

$$\nu = R_H \left(\frac{1}{n_f^2} - \frac{1}{n_i^2} \right) \quad (3.5)$$

où les entiers n_i et n_f vérifient $1 \leq n_f < n_i$. Non seulement cette formule permet de retrouver les raies connues à l'époque, mais elle conduisit à en prédire de nouvelles, dans les infrarouges et les ultraviolets, qui furent progressivement découvertes. La valeur $n_f = 1$ donne la série de Lyman, dont la première raie fut découverte en 1906, $n_f = 2$ donne la série de Balmer, $n_f = 3$ donne la série de Paschen, observée en 1908, etc. On avait donc une formule très simple, qui marchait formidablement bien, mais qu'on ne comprenait pas du tout ! Remarquons également qu'on peut étendre cette formule, à condition de modifier la constante, aux ions possédant un seul électron, comme par exemple l'ion He^+ , mais pas aux atomes ou aux ions possédant plus d'un électron autour du noyau.

Un autre principe dont on ne comprenait pas l'origine avait été indentifié en 1908 par Ritz. Ce dernier a montré qu'il était possible d'établir pour tout atome une suite de fréquences, les termes spectraux, telle que les fréquences des raies spectrales de l'atome soient toujours égales à la somme ou à la différence de deux termes spectraux. On appelle ça le principe de combinaison. La formule de Rydberg en est un cas particulier, la suite des termes spectraux de l'hydrogène étant simplement la suite R_H/n^2 , avec $n \in \mathbb{N}^*$. Notons que la réciproque du principe de combinaison est fautive : toute différence ou somme de nombres de la suite n'est pas forcément une fréquence du spectre.

En 1913, Niels Bohr propose de résoudre ces énigmes en supposant que l'électron ne puisse résider que sur certaines orbites circulaires, ou « couches », correspondant à des énergies bien précises. Autrement dit, l'énergie de l'électron et donc de l'atome est quantifiée. Comme dans l'hypothèse de Planck, les échanges d'énergie se font par paquets discrets. Plus précisément, si un électron passe de la couche j à la couche i , avec $i < j$, il émet une énergie $E_j - E_i$ sous forme de rayonnement de fréquence ν telle que

$$E_j - E_i = h\nu. \quad (3.6)$$

Pour passer de la couche i à la couche j , il doit absorber la même quantité de rayonnement.

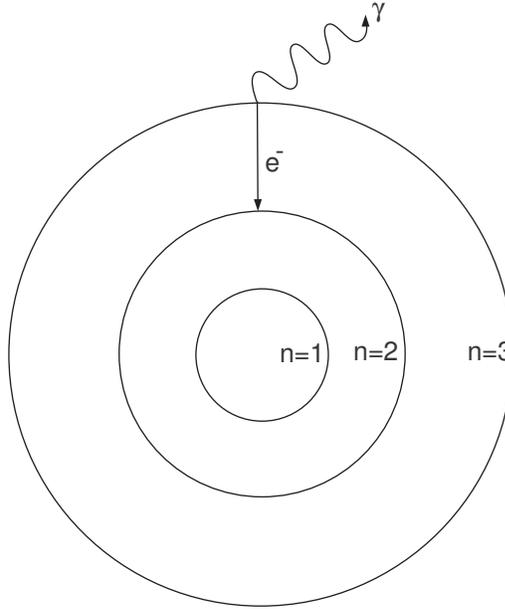


FIGURE 3.3 – Modèle de Bohr. On a représenté un électron passer de la couche 3 à la couche 2 en émettant un photon de fréquence ν telle que $h\nu = E_3 - E_2$.

Remarquons que dans le modèle de Bohr, électrons suivent des trajectoires classiques. Ce qui est en rupture avec la mécanique classique, c'est que seules certaines d'entre elles sont autorisées. Notons également qu'il serait plus simple de décrire ce modèle en parlant d'émission ou de réception de photon par l'électron. C'est d'ailleurs ce qu'on fait très souvent. Mais Bohr ne croyait pas au photon ! En tout cas pas en 1913. Le modèle de Bohr rend compte de l'existence de raies spectrales discrètes satisfaisant au principe de combinaison de Ritz, les termes spectraux étant égaux aux niveaux d'énergie E_n . Dans le cas de l'atome d'hydrogène, on retrouve la loi de Balmer en supposant que

$$E_n = -h \frac{R_H}{n^2}, \quad i \in \mathbb{N}^* \quad (3.7)$$

Notons que dans l'état fondamental ($n = 1$), l'électron ne peut pas émettre de rayonnement, et il n'a donc « pas le droit » de spiraler vers le noyau.

En comparant les niveaux d'énergie E_n à l'énergie mécanique d'un électron sur une orbite circulaire autour d'un proton, égale à⁷

$$E(r) = -\frac{e^2}{8\pi\epsilon_0 r} \quad (3.8)$$

on en déduit le rayon de la n -ième orbite du modèle de Bohr :

$$r(n) = \frac{e^2 n^2}{8\pi\epsilon_0 h R_H} \quad (3.9)$$

7. Ce résultat peut se généraliser aux orbites elliptiques à l'aide du théorème du viriel, voir par exemple [1] p. 161

À l'aide du principe fondamental de la dynamique on calcule facilement la vitesse et la fréquence de la n -ième orbite comme étant :

$$\nu(n) = \frac{e}{4\sqrt{m_e\pi^3\epsilon_0}r(n)^3} = \frac{4\sqrt{2}\epsilon_0 R_H^{3/2} h^{3/2}}{n^3 e^2 m_e^{1/2}} \quad (3.10)$$

Pour n assez grand, la fréquence d'un photon émis lors d'une transition d'un électron entre la couche n et la couche $n - 1$ doit être proche de la fréquence de l'harmonique fondamentale calculée d'après la théorie classique, qui n'est autre que la fréquence $\nu(n)$ de la n -ième orbite. Autrement dit, pour n grand, on doit avoir

$$\frac{1}{h}(E_n - E_{n-1}) \approx \nu(n) \quad (3.11)$$

Or d'après (3.7) on a, pour n grand :

$$\frac{1}{h}(E_{n+1} - E_n) \approx \frac{2R_H}{n^3} \quad (3.12)$$

En comparant (3.11) et (3.12) à l'aide de (3.10), on trouve une relation satisfaite par R_H , d'où l'on peut déduire⁸ :

$$R_H = \frac{m_e e^4}{8\epsilon_0^2 h^3} \quad (3.13)$$

Autrement dit, le modèle de Bohr permet de calculer la constante de Rydberg à partir de constantes fondamentales. La valeur donnée par cette formule⁹ correspond à 10^{-4} près avec la valeur déterminée expérimentalement. Pour être plus précis encore, il faut tenir compte des effets relativistes dus à la vitesse de l'électron. C'est un grand succès pour le modèle de Bohr!

Une fois calculée la constante R_H , il est possible d'exprimer le rayon de la n -ième orbite à l'aide de la formule (3.9) :

$$r(n) = n^2 \frac{h^2 \epsilon_0}{\pi m_e e^2} \quad (3.14)$$

Pour $n = 1$, on parle d'orbite fondamentale et le rayon, qui se note a_0 , s'appelle « rayon de Bohr » et vaut :

$$a_0 = \frac{h^2 \epsilon_0}{\pi m_e e^2} \approx 5,3 \times 10^{-11} \text{m} \approx 0,53 \text{ \AA} \quad (3.15)$$

Ce rayon est typique des échelles de grandeurs en physique atomique. On peut de même calculer l'énergie de l'état fondamental :

$$E_1 = -hR_H = -\frac{m_e e^4}{8\epsilon_0^2 h^2} \approx -13,6 \text{ eV} \quad (3.16)$$

On trouve donc l'énergie de liaison de l'atome d'hydrogène, ce qui peut être confirmé expérimentalement¹⁰. Du reste, la quantification de l'énergie des atomes est confirmée par Frank et Hertz¹¹ dès 1914. Dans cette expérience, on bombarde des atomes (des atomes de mercure dans l'expérience initiale) avec des électrons possédant une énergie cinétique T bien précise. On observe un effet de seuil : si T est au dessous d'une certaine valeur ΔE , les collisions entre électrons et atomes sont toutes élastiques, autrement dit l'énergie

8. Ce résultat s'obtient également à partir de la quantification du moment cinétique, voir [1] p. 162

9. On a supposé que la masse du proton était infinie devant celle de l'électron. Pour plus de précision il faut remplacer la masse de l'électron par la masse réduite $m' = m_e m_p / (m_e + m_p)$.

10. Ce dernier résultat est un peu trop beau pour être vrai étant donné la simplicité du modèle, et apparaît aujourd'hui comme une coïncidence numérique.

11. Il s'agit de Gustav Hertz, neveu de Heinrich Hertz.

cinétique du système atome + électron est conservée. Si $T \geq \Delta E$, certaines collisions sont inélastiques, avec une perte d'énergie cinétique totale exactement égale ΔE . L'interprétation en terme de niveaux d'énergie atomiques est très claire : $\Delta E = E_2 - E_1$ est la différence d'énergie entre le niveau fondamental E_1 de l'atome et le premier niveau accessible. Au départ les atomes ne sont pas excités, ils sont dans leur niveau d'énergie fondamental E_1 . Si $T < E_2 - E_1$, l'énergie de l'électron est insuffisante pour faire passer l'atome dans son premier état excité, et il n'y a aucun transfert de l'énergie cinétique de l'électron vers l'énergie interne de l'atome. Dès que $T \geq E_2 - E_1$ un tel transfert est possible. Cette expérience valut à ses auteurs le prix Nobel en 1925.

Malgré ses succès, le modèle de Bohr souffrait de nombreuses insuffisances. Tout d'abord, il ne rendait pas compte de la structure fine du spectre de l'hydrogène. En effet, une observation minutieuse montre que les raies spectrales sont dédoublées : elles sont constituées de deux raies très proches mais distinctes. On a d'abord cru pouvoir expliquer cette structure fine à l'aide de la théorie de Sommerfeld, qui étend le modèle de Bohr à des orbites elliptiques et incorpore les corrections relativistes, mais il s'est avéré par la suite que cette théorie était erronée ([5] p. 140). Le modèle de Bohr n'expliquait pas non plus l'effet Zeeman¹² (démultiplication des raies spectrales sous l'effet d'un champ magnétique), ni l'effet Stark¹³ (un effet similaire mais dû cette fois à un champ électrique). Plus ennuyeux encore, il s'avère impossible de généraliser ce modèle aux systèmes plus complexes que l'hydrogène ou les ions hydrogénoïdes.

Mais par-dessus tout, le modèle de Bohr appelle lui-même une explication car il ne fait que postuler la quantification de l'énergie. Pourquoi certaines orbites sont-elles autorisées et pas les autres ? Tout cela semble vraiment mystérieux.

3.3 Autres arguments

On peut citer d'autres problèmes que la physique classique ne parvenait pas à expliquer, et qui ont joué un rôle dans la maturation des esprits nécessaire à la mise en place d'une nouvelle théorie :

- l'effet Compton (Compton, 1923), qui montra définitivement que la lumière pouvait avoir des aspects corpusculaire (cf [1] p. 131).
- la théorie des chaleurs spécifiques des solides, dont Einstein a jeté les bases en 1907 (la mécanique classique ne permet pas de comprendre pourquoi la chaleur spécifique tend vers zéro avec la température absolue, cf [7] p. 629).
- le problème du calcul de l'entropie d'un gaz parfait (paradoxe de Gibbs) : le problème provient de l'indiscernabilité des particules quantiques, un sujet sur lequel nous reviendrons (voir [9], p. 132).

On pourrait ajouter le théorème de Bohr-van Leeuwen (cf [1] p. 359) : en physique classique il ne peut pas y avoir de magnétisme ! Ce résultat surprenant ne fut cependant obtenu qu'après l'avènement de la mécanique quantique, et ne joua donc pas de rôle historiquement.

12. En 1896, Pieter Zeeman désobéit à son supérieur en utilisant l'équipement du laboratoire dans lequel il travaillait pour mesurer l'effet d'un champ magnétique sur les spectres atomiques. Il est renvoyé. Il prendra sa revanche en recevant le prix Nobel de physique en 1902 pour sa découverte.

13. Stark fit sa découverte en 1913, qui lui valut un prix Nobel en 1919. Un peu plus tard il s'inscrivit au parti nazi et se lança dans une croisade pour la « physique allemande », qui s'opposait à la relativité du « juif Einstein », et à la mécanique quantique du « juif blanc » Heisenberg (les nazis désignaient ainsi des personnes qui, sans être juives, méritaient d'être traitées comme telles). Stark publia divers pamphlets contre la « physique juive », où il écrivit entre autres bêtises que « les juifs n'ont pas les aptitudes à la vraie créativité dans les sciences de la nature ». Il réclama que les postes universitaires soient réservés aux vrais allemands, et reçut du régime nazi différentes gratifications pour ses bons et loyaux services. En dépit des vociférations de Stark, son effet fut complètement expliqué en 1926 par « le juif » Paul Sophus Epstein, à l'aide de la mécanique quantique. En 1947, le comité de dénazification décida de donner à Stark l'occasion de réfléchir à tout ça en le condamnant à quatre ans de prison.

Chapitre 4

La transition vers une nouvelle mécanique : 1923–1926

4.1 Les ondes de matière

Le 29 novembre 1924, en Sorbonne (car pour faire chic on dit « en Sorbonne » comme on dit « en Avignon »), Louis de Broglie soutient une thèse de physique. Le contenu en est si peu banal que le jury, comprenant entre autres Paul Langevin et Jean Perrin, hésite : est-ce l'œuvre d'un fou ou d'un génie ? Langevin prend l'avis d'Einstein. Le père de la relativité est immédiatement emballé : « Il a levé un coin du grand voile », déclare-t-il.

De Broglie obtient sa thèse, mais que contenait-elle de si extraordinaire ? Profondément influencé par Einstein, qui avait proposé en 1905 que la lumière avait des aspects corpusculaires, de Broglie suppose que, réciproquement, la matière puisse se comporter comme une onde. Cette hypothèse ne provient pas seulement d'une attirance pour la symétrie dans les lois de la Nature, encore que de telles considérations aient effectivement eu une influence sur le physicien français, mais d'une analogie poussée entre mécanique et optique. Cette analogie, remarquée par Hamilton dès 1827, repose sur la reformulation de la mécanique à l'aide du principe de moindre action : la trajectoire suivie par une particule est telle que la variation de l'action S est nulle (voir l'appendice B). Dans le cas d'une particule dans \mathbb{R}^3 , la quantité de mouvement et l'énergie dérivent de l'action :

$$\vec{p} = \vec{\nabla} S, \quad \frac{\partial S}{\partial t} = -E \quad (4.1)$$

D'un autre côté, en utilisant la relation

$$\phi = \vec{k} \cdot \vec{r} - \omega t$$

entre vecteur d'onde, pulsation et phase d'une onde, on trouve

$$\vec{k} = \vec{\nabla} \phi, \quad \frac{\partial \phi}{\partial t} = -\omega \quad (4.2)$$

La relation d'Einstein $E = h\nu = \hbar\omega$ pour le photon, suggère fortement de renforcer l'analogie entre (4.1) et (4.2) et de poser $\vec{p} = \hbar\vec{k}$. Mais pour l'instant ceci n'est qu'un jeu formel avec des équations. Encore faut-il interpréter et justifier ce que l'on a fait.

L'interprétation est la suivante : à une particule de quantité de mouvement \vec{p} est associée une onde de vecteur \vec{k} tel que $\vec{p} = \hbar\vec{k}$, ce qui, compte tenu de la relation $k = \frac{2\pi}{\lambda}$ donne la relation de de Broglie¹ :

1. On peut objecter que la relation de de Broglie n'est pas invariante par changement de référentiel. Du point de vue de la relativité restreinte, il faut considérer les quadrivecteurs (E, \vec{p}) et (ω, \vec{k}) . C'est la proportionnalité de ces deux quadrivecteurs qui a un sens indépendant du référentiel.

$$\lambda = \frac{h}{p} \quad (4.3)$$

La mécanique classique des particules ponctuelles doit donc être remplacée, dans certaines circonstances², par une nouvelle *mécanique ondulatoire*, qui entretient avec la mécanique classique le même rapport que l'optique ondulatoire avec l'optique géométrique. Ainsi, pour les longueurs d'onde très petites devant les dimensions caractéristiques du problème considéré, la mécanique ondulatoire doit donner des résultats interprétables en terme de particules suivant les lois de la mécanique classique. À titre d'exemple, calculons la longueur d'onde de de Broglie associée à un moustique en vol. La masse d'un moustique est de l'ordre du milligramme, et on peut estimer sa vitesse à $0,1 \text{ ms}^{-1}$. On trouve alors une longueur d'onde de l'ordre de 10^{-27} m , ce qui se passe de commentaire!

Pour que l'interprétation soit complète, il faudrait que l'on explique ce que représentent exactement ces ondes de matière. Ce sont des ondes, fort bien, mais des *ondes de quoi*? Ce problème est épineux, nous en reparlerons, mais pour l'instant nous allons le laisser en suspens. Notons que la théorie est utilisable telle quelle, dans certaines limites, sans interprétation physique claire de l'onde ni équation d'onde, exactement comme on peut faire de l'optique ondulatoire sans savoir que la lumière est une onde du champ électromagnétique et sans connaître les équations de Maxwell.

Quant à la justification, comme l'écrit de Broglie dans sa thèse : « Cette hypothèse [...] vaut, comme toutes les hypothèses, ce que valent les conséquences qu'on peut en déduire. » À la fin de sa soutenance, de Broglie suggère d'étudier la diffraction des électrons afin de vérifier expérimentalement sa théorie. Cela sera fait par Davisson et Germer qui observeront la diffraction des électrons par un réseau cristallin en 1927. La théorie de de Broglie permet également d'interpréter les orbites de Bohr comme un phénomène d'ondes stationnaires.

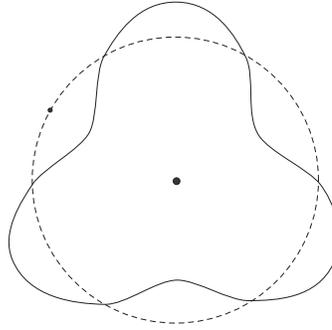


FIGURE 4.1 – Représentation intuitive d'une onde stationnaire associée à une orbite de Bohr.

En effet, à chaque révolution, l'onde associée à un électron en orbite circulaire acquiert une phase qui doit être un multiple de 2π pour que l'onde soit stationnaire, d'où la condition :

$$\phi = \oint \vec{k} \cdot \vec{dr} = 2\pi n$$

À l'aide de la relation $\vec{p} = \hbar\vec{k}$ on obtient la contrainte

$$\oint \vec{p} \cdot \vec{dr} = nh$$

On calcule facilement l'intégrale pour un mouvement circulaire uniforme³, et on obtient :

2. De même qu'il est impossible d'abandonner complètement la théorie ondulatoire de la lumière, on ne peut pas envisager une théorie purement ondulatoire de la matière, à ce sujet, consulter [12] p. 57.

3. Cela se généralise également aux orbites elliptiques de Sommerfeld.

$$\frac{4\pi^2 m_e r^2}{T} = nh$$

En calculant enfin la période T de l'orbite de rayon r sous l'action du champ électrostatique à l'aide du PFD, on trouve finalement :

$$r(n) = n^2 \frac{\hbar^2 \epsilon_0}{\pi m_e e^2} \quad (4.4)$$

C'est-à-dire exactement la formule (3.14) !

On voit que l'hypothèse de de Broglie est très séduisante. Elle explique en partie le modèle de Bohr, prédit des phénomènes nouveaux qui sont effectivement observés⁴, et introduit une symétrie entre matière et onde qui étend l'hypothèse du photon d'Einstein, la fameuse *dualité onde-corpuscule*. Pour autant, cette théorie n'a pas été très bien accueillie à ses débuts, en partie à cause de son étrangeté, en partie à cause de la piètre estime dans laquelle Bohr tenait de Broglie, suite à une ancienne querelle scientifique dont Bohr avait triomphé. Or c'est Bohr faisait la loi, alors tout le monde répétait que la théorie de de Broglie était absurde... Tout le monde sauf Einstein. Et c'est ce dernier finit par persuader un certain Schrödinger de s'intéresser aux ondes de matière...

4.2 L'équation de Schrödinger

C'est donc à Schrödinger⁵ qu'il incombait de déterminer l'équation des ondes de la nouvelle mécanique, qu'il qualifia d'ailleurs le premier d'ondulatoire.

Le paragraphe précédent nous indique qu'à une particule de masse m , de quantité de mouvement \vec{p} et d'énergie E est associée une onde plane monochromatique du type :

$$\Psi(\vec{r}, t) = e^{\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - Et)}$$

Un calcul immédiat montre que l'on a :

$$i\hbar \frac{\partial}{\partial t} \Psi = E\Psi \quad \text{et} \quad -\frac{\hbar^2}{2m} \Delta \Psi = \frac{p^2}{2m} \Psi$$

Compte tenu de la relation entre énergie et quantité de mouvement pour une particule libre :

$$E = E_c = \frac{p^2}{2m}$$

On obtient

$$i\hbar \frac{\partial}{\partial t} \Psi = -\frac{\hbar^2}{2m} \Delta \Psi$$

Si la particule est plongée dans un potentiel ne dépendant pas du temps, on aura la relation plus générale :

4. Ce qui engendra le lot de prix Nobel habituels : de Broglie en 1929, Davisson en 1937, co-lauréat avec G.P. Thomson.

5. Erwin Schrödinger (1887–1961), physicien autrichien. Il reçut le prix Nobel en 1933 pour l'équation qui porte son nom. La même année son opposition au régime nazi le pousse à quitter Berlin pour Oxford, puis pour Princeton. Il n'y demeure cependant pas très longtemps, sa volonté de fonder un foyer avec sa femme et sa maîtresse étant assez mal vue... Après être retourné quelques temps en Autriche, il finira par s'établir en Irlande. En plus de sa contribution majeure à la formulation de la mécanique quantique, il découvrit un effet paradoxal dans la mécanique quantique relativiste, le « Zitterbewegung », ou tremblement, qui contribua à l'abandon de cette théorie, plus tard remplacée par la théorie quantique des champs. Il s'opposa toute sa vie à l'interprétation standard de la mécanique quantique, contre laquelle il mis au point la célèbre expérience de pensée du « chat de Schrödinger ». Il étudia également la perception des couleurs, et s'intéressa à la fin de sa vie à la biochimie. Son livre *Qu'est-ce que la vie ?* eut une certaine influence sur les découvreurs de la structure de l'ADN, Crick et Watson. Il mourut à 73 ans d'une tuberculose dont il souffrait depuis des décennies. C'est d'ailleurs pendant son séjour au sanatorium d'Arosa – avec sa maîtresse – qu'il conçut sa célèbre équation.

$$E = E_c + V(\vec{r}) = \frac{p^2}{2m} + V(\vec{r})$$

d'où la fameuse équation de Schrödinger :

$$i\hbar \frac{\partial}{\partial t} \Psi = -\frac{\hbar^2}{2m} \Delta \Psi + V \Psi \quad (4.5)$$

Revenons sur nos pas : nous avons établi cette équation dans le cas d'une « onde de matière » Ψ de type onde plane monochromatique. Autrement dit, nous avons trouvé une équation dont Ψ est une solution. Mais il y en a beaucoup d'autres ! À ce titre, l'équation de Schrödinger ne saurait être démontrée, elle s'obtient par un raisonnement de type inductif à la fin duquel nous *postulons* que (4.5) est l'équation générale des ondes de matière. On peut justifier cette équation de façon un peu plus convaincante, donner toute sorte d'arguments, mais *in fine* cette équation se justifie... parce qu'elle donne de bons résultats. En particulier elle a permis à Schrödinger de retrouver le spectre de l'hydrogène, la dégénérescence des niveaux d'énergie et les trois nombres quantiques n, l, m , et l'effet Zeeman « normal ». Pour ne pas introduire une trop longue digression mathématique, nous renvoyons l'étude de ces phénomènes à plus tard (paragraphe 6.7.2).

Notons cependant deux propriétés qui justifient que l'on s'intéresse à (4.5), avant même d'en déduire des prédictions expérimentalement vérifiables. La première est qu'il s'agit d'une équation linéaire. Par conséquent, toute combinaison linéaire de deux solutions est une solution : or c'est un des grands principes de la physique quantique, que l'on peut vérifier avec une grande précision, et dont on a déjà rencontré une incarnation en spectroscopie. De plus, il s'agit d'une équation du premier ordre en t , ce qui montre que l'on peut calculer la fonction d'onde pour tout temps dès lors qu'on la connaît à un instant fixé t_0 . Ceci permet d'interpréter l'équation de Schrödinger comme une équation d'évolution.

Remarquons que, pour la première fois dans l'histoire de la physique, le nombre complexe i apparaît dans une équation. Cela nous montre que la fonction d'onde complexe Ψ n'est pas qu'un artifice mathématique : il n'est pas question de prendre la partie réelle à la fin. On constate d'ailleurs la confusion de Schrödinger à ce propos dans ses articles originaux (cf [1], p. 283) : il écrit au tout début que l'on doit revenir à la forme réelle... mais il ne le fait pas, et finit par conclure que Ψ doit être « *considérée comme essentiellement complexe* » !

Enfin nous avons des ondes, et une équation pour ces ondes, mais nous ne savons toujours pas ce que représentent ces ondes ! Schrödinger a remarqué que si l'on posait :

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

alors son équation impliquait :

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0$$

On reconnaît la forme locale d'une équation de conservation, où ρ joue le rôle d'une densité et \vec{j} celui d'un courant. Mais, une densité et un courant de quoi ? Après avoir multiplié ces deux quantités par e , Schrödinger avait imaginé qu'il s'agissait d'une densité et d'un courant de charge. Nous allons voir un peu plus loin qu'il se trompait : on peut interpréter $\rho = |\psi(x)|^2$ comme une densité de probabilité⁶. Il est d'ailleurs important de noter que la quantification, par exemple des niveaux d'énergie d'un atome, provient du fait que l'intégrale de $|\psi^2(x)|$ soit égale⁷ à 1.

4.3 Le hasard et l'incertitude

Le phénomène de la radioactivité montre un aspect aléatoire indéniable, qui pourrait indiquer que le hasard joue un rôle dans la description des phénomènes à l'échelle atomique et subatomique. Néanmoins,

6. Au moins dans certains cas. Nous reviendrons sur cette question.

7. Il suffit en fait que cette intégrale soit finie. On peut alors la rendre égale à 1 en divisant par un facteur de normalisation.

il se pourrait également que ce caractère aléatoire soit la conséquence d'une dynamique complexe mais déterministe, comme celle qui régit le lancer d'un dé ou le brassage des boules de loterie (ce que nous appellerons le *hasard classique*).

En réalité, le caractère probabiliste des lois quantiques est d'une nature très différente du hasard classique, et il se révèle de la façon la plus claire dans une célèbre expérience, dont Richard Feynman disait ([8], p 2), en exagérant un peu, qu'elle contenait « *l'unique mystère de la mécanique quantique.* »

4.3.1 L'expérience des fentes d'Young

C'est une expérience bien connue. Elle permet de démontrer définitivement, croyait-on, le caractère ondulatoire de la lumière.

On dispose d'une source de lumière monochromatique de longueur d'onde λ , d'un écran percé de deux fentes, et d'une plaque photographique. Le dispositif expérimental est montré sur la figure 4.2.

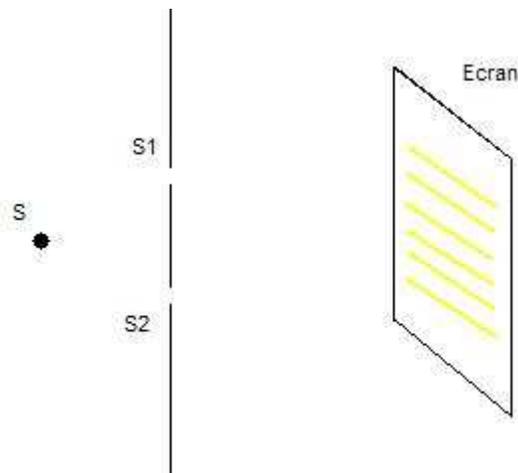


FIGURE 4.2 – Expérience des fentes d'Young

Pour peu que la longueur d'onde λ soit du même ordre de grandeur que la distance d entre les trous, on observe de magnifiques franges d'interférence (figure 4.3).

L'intensité $I(x)$ de la lumière reçue par la plaque a l'allure de la courbe représentée sur la figure 4.4.

L'amortissement des oscillations est due à la diffraction de la lumière par les fentes, un phénomène qui ne joue pas de rôle particulier dans cette discussion. Tout ceci est parfaitement bien expliqué par l'optique ondulatoire.

Cependant, si on abaisse fortement l'intensité de la source, on observe que les franges ne se forment pas immédiatement (figure 4.5). On voit plutôt apparaître de façon aléatoire de petites taches brillantes. Si on attend suffisamment longtemps, les taches finissent par former une figure qui ressemble de plus en plus à la figure bien lisse des franges obtenues avec de la lumière plus intense. De plus, on peut constater que les taches ne deviennent pas plus grandes lorsqu'on recule l'écran.

Il semble bien que l'on puisse expliquer ce phénomène en termes de particules de lumière, de photons, heurtant la plaque photographique. Si la lumière est intense, le nombre gigantesque de photons émis par la source ne permet pas de distinguer le caractère discret de la figure d'interférence. Cependant, avec de la lumière très peu intense, on s'aperçoit que les photons interagissent un par un avec la plaque. Si on attend assez longtemps pour que le nombre de photons émis depuis le début de l'expérience avec la source faible soit très grand, on retrouve ce qu'on obtient presque instantanément avec la source intense⁸.

8. Dès 1909, G.I. Taylor a réalisé une expérience avec de la lumière atténuée par des plaques de verre fumé. L'intensité obtenue était si faible qu'il en a conclu qu'un seul photon à la fois traversait l'appareil. Il a attendu 3 mois pour voir

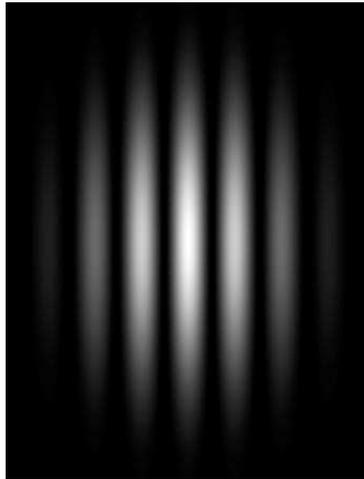


FIGURE 4.3 – Franges d’interférences

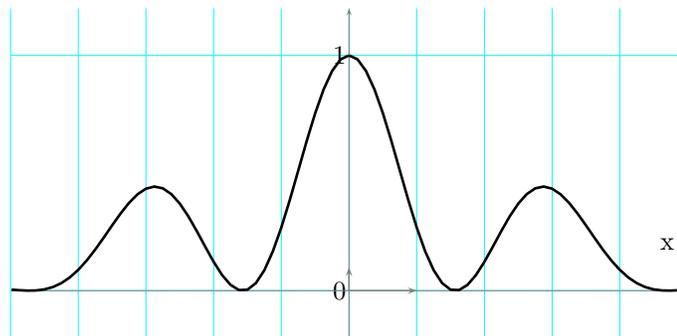


FIGURE 4.4 – Intensité lumineuse en fonction de la distance par rapport au centre de la plaque

Dans cette interprétation, l’intensité lumineuse reçue dans un intervalle $[x; x+dx]$ n’est rien d’autre qu’une mesure du nombre de photons $N(x)$ ayant été interceptés dans cet intervalle. En divisant par le nombre total N de photons émis, on obtient la fréquence $f(x)$ de l’interception d’un photon dans $[x; x+dx]$. Le nombre N étant supposé assez grand, la loi des grands nombres nous autorise à remplacer cette fréquence par une probabilité⁹. On obtient donc que la probabilité qu’un photon arrive dans $[x; x+dx]$ est proportionnelle à l’intensité reçue dans cet intervalle, soit en formule :

$$I(x) \propto P(x) \tag{4.6}$$

Cette première analyse tendrait à montrer que la lumière est fondamentalement corpusculaire, et que son caractère ondulatoire n’est qu’une limite obtenue lorsque le nombre de corpuscules est très grand. Or, il n’en est rien !

En effet, si on bouche l’une des deux fentes, on obtient une distribution de photons centrée autour de la position alignée avec la source et la fente ouverte, exactement comme on s’y attend.

Maintenant, si on ouvre les deux fentes et qu’on raisonne en terme de particules passant par l’une ou

apparaître les franges d’interférences. Néanmoins, on sait aujourd’hui que sa source de lumière émettait des photons par « paquets », et qu’il y a une probabilité non négligeable d’avoir plus d’un photon à la fois dans l’appareil avec ce genre de dispositif. Ce n’est que récemment que l’on a pu obtenir une source de lumière vraiment « monophotonique », grâce aux travaux de P. Grangier à l’institut d’optique d’Orsay.

9. Il s’agit en fait de la densité d’une loi de probabilité.

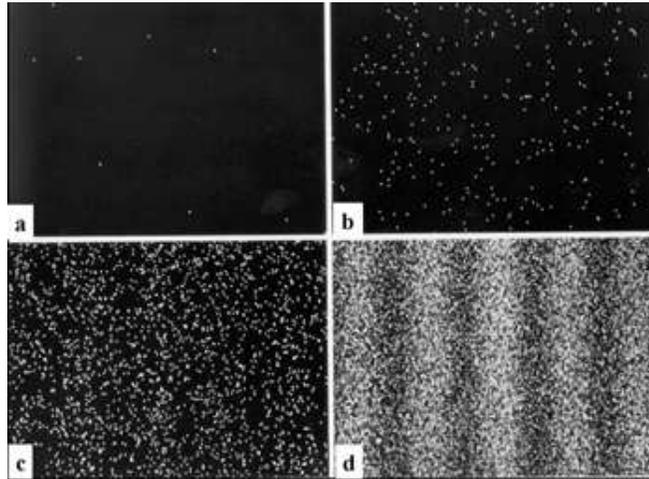


FIGURE 4.5 – Apparition progressive des franges d’interférence avec une source de faible intensité.

l’autre des fentes, on obtient que le nombre de photons reçus dans $[x; x + dx]$ est simplement le nombre $N_1(x)$ de photons venant de la fente 1, plus le nombre $N_2(x)$ de ceux qui viennent de la fente 2. Ce nombre est bien sûr le même que celui qu’on obtenait en bouchant l’une des deux fentes. En passant aux probabilités, on obtient naturellement

$$P(x) = P_1(x) + P_2(x) \quad (4.7)$$

Mais il suffit de jeter un coup d’oeil aux figures 4.4 et 4.6 pour voir que la relation (4.7) n’est pas du tout vérifiée.

$$P(x) \neq P_1(x) + P_2(x) ! \quad (4.8)$$

On voit qu’il y a des endroits où la probabilité de trouver un photon devient plus petite avec deux trous ouverts qu’avec un seul, ce qui est totalement incompréhensible si on raisonne en terme de particules passant par l’un des deux trous. D’une manière générale, on ne voit pas bien comment on pourrait obtenir des interférences destructives avec des particules : une particule plus une particule ça ne peut pas se détruire, tandis qu’un creux plus une bosse ça peut faire quelque chose de plat. C’est bien la raison pour laquelle l’expérience d’Young avait en son temps éliminé la théorie corpusculaire de la lumière due à Newton.

Conclusion de cette seconde analyse : la lumière ne peut pas être fondamentalement constituée de particules.

Il semble bien que l’on soit face à une contradiction. Les choses ne s’arrangent pas si on considère une source d’électrons à la place d’une source de lumière. Il n’avait jamais fait aucun doute pour personne que les électrons étaient bien des particules, de multiples faits expérimentaux en témoignent. Mais on peut réaliser un dispositif similaire à celui des fentes d’Young pour des électrons, en choisissant d de l’ordre de la longueur d’onde de de Broglie pour ces particules (qui dépend de leur énergie), et on obtient un résultat en tout point identique à celui que nous venons de décrire pour la lumière. C’est précisément ce qu’ont réalisé Davisson et Germer.

On a longtemps décrit la situation comme relevant d’une dualité entre la nature ondulatoire et corpusculaire de la lumière, de l’électron, ou de tout autre objet quantique. Certains types de questions relèvent de l’aspect corpusculaire, d’autres de l’aspect ondulatoire. On évite la contradiction parce qu’aucune question ne nécessite d’avoir recours aux deux aspects en même temps. Le champion de cette interprétation était Niels Bohr.

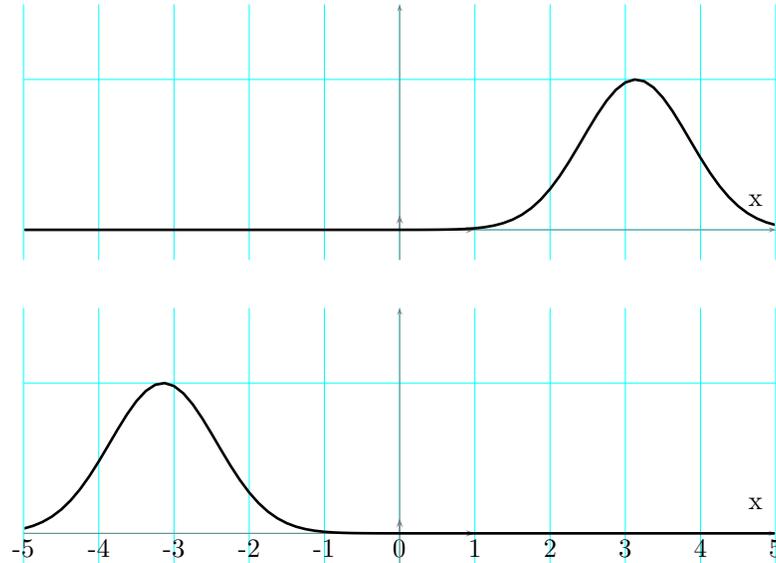


FIGURE 4.6 – Distribution de photons avec un trou bouché.

Je crois pouvoir dire que ce point de vue est aujourd’hui obsolète. Il est beaucoup plus simple de dire que les objets quantiques ne sont ni des ondes, ni des particules. Pourquoi le seraient-ils ? Les notions d’ondes et de particules sont intimement liées à l’intuition que nous avons d’une matière faite de petites billes dures, bien localisées, et suivant des trajectoires bien déterminées. Cette vision des choses nous est chevillée au corps, car elle correspond à ce que nous pouvons observer à notre échelle, mais il n’y a aucune raison de considérer qu’elle « descende » sans modification jusqu’à l’échelle atomique. D’ailleurs, les tentatives pour faire fonctionner une vision purement mécaniste du monde ont toujours abouti à des contradictions. Car si on parle d’onde, il faut dire quel est le support matériel de cette onde. . . Il faut bien se convaincre que les notions d’onde et de particules ne sont que des images, qui nous aident à réfléchir dans certains cas, mais nous trompent dans d’autres. Les objets quantiques sont différents. Certains ont proposé de les nommer « quantons », ou « particules quantiques ». L’usage et la paresse intellectuelle nous poussent à continuer à les appeler « particules ». On dira qu’un photon ou un électron est une particule, mais il faudra se souvenir qu’il s’agit en fait d’un quanton. Mais alors que sont les quantons ? Nous ne devons bien sûr rien rechercher à notre échelle qui nous permette d’avoir une intuition de leur nature, mais fort heureusement, on peut faire une théorie mathématiquement cohérente de leur comportement. C’est précisément ça, la mécanique quantique : tout le monde est d’accord sur les calculs, tout le monde est d’accord sur les résultats que doivent donner les expériences¹⁰, mais les problèmes commencent quand il s’agit de les interpréter.

L’expérience des fentes d’Young va d’ailleurs nous aider à comprendre de quoi va être fait, sinon les objets quantiques eux-mêmes, du moins le formalisme qui les décrit. En particulier, comme il y a un phénomène d’interférence, on peut se rapporter à la façon dont on interprète ce type de phénomène en physique ondulatoire classique. L’intensité I est reliée à une amplitude complexe h du type $h = Ae^{i\omega t}$, par la relation $I = |h|^2$. Les interférences se produisent car les amplitudes pour les sources secondaires s’additionnent. Ceci est une conséquence du fait que les équations des ondes sont linéaires. On a donc :

$$h = h_1 + h_2 \Rightarrow I = |h_1 + h_2|^2 \quad (4.9)$$

Or les « ondes de matière » Ψ de la mécanique ondulatoire suivent aussi une équation linéaire. En

10. À commencer par les expériences elles-mêmes puisque cette théorie n’a jamais été prise en défaut.

rapprochant les équations (4.6) et (4.9), nous sommes irrésistiblement invités à postuler :

$$P(x) \propto |\Psi(x)|^2 \quad (4.10)$$

Le coefficient de proportionnalité sera imposé par le fait que la probabilité totale doit valoir 1. Nous renonçons maintenant à appeler Ψ une « onde de matière ». En effet, elle représente une « amplitude de probabilité ». Cette expression n'a d'ailleurs pas vraiment de sens en elle-même, elle est juste une façon commode de rappeler que le carré de son module est une (densité de) probabilité. On appelle souvent Ψ une « fonction d'onde ». Nous verrons plus loin une terminologie meilleure et plus générale. Cette interprétation de la fonction d'onde en terme d'amplitude de probabilité est due à Max Born. C'est celle que nous allons commencer par développer, mais il en existe d'autres, que nous survolerons dans le chapitre 7. Pour l'instant, contentons nous de nous réjouir que cette interprétation nous permette d'y voir un peu plus clair dans l'expérience des deux fentes. En effet, on a

$$P(x) \propto |\Psi(x)|^2 = |\Psi_1(x) + \Psi_2(x)|^2 \neq |\Psi_1(x)|^2 + |\Psi_2(x)|^2 \propto P_1(x) + P_2(x) \quad (4.11)$$

Ce qui explique, au moins qualitativement l'équation (4.8). La différence entre les deux expressions est exactement égale à $2\text{Re}(\Psi_1\bar{\Psi}_2)$, qu'on appelle le terme d'interférence.

On pourrait, en théorie, résoudre l'équation de Schrödinger avec des conditions initiales adaptées au problème pour déterminer la fonction d'onde de l'électron (au moins en ignorant les questions de spin), et retrouver exactement le profil d'interférence à l'aide de (4.10). Une telle démarche aurait naturellement une place de choix dans n'importe quel cours de mécanique quantique, malheureusement, elle est impossible à suivre. En effet, on ne connaît pas de solution exacte à l'équation de Schrödinger dans ce cas, et tout ce qu'on peut faire est une résolution numérique approchée sur ordinateur, ce qui est naturellement un peu frustrant.

Notons un autre phénomène étonnant. Si on essaie de savoir par où est passée la particule (par exemple, si c'est un électron, on peut placer une source de lumière juste après un trou), on détruit la figure d'interférence! Il se produit dans ce cas une modification des fonctions d'onde qui annule le terme d'interférence, et on retrouve alors $P(x) = P_1(x) + P_2(x)$! On peut en tirer une conclusion importante : on ne peut pas négliger la perturbation engendrée par l'observation, il faut incorporer cette perturbation d'une façon ou d'une autre dans la théorie.

4.4 La mécanique des matrices

Quelques mois avant que Schrödinger ne trouve son équation¹¹, Werner Heisenberg¹² était parvenu à une théorie apparemment très différente mais tout aussi capable de décrire les phénomènes quantiques. Cette théorie fut connue pendant un certain temps sous le nom de mécanique des matrices.

Heisenberg tira son inspiration d'un principe qu'on peut qualifier de philosophique : si une quantité est inobservable, elle ne doit pas intervenir dans la formulation de la théorie. Heisenberg s'intéresse en premier lieu à l'atome d'hydrogène. Pour lui, les trajectoires de l'électron autorisées par le modèle de Bohr sont des notions classiques sans pertinence.

Il faut bien comprendre qu'il ne s'agit pas d'une question pratique (nul ne dispose d'un microscope assez puissant pour observer la trajectoire d'un électron), mais d'une question de principe. En effet, pour observer la position d'une particule, il faut par exemple l'éclairer¹³. Mais éclairer un objet c'est le soumettre à un bombardement de photons. Chacun de ces photons donne un petit « coup » à l'objet,

11. Décembre 1925 pour Schrödinger, Juillet 1925 pour Heisenberg

12. Werner Heisenberg (1901–1976), physicien allemand. Il reçut le prix Nobel 1932 pour la mécanique des matrices. Il contribua avec Pauli à jeter les bases de la théorie quantique des champs. Son attitude pendant la guerre fut controversée. Il fit d'abord l'objet d'attaques de la part des partisans de la « science allemande », puis fut réhabilité par le régime, et prit la tête du projet allemand d'élaboration d'une arme atomique. Son degré d'implication dans cette entreprise n'est pas connu avec certitude.

13. On peut aussi la faire passer par un diaphragme, auquel cas la diffraction de l'onde de de Broglie par le diaphragme permet d'obtenir la même estimation de l'ordre de grandeur de $\Delta x \Delta p_x$.

ce qui modifie sa quantité de mouvement. Pour un objet macroscopique, l'effet de ce bombardement est totalement négligeable, mais ce n'est pas du tout le cas pour un électron. On peut même estimer que pour obtenir une précision de l'ordre de Δx sur la position, on doit utiliser un photon (un seul photon, c'est la lumière la moins intense qu'on puisse obtenir!) de longueur d'onde $\lambda \sim \Delta x$. Ce photon possède une quantité de mouvement $p = h/\lambda$, il va donc entraîner une modification $\Delta p_x \sim h/\lambda$ de la quantité de mouvement de la particule observée, soit finalement, en ordre de grandeur :

$$\Delta x \Delta p_x \sim h \quad (4.12)$$

Cette relation serait violée si on pouvait connaître parfaitement la trajectoire de la particule, puisqu'on pourrait obtenir la vitesse et donc la quantité de mouvement en dérivant¹⁴.

L'argument ci-dessus n'est bien sûr pas entièrement rigoureux, et nous donnerons plus loin une relation plus précise que (4.12). Néanmoins, il suffit à nous montrer le bien-fondé de la volonté de Heisenberg d'évacuer la notion de trajectoire d'une particule de la description des phénomènes quantiques¹⁵.

Pour se débarrasser de la trajectoire de l'électron dans l'atome d'hydrogène, Heisenberg commence par développer celle-ci en série de Fourier¹⁶. Si $q(t)$ représente la projection de la position de l'électron sur un axe quelconque, on peut écrire :

$$q(t) = \sum_{k \in \mathbb{Z}} q_k e^{i\omega_k t} \quad (4.13)$$

où $\omega_k = 2k\pi/T$. L'idée est que les modes de Fourier q_k sont liés au spectre d'émission de l'atome d'hydrogène, et que le spectre c'est ce qu'on peut observer. Heisenberg entend donc s'exprimer dans le langage des modes de Fourier, qui sont plus proches de l'observation. En appliquant ceci à la n -ième orbite de Bohr, on obtient

$$q_n(t) = \sum_{k \in \mathbb{Z}} q(n, k) e^{i\omega(n, k)t} \quad (4.14)$$

Le point crucial est que la connaissance de la position de l'électron à l'instant t est équivalente à la connaissance de la suite double $(q(n, k) e^{i\omega(n, k)t})$. À la suite d'un bricolage inspiré, que Born qualifia de « mystique », et que nous nous garderons de reproduire ici (voir [1] p. 261 pour une discussion plus détaillée), Heisenberg est amené à remplacer cette suite double par une autre, notée $(Q_{n, k}(t))$. Ce qui nous importe, c'est que $Q_{n, k}(t)$ est de la forme $q_{n, k} e^{i\omega_{n, k} t}$, et que les pulsations satisfont la relation

$$\omega_{n, k} + \omega_{k, l} = \omega_{n, l} \quad (4.15)$$

Cette relation provient du fait que $\omega_{n, k}$ s'interprète comme la pulsation associée au saut du niveau d'énergie E_n au niveau d'énergie E_k :

$$\omega_{n, k} = \frac{1}{\hbar} (E_n - E_k) \quad (4.16)$$

Donc, en dernière analyse, (4.15) provient de la formule de Balmer-Rydberg, c'est-à-dire d'un fait expérimental.

Heisenberg décide donc que le tableau de nombre $Q(t) = (Q_{n, k}(t))_{n, k \in \mathbb{Z}}$ est ce qui doit prendre la place de la quantité classique inobservable $q(t)$.

Maintenant, Heisenberg se demande ce qui doit remplacer $q(t)^2$. Ce doit être un autre tableau de nombre, que nous noterons $Q^{(2)}$. Il montre que pour pouvoir faire apparaître (4.15), il est naturel de poser :

$$Q_{n, k}^{(2)} = \sum_{m \in \mathbb{Z}} Q_{n, m} Q_{m, k} \quad (4.17)$$

14. Le mieux que l'on puisse envisager, c'est que la trajectoire de la particule, si elle existait, soit une fonction nulle part dérivable.

15. Nous aurions pu nous fonder sur l'analyse de l'expérience des fentes d'Young pour arriver à cette conclusion, mais cela n'aurait pas été historiquement cohérent.

16. Les orbites de Bohr sont bien des trajectoires périodiques.

où, pour simplifier, on n'a pas écrit la dépendance en t .

Bien sûr, il est facile aujourd'hui de reconnaître dans (4.17) un produit de matrices (encore qu'il s'agisse de matrices infinies, ce qui engendre des problèmes de convergence qu'en bon physicien Heisenberg ne se pose pas). Cependant, les matrices n'étaient pas du tout d'un usage répandu en physique en 1925, ni même en mathématiques d'ailleurs, et Heisenberg ne reconnût pas ce qu'il avait sous les yeux ! C'est Born qui réalisa que Heisenberg proposait tout simplement de remplacer les quantités classiques comme $q(t)$ par des matrices.

La nouvelle mécanique prit donc le nom de « mécanique des matrices ». Le problème, c'est que si Heisenberg avait renoncé à toute description de l'atome en termes classiques, il n'avait aucune visualisation nouvelle à proposer. La nouvelle mécanique semblait très formelle, et au début, Heisenberg lui-même fut effrayé par le fait que deux matrices ne commutent pas toujours. C'est encore Born qui comprit que les matrices Q et P représentant la position et la quantité de mouvement « quantiques » devaient satisfaire la relation de commutation :

$$QP - PQ = i\hbar I \quad (4.18)$$

où I représente la matrice unité. Les physiciens se retrouvèrent englués dans une algèbre non commutative, avec (4.18) comme seul moyen de produire des calculs ! Pauli réussit par un tour de force à en déduire le spectre de l'hydrogène, mais il faut bien admettre que la majorité des physiciens étaient très réticents, et ils furent soulagés quand Schrödinger découvrit son équation. En effet, même avec l'étrange apparition du i , l'équation de Schrödinger est une équation aux dérivées partielles, ce qui était beaucoup plus parlant pour les physiciens que les matrices de Heisenberg. De plus, les fonctions d'ondes permettaient d'utiliser une intuition de physique ondulatoire. Schrödinger lui-même pensa pendant un temps qu'on pouvait se débarrasser complètement de la notion de particule et de sauts quantiques. Mais d'une part Born mis un terme à cet espoir en interprétant la fonction d'onde comme une amplitude de probabilité, et d'autre part Schrödinger lui-même, puis Dirac, réalisèrent que, loin d'être concurrentes, la mécanique ondulatoire et la mécanique des matrices étaient équivalentes¹⁷ !

4.5 La fusion des deux nouvelles mécaniques

Nous voici dans une situation étrange : nous avons deux théories qui donnent les mêmes résultats et semblent pourtant très différentes. Dans les années 1926–1927, Dirac, Hilbert et Schrödinger vont tirer la situation au clair, et dégager la structure générale de la mécanique quantique, dont la mécanique ondulatoire de Schrödinger et la mécanique des matrices de Heisenberg sont des cas particuliers. Dans la suite, nous suivrons en gros le cheminement de Dirac¹⁸.

À l'été 1925, alors qu'il n'est encore qu'un étudiant, Dirac prend connaissance de la théorie de Heisenberg. Il comprend que la non-commutativité des quantités observables est la nouveauté fondamentale. Pour Dirac, le passage de la physique classique à la physique quantique revient simplement à passer d'une théorie où les quantités physiques sont des nombres qui commutent, ce qu'il appelle des c -nombres (le c pouvant signifier commutatif ou classique), à une théorie où celles-ci sont des nombres qui ne commutent pas, des q -nombres (q pour quantum).

17. Cependant, Schrödinger n'accepta jamais l'interprétation probabiliste de la mécanique quantique. Nous en reparlerons à l'occasion du paragraphe sur la fameuse « expérience du chat ».

18. Paul Adrien Maurice Dirac (1902–1984), physicien britannique. Il reçut le prix Nobel en 1933. Il fut le premier à remarquer l'analogie entre les commutateurs quantiques et les crochets de Poisson. Il établit l'équivalence des formulations de Schrödinger et Heisenberg de la mécanique quantique et contribua à la formulation mathématique précise de la théorie. Il découvrit l'équation relativiste pour une particule de spin 1/2, qui porte son nom, grâce à laquelle il prédit l'existence de l'antimatière. Il fit également de très importantes contributions à l'électrodynamique quantique et aux problèmes généraux de la quantification des systèmes dynamiques. Sur le plan personnel, Dirac était très introverti et travailleur. Il ne s'arrêtait que le dimanche, où il s'autorisait de longues promenades. Il est possible qu'il ait souffert d'une forme atténuée d'autisme. Il se maria à 35 ans, au grand étonnement de ses collègues, avec Margit Wigner, la soeur du physicien hongrois. Toute sa vie Dirac présenta sa femme comme « la soeur d'Eugène Wigner ». Il dit d'elle qu'elle le rendit humain... Néanmoins on ne le vit pleurer qu'une fois : le jour de la mort d'Albert Einstein.

Les q -nombres sont donc des éléments d'une algèbre (il est important de pouvoir les additionner et les multiplier) non commutative. Dans une telle algèbre, la mesure de la non-commutativité entre deux éléments x et y est donnée par le *commutateur*, défini par :

$$[x, y] := xy - yx \quad (4.19)$$

L'équation (4.18) de la mécanique des matrices nous fournit un exemple d'apparition de ce commutateur. On peut y remarquer deux choses importantes : premièrement, l'apparition de i nous indique que le corps de base de notre algèbre est le corps des complexes, et deuxièmement, le commutateur est proportionnel à \hbar , qui est un nombre très petit. Ceci explique que la non-commutativité des quantités physiques ait pu passer inaperçue pendant très longtemps.

Ensuite, Dirac remarque une analogie qui doit sauter aux yeux si on connaît la mécanique analytique. Cette dernière est une formulation très générale de la mécanique classique établie par Lagrange et Hamilton (voir appendice B). Prenons le cas le plus simple d'une particule dans \mathbb{R} . L'état de ce système est complètement déterminé à chaque instant t par 2 nombres : sa position q et son impulsion p (qui est ici égale à la quantité de mouvement). On peut voir la position et l'impulsion comme des fonctions (des projections en fait) sur l'espace des phases (ici $\mathbb{R} \times \mathbb{R}$). Ces fonctions vérifient

$$\{q, p\} = 1 \quad (4.20)$$

En comparant avec (4.18), on s'aperçoit que dans l'algèbre des q -nombres, le crochet de Poisson est remplacé par $1/i\hbar$ fois le commutateur, au moins dans le cas particulier le plus simple de la position et de l'impulsion. Si on prend l'analogie au sérieux et qu'on la généralise, on obtient une équation d'évolution pour les q -nombres. En effet, l'évolution d'une fonction f définie sur l'espace des phases d'un système classique est donnée en mécanique hamiltonienne par

$$\frac{df}{dt} = \{f, H\} \quad (4.21)$$

où H est le hamiltonien, c'est-à-dire l'énergie écrite comme une fonction de p et q . D'où par analogie :

$$\frac{d\hat{f}}{dt} = \frac{1}{i\hbar}[\hat{f}, \hat{H}] \quad (4.22)$$

où on a mis des chapeaux sur les q -nombres pour les distinguer des c -nombres. Nous montrerons plus loin que (4.22) est une autre façon d'écrire l'équation de Schrödinger !

Le programme de Dirac est donc clair : il s'agit (pour un système physique donné), de déterminer une \mathbb{C} -algèbre non commutative, engendrée par les analogues quantiques des positions et des impulsions, vérifiant des relations de commutations telles que (4.18), et dont l'évolution est soumise à l'équation (4.22). Il y a une autre structure importante dont nous n'avons pas parlé : il doit y avoir une notion de « réalité » dans notre algèbre. En effet, les quantités physiques observables en physique classique, et les résultats des mesures, sont toujours des nombres réels.

À ce stade, on peut remarquer que les algèbres de matrices $\mathcal{M}_n(\mathbb{C})$ sont des candidats possibles. Ce sont des \mathbb{C} -algèbres non commutatives, et il y existe une notion de réalité : le passage à l'adjoint est analogue à la conjugaison complexe, et une matrice autoadjointe se comporte un peu comme un nombre réel. Ce cas peut effectivement se produire en mécanique quantique.

Cependant, en général, les algèbres de matrices ne suffisent pas. À vrai dire, on peut constater que les matrices de Heisenberg n'en sont pas vraiment, puisqu'elles sont infinies. Mais il existe en mathématique une généralisation des algèbres de matrices qui convient parfaitement. On sait que les matrices carrées $n \times n$ représentent des endomorphismes d'un \mathbb{C} -espace vectoriel de dimension n . Pour obtenir des tableaux infinis, il faut donc que l'on considère des endomorphismes d'un \mathbb{C} -espace vectoriel de dimension infinie. Il s'avère qu'on a une théorie qui a des bonnes propriétés et qui est capable d'englober à la fois les théories de Heisenberg et de Schrödinger : c'est celle des opérateurs dans les espaces de Hilbert.

Voyons comment ceci permet de retrouver l'équation de Schrödinger.

Le point délicat est de trouver les opérateurs (q -nombres) qui représentent les observables classiques. C'est le problème (en général non trivial) de la *quantification*. Dans le cas d'une particule libre dans \mathbb{R} c'est assez facile. On considère l'espace vectoriel $S = \mathcal{C}_c^\infty(\mathbb{R}, \mathbb{C})$ des fonctions infiniment dérivables à support borné¹⁹ et à valeurs complexes. Soit ψ une telle fonction. Notons :

$$P\psi = \frac{\hbar}{i} \frac{d}{dx} \psi \text{ et } Q\psi = (x \mapsto x\psi(x)) \quad (4.23)$$

On vérifie trivialement que P et Q sont des endomorphismes de S , ce qu'on appelle parfois des *opérateurs*. Il est d'usage de noter $P\psi$ plutôt que $P(\psi)$. De plus, on a, pour tout $x \in \mathbb{R}$ et pour tout $\psi \in S$:

$$\begin{aligned} ((QP - PQ)\psi)(x) &= \frac{\hbar}{i}(x\psi'(x) - x\psi'(x) - \psi(x)) \\ &= i\hbar\psi(x) \end{aligned} \quad (4.24)$$

On a donc montré que Q et P vérifient

$$[Q, P] = i\hbar \text{Id}_S \quad (4.25)$$

Maintenant, supposons qu'on ait une particule de masse m dans \mathbb{R} , soumise à un potentiel V ne dépendant que de la position. L'énergie de la particule classique est donnée, en fonction de la position et de l'impulsion par le hamiltonien²⁰ :

$$H(p, q) = \frac{p^2}{2m} + V(q) \quad (4.26)$$

En faisant l'hypothèse, à l'aide de (4.25), que les q -nombres de position et d'impulsion sont donnés par Q et P , respectivement, l'équation (4.26) donne :

$$\hat{H} = \frac{P^2}{2m} + V(Q) \quad (4.27)$$

où \hat{H} est le « hamiltonien quantique ». En appliquant ceci à un élément ψ de S , on trouve :

$$\hat{H}\psi = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi + V\psi \quad (4.28)$$

On généralise immédiatement à trois dimensions, en passant des c -nombres aux q -nombres de la façon suivante :

$$x_i \rightarrow \text{multiplication par } x_i; \quad p_i \rightarrow \frac{\hbar}{i} \frac{\partial}{\partial x_i} \quad (4.29)$$

On trouve alors exactement l'équation de Schrödinger, à condition de poser²¹

$$\hat{H}\psi = i\hbar \frac{\partial}{\partial t} \psi \quad (4.30)$$

Nous pouvons tirer de tout ceci les conséquences suivantes :

- Dans l'approche de Schrödinger, les q -nombres ne sont rien d'autre que des opérateurs différentiels.
- Ceux-ci agissent sur la fonction d'onde.
- La mécanique ondulatoire est donc un cas particulier de la théorie de Dirac, dans laquelle les q -nombres sont des opérateurs linéaires agissant sur un certain espace vectoriel de fonctions.

19. Cela signifie que la fonction s'annule en dehors d'un intervalle borné.

20. Potentiel et énergie potentielle sont synonymes en mécanique quantique. Il s'agit en fait d'une énergie potentielle.

21. Il y a une légère difficulté ici, due à une dépendance dans le temps qui n'est pas la même dans la théorie de Dirac et celle de Schrödinger, nous reviendrons plus loin sur ce point.

- Comme la fonction d’onde représente une amplitude de probabilité, son module au carré doit être intégrable sur \mathbb{R}^3 . Par conséquent, l’espace de fonctions que l’on considère est probablement du type $L^2(\mathbb{R}^3)$, qui se trouve être un espace de Hilbert.

Sans rentrer dans les détails, dans la mécanique des matrices de Heisenberg, les matrices sont en réalité des opérateurs sur l’espace des suites doubles de carré sommable, un autre espace de Hilbert.

Il y a quand même une différence importante entre l’approche de Heisenberg et celle de Schrödinger, sur laquelle nous n’avons pas encore attiré l’attention pour ne pas engendrer de complications inutiles : pour Schrödinger, la fonction d’onde dépend du temps, tandis que pour Heisenberg (et Dirac) ce sont les observables qui dépendent du temps. Il est possible de passer d’un point de vue à l’autre, nous reviendrons sur cette question.

Chapitre 5

La formulation mathématique de la mécanique quantique

5.1 Introduction

Au cours du chapitre précédent, nous avons vu émerger les caractéristiques fondamentales d'une nouvelle théorie, destinée à remplacer la mécanique classique. Nous avons peu ou prou suivi le cheminement historique. À présent, il nous apparaît inutile de suivre dans les détails les contributions de Dirac, Von Neumann, Hilbert, etc. à la formulation précise de la mécanique quantique. Nous allons énoncer les postulats sous la forme définitive qu'ils ont fini par prendre, et nous allons voir qu'ils éclairent de façon spectaculaire les difficultés rencontrées par la physique classique et par « l'ancienne théorie des quanta ». Nous devons avertir le lecteur que certains de ces postulats relèvent purement de la structure mathématique de la physique quantique, sur laquelle tous les physiciens s'accordent. D'autres en revanche, impliquent une interprétation particulière de la mécanique quantique, appelée *interprétation de Copenhague*, qui est la plus courante. Nous dirons un mot par la suite de certaines autres interprétations possibles.

Avant tout, il nous faut procéder à quelques rappels de mathématiques. Il s'agit essentiellement de la théorie des opérateurs dans un espace de Hilbert.

5.2 Rappels de mathématiques

5.2.1 Bras et Kets

En mécanique quantique le corps des scalaires est toujours \mathbb{C} , c'est pourquoi nous nous placerons dans ce cas même si les notions rappelées ci-dessous sont plus générales.

Soit V un \mathbb{C} -espace vectoriel. On appelle *forme linéaire* sur V toute application linéaire de V dans \mathbb{C} .

Si V est l'ensemble des fonctions intégrables à valeurs complexes définies sur \mathbb{R} , l'application $\phi : f \mapsto \int_{-\infty}^{+\infty} f(t) dt$ est un exemple de forme linéaire. Si $V = \mathbb{C}^3$, l'application $(z_1, z_2, z_3) \mapsto z_1$ est un autre exemple.

L'ensemble des formes linéaires sur V est appelé le *dual* de V . C'est également un \mathbb{C} -espace vectoriel, et on le note V^* . Si V est de dimension finie, alors V et V^* ont la même dimension. Nous verrons plus bas qu'on peut construire un isomorphisme canonique entre V et V^* à l'aide d'un produit scalaire.

En physique quantique il est d'usage d'appeler « produit scalaire » sur V ce qu'on appelle en mathématique un produit scalaire hermitien. Rappelons qu'il s'agit d'une forme sesquilinéaire $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$, qui à tout couple de vecteurs (v, w) associe le complexe $\langle v, w \rangle$, et vérifiant les propriétés suivantes¹ :

1. Notons que la linéarité se fait par rapport à la seconde variable, et l'antilinearité par rapport à la première, à l'inverse de la convention la plus courante en mathématiques

1. $\langle \lambda v, w \rangle = \bar{\lambda} \langle v, w \rangle$ et $\langle v, \lambda w \rangle = \lambda \langle v, w \rangle$,
2. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$,
3. $\langle v, w \rangle = \overline{\langle w, v \rangle}$,
4. $\langle v, v \rangle \in \mathbb{R}^+$,
5. $\langle v, v \rangle = 0 \Rightarrow v = 0$.

Grâce à la dernière propriété on peut définir la norme d'un vecteur v de V par $\|v\| = \sqrt{\langle v, v \rangle}$. On peut montrer que le couple $(V, \|\cdot\|)$ vérifie les axiomes d'un espace vectoriel normé. Si V est complet² pour la norme $\|\cdot\|$, on dit que $(V, \langle \cdot, \cdot \rangle)$ est un *espace de Hilbert*.

Un produit scalaire tel que $\langle \cdot, \cdot \rangle$ permet de définir à partir de tout vecteur de V la forme linéaire « produit scalaire par v ». Plus précisément, on peut définir une application $\epsilon : V \rightarrow V^*$ de la manière suivante : pour tout v , on note $\epsilon(v)$ la forme $\epsilon(v) = \langle v, \cdot \rangle$. Autrement dit, $\epsilon(v)$ est l'application qui à tout $w \in V$ associe $\langle v, w \rangle$.

Exercice 5.2.1 Démontrer que ϵ est une application antilinéaire injective. Montrer que si V est de dimension finie, ϵ est bijective.

Dirac a imaginé une notation très pratique qui a été universellement adoptée par les spécialistes de physique quantique. Nous verrons que l'espace des états d'un système quantique est un espace de Hilbert V . Ses éléments sont appelés des « kets », et sont notés comme dans l'exemple suivant : $|v\rangle \in V$. Ceci peut se lire : « le ket $|v\rangle$ appartient à V . De plus, l'élément $\epsilon(v)$ de V^* est appelé un « bra », et se note $\langle v|$. On peut appliquer un bra sur un ket, et on obtient alors un « bracket » c'est-à-dire un produit scalaire, selon la formule suivante :

$$\langle v|w\rangle = \epsilon(v)(w) = \langle |v\rangle, |w\rangle \quad (5.1)$$

On peut également multiplier un ket par un bra : le symbole $|u\rangle\langle v|$ désigne en effet l'application linéaire $|u\rangle\langle v|\cdot$, qui va de V dans lui-même et qui au ket $|w\rangle$ associe le ket $|u\rangle\langle v|w\rangle$. Il est clair que l'image de l'application linéaire $|u\rangle\langle v|$ est la droite vectorielle dirigée par $|u\rangle$ et que son noyau est l'orthogonal de $|v\rangle$, c'est-à-dire le sous-espace de V constitué des kets $|w\rangle$ tels que $\langle v|w\rangle = 0$.

Ainsi nous voyons que ket est un autre mot pour « vecteur de V », tandis que le mot bra désigne un certain type de forme linéaire (en fait n'importe quelle forme linéaire si V est de dimension finie, voir l'exercice 5.2.1). Pour notre part, nous nous en tiendrons à un usage plus mathématique, en appelant les ket des vecteurs, et les bras des formes.

Supposons maintenant que V soit de dimension finie n , et notons $|e_1\rangle, \dots, |e_n\rangle$ une base orthonormée de V . Soit $|v\rangle$ un vecteur quelconque de V , et soit (v_1, \dots, v_n) le n -uplet des coordonnées de $|v\rangle$ dans la base $|e_1\rangle, \dots, |e_n\rangle$. En mathématique on appelle i -ème forme coordonnée l'application $|e_i\rangle^* : V \rightarrow \mathbb{C}$, définie par $|e_i\rangle^*(v) = v_i$. Il est clair que $|e_i\rangle^*$ n'est rien d'autre que le bra $\langle e_i|$. De plus, on appelle « projection sur le i -ème vecteur de base » l'application $p_i : V \rightarrow V$, telle que $p_i(|v\rangle) = v_i|e_i\rangle$.

Exercice 5.2.2 Montrer que $p_i = |e_i\rangle\langle e_i|$, et que l'application identité peut se décomposer de la façon suivante :

$$\text{Id}_V = \sum_{i=1}^n |e_i\rangle\langle e_i|$$

Plus généralement, soit f un endomorphisme diagonalisable de V et soit $|e_1\rangle, \dots, |e_n\rangle$ une base de vecteurs propres de f associés respectivement aux valeurs propres $\lambda_1, \dots, \lambda_n$. En employant les mêmes notations qu'au paragraphe précédent, il est clair que pour tout $|v\rangle \in V$ on a :

$$|v\rangle = \sum_{i=1}^n v_i|e_i\rangle$$

2. Rappelons que cela signifie que toutes les suites de Cauchy convergent. Tout \mathbb{C} -e.v. normé de dimension finie est complet. L'espace des fonctions continues à support borné n'est pas complet pour la norme L^2 .

d'où l'on déduit :

$$f(|v\rangle) = \sum_{i=1}^n \lambda_i v_i |e_i\rangle = \sum_{i=1}^n \lambda_i p_i(|v\rangle)$$

Soit finalement l'égalité entre endomorphismes :

$$f = \sum_{i=1}^n \lambda_i p_i$$

qui peut encore s'écrire :

$$f = \sum_{i=1}^n \lambda_i |e_i\rangle \langle e_i| \quad (5.2)$$

Il est temps d'introduire une notation importante que les physiciens ont l'habitude d'employer, et qui de prime abord est un peu déconcertante. Il est possible d'utiliser la valeur propre λ_i elle-même pour indexer le vecteur propre correspondant, qui se notera alors $|\lambda_i\rangle$. Ainsi l'équation (5.2) peut se réécrire :

$$f = \sum_{i=1}^n \lambda_i |\lambda_i\rangle \langle \lambda_i| \quad (5.3)$$

Si les valeurs propres sont deux à deux distinctes, il n'est même pas nécessaire de numéroter les valeurs propres de 1 à n : il suffit d'écrire une somme sur les éléments du spectre de f :

$$f = \sum_{\lambda \in \sigma(f)} \lambda |\lambda\rangle \langle \lambda| \quad (5.4)$$

Dans l'équation (5.4), λ parcourt tous les éléments du spectre $\sigma(f)$ de l'opérateur f .

5.2.2 Opérateurs sur un espace de Hilbert. Théorème spectral.

Les physiciens ont tendance à employer le mot « opérateur » plutôt que « endomorphisme de V ». La formule (5.3) exprime donc la diagonalisation d'un opérateur diagonalisable. L'application d'un opérateur à un ket se note en juxtaposant simplement les symboles. Là où un mathématicien noterait $f(v)$, un physicien notera $f|v\rangle$.

Une classe particulièrement importante d'opérateurs diagonalisables est celle des opérateurs hermitiens (aussi appelés autoadjoints). Rappelons que si O est un opérateur sur V , son adjoint O^* (les physiciens notent souvent O^\dagger , lire « O dague ») est l'unique opérateur sur V tel que, pour tous $|u\rangle, |v\rangle$ on ait :

$$\langle |u\rangle, O|v\rangle \rangle = \langle O^*|u\rangle, |v\rangle \rangle \quad (5.5)$$

On peut faire disparaître toute référence au produit hermitien \langle, \rangle dans le membre de gauche de la formule précédente en utilisant les conventions sur les bras et les kets. En effet on a : $\langle |u\rangle, O|v\rangle \rangle = \langle u|O|v\rangle$. Dans cette écriture, on comprend que O agit sur $|v\rangle$, mais on peut aussi convenir que O agit sur $\langle u|$ en stipulant que $\langle u|O$ représente le bra associé au vecteur $O^*|u\rangle$.

Exercice 5.2.3 Soient O_1 et O_2 deux opérateurs, et λ_1, λ_2 deux complexes. Montrer que $(\lambda_1 O_1 + \lambda_2 O_2)^* = \bar{\lambda}_1 O_1^* + \bar{\lambda}_2 O_2^*$.

Exercice 5.2.4 Montrer que $(O^*)^* = O$ et que $(O_1 O_2)^* = O_2^* O_1^*$.

Exercice 5.2.5 Montrer que $\overline{\langle u|O|v\rangle} = \langle v|O^*|u\rangle$.

On dit que O est *hermitien* ssi $O = O^*$, *antihermitien* ssi $O = -O^*$, et *normal* ssi $OO^* = O^*O$ (O commute avec son adjoint). Notons que si O est hermitien ou antihermitien alors il est normal. Enfin, on dit que U est *unitaire* ssi $UU^* = U^*U = \text{Id}_V$.

Exercice 5.2.6 Montrer que O est antihermitien ssi iO est hermitien.

On démontre en mathématiques que tout opérateur normal est diagonalisable, qui plus est en base orthonormée. Ce théorème est parfois appelé « théorème spectral ». Comme nous allons le voir, il joue un rôle très important en physique quantique.

Théorème 5.2.1 (théorème spectral) Soit V un espace de Hilbert de dimension finie, et $f : V \rightarrow V$ un opérateur normal. Alors il existe une base orthonormée de vecteurs propres pour f .

De plus, si f est hermitien, alors le spectre de f est inclus dans \mathbb{R} . Si f est antihermitien, les valeurs propres de f sont imaginaires pures. Enfin si f est unitaire, les valeurs propres de f sont des complexes de module 1.

Exercice 5.2.7 Démontrer la deuxième partie du théorème spectral.

En fixant une base orthonormée de V , et en exprimant les opérateurs dans cette base, nous obtenons une version matricielle de ce théorème.

Théorème 5.2.2 (théorème spectral, version matricielle)

Soit $A \in M_n(\mathbb{C})$ une matrice normale. Alors il existe une matrice unitaire U et une matrice diagonale D telles que :

$$A = UDU^{-1}$$

5.2.3 Codiagonalisation

En physique quantique il est très important de savoir si deux opérateurs diagonalisables admettent une base commune de vecteurs propres (on dit alors qu'ils sont *codiagonalisables*). La réponse à cette question est contenue dans le théorème suivant.

Théorème 5.2.3 Soient a et b deux opérateurs diagonalisables. Alors a et b sont codiagonalisables ssi ils commutent.

Démonstration: Supposons tout d'abord que a et b soient codiagonalisables. Alors ils admettent une base commune de vecteurs propres, et dans cette base, a et b sont représentés par des matrices diagonales. Or les matrices diagonales commutent entre elles, donc a et b commutent.

Réciproquement, supposons que a et b commutent, et soit $(e_i)_{i \in I}$ une base de vecteurs propres de a . Notons λ_i la valeur propre de a associée au vecteur propre e_i . Ainsi on a $b \circ a(e_i) = b(a(e_i)) = b(\lambda_i e_i) = \lambda_i b(e_i)$, d'une part, et d'autre part on a $b \circ a(e_i) = a \circ b(e_i) = a(b(e_i))$. On en déduit que $a(b(e_i)) = \lambda_i b(e_i)$. Ceci montre que le vecteur $b(e_i)$ appartient à l'espace propre de a associé à la valeur propre λ_i .

- premier cas : si les valeurs propres de a sont deux à deux distinctes. Dans ce cas les espaces propres de a sont tous de dimension 1, et on peut conclure que pour tout $i \in I$, $b(e_i)$ est proportionnel à e_i . Ainsi, l'opérateur b est représenté par une matrice diagonale dans la base $(e_i)_{i \in I}$ qui est donc diagonalisante à la fois pour a et pour b .
- second cas : le spectre de a admet des multiplicités. Supposons que a possède k valeurs propres distinctes en notons $\lambda_{i_1}, \dots, \lambda_{i_k}$ ces k valeurs. Notons encore $E_j = \ker(a - \lambda_j \text{Id})$ l'espace propre de a associé à la valeur propre λ_j . Nous avons démontré plus haut que $b(E_j) \subset E_j$. Soit $b_j : E_j \rightarrow E_j$ l'endomorphisme de E_j défini par restriction de b à E_j . On démontre en algèbre, à l'aide de la notion de polynôme minimal, que b_j est nécessairement diagonalisable puisque b l'est. Ainsi E_j admet une base \mathcal{B}_j de vecteurs propres pour b_j . Mais comme E_j est un espace propre de a , cette base est aussi une base de vecteurs propres pour a . Comme les E_j sont supplémentaires (puisque a est diagonalisable), on en déduit que la réunion des bases $\mathcal{B}_1, \dots, \mathcal{B}_k$ est une base de l'espace tout entier, et cette base diagonalise à la fois a et b , par construction. ¶

5.2.4 Norme d'opérateur. Exponentielle d'opérateur.

Terminons ces rappels par la notion d'exponentielle d'opérateur. Rappelons tout d'abord que la *norme infinie* d'un opérateur a sur un espace de Hilbert V est :

$$\|a\|_\infty = \sup_{v \in V \setminus \{0\}} \frac{\|a(v)\|}{\|v\|} \quad (5.6)$$

Remarque importante : si V est de dimension finie, le nombre $\|a\|$ est toujours un réel, mais si V est de dimension infinie, il se peut que $\|a\|$ soit infinie. Si $\|a\|$ est finie, on dit que a est borné, et on peut démontrer que ceci équivaut à dire que a est une application continue. Si $\|a\|$ est infinie, on dit que a est un opérateur non borné. L'ensemble des opérateurs bornés sur un espace de Hilbert V se note $\mathcal{B}(V)$. On montre que $(\mathcal{B}(V), \|\cdot\|_\infty)$ est un espace vectoriel normé complet, stable par composition et par passage à l'adjoint. Si V est de dimension finie, $\mathcal{B}(V) = \text{End}(V)$.

Cette norme est parfois appelée norme d'opérateur, ou norme triple. Comme nous ne considérerons que cette norme sur les opérateurs, nous omettrons dans la suite d'écrire le symbole ∞ .

Cette norme possède deux propriétés fondamentales :

1. $\|a\| = \|a^*\|$,
2. $\|a \circ b\| \leq \|a\| \times \|b\|$

Exercice 5.2.8 Montrer que $a \mapsto a^*$ est une application continue de $\mathcal{B}(V)$ dans lui-même.

Grâce à cette norme, nous pouvons donner un sens à la notion de convergence d'une suite d'opérateurs. Une suite qui converge toujours est celle qui définit l'exponentielle.

Théorème 5.2.4 *Pour tout opérateur borné a sur l'espace de Hilbert V , la suite de terme général $s_n(a) = \sum_{k=0}^n \frac{a^k}{k!}$ converge. La limite se note $\exp(a)$ ou e^a .*

Si V est de dimension finie, et si A est une matrice représentant a dans une certaine base, alors la matrice $\exp(A)$ représente $\exp(a)$ dans cette même base.

Proposition 5.2.1 1. Si a et b commutent, alors $\exp(a + b) = \exp(a) \exp(b) = \exp(b) \exp(a)$.

2. En particulier $\exp(a) \exp(-a) = \exp(-a) \exp(a) = \exp(0) = \text{Id}$. Ainsi l'exponentielle d'un opérateur a est toujours inversible, et son inverse est $\exp(-a)$.

Exercice 5.2.9 1. Montrer que pour tout $a \in \mathcal{B}(V)$, $\exp(a^*) = \exp(a)^*$.

2. Montrer que l'exponentielle d'un opérateur hermitien est un opérateur hermitien.

3. Montrer que l'exponentielle d'un opérateur antihermitien est un opérateur unitaire.

5.2.5 Avertissement sur la dimension infinie

Le cas de la dimension finie, que nous venons de voir, nous donne un cadre juste suffisant pour comprendre le sens des postulats de la mécanique quantique.

Certains des théorèmes que nous n'avons énoncés qu'en dimension finie admettent des généralisations au cas où l'espace de Hilbert est de dimension infinie. Ce cas n'est pas juste un amusement de mathématicien : il apparaît de lui-même en physique, et même les systèmes les plus simples, comme l'étude d'une particule se déplaçant sur une droite, ou celle d'un oscillateur harmonique, nécessitent l'utilisation d'un espace d'états de dimension infinie. Malheureusement, certaines des généralisations évoquées plus haut sont assez subtiles sur le plan mathématique, et sont souvent traitées « par-dessus la jambe » dans les ouvrages de physique. Il arrive même que l'on ne puisse pas donner un sens mathématique précis à certaines manipulations formelles pratiquées par les physiciens, qui se réduisent alors à de simples raisonnements par analogie ! Dans la suite de cette introduction, nous serons obligés de nous résoudre à maltraiter ainsi la rigueur mathématique, car il serait antipédagogique d'alourdir une introduction à une théorie déjà complexe par des considérations plus ou moins claires d'analyse fonctionnelle ! Néanmoins, nous tenterons d'indiquer les passages délicats, et de préciser les analogies utilisées.

5.2.6 c -nombres et q -nombres

On peut remarquer une analogie troublante entre les opérateurs et les nombres complexes. Le passage à l'adjoint correspond à la conjugaison. Un opérateur hermitien étant son propre adjoint, il se comporte un peu comme un nombre réel. Il est possible d'additionner, de multiplier, et même de passer à l'exponentielle des opérateurs, tout comme on le peut avec des nombres complexes. Les propriétés de ces opérations sont les mêmes sauf pour ce qui concerne la commutativité : deux complexes commutent toujours, mais pas deux opérateurs.

Nous tenons-là nos c -nombres (les nombres complexes) et nos q -nombres (les opérateurs dans un espace de Hilbert).

5.3 Les postulats de la mécanique quantique

5.3.1 Vecteurs d'états, Espace des états

Premier postulat

Postulat 1 À tout système physique correspond un espace de Hilbert, appelé *l'espace des états*.

Par ce postulat nous déclarons que tout système physique est soumis aux lois quantiques. Or nous devons noter que les systèmes physiques envisagés sous l'angle de la mécanique quantique sont presque toujours des systèmes microscopiques. Bien que les lois quantiques s'appliquent a priori aussi aux systèmes macroscopiques, ceux-ci sont généralement constitués de beaucoup trop de parties pour qu'on puisse déterminer quel espace de Hilbert leur est associé. Cependant, on considère que dans la limite où un système devient macroscopique, les lois quantiques se réduisent aux lois de la mécanique classique. Cette croyance est justifiée rigoureusement dans de nombreux cas. Néanmoins, notons qu'il existe deux domaines qui font encore l'objet de recherches : celui de l'interaction entre un système macroscopique (typiquement un appareil de mesure) et un système microscopique, et celui d'un système faisant intervenir la gravitation. En effet, la force de gravitation est décrite par la théorie de la relativité générale qui est en un certain sens incompatible avec la théorie quantique. Cette situation est ennuyeuse dans la mesure où la force de gravitation possède une portée infinie, ce qui fait qu'en principe tout système est soumis à cette force ! Néanmoins, dans les systèmes microscopiques étudiés à ce jour dans le cadre quantique, l'interaction gravitationnelle entre les constituants est si faible qu'elle peut être négligée. Il n'en reste pas moins que l'unification des lois de la physique, ainsi que l'étude de certains systèmes extrêmes, comme les trous noirs ou les premiers instants de l'univers, nécessiteraient une théorie généralisant à la fois la relativité générale et la théorie quantique. Une telle théorie a de bonne chance de mettre à bas certains des postulats quantiques ou relativistes, si bien que, malgré tous les succès de la mécanique quantique, nous devons admettre que les postulats sur lesquels elle est fondée ne sont peut-être que des approximations de principes plus fondamentaux.

Notons également que la correspondance entre système physique et espace de Hilbert n'est pas bijective : il est possible qu'un même espace de Hilbert corresponde à plusieurs systèmes physiques différents.

Deuxième postulat

Postulat 2 À chaque instant t , l'état d'un système physique est décrit par un vecteur non nul $|\psi(t)\rangle$ de l'espace des états, appelé *vecteur d'état*. Deux vecteurs non nuls représentent le même état si, et seulement si, ils sont proportionnels.

Ce postulat amène plusieurs remarques. La première est qu'il fait référence au temps comme si ce dernier était un paramètre extérieur à tout système physique et indépendant de l'observateur. Or nous savons (voir [6]) qu'il n'en est rien : la théorie de la relativité restreinte, validée par de nombreuses expériences,

a démontré que la variable de temps n'a pas de sens physique intrinsèque, et que sa définition dépend du mouvement relatif de l'observateur et du système physique considéré. De plus, la théorie de la relativité générale, également corroborée par l'observation, a établi que le champ gravitationnel influe également sur l'écoulement du temps. D'un point de vue relativiste, le postulat 2 n'est tout simplement pas acceptable. Il faut donc comprendre que la mécanique quantique, fondée sur ce principe, ne peut être valide que dans le domaine non relativiste. Cela signifie que les vitesses (relatives à un observateur galiléen) des particules considérées doivent être très inférieures à celle de la lumière. Dans la plupart des applications de la mécanique quantique (chimie, électronique), cette condition est bien vérifiée. Nous dirons un mot à la fin de ce cours de ce qu'il advient lorsque l'on quitte le domaine de validité de la mécanique quantique. La deuxième remarque est que l'on peut reformuler le second postulat d'une façon mathématiquement plus élégante en disant que les états d'un système physique correspondent bijectivement aux sous-espaces vectoriels de dimension 1 de l'espace des états³. Autrement dit, un état est exactement décrit par une droite (complexe). Cette formulation fait jouer aux droites le rôle fondamental, et il est en effet possible d'édifier toute la mécanique quantique sans parler de vecteur d'état. Néanmoins, cette vision des choses est peu adaptée aux calculs et n'est pas très répandue chez les physiciens.

Enfin, il importe de signaler que l'énoncé du postulat 2 utilise ce que l'on appelle « le point de vue de Schrödinger ». Il existe une formulation équivalente de la mécanique quantique : « le point de vue de Heisenberg ». Selon ce dernier, les vecteurs d'états sont indépendants du temps, ce sont les observables (que nous définirons plus loin) qui en dépendent. Nous reviendrons plus loin sur l'équivalence entre les deux points de vue. Nous optons pour le point de vue de Schrödinger car il est plus intuitif et plus répandu dans les traités élémentaires.

Les deux premiers postulats introduisent une caractéristique de la mécanique quantique qui n'a pas d'équivalent classique : la \mathbb{C} -linéarité, ou principe de superposition quantique. Si $|\psi(t)\rangle$ et $|\phi(t)\rangle$ sont deux vecteurs d'états d'un certain système physique, alors pour n'importe quels nombres complexes c_1 et c_2 , la combinaison linéaire $c_1|\psi(t)\rangle + c_2|\phi(t)\rangle$ représente aussi un état du système physique⁴. Bien entendu, on ne peut comprendre la portée de ce principe de superposition quantique que si l'on sait interpréter les vecteurs d'états. Comment peut-on les utiliser concrètement ? C'est l'objet des deux postulats suivants.

5.3.2 Observables

Troisième postulat

Postulat 3 À chaque propriété observable d'un système physique correspond un opérateur hermitien sur l'espace des états. Un tel opérateur s'appelle *une observable*.

L'énergie, la projection sur un axe de la quantité de mouvement ou de la position d'une particule sont des exemples de propriétés observables. Une propriété observable définie sur un système physique est parfois appelée une *variable dynamique*.

Notons qu'il n'existe pas de règle rigoureuse permettant de déterminer pour chaque variable quel est l'opérateur correspondant. Il existe seulement des procédés heuristiques permettant de deviner la définition correcte des observables à partir de la description par la mécanique classique du système considéré. La validité de chacune de ces recettes de « quantification » doit être vérifiée a posteriori.

Le troisième postulat peut sembler assez étrange, et nous ne pourrions le comprendre pleinement qu'avec la pratique. Notons néanmoins qu'il est assez naturel si l'on adopte le point de vue de Dirac, évoqué plus haut. En effet, la physique classique décrit les propriétés des objets à l'aide de c -nombres réels : la position d'une particule le long d'un axe est classiquement un tel nombre. Si l'on admet que la physique quantique s'obtient en passant des c -nombres aux q -nombres, alors une telle propriété doit être décrite par un q -nombre réel, autrement dit un opérateur hermitien.

3. Il ne faut pas oublier que le corps des scalaires est le corps des complexes : il s'agit donc de sous-espaces vectoriels de dimension 1 sur \mathbb{C} .

4. sauf si elle est nulle

Quatrième postulat

Postulat 4 Les résultats possibles de la mesure d'une variable dynamique sont les valeurs propres de l'observable correspondante.

Notons que ces valeurs propres sont bien des nombres réels. Grâce à ce postulat, nous pouvons enfin commencer à comprendre le lien entre les objets mathématiques que nous avons introduits et leur interprétation concrète. Il nous manque encore un postulat pour compléter ce lien.

5.3.3 Interprétation probabiliste

Cinquième postulat

Étant donné un système physique pouvant être décrit par la mécanique classique, il est possible de calculer exactement la valeur prise par une variable dynamique connaissant l'état du système à l'instant t . La théorie quantique renonce à ce type de prédiction : elle permet seulement de calculer la loi de probabilité des valeurs que peut prendre une variable dynamique lorsque le système est dans un certain état.

Avant d'énoncer précisément cette loi, nous allons commencer par étudier un cas particulier. Soit a une variable dynamique, et supposons que l'observable A qui correspond à cette variable ait un spectre sans multiplicité. Pour tout valeur propre α nous notons $|\alpha\rangle$ un vecteur propre quelconque associé à cette valeur propre. Remarquons que dans le cas particulier où nous sommes, l'espace propre associé à α est de dimension 1, ainsi les différents choix possibles pour le vecteur $|\alpha\rangle$ sont colinéaires entre eux. Le postulat 5 prend alors la forme suivante : la probabilité $P(a \rightarrow \alpha|\psi)$ qu'une mesure de la variable a donne la valeur α sachant que le système est dans l'état $|\psi\rangle$ est :

$$P(a \rightarrow \alpha|\psi) = \frac{|\langle \alpha|\psi\rangle|^2}{\langle \alpha|\alpha\rangle\langle \psi|\psi\rangle} \quad (5.7)$$

Dans le cas où $|\alpha\rangle$ et $|\psi\rangle$ sont tous les deux normalisés, la formule (5.7) se simplifie en :

$$P(a \rightarrow \alpha|\psi) = |\langle \alpha|\psi\rangle|^2 \quad (5.8)$$

Le complexe $\langle \alpha|\psi\rangle$ est alors appelé *une amplitude de probabilité*. Notons que l'amplitude de probabilité, aussi bien que la probabilité elle-même, dépendent du temps par l'intermédiaire de ψ .

Prenons un exemple, sans signification physique particulière pour l'instant, afin de fixer les idées. Supposons qu'un certain système physique puisse être associé à l'espace de Hilbert \mathbb{C}^3 muni du produit scalaire canonique. Les vecteurs de la base canonique de \mathbb{C}^3 seront notés $|1\rangle, |2\rangle, |3\rangle$. Supposons qu'une observable O représentant une variable dynamique o soit associée à ce système, et s'écrive dans la base canonique :

$$O = \begin{pmatrix} 0 & i & 0 \\ -i & 0 & -i \\ 0 & i & 0 \end{pmatrix}$$

C'est un exercice très simple d'algèbre linéaire de vérifier que le spectre de O est $\{-\sqrt{2}; 0; \sqrt{2}\}$.

Une base orthonormée de vecteurs propres est donnée par : $|\sqrt{2}\rangle = \begin{pmatrix} 1/2 \\ i/\sqrt{2} \\ 1/2 \end{pmatrix}$, $|0\rangle = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix}$,

$$|-\sqrt{2}\rangle = \begin{pmatrix} -1/2 \\ i/\sqrt{2} \\ -1/2 \end{pmatrix}.$$

Supposons qu'à un certain instant t , le vecteur d'état du système soit la superposition suivante des états de base : $|\psi\rangle = \sqrt{2}|1\rangle - 2i|2\rangle + \sqrt{2}|3\rangle$. On voit facilement que $|\psi\rangle = -2\sqrt{2}|\sqrt{2}\rangle$. Il est donc certain qu'une mesure à l'instant t de la variable o donnera le résultat « $\sqrt{2}$ ».

Supposons maintenant que le vecteur d'état à l'instant t soit l'état de base $|1\rangle$. La probabilité qu'une mesure à l'instant t de la variable o donne le résultat « $\sqrt{2}$ » vaut :

$$P(o \rightarrow \sqrt{2}, t) = |\langle \sqrt{2} | 1 \rangle|^2 = |(-1/2)|^2 = \frac{1}{4}$$

On vérifie de même que $P(o \rightarrow 0, t) = 1/2$ et $P(o \rightarrow -\sqrt{2}, t) = 1/4$. Nous sommes ainsi capable de calculer la distribution de probabilité des différents résultats possibles pour une mesure de o à l'instant t . C'est tout ce que la théorie quantique nous permet de faire. Pour vérifier cette prédiction, il faut être capable de reproduire de nombreuses fois les mêmes conditions expérimentales pour le système étudié. Supposons par exemple que nous soyons capables de nous assurer que le système⁵ soit dans l'état $|1\rangle$ pour $t = 0, 1, 2, \dots$ ms, et que nous mesurions la variable o à chaque milliseconde. La mécanique quantique prédit alors que o aura pour valeur $\sqrt{2}$ environ une fois sur quatre en moyenne. Cette prédiction pourra se vérifier en attendant quelques secondes, de façon à disposer de suffisamment de résultats pour calculer une moyenne significative.

Passons maintenant à l'énoncé général du postulat 5. Il s'agit de traiter le cas où le spectre d'une observable A possède des multiplicités. Dans le cas le plus courant d'une valeur propre α de multiplicité finie, c'est-à-dire d'un sous-espace propre de dimension finie m , nous pouvons considérer $|\alpha^1\rangle, \dots, |\alpha^m\rangle$ une base **orthonormée** de l'espace propre associé à la valeur propre α . Le postulat 5 s'énonce alors comme suit.

Postulat 5 La probabilité qu'une mesure de la variable dynamique a donne le résultat α lorsque le système est dans l'état $|\psi\rangle$ est :

$$P(a \rightarrow \alpha | \psi) = \sum_{j=1}^m \frac{|\langle \alpha^j | \psi \rangle|^2}{\|\psi\|^2} \quad (5.9)$$

Notons que si l'on introduit le sous-espace propre E_α associé à la valeur propre α , et le projecteur orthogonal π_α sur l'espace W , alors le postulat 5 s'exprime plus simplement :

$$P(a \rightarrow \alpha | \psi) = \frac{\|\pi_\alpha |\psi\rangle\|^2}{\|\psi\|^2} \quad (5.10)$$

Cette formule est encore valable dans le cas d'un sous-espace propre de dimension infinie.

5.3.4 L'équation d'évolution

Sixième postulat

Il s'agit maintenant de déterminer la façon dont un vecteur d'état évolue avec le temps.

Postulat 6 Soit H le hamiltonien d'un système quantique et $|\psi(t)\rangle$ son vecteur d'état à l'instant t . Alors, en l'absence de toute opération de mesure, $|\psi(t)\rangle$ satisfait l'équation d'évolution de Schrödinger :

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle \quad (5.11)$$

Remarquons que ce postulat ne nous dit pas comment trouver le hamiltonien. Comme pour toute autre observable, la définition précise du hamiltonien quantique est en elle-même un postulat. La cohérence de ce postulat est soumise à deux conditions : le hamiltonien doit être hermitien, et il doit redonner à

5. ou, plus vraisemblablement, un système en tout point identique

la limite classique l'expression connue en mécanique classique du hamiltonien du système considéré... pour peu qu'il ait un équivalent classique. Bien sûr, l'expression du hamiltonien quantique se justifie a posteriori par les prédictions expérimentales qu'il permet.

Signalons également que H peut dépendre du temps. Ce sera le cas si le système n'est pas isolé.

Nous avons vu que selon les postulats de la mécanique quantique, le résultat de la mesure d'une variable dynamique n'est pas, sauf exception, entièrement déterminée : seule sa loi de probabilité l'est. C'est pourquoi l'on dit que la mécanique quantique n'est pas déterministe. Néanmoins, il faut insister sur le fait que, tant qu'on ne mesure rien, la façon dont le vecteur d'état évolue avec le temps est parfaitement déterministe. D'ailleurs, dans le cas où H ne dépend pas du temps, l'équation d'évolution (5.11) s'intègre immédiatement en

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle, \text{ où } U(t, t_0) = e^{-\frac{i}{\hbar}(t-t_0)H} \quad (5.12)$$

On appelle $U(t, t_0)$ l'opérateur d'évolution. L'opérateur H étant hermitien, l'opérateur iH est antihermitien, et il en résulte que l'opérateur d'évolution est unitaire (voir exercice 5.2.9). De ce fait, $U(t, t_0)$ conserve les produits scalaires :

$$\langle U(t, t_0)\phi(t_0)|U(t, t_0)\psi(t_0)\rangle = \langle \phi(t_0)|U(t, t_0)^{-1}U(t, t_0)\psi(t_0)\rangle$$

En utilisant (5.12) on obtient l'invariance des amplitudes par rapport à une translation du temps⁶

$$\langle \phi(t)|\psi(t)\rangle = \langle \phi(t_0)|\psi(t_0)\rangle \quad (5.13)$$

d'où découle l'invariance des probabilités par rapport à une telle translation. On pourrait, en procédant de manière inverse, déduire l'existence d'un opérateur d'évolution unitaire à partir de cette exigence d'invariance. Le postulat 6 serait tout-de-même nécessaire pour identifier cet opérateur unitaire avec l'exponentielle d'un certain multiple du hamiltonien.

En prenant $\phi = \psi$ dans (5.13) on voit également que la norme d'un vecteur d'état est invariante par translation du temps.

Indiquons pour terminer que l'existence d'un opérateur d'évolution unitaire et tous les résultats qui en découlent, est également vraie dans le cas où le hamiltonien est dépendant du temps (consulter par exemple [7] p. 308 ou [10] p. 40). Seule l'expression (5.12) de l'opérateur d'évolution en fonction du hamiltonien est modifiée.

États stationnaires

On suppose dans cette section que le hamiltonien est indépendant du temps.

Supposons que $|\psi(t_0)\rangle$ soit un état propre du hamiltonien associé à la valeur propre E . Cela revient à dire qu'une mesure de l'énergie du système donne la valeur E avec une probabilité de 1. Autrement dit, $|\psi(t_0)\rangle$ est un état d'énergie bien définie. Alors l'application de l'opérateur donne, d'après (5.12) :

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}(t-t_0)E}|\psi(t_0)\rangle \quad (5.14)$$

On peut également trouver ce résultat en résolvant directement l'équation différentielle du premier ordre (5.11).

L'équation (5.14) montre que les vecteurs $|\psi(t)\rangle$ et $|\psi(t_0)\rangle$ sont proportionnels, donc qu'ils représentent le même état du système. Nous pouvons donc conclure qu'un état dans lequel l'énergie est bien définie est *stationnaire*. Il reste le même au cours du temps.

Remarque : Ce résultat semble impliquer qu'un état excité d'un atome *isolé*... ne se désexcite jamais ! En réalité, les états d'énergie d'un atome peuvent se comparer à des états d'équilibre d'un système classique dont un seul, l'état fondamental (ground state) serait stable. Ainsi, la moindre perturbation extérieure tend à provoquer une désexcitation. Même en supposant l'atome complètement isolé, le traitement complet du phénomène nécessite la description quantique du champ électromagnétique de l'atome. Sans

6. Rappelons-le, tout ceci est vrai tant qu'il n'y a pas d'opération de mesure.

rentrer dans les détails, fort complexes, de l'électrodynamique quantique, on peut dire de façon imagée que le champ est soumis à de petites fluctuations aléatoires qui suffisent à « décrocher » les électrons éventuellement excités de leur état d'équilibre instable.

Passage du point de vue de Schrödinger au point de vue de Heisenberg

Dans le point de vue de Heisenberg, on travaille avec un vecteur d'état à temps fixé. Autrement dit, l'état « de Heisenberg » n'est autre que l'état « de Schrödinger » pour un temps t_0 fixé, ce que nous écrivons :

$$|\psi_H\rangle = |\psi_S(t_0)\rangle$$

La connaissance de $|\psi_H\rangle$ suffit à déterminer l'état du système pour tout temps t tant que celui-ci suit l'évolution unitaire, puisque

$$|\psi_S(t)\rangle = U(t, t_0)|\psi_S(t_0)\rangle = U(t, t_0)|\psi_H\rangle$$

Ainsi, toute l'évolution du système est inscrite dans $|\psi_H\rangle$ de même qu'un film (objet dynamique) est inscrit dans un DVD (objet statique).

Soit O_S une observable dans le point de vue de Schrödinger. Elle est indépendante du temps. Mais le produit scalaire⁷ $\langle\psi_S(t)|O|\psi_S(t)\rangle$ dépend du temps par l'intermédiaire de ψ_S . Dans le point de vue de Heisenberg, cette dépendance va se reporter sur l'observable :

$$\langle\psi_S(t)|O_S|\psi_S(t)\rangle = \langle\psi_S(t_0)|U^*(t, t_0)O_S U(t, t_0)|\psi_S(t_0)\rangle = \langle\psi_H|O_H(t)|\psi_H\rangle$$

où l'on a défini dans la dernière égalité l'observable au sens de Heisenberg, dépendante du temps :

$$O_H(t) = U^*(t, t_0)O_S U(t, t_0)$$

En dérivant cette équation par rapport au temps, on obtient l'équation d'évolution des observables de Heisenberg :

$$\frac{dO_H}{dt} = \frac{1}{i\hbar}[O_H, H]$$

où H est le hamiltonien.

Exercice 5.3.1 Démontrer cette dernière affirmation.

5.3.5 Réduction du paquet d'onde

Nous voici arrivés à l'un des aspects les plus problématiques de la théorie quantique. Nous avons indiqué dans le postulat 6 comment évoluait le vecteur d'état tant qu'on s'abstient de toute mesure. Il nous faut maintenant expliquer ce qu'il advient quand on effectue une mesure.

Hypothèse (réduction du paquet d'onde) Soit a une variable dynamique, α une valeur propre associée à cette variable, et E_α le sous-espace propre associé à cette valeur propre. Alors si on effectue une mesure de a entre les instants t et $t + \epsilon$ sur un système, le vecteur d'état de ce système vérifie :

$$|\psi(t + \epsilon)\rangle = \frac{\pi_\alpha |\psi(t)\rangle}{\|\pi_\alpha |\psi(t)\rangle\|}$$

où π_α est le projecteur orthogonal sur E_α .

7. Nous verrons plus loin que ce produit scalaire représente une moyenne. C'est ce qu'on observe en pratique.

Pour exprimer le contenu intuitif de cette hypothèse, plaçons nous dans le cas où α est une valeur propre simple, de vecteur propre normalisé $|\alpha\rangle$. Dans ce cas, l'hypothèse s'écrit :

$$|\psi(t + \epsilon)\rangle = e^{i\phi}|\alpha\rangle$$

Autrement dit, à l'instant où la mesure de a est terminée, le système se retrouve dans un état propre normalisé de l'observable A associée à a . Ceci est bien naturel, puisque si on mesure de nouveau a juste après l'avoir mesuré, c'est-à-dire à l'instant $t + \epsilon$, on s'attend à retrouver exactement la même valeur avec une probabilité 1, ce qui n'est possible que si le système est dans l'état propre correspondant à cette valeur.

On peut donc dire que le processus de mesure projète le système dans un état propre. De ce point de vue, mesurer une variable revient à préparer le système dans un certain état... à ceci près qu'on ne peut pas prévoir dans quel état on prépare le système puisque la valeur mesurée α est aléatoire.

L'évolution du vecteur d'état d'un système quantique peut donc être schématiquement résumée de la façon suivante (voir [15] p. 527 et suivantes). Il existe une évolution unitaire, que nous noterons U , et qui est complètement déterministe : si on connaît $|\psi(t_0)\rangle$, alors pour n'importe quel instant t on connaît $|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle$, tant que le système suit l'évolution unitaire, c'est-à-dire tant qu'on⁸ ne mesure rien. Mais il existe aussi une évolution non déterministe, notée R , qui se produit dès que l'on mesure une variable. Cette évolution sélectionne un état propre au hasard (selon une loi de probabilité qui, elle, est bien déterminée) et projète le système dans cet état. Une succession de telles évolutions est représentée sur la figure 5.1.

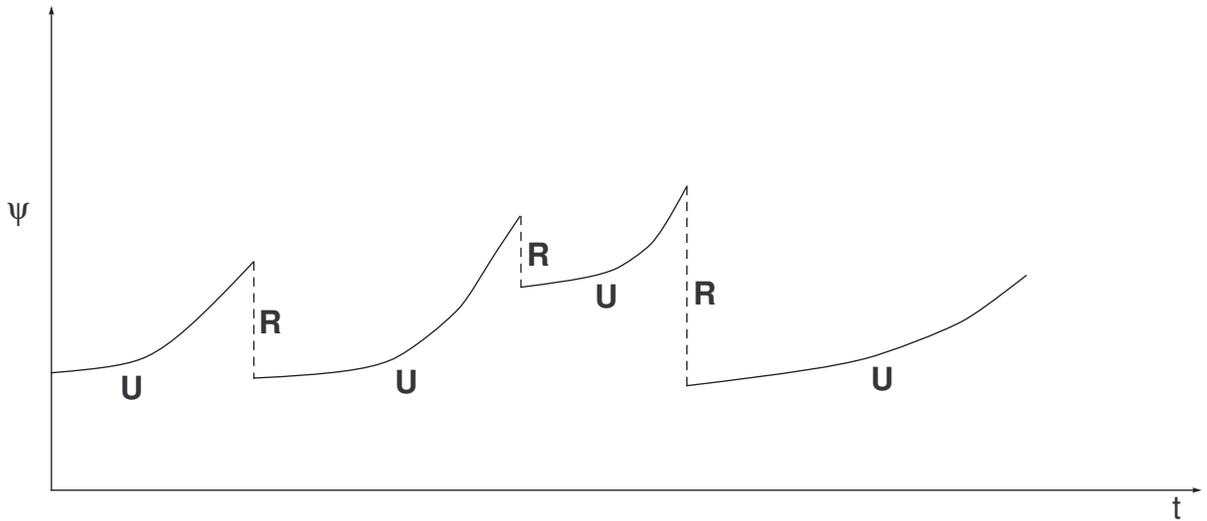


FIGURE 5.1 – Succession d'évolutions unitaires et de réductions du paquet d'onde.

Il peut paraître très étonnant que le simple fait de mesurer quelque chose introduise une telle discontinuité dans l'évolution d'un système. Pour beaucoup de spécialistes de la physique quantique, l'évolution R n'est qu'un moyen commode de représenter une interaction complexe entre un système quantique et un appareil de mesure typiquement macroscopique, qui lui aussi, est en principe régi par les lois quantiques. De ce point de vue, l'évolution du système complet « système mesuré + appareil de mesure » doit toujours être unitaire, et l'hypothèse de réduction du paquet d'onde n'est qu'une définition d'un processus de mesure idéalisé. La théorie de la *décohérence* est un programme de recherche qui se propose de décrire plus

8. Nous parlerons un peu plus de ce « on » à l'occasion de l'expérience du chat de Schrödinger...

précisément ce qui se passe pendant la réduction du paquet d'onde. Il s'est beaucoup développé durant ces dernières années (voir [13]).

Cependant, d'autres auteurs confèrent un statut de postulat à l'évolution R , et considèrent qu'il s'agit d'un processus réel, irréductible à l'évolution U , bien que la nature de ce processus diffère selon les interprétations. Nous reviendrons sur cette question à propos du chat de Schrödinger.

Étant donné l'absence de consensus sur le statut de la réduction du paquet d'onde, nous ne l'avons pas incorporé aux postulats fondamentaux de la mécanique quantique, et en avons préféré la qualifier d'hypothèse.

Chapitre 6

Premières applications des postulats

6.1 Espérance et écart-type d'une variable dynamique

À partir des postulats du chapitre précédent, il est possible de déterminer l'espérance et l'écart-type du résultat d'une mesure d'une certaine variable a lorsque le système considéré est dans un état $|\psi\rangle$ donné. Notons $\langle A \rangle_\psi$ l'espérance et $(\Delta A)_\psi$ l'écart-type ainsi définis. On a :

$$\langle A \rangle_\psi = \frac{\langle \psi | A | \psi \rangle}{\|\psi\|^2} \quad (6.1)$$

Pour éviter des complications mathématiques, nous donnons la preuve en dimension finie seulement. Dans ce cas, l'observable A possède k valeurs propres $\alpha_1, \dots, \alpha_k$, de multiplicités respectives m_1, \dots, m_k . On note $|\alpha_i^1\rangle, \dots, |\alpha_i^{m_i}\rangle$ une base orthonormée de vecteurs propres pour la valeur propre α_i . On a :

$$\begin{aligned} \langle A \rangle_\psi &= \sum_{i=1}^k \alpha_i P(a \rightarrow \alpha_i | \psi) \\ &= \frac{1}{\|\psi\|^2} \sum_{i=1}^k \alpha_i \sum_{j=1}^{m_i} |\langle \alpha_i^j | \psi \rangle|^2 \\ &= \frac{1}{\|\psi\|^2} \sum_{i=1}^k \sum_{j=1}^{m_i} \alpha_i \langle \psi | \alpha_i^j \rangle \langle \alpha_i^j | \psi \rangle \\ &= \frac{1}{\|\psi\|^2} \langle \psi | \left(\sum_{i=1}^k \alpha_i \sum_{j=1}^{m_i} \alpha_i |\alpha_i^j\rangle \langle \alpha_i^j| \right) | \psi \rangle, \text{ en utilisant (5.3)} \\ &= \frac{\langle \psi | A | \psi \rangle}{\|\psi\|^2} \end{aligned}$$

La preuve précédente n'est pas très complexe d'un point de vue mathématique, mais elle peut rebuter au premier abord en raison des notations utilisées. Le lecteur est invité à reprendre la démonstration dans un cas particulier (avec par exemple $k = 3$ et sans multiplicité) afin de se familiariser avec les habitudes d'écriture des physiciens. Avec un peu de pratique on se rend compte qu'elles sont très commodes. On peut cependant arriver au résultat de façon beaucoup plus élégante à l'aide des projecteurs.

Exercice 6.1.1 Redémontrer la formule précédente à l'aide des projecteurs π_i sur les espaces propres.

En appliquant ce qui précède à A^2 on voit que l'espérance de a^2 vaut $\frac{\langle \psi | A^2 | \psi \rangle}{\|\psi\|^2}$. Il en résulte que l'écart-type de a dans l'état ψ est :

$$(\Delta A)_\psi = \frac{1}{\|\psi\|} (\langle \psi | A^2 | \psi \rangle - \langle \psi | A | \psi \rangle^2)^{1/2} \quad (6.2)$$

6.2 Exemple : l'ion H_2^+

Nous voudrions maintenant appliquer les postulats énoncés à une situation concrète. Malheureusement, le traitement rigoureux des situations même les plus simples entraîne le plus souvent l'utilisation d'espaces d'états de dimension infinie, et nous voudrions démarrer avec un exemple où l'on puisse tout calculer à l'aide de l'algèbre linéaire élémentaire. Il existe un système très important qui répond à ce besoin : la description des états de spin d'une particule telle que l'électron. Cependant nous n'avons pas encore parlé de spin, et nous ne voulons pas introduire maintenant cette nouvelle difficulté. C'est pourquoi nous optons pour l'étude d'un système qui peut, en première approximation, se traiter à l'aide d'un espace des états de dimension 2 (on dit parfois « un système à deux niveaux »). Il s'agit de l'ion positif H_2^+ . Notre traitement suit essentiellement [10] p. 45-48. Voir aussi [8], p. 174-181.



FIGURE 6.1 – Les deux états de l'ion H_2^+

Ce système est constitué de deux protons p_1^+ et p_2^+ et d'un électron e^- . Nous supposons que la distance d entre les deux protons n'évolue pas au cours du temps, et qu'au niveau d'approximation où nous nous plaçons, l'ion H_2^+ puisse être observé dans l'un des deux états suivants : $H - H^+$ ou $H^+ - H$. La signification précise de cet énoncé est la suivante : si l'on mesure la position de l'électron, on trouvera qu'il est soit proche de p_1^+ , soit proche de p_2^+ . À ces deux états nous associons respectivement des vecteurs d'états notés $|1\rangle$ et $|2\rangle$, formant une base orthonormée d'un espace de Hilbert de dimension 2.

Cette démarche respecte les postulats que nous avons énoncés. En effet, ce que nous affirmons en réalité c'est que nous considérons que le système est entièrement décrit par une variable dynamique, la position x de l'électron par rapport au centre de masse des deux protons, ne pouvant prendre que deux valeurs. À cette variable dynamique est associée une observable X conformément au postulat 3. Les résultats d'une mesure de x donnent toujours $-d/2$ ou $d/2$ selon notre hypothèse. Il en résulte que le spectre de X est constitué de ces deux valeurs. L'opérateur le plus simple ayant ce spectre a pour matrice $\begin{pmatrix} -d/2 & 0 \\ 0 & d/2 \end{pmatrix}$

dans une base de vecteurs propres. L'état $|1\rangle$ introduit plus haut est simplement un vecteur propre normalisé de X associé à la valeur propre $-d/2$, et $|2\rangle$ est un vecteur propre normalisé associé à l'autre valeur propre. Le point sur lequel nous voulons insister est le suivant : quand nous disons que le système ne peut s'observer dans l'un ou l'autre de deux états *lorsque l'on regarde où se trouve l'électron*, nous ne voulons pas dire que le système ne peut pas être dans un autre état. En fait, le principe de superposition quantique indique que toutes les combinaisons linéaires complexes du type $\alpha|1\rangle + \beta|2\rangle$, c'est-à-dire tous les éléments de l'espace vectoriel complexe engendré par $|1\rangle$ et $|2\rangle$, représentent des états du système. Simplement, dans un état « superposé » la position de l'électron n'est pas déterminée. Elle ne l'est que si α ou β est nul (le vecteur d'état est alors un vecteur propre de X).

Nous nous demandons maintenant ce qui se passe si nous mesurons l'énergie du système. On sait que l'observable correspondant à l'énergie est le hamiltonien H . Comme il s'agit d'un opérateur hermitien,

sa matrice dans la base $|1\rangle, |2\rangle$ est du type $\begin{pmatrix} H_{11} & H_{12} \\ \bar{H}_{12} & H_{22} \end{pmatrix}$ où H_{11} et H_{22} sont réels. En fait on a $H_{11} = \langle 1|H|1\rangle = \langle H \rangle_1$: c'est la valeur moyenne de l'énergie dans l'état $|1\rangle$. La symétrie évidente entre les états $|1\rangle$ et $|2\rangle$ indique que la valeur moyenne de l'énergie mesurée dans ces deux états doit être la même. Ainsi on a $H_{11} = H_{22}$. Notons E cette valeur commune. Notons d'autre part $H_{12} = ae^{i\theta}$ la décomposition polaire du complexe H_{12} . La diagonalisation de H se fait sans problème. Le polynôme caractéristique de H est $\chi_H(\lambda) = \lambda^2 - 2E\lambda + E^2 - a^2$ et son spectre est l'ensemble $\{E - a; E + a\}$.

Dans le cas où $a = 0$, le hamiltonien n'est rien d'autre qu'un multiple de l'identité. Ce cas correspond donc à un système où tous les états sont stationnaires. Ici en particulier, cela signifie que si l'on place le système dans l'état $|1\rangle$ (respectivement dans l'état $|2\rangle$), il y reste ad vitam aeternam. Autrement dit, si $a = 0$ l'électron n'a aucune chance de sauter d'un proton à l'autre.

Gardons ceci en tête, et analysons ce qui se passe si $a \neq 0$. Dans ce cas on voit que H est diagonalisable, et admet comme base orthonormée de vecteurs propres la base $|e_1\rangle, |e_2\rangle$, avec $|e_1\rangle = \frac{1}{\sqrt{2}}(e^{i\theta}|1\rangle - |2\rangle)$ correspondant à la valeur propre $E - a$, et $|e_2\rangle = \frac{1}{\sqrt{2}}(e^{i\theta}|1\rangle + |2\rangle)$ correspondant à la valeur propre $E + a$. Notons que $E - a$ et $E + a$ sont les seules valeurs que l'on peut obtenir lorsque l'on mesure l'énergie du système. Ce sont donc nécessairement des réels positifs, donc on a $0 < a \leq E$. L'état $|e_1\rangle$ est l'état de plus basse énergie : on dit que c'est l'état fondamental.

Que se passe-t-il si l'on mesure la position de l'électron quand le système est dans l'état $|e_1\rangle$? En vertu du postulat 5, la probabilité de trouver l'électron du côté de p_1^+ est $P(H - H^+|e_1) = |\langle 1|e_1\rangle|^2 = 1/2$. Si maintenant on mesure l'énergie du système quand il est dans l'état $|1\rangle$, on trouve qu'il y a une probabilité $1/2 = |\langle e_1|1\rangle|^2$ de trouver $E - a$ et une probabilité $1/2 = |\langle e_2|1\rangle|^2$ de trouver $E + a$.

Nous reviendrons plus loin sur la signification de ces calculs pour l'ion H_2^+ .

Exercice 6.2.1 Calculer la moyenne et l'écart-type de la position de l'électron dans l'état $|e_1\rangle$, puis dans l'état $|e_2\rangle$.

6.3 Relations d'incertitude et conséquences

6.3.1 Énoncé des relations

On dit parfois un peu rapidement qu'une observation perturbe toujours le système observé. Ceci serait l'essence du « principe d'incertitude de Heisenberg », et de la mécanique quantique. Nous avons vu que des considérations de ce type pouvaient bien servir de justification à l'abandon du concept de trajectoire bien définie d'une particule, et mener *in fine* à la mécanique quantique. Pourtant, on ne saurait parler de « principe d'incertitude » au sens où il s'agirait d'un principe physique sur lequel la mécanique quantique serait fondé. Il est préférable de parler de relations d'incertitudes, et celles-ci se démontrent à l'aide des postulats que nous avons énoncés. Il s'agit en fait simplement d'une application de l'inégalité de Cauchy-Schwarz!

Théorème 6.3.1 Soient A et B deux observables, $[A, B] = AB - BA$ leur commutateur, et ψ un vecteur d'état normalisé quelconque. Alors on a

$$(\Delta A)_\psi (\Delta B)_\psi \geq \frac{1}{2} |\langle [A, B] \rangle_\psi| \quad (6.3)$$

Démonstration: Commençons par remarquer que l'on peut remplacer les observables A et B par les observables centrées $A' = A - \langle A \rangle_\psi \text{Id}$ et $B' = B - \langle B \rangle_\psi \text{Id}$, puisqu'on a $(\Delta A)_\psi = (\Delta A')_\psi$, $(\Delta B)_\psi = (\Delta B')_\psi$ et $[A, B] = [A', B']$. Dans la suite on peut supposer que A et B sont centrées.

On a :

$$\begin{aligned} \langle \psi | (AB - BA) \psi \rangle &= \langle A\psi | B\psi \rangle - \langle B\psi | A\psi \rangle \\ &= \langle A\psi | B\psi \rangle - \overline{\langle A\psi | B\psi \rangle} \\ &= 2i \operatorname{im} \langle A\psi | B\psi \rangle \end{aligned}$$

Or $|\operatorname{Im}\langle A\psi|B\psi\rangle| \leq |\langle A\psi|B\psi\rangle|$, et par Cauchy-Schwarz :

$$\begin{aligned} |\langle A\psi|B\psi\rangle| &\leq \sqrt{\langle A\psi|A\psi\rangle}\sqrt{\langle B\psi|B\psi\rangle} \\ &\leq \sqrt{\langle A^2\rangle_\psi}\sqrt{\langle B^2\rangle_\psi} \\ &\leq (\Delta A)_\psi(\Delta B)_\psi \end{aligned}$$

d'où

$$(\Delta A)_\psi(\Delta B)_\psi \geq |\langle A\psi|B\psi\rangle| \geq |\operatorname{Im}\langle A\psi|B\psi\rangle| \geq \frac{1}{2}|\langle[A, B]\rangle_\psi|$$

¶

Soit une particule dans l'espace, représentée classiquement par les projections q_1 , q_2 et q_3 de sa position selon trois axes orthonormés. Définissons également les projections p_1 , p_2 et p_3 de son impulsion. En mécanique analytique on a les relations

$$\{p_i, p_j\} = \{q_i, q_j\} = 0, \quad \{q_i, p_j\} = \delta_{ij} \quad (6.4)$$

qui se traduisent pour les observables correspondantes en mécanique quantique par

$$[P_i, P_j] = [Q_i, Q_j] = 0, \quad [Q_i, P_j] = i\hbar\delta_{ij} \quad (6.5)$$

Ces relations sont appelées *relations de commutation canoniques*¹.

En appliquant le théorème précédent aux observables Q_1 et P_1 , nous obtenons

$$(\Delta Q_1)_\psi(\Delta P_1)_\psi \geq \frac{1}{2}\hbar \quad (6.6)$$

En supposant qu'on note x la coordonnée sur le premier axe, et p_x la projection correspondante de l'impulsion, alors $(\Delta Q_1)_\psi$ représente l'écart-type d'une série de mesures de la coordonnées x de la particule quand celle-ci est dans l'état ψ . Notons Δx cet écart-type, la dépendance en ψ étant ici sans importance. De même, notons Δp_x l'écart-type correspondant pour les impulsions selon (Ox) . On obtient alors l'inégalité de Heisenberg :

$$\Delta x \Delta p_x \geq \frac{\hbar}{2} \quad (6.7)$$

Comme Δx et Δp_x peuvent s'interpréter comme des incertitudes sur, respectivement, la position et l'impulsion selon un certain axe, cette inégalité et les inégalités de même type obtenues en utilisant les axes (Oy) et (Oz) se nomment les *relations d'incertitude de Heisenberg*.

Nous voyons que ces relations proviennent de la non-commutation des observables correspondantes. La non-commutativité de l'algèbre des observables apparaît donc bien comme une des caractéristiques principales de la mécanique quantique. Remarquons également qu'il est faux de dire que toute observation perturbe le système observé. En effet, si un système est dans un état propre $|\psi\rangle$ d'une observable A , et que B est une observable qui commute² avec A , alors $|\psi\rangle$ est également un état propre de B . On peut donc faire autant de fois qu'on veut des mesures de A et de B , le vecteur $|\psi\rangle$ est toujours projeté sur lui-même et l'état du système n'est donc absolument pas perturbé.

Nous pouvons généraliser sans difficulté les relations d'incertitude de Heisenberg. En effet, nous venons de voir qu'il suffit de trouver deux observables qui ne commutent pas pour obtenir de telles relations. Mais la non-commutation des observables est elle-même directement liée à la non-annulation des crochets de Poisson des variables classiques correspondantes. On aura donc une relation d'incertitude de type Heisenberg pour tout couple d'observables dont les variables classiques correspondantes ont des crochets

1. Notez qu'il y a deux i qui n'ont rien à voir : le nombre complexe et l'indice allant de 1 à 3. Nous espérons qu'aucune confusion n'est possible.

2. On peut prendre comme exemple la position selon un axe et l'impulsion selon un autre axe.

de Poisson non nuls. C'est le cas par exemple des moments cinétiques relativement à des axes différents. Nous reviendrons sur ce point dans la discussion du spin.

Il existe cependant une exception importante à l'analyse qui précède. En effet, la coordonnée de temps n'est pas représentée en mécanique quantique par un opérateur. On peut toutefois déterminer une relation d'incertitude temps-énergie en procédant directement à partir de l'équation d'évolution de Schrödinger. En effet on a, d'après (5.11) :

$$\hbar^2 \left\| \frac{d\psi}{dt} \right\|^2 = \|H\psi\|^2 = \langle H^2 \rangle_\psi := (\Delta E)_\psi^2$$

Le terme $\left\| \frac{d\psi}{dt} \right\|$ est une mesure de l'évolution du vecteur d'état. Comme on peut toujours supposer le vecteur d'état normalisé³, il faut imaginer que celui-ci se déplace à la surface d'une sphère. Si on désigne par τ un temps d'évolution caractéristique du système pendant lequel le vecteur d'état s'écarte de sa position initiale d'une « distance »⁴ de l'ordre de l'unité, on aura, en ordre de grandeur :

$$\tau \Delta E \sim \hbar \tag{6.8}$$

Il est important de noter que τ ne désigne pas une incertitude sur une mesure du temps : une telle mesure peut être rendue arbitrairement précise en mécanique quantique.

6.3.2 Conséquences

Stabilité de la matière

Considérons un électron dans un atome. Pour ne pas changer, nous allons prendre le plus simple des atomes, c'est-à-dire l'atome d'hydrogène. Notons r sa distance au noyau et p sa quantité de mouvement. Par (6.6) on obtient :

$$(\Delta R)(\Delta P) \gtrsim \hbar$$

On peut dire que plus l'électron va s'approcher du noyau, mieux il sera localisé. Sa quantité de mouvement va donc subir de grandes variations, qui vont avoir tendance à l'éloigner du noyau. Intuitivement, il rencontre une sorte de « répulsion de Heisenberg » quand il se rapproche du noyau, répulsion qui entre en compétition avec l'attraction coulombienne. On va essayer de préciser un peu cet argument, mais n'exagérons rien, cela reste un raisonnement très heuristique.

La moyenne $\langle E_c \rangle$ de l'énergie cinétique de l'électron est $\frac{1}{2m_e} \langle P^2 \rangle = \frac{1}{2m_e} (\Delta P)^2$, puisque, par symétrie, $\langle P \rangle = 0$. On a donc, en ordre de grandeur :

$$\langle E_c \rangle \sim \frac{\hbar^2}{2m_e (\Delta R)^2} \gtrsim \frac{\hbar^2}{2m_e \langle R \rangle^2}$$

En notant q la charge fondamentale, l'énergie potentielle coulombienne est de l'ordre

$$V \sim -\frac{q^2}{4\pi\epsilon_0 \langle R \rangle}$$

Bien sûr, l'inverse de la moyenne n'est pas la moyenne des inverses ! Encore une fois nous raisonnons seulement sur les ordres de grandeurs. L'énergie totale vérifie donc

$$E \sim \frac{\hbar^2}{2m_e \langle R \rangle^2} - \frac{q^2}{4\pi\epsilon_0 \langle R \rangle}$$

Il se trouve que cette expression est minorée. Une simple étude de fonction (voir figure 6.2) montre qu'elle atteint un minimum pour $\langle R \rangle \simeq 0,53 \text{ \AA}$, et qu'elle y prend une valeur minimum de... -13,6 eV ! Sans

3. Sa norme est de toute façon constante puisque l'évolution est unitaire.

4. Il s'agit d'une distance sans unité dans l'espace de Hilbert des états.

prendre cette coïncidence numérique trop au sérieux, on peut quand même conclure grâce aux ordres de grandeur que l'énergie de l'électron ne peut pas tendre vers $-\infty$, ce qu'elle fait en mécanique classique. Ici, elle doit être bornée inférieurement, et donc posséder un minimum de l'ordre de $-13,6$ eV.

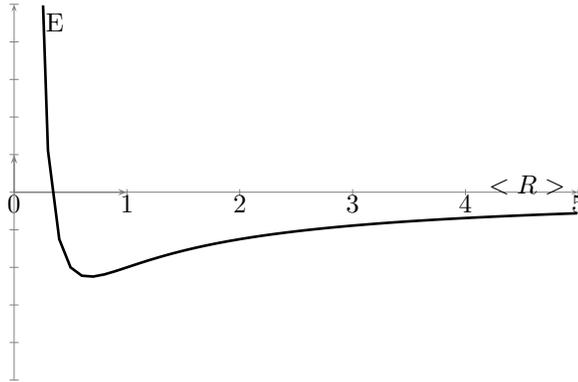


FIGURE 6.2 – Sur la partie gauche de la courbe, la contribution de la pression de Heisenberg est dominante, tandis que sur la partie droite, elle est négligeable devant l'énergie potentielle coulombienne.

La mécanique quantique explique donc la stabilité de la matière, via les relations d'incertitude, par la non-commutativité d'une certaine algèbre!

6.4 Le théorème d'Ehrenfest

Soit O une observable et ψ un vecteur d'état. On désire calculer la dérivée de l'espérance $\langle O \rangle_\psi$ de O quand le système est dans l'état ψ . Notons que, comme nous utilisons le point de vue de Schrödinger, l'observable O ne dépend pas du temps, mais l'état ψ en dépend. On a :

$$\begin{aligned}
 \frac{d}{dt} \langle O \rangle_\psi &= \frac{d}{dt} \langle \psi(t) | O | \psi(t) \rangle \\
 &= \left\langle \frac{d\psi}{dt} | O \psi(t) \right\rangle + \langle \psi(t) | O \frac{d\psi}{dt} \rangle \\
 &= \left\langle \frac{1}{i\hbar} H \psi(t) | O \psi(t) \right\rangle + \langle \psi(t) | \frac{1}{i\hbar} O H \psi(t) \rangle \\
 &= -\frac{1}{i\hbar} \langle \psi(t) | H O \psi(t) \rangle + \frac{1}{i\hbar} \langle \psi(t) | O H \psi(t) \rangle \\
 &= \frac{1}{i\hbar} \langle [O, H] \rangle_\psi
 \end{aligned} \tag{6.9}$$

Appliquons ceci au cas d'une particule de masse m dans \mathbb{R}^3 soumise à un potentiel $V(r)$. Le hamiltonien classique s'écrit :

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

En notant avec un chapeau les observables correspondantes, on postule⁵ que le hamiltonien quantique est

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}, \hat{y}, \hat{z}), \text{ où } \hat{p}^2 := \hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2 \tag{6.10}$$

5. Il s'agit bien d'un postulat : on pourrait ajouter à H un terme en $pr - rp$ qui serait classiquement invisible, puisque les c -nombres commutent. Néanmoins le postulat que nous utilisons est justifié par les bons résultats qu'il donne. En physique, c'est souvent la solution la plus simple qui est la bonne!

Nous aurons besoin dans la suite d'un petit lemme d'algèbre très simple.

Lemme 6.4.1 Soient a, b des q -nombres (des éléments d'une \mathbb{C} -algèbre quelconque), tels que leur commutateur soit une constante $\alpha \in \mathbb{C}$, i.e. $[a, b] = ab - ba = \alpha$. Alors pour tout polynôme P , on a :

$$[a, P(b)] = \alpha P'(b)$$

Preuve du Lemme: La démonstration se fait par récurrence sur le degré du polynôme. Par linéarité du commutateur, on peut supposer que P est un monôme. Si $P(b) = b$, le résultat est trivial. Supposons qu'on ait montré que $[a, b^n] = \alpha n b^{n-1}$. Alors on a :

$$\begin{aligned} [a, b^{n+1}] &= ab^{n+1} - b^{n+1}a \\ &= abb^n - bab^n + bab^n - bb^n a \\ &= [a, b]b^n + b[a, b^n] \\ &= \alpha b^n + b(n\alpha b^{n-1}), \text{ par hypothèse de récurrence} \\ &= \alpha(n+1)b^n \end{aligned}$$

♣

Nous avons les relations de commutation canoniques (6.5) entre les observables de position et d'impulsions. Nous pouvons en déduire grâce au lemme que :

$$[\hat{x}, \hat{p}^2] = [\hat{x}, p_x^2] = 2i\hbar \hat{p}_x$$

et de même

$$[\hat{y}, \hat{p}^2] = 2i\hbar \hat{p}_y, \quad [\hat{z}, \hat{p}^2] = 2i\hbar \hat{p}_z$$

Si le potentiel est un polynôme en x, y, z , le lemme donne directement

$$[\hat{p}_x, V(\hat{x}, \hat{y}, \hat{z})] = -i\hbar \frac{\partial V}{\partial x}(\hat{x}, \hat{y}, \hat{z})$$

et des relations similaires pour $[\hat{p}_y, V]$ et $[\hat{p}_z, V]$. On peut généraliser ces relations au cas où on peut approcher suffisamment bien la fonction V par des polynômes. Nous admettrons dans la suite qu'elles sont valables.

En regroupant les trois observables d'impulsion dans un vecteur $\hat{p} = \begin{pmatrix} \hat{p}_x \\ \hat{p}_y \\ \hat{p}_z \end{pmatrix}$ et les trois observables

de position dans un vecteur $\hat{r} = \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{pmatrix}$, on peut écrire symboliquement les relations de commutations démontrées ci-dessus :

$$[\hat{r}, \hat{p}^2] = 2i\hbar \hat{p}, \text{ et } [\hat{p}, V(\hat{r})] = -i\hbar \vec{\nabla} V(\hat{r}) \quad (6.11)$$

Par ailleurs il est clair que

$$[\hat{r}, V(\hat{r})] = 0, \text{ et } [\hat{p}, \hat{p}^2] = 0 \quad (6.12)$$

Il faut bien garder à l'esprit que les relations (6.11) et (6.12) représentent en fait chacune 2×3 relations, soit 12 relations en tout.

L'application de (6.9) donne donc

$$\begin{aligned} \frac{d}{dt} \langle \hat{r} \rangle &= \frac{1}{i\hbar} \langle [\hat{r}, \frac{\hat{p}^2}{2m} + V(\hat{r})] \rangle \\ &= \frac{1}{m} \langle \hat{p} \rangle \end{aligned} \quad (6.13)$$

et

$$\begin{aligned}\frac{d}{dt}\langle\hat{p}\rangle &= \frac{1}{i\hbar}\langle[\hat{p}, \frac{\hat{p}^2}{2m} + V(\hat{r})]\rangle \\ &= -\langle\vec{\nabla}V(\hat{r})\rangle\end{aligned}\tag{6.14}$$

Les formules (6.13) et (6.14) ressemblent beaucoup aux relations classiques

$$\vec{p} = m\frac{d\vec{r}}{dt}, \text{ et } \frac{d\vec{p}}{dt} = -\nabla V$$

Dans la mesure où l'on peut écrire l'approximation

$$\langle V(\hat{r}) \rangle \approx V(\langle \hat{r} \rangle)\tag{6.15}$$

alors on peut dire que la moyenne des observables de position et d'impulsion satisfont les lois de la mécanique classique. En particulier, si les écart-types Δr et Δp sont assez petits pour être négligés, et si le potentiel varie assez lentement sur un intervalle de longueur Δr , alors on retrouve une particule classique dont la position et l'impulsion sont bien définies et évoluent selon les lois classiques du mouvement.

Nous voyons donc que, au moins dans certaines conditions, les équations de la mécanique classique découlent de l'équation d'évolution quantique. Il s'agit bien sûr d'une nécessité : on sait bien que la mécanique classique est valide pour les boules de billard ! Mais ce n'était pas acquis d'avance, car l'équation de Schrödinger devient singulière quand on fait tendre \hbar vers 0. La question de la limite classique de la mécanique quantique est donc assez subtile, et il s'agit d'ailleurs toujours d'un sujet de recherche. Le théorème d'Ehrenfest doit donc être vu comme un résultat de cohérence interne, qui éclaire le rapport entre mécanique classique et quantique, sans complètement clore le sujet.

6.5 L'oscillateur harmonique

Dans cette section nous allons procéder à la quantification d'un système physique aussi simple qu'important : l'oscillateur harmonique. Nous voulons dire par là que nous cherchons un système quantique qui puisse s'interpréter à la limite classique comme une particule soumise à une force de rappel proportionnelle à l'écart à la position d'équilibre. Pour simplifier nous considérerons le cas d'une particule restreinte à un axe. Nous notons q la position sur l'axe, la position d'équilibre définissant l'origine. Classiquement, les équations du mouvement s'écrivent :

$$p = m\frac{dq}{dt}, \quad \frac{dp}{dt} = -kq\tag{6.16}$$

La force de rappel $-kq$ dérive de l'énergie potentielle $V = \frac{1}{2}kq^2$. L'expression de l'énergie mécanique totale en fonction de p et q permet d'écrire le hamiltonien classique⁶

$$H = \frac{p^2}{2m} + \frac{1}{2}kq^2\tag{6.17}$$

La constante k étant positive, on peut l'écrire sous la forme $k = m\omega^2$. La résolution des équations du mouvement montre des solutions sinusoidales, et ω est leur pulsation.

Un hamiltonien quantique qui redonne à l'évidence (6.17) quand on remplace les q -nombres par des c -nombres est :

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{q}^2\tag{6.18}$$

6. Tout potentiel tel que 0 soit un point d'équilibre stable peut se développer sous la forme $V(0) + kq^2 + o(q^2)$, avec $k \geq 0$, ce qui explique l'omniprésence de l'oscillateur harmonique.

Ce n'est bien sûr pas la seule possibilité : on pourrait ajouter un terme proportionnel à $\hat{p}\hat{q} - \hat{q}\hat{p}$, par exemple. Mais c'est la plus simple. Ainsi, il faut voir (6.18) comme une définition de ce qu'est un oscillateur harmonique quantique.

En posant

$$Q = \sqrt{\frac{m\omega}{\hbar}} \hat{q}, \text{ et } P = \frac{\hat{p}}{\sqrt{m\omega\hbar}} \quad (6.19)$$

On obtient le hamiltonien quantique sous une forme plus agréable :

$$\hat{H} = \frac{1}{2}\hbar\omega(Q^2 + P^2) \quad (6.20)$$

Enfin, il va s'avérer judicieux de poser :

$$a = \frac{1}{\sqrt{2}}(Q + iP) \quad (6.21)$$

Les observables P et Q vérifient $P^* = P$ et $Q^* = Q$, d'où l'on déduit que

$$a^* = \frac{1}{\sqrt{2}}(Q - iP) \quad (6.22)$$

À l'aide de la relation de commutation canonique $[\hat{q}, \hat{p}] = i\hbar$, on montre que (exercice)

$$[a, a^*] = 1 \quad (6.23)$$

Les q -nombres a et a^* s'appellent respectivement les opérateurs d'annihilation et de création. Nous verrons plus loin la justification de cette terminologie.

On vérifie enfin sans difficulté que

$$\hat{H} = \hbar\omega(N + \frac{1}{2}), \text{ avec } N = a^*a \quad (6.24)$$

L'observable N se nomme « opérateur de nombre ».

À l'aide de (6.23), on voit que

$$[N, a] = -a, \text{ et } [N, a^*] = a^* \quad (6.25)$$

L'opérateur de nombre est hermitien. Sous certaines hypothèses mathématiques, on peut en déduire qu'il possède une base orthonormée de vecteurs propres. Ceci pourrait se justifier en passant « en représentation r » et en résolvant l'équation de Schrödinger. Notre but cependant est de montrer que l'on peut dire beaucoup de choses en raisonnant de façon complètement algébrique à partir des relations de commutation (6.25) et des postulats. Nous allons donc montrer le théorème suivant, sous l'hypothèse que N possède bien une base orthonormée de vecteurs propres.

Théorème 6.5.1 1. Le spectre de N est \mathbb{N} . Il en résulte que le spectre de \hat{H} est $\{(n+1/2)\hbar\omega | n \in \mathbb{N}\}$.

2. Si $|\psi_n\rangle$ est un vecteur propre de N associé à la valeur propre n , alors $a^*|\psi_n\rangle$ est un vecteur propre de N associé à la valeur propre $n+1$.

3. Si $|\psi_n\rangle$ est un vecteur propre de N associé à la valeur propre n , alors si $n \neq 0$, $a|\psi_n\rangle$ est un vecteur propre de N associé à la valeur propre $n-1$, et si $n=0$, $a|\psi_0\rangle = 0$.

Démonstration: Soit $|\psi_n\rangle$ un vecteur propre de N associé à la valeur propre n . On ne présume pas pour l'instant du fait que n soit un entier. Calculons la norme de $a|\psi_n\rangle$:

$$\|a|\psi_n\rangle\|^2 = \langle\psi_n|a^*a|\psi_n\rangle = n\|\psi_n\|^2$$

Ceci montre que $n \geq 0$ et que $a|\psi_n\rangle = 0$ ssi $n=0$. Montrons maintenant le troisième point de l'énoncé.

$$\begin{aligned}
Na|\psi_n\rangle &= (aN - a)|\psi_n\rangle, \text{ d'après (6.25)} \\
&= an|\psi_n\rangle - a|\psi_n\rangle \\
&= (n - 1)a|\psi_n\rangle
\end{aligned}$$

On a donc $N(a|\psi_n\rangle) = (n - 1)(a|\psi_n\rangle)$. De plus on sait que $a|\psi_n\rangle$ est non nul ssi $n \neq 0$. Donc si $n \neq 0$, alors $a|\psi_n\rangle$ est bien un vecteur propre de N associé à la valeur propre $n - 1$.

Le deuxième point se montre de façon très similaire, et est laissé à titre d'exercice.

Il reste à montrer que n est un entier. En appliquant a suffisamment de fois au vecteur $|\psi_n\rangle$, on peut trouver des vecteurs propres associés aux valeurs propres $n, n - 1, n - 2, \dots$. Or si n n'est pas un entier, alors cette suite ne passe pas par 0, et peut se prolonger jusqu'à atteindre des valeurs négatives. Cela contredirait ce qu'on a déjà démontré. \blacksquare

Pour déterminer les vecteurs propres de N , il faut d'abord montrer que les valeurs propres sont non dégénérées. Pour cela passons en représentation r . Rappelons que dans cette représentation, \hat{p} s'interprète comme l'opérateur $\frac{\hbar}{i} \frac{d}{dx}$ et \hat{q} comme la multiplication par x . L'équation $a|\psi_0\rangle = 0$ devient alors :

$$\begin{aligned}
\sqrt{\frac{m\omega}{\hbar}} x \psi_0(x) + \frac{\hbar}{\sqrt{m\hbar\omega}} \psi_0'(x) &= 0 \\
\Leftrightarrow \frac{m\omega}{\hbar} x \psi_0(x) + \psi_0'(x) &= 0
\end{aligned}$$

Cette équation différentielle linéaire du premier ordre s'intègre sans problème. Les solutions sont :

$$\psi_0(x) = C_0 e^{-\frac{m\omega}{2\hbar} x^2} \quad (6.26)$$

où C_0 est une constante. Donc il y a une droite vectorielle de solutions. L'espace propre associé à la valeur propre 0 est de dimension 1, et en faisant agir $(a^*)^n$ sur cet espace propre, on trouve l'espace propre associé à la valeur propre n . Ainsi, les vecteurs propres $|\psi_n\rangle$ sont de la forme

$$|\psi_n\rangle = (a^*)^n |\psi_0\rangle$$

Soit, en représentation r :

$$\psi_n(x) = C_n \left[\sqrt{\frac{m\omega}{\hbar}} x - \sqrt{\frac{\hbar}{m\omega}} \frac{d}{dx} \right]^n e^{-\frac{m\omega}{2\hbar} x^2} \quad (6.27)$$

Les vecteurs propres de N sont aussi ceux de \hat{H} , puisque $\hat{H} = \hbar\omega(N + 1/2)$. À titre de comparaison, essayons de déterminer directement en représentation r les vecteurs propres de \hat{H} . L'équation

$$H|\psi_n\rangle = E|\psi_n\rangle$$

devient donc

$$-\frac{\hbar^2}{2m} \psi_n''(x) + \frac{1}{2} m\omega^2 x^2 \psi_n(x) = E\psi_n(x) \quad (6.28)$$

L'équation (6.28) est linéaire du second ordre à coefficients non constants : il n'y a pas de méthode générale pour résoudre ce type d'équation. Il se trouve que celle-ci est connue en mathématique, et a été résolue par Hermite (voir [1] p 594 et suivantes pour une résolution analytique et des détails sur les polynômes de Hermite). Elle a des solutions pour toute valeur de $E > 0$, mais celle-ci tendent génériquement vers l'infini à l'infini, ce qui n'est pas compatible avec l'interprétation probabiliste de la fonction d'onde. En fait, on peut montrer que c'est seulement pour les valeurs $E_n = \hbar\omega(n + 1/2)$, où $n \in \mathbb{N}$, que les solutions de (6.28) sont de carré sommables, et peuvent donc être normalisées grâce à un choix de constante.

Nous avons donc retrouvé les solutions normalisables de façon automatique grâce à notre petit détour dans l'algèbre!

Mais trêve de publicité pour l’algèbre. Le plus important dans tout ceci c’est que le spectre de l’énergie est :

$$E_n = \hbar\omega(n + 1/2), \quad n \in \mathbb{N}$$

Ce spectre possède trois particularités remarquables : il est discret, deux valeurs propres successives sont toujours séparées par une même quantité $\hbar\omega$, et il possède une plus petite valeur

$$E_0 = \frac{1}{2}\hbar\omega$$

qui est non nulle!

Rappelons-nous que pour expliquer le spectre du corps noir, Planck avait postulé l’existence d’oscillateurs harmoniques dans les parois du four, qui pouvait échanger avec le champ de l’énergie uniquement par paquet discret. Le spectre de l’oscillateur harmonique que nous venons de déterminer nous montre que cette hypothèse était justifiée.

Le vecteur propre normalisé associé est noté $|0\rangle$. Il ne faut pas le confondre avec le vecteur nul. On l’appelle l’état fondamental (ground state). L’énergie associée est appelée énergie de point zéro. Le vecteur propre normalisé associé à la valeur propre $\hbar\omega(n + 1/2)$ est noté $|n\rangle$. On voit que l’action de l’opérateur a^* crée un quantum d’énergie $\hbar\omega$, tandis que l’action de a en détruit un, justifiant ainsi la terminologie.

Exercice 6.5.1 Calculer à l’aide de (6.26) la constante de normalisation pour l’état fondamental.

On peut comprendre ce phénomène à l’aide des relations d’incertitude. Celles-ci interdisent en effet de minimiser simultanément la partie cinétique et la partie potentielle (qui ne dépend ici que de la position) de l’énergie. Plus précisément, on peut voir que \hat{p} et \hat{q} ne commutent pas avec \hat{H} . Il en résulte que dans les états stationnaires (i.e. les états d’énergie bien définie), la position et l’impulsion prennent des valeurs aléatoires. Les écart-types Δq et Δp vérifient bien sûr la relation de Heisenberg. Mais on peut montrer que dans l’état fondamental, le produit $\Delta q \Delta p$ est le plus petit possible.

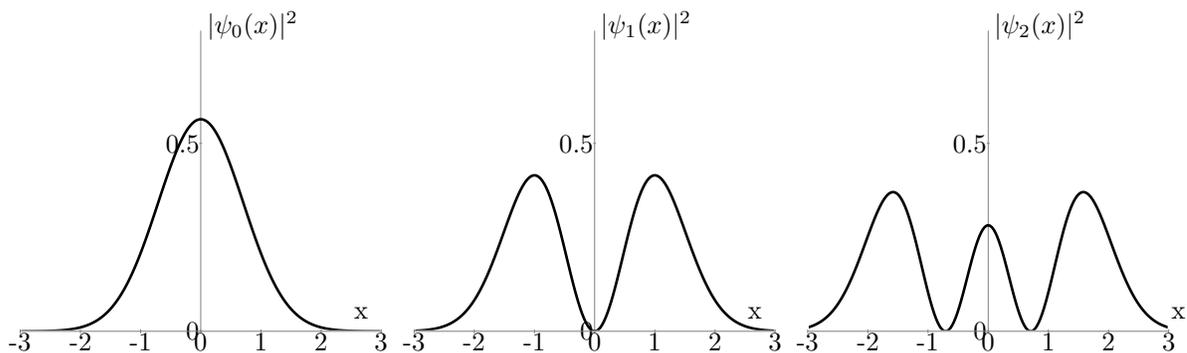
Exercice 6.5.2 Montrer que

$$(\Delta \hat{q})_{|0\rangle} (\Delta \hat{p})_{|0\rangle} = \frac{\hbar}{2}$$

Il est par ailleurs très instructif de calculer les lois de probabilités de la position dans les premiers états d’énergie.

Exercice 6.5.3 Déterminer la densité de probabilité de présence de la particule lorsque le système est dans, respectivement, l’état d’énergie $|0\rangle$, $|1\rangle$, et $|2\rangle$.

Les courbes correspondantes sont représentées ci-dessous.



Il y aurait encore beaucoup à dire sur l'oscillateur harmonique. De nombreuses applications sont données au chapitre V de [7]. Terminons simplement en signalant qu'en théorie quantique des champs, un champ est représenté par une assemblée de tels oscillateurs (un en chaque point de l'espace).

6.6 Effet tunnel

Après l'oscillateur harmonique, nous allons considérer une autre situation très simple où l'équation de Schrödinger peut être résolue exactement : celle d'une particule sur une droite, soumise à un potentiel constant par morceaux. Bien que ce cas soit peu réaliste physiquement, il peut servir à modéliser une situation où le potentiel varie très vite sur un petit intervalle.

Commençons par rappeler ce qui se passe classiquement. Supposons qu'une particule arrive de $-\infty$ avec une énergie E . En 0, elle rencontre une marche de potentiel (voir figure 6.3) de hauteur V_0 . Elle est alors soumise à une force très violente $F = -\frac{\partial}{\partial x}V$ qui la repousse vers les x négatifs. Si la particule a une énergie $E > V_0$ alors elle peut passer la barrière et continue sa course du côté des x positifs (où, le potentiel étant constant, elle n'est plus soumise à aucune force). Au contraire, si $E < V_0$, elle est repoussée vers la gauche.

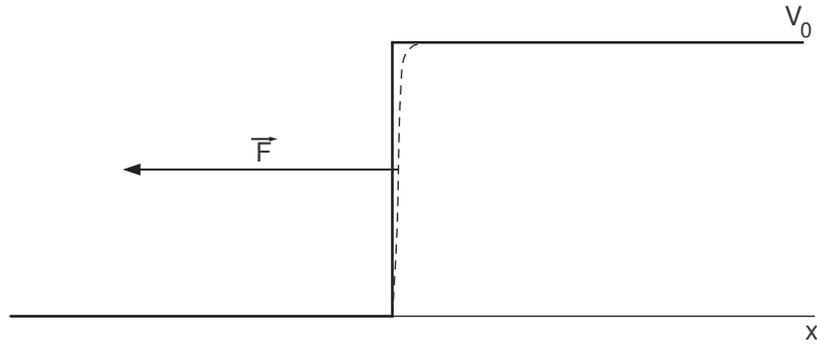


FIGURE 6.3 – La particule classique prend un « coup de pied » quand elle arrive sur la marche. La courbe en pointillé représente le vrai potentiel, dont la marche est une approximation. La force $\vec{F} = -\vec{\nabla}V$ est très grande en norme.

Plaçons-nous maintenant dans le cadre quantique, et considérons une particule d'énergie bien définie E . Soit ψ sa fonction d'onde. Comme $H\psi = E\psi$, l'équation de Schrödinger s'écrit :

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi + V\psi = E\psi$$

Comme on est dans un état stationnaire, on sait que $\psi(x, t) = e^{-\frac{i}{\hbar}Et}\psi(x, 0)$. En posant $\phi(x) = \psi(x, 0)$, l'équation de Schrödinger devient

$$\phi'' + \lambda\phi = 0, \text{ avec } \lambda = \frac{2m}{\hbar^2}(E - V) \quad (6.29)$$

Sur un intervalle où V est constant, il vient deux cas⁷ :

1. Si $E > V$, alors $\lambda > 0$ et on pose $k = \sqrt{\lambda}$. Les solutions sont

$$\phi(x) = Ae^{ikx} + Be^{-ikx} \quad (6.30)$$

2. Si $E < V$, alors $\lambda < 0$, on pose $\kappa = \sqrt{-\lambda}$. Les solutions sont

$$\phi(x) = Ae^{\kappa x} + Be^{-\kappa x} \quad (6.31)$$

Supposons pour commencer que le potentiel V soit identiquement nul sur \mathbb{R} . Il est alors clair que le cas 2 est non physique. En effet, la densité de probabilité de trouver la particule autour de x est $|\psi(x, t)|^2 = |\phi(x)|^2 \rightarrow +\infty$ en $+\infty$ ou en $-\infty$. Ceci est rassurant : les énergies autorisées sont forcément positives⁸. Dans le cas 1, on a $E > 0$, mais les solutions posent tout de même un problème d'interprétation. En effet, ϕ peut se réécrire sous la forme $C \cos(kx + \phi)$, et $|\phi(x)|^2$ a une intégrale divergente, d'où

$$\int_{\mathbb{R}} |\psi(x)|^2 = +\infty$$

Autrement dit, on ne peut pas normaliser ψ pour faire en sorte que $|\psi(x)|^2$ s'interprète comme une densité de probabilité. On dit que ψ est un *état de diffusion* (on dit aussi *non lié*). Un tel état viole les postulats de la mécanique quantique : il ne faut pas le considérer comme un état physique, mais comme une idéalisation d'une situation physique. Voyons pourquoi en nous concentrant sur la solution la plus simple : $\phi(x) = e^{ikx}$. On voit que, pour tout x , $|\phi(x)|^2 = 1$. D'une certaine façon, cet état nous dit que la particule a autant de chance de se trouver en n'importe quel point de \mathbb{R} . Ceci n'est pas sans poser des problèmes physiques (est-il bien raisonnable de considérer que la particule a autant de chance de se trouver dans le laboratoire ou à l'autre bout de la galaxie?), et mathématiques (il n'existe pas de probabilité uniforme sur \mathbb{R}). Le problème se comprend très bien du point de vue des relations d'incertitude. La particule étant dans un état d'énergie E bien définie, et l'énergie d'une particule libre étant proportionnelle au carré de son impulsion, p est bien définie au signe près. Effectivement, la solution générale est une superposition de e^{ikx} et e^{-ikx} qui sont des vecteurs propres de \hat{p} associés respectivement aux valeurs propres $\pm \hbar k$. Dans le cas particulier de $\phi(x) = e^{ikx}$ la particule est dans un état d'impulsion parfaitement définie, et de position complètement indéterminée, conformément aux relations d'incertitude. Ce type d'état peut être traité de façon rigoureuse sur le plan mathématique, mais c'est assez subtil. Ici, la sagesse du physicien consiste à garder en tête qu'une particule n'aura jamais une impulsion parfaitement bien définie, et que les états de diffusion représentent des limites idéales de situations physiques. Encore faut-il savoir correctement les interpréter.

Pour cela il est très instructif de regarder un état physique qui approche d'aussi près que possible un état de diffusion. Un tel état n'aura pas une valeur bien définie de l'impulsion, mais sera une superposition d'états de diffusions d'impulsions très proches d'une valeur moyenne p_0 . Ce qu'on appelle un paquet d'ondes. L'exemple le plus simple qu'on puisse imaginer est de poser

$$\psi_\epsilon(x, t) = \frac{1}{2\epsilon} \int_{p_0-\epsilon}^{p_0+\epsilon} e^{\frac{i}{\hbar}(px-Et)} dp \quad (6.32)$$

Notons que nous ne pouvons plus séparer la dépendance en temps et en x . En effet, l'état n'ayant pas une impulsion bien définie n'a donc pas non plus d'énergie parfaitement définie, et n'est donc pas stationnaire. On a $E = p^2/2m$ dans l'exponentielle, et il faut bien tenir compte de ce terme. En introduisant le petit paramètre $\hbar = p - p_0$, on peut faire un développement limité sous l'intégrale, et on trouve :

$$\psi_\epsilon(x, t) = e^{\frac{i}{\hbar}(p_0x-E_0t)} \left(1 - \frac{\epsilon^2}{3\hbar^2} \left[\left(x - \frac{p_0t}{m} \right)^2 + i\hbar \frac{t}{2m} \right] + \mathcal{O}(\epsilon^4) \right)$$

7. On ne va pas s'intéresser au cas où $E = V$.

8. Dans le cas d'une particule libre, le hamiltonien quantique est $\hat{p}^2/2m$, et ce q -nombre est, en un certain sens, nécessairement positif.

On remarque bien que ψ_ϵ tend vers l'état de diffusion $\psi_0 = e^{\frac{i}{\hbar}(p_0x - E_0t)}$ quand ϵ tend vers 0. De plus, dans la partie réelle du terme en ϵ^2 , on voit apparaître $x - v_0t$, où v_0 est la vitesse moyenne de la particule : on voit une déformation de la fonction d'onde qui se déplace à la vitesse v_0 . Précisons cela en regardant la densité de probabilité $|\psi_\epsilon|^2$:

$$|\psi_\epsilon(x, t)|^2 = 1 - \frac{2\epsilon^2}{3\hbar^2}(x - v_0t)^2 + \mathcal{O}(\epsilon^4)$$

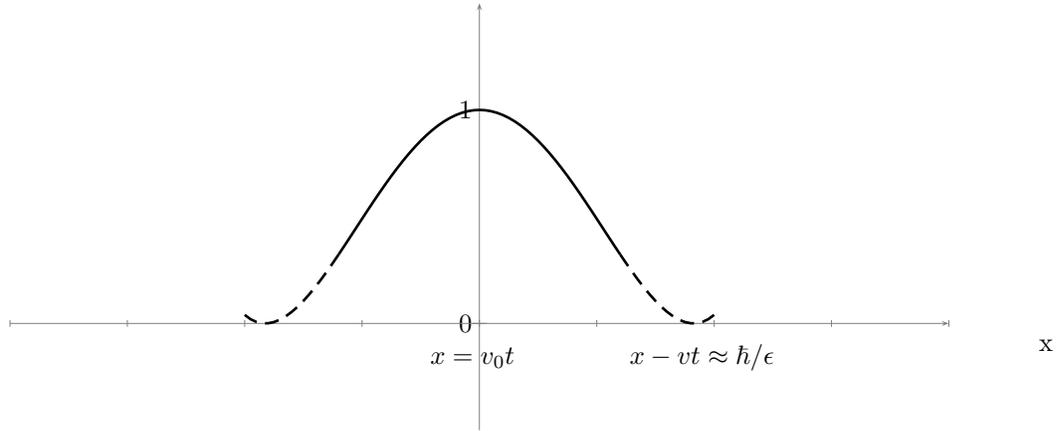


FIGURE 6.4 – Représentation du paquet d'onde ψ_ϵ . Pour la partie en pointillés, le développement limité n'est plus correct.

On voit qu'il s'agit d'une courbe ayant une bosse en $x = v_0t$, dont l'extension est de l'ordre de \hbar/ϵ . Notre méthode ne nous permet pas de dire ce qui se passe quand $x - v_0t$ devient de l'ordre de \hbar/ϵ , puisque c'est précisément à ce moment-là que les termes négligés dans le développement limité deviennent prépondérants. Néanmoins, on peut démontrer par des méthodes plus précises de transformation de Fourier que le paquet d'ondes ψ_ϵ prend effectivement des valeurs non négligeables sur une zone d'extension \hbar/ϵ , d'autant plus grande que ϵ est petit, et que la « bosse » se déplace à la vitesse v_0 .

Le lecteur est invité à imaginer la limite $\epsilon \rightarrow 0$ (la bosse devient de plus en plus large, jusqu'à remplir \mathbb{R}).

Passons maintenant au cas d'une barrière de potentiel d'extension finie L et de hauteur V_0 . Nous allons observer un phénomène typiquement quantique, l'effet tunnel, en envoyant une particule d'énergie $0 < E < V_0$ depuis la gauche sur la barrière.

On a trois régions. Dans la région $]-\infty; 0[$, le potentiel vaut 0, et d'après (6.30), on a dans cette région

$$\phi(x) = Ae^{ikx} + Be^{-ikx}$$

Dans la région $[0; L]$, le potentiel vaut $V_0 > E$, et d'après (6.31), on a dans cette région

$$\phi(x) = Ce^{\kappa x} + De^{-\kappa x}$$

Enfin, dans la région $]L; +\infty[$, le potentiel vaut à nouveau 0 et

$$\phi(x) = Re^{ikx} + Se^{-ikx}$$

Or on envoie une particule depuis la gauche : on peut donc en déduire que $S = 0$. Comme ce n'est pas forcément évident de comprendre que c'est bien S et pas B qui doit être nul, donnons l'explication détaillée, à partir des paquets d'onde. L'état que nous décrivons est une limite de paquet d'onde. On veut modéliser une particule qui arrive depuis la gauche, donc quand $t \rightarrow -\infty$, il doit y avoir une seule bosse, se déplaçant de gauche à droite. Or si nous regardons les états de diffusions dans les régions décrites plus

haut, les deux seuls qui « existent » quand $t \rightarrow -\infty$ sont Ae^{ikx} et Se^{-ikx} . En effet, le terme Re^{ikx} , par exemple, est défini seulement sur $]L; +\infty[$. Il représente une limite de paquet d'ondes arrivant du type de (6.32), mais multiplié par la fonction indicatrice de l'intervalle $]L; +\infty[$. Il s'agit donc d'une bosse qui émerge à un certain instant temps τ du point d'abscisse L . En analysant ainsi chaque terme, on voit que Ae^{ikx} représente une particule arrivant depuis $-\infty$, B^{ikx} une particule rebondissant sur la barrière de potentiel, $Ce^{\kappa x}$ une particule traversant la barrière, $De^{-\kappa x}$, une particule renvoyée par le « puits de potentiel » en L (phénomène étrange mais dont nous verrons qu'il existe bien), Re^{ikx} une particule qui ressort après avoir traversé la barrière, et Se^{-ikx} une particule envoyée de $+\infty$. Nous devons bien prendre $S = 0$.

L'existence de ϕ et de sa dérivée en L (conditions nécessaires pour que l'équation de Schrödinger soit bien définie en ce point) donne le système suivant :

$$\begin{cases} Ce^{\kappa L} + De^{-\kappa L} = Re^{ikL} \\ \kappa Ce^{\kappa L} - \kappa De^{-\kappa L} = ikRe^{ikL} \end{cases}$$

qui conduit immédiatement aux solutions :

$$C = R \frac{\kappa + ik}{2\kappa} e^{(-\kappa + ik)L}; \quad D = R \frac{\kappa - ik}{2\kappa} e^{(\kappa + ik)L} \quad (6.33)$$

De même, l'existence de $\phi(0)$ et de $\phi'(0)$ donnent

$$\begin{cases} A + B = C + D \\ ikA - ikB = \kappa C - \kappa D \end{cases}$$

qui équivaut à

$$A = \frac{(\kappa + ik)C + (-\kappa + ik)D}{2ik}; \quad B = \frac{(-\kappa + ik)C + (\kappa + ik)D}{2ik} \quad (6.34)$$

Nous voyons immédiatement à l'aide de (6.33) et (6.34) que si $R = 0$, c'est-à-dire si une particule d'énergie $E < V_0$ ne peut pas traverser la barrière, comme c'est le cas en mécanique classique, alors $A = B = C = D = 0$. Autrement dit, la particule ne peut pas avoir été envoyée, c'est absurde ! Par contraposée, si $A \neq 0$, alors $R \neq 0$, d'où il suit que C et D sont non nuls. Par conséquent la fonction d'onde doit nécessairement s'étendre dans la barrière et au-delà. Il existe donc une probabilité de trouver la particule dans la barrière (comme si elle avait creusé un tunnel, d'où le nom de cet effet), et au-delà. Regardons comment se comporte la densité de probabilité⁹ de trouver la particule dans la barrière. En entre 0 et L on a, tout calcul fait :

$$|\phi(x)|^2 = |R|^2 (\cosh^2(\kappa(x - L)) + \frac{E}{V_0 - E} \sinh^2(\kappa(x - L)))$$

Cette fonction décroît très vite avec x quand $x \rightarrow L$ (voir figure 6.5). Ce phénomène est similaire aux ondes évanescences en électromagnétisme.

On peut montrer que $\frac{|B|^2}{|A|^2}$ s'interprète comme un coefficient de réflexion, et $\frac{|R|^2}{|A|^2}$ comme un coefficient de transmission.

Les conséquences de l'effet tunnel sont multiples. Dans le cas d'un système qui a classiquement plusieurs configurations stables (le potentiel a plusieurs minima), il y aura une probabilité non nulle pour passer de l'une à l'autre. Par exemple dans le cas de l'ion H_2^+ traité en 6.2, on peut affirmer que $a \neq 0$. Cette possibilité pour l'électron de sauter d'un proton à l'autre est à l'origine de la stabilité de l'ion. Il s'agit d'un type de liaison chimique, appelé liaison π , qui est à l'œuvre dans bien d'autres molécules. Autre exemple de système où ce principe se manifeste : la molécule d'ammoniac. Cette molécule possède deux configurations symétriques de même énergie. L'effet tunnel se manifeste ici par un retournement de la molécule. Dans ces deux cas, on constate que les états d'énergie bien définis sont des superpositions des

9. Ce terme est impropre, puisqu'on est dans un état non normalisable.

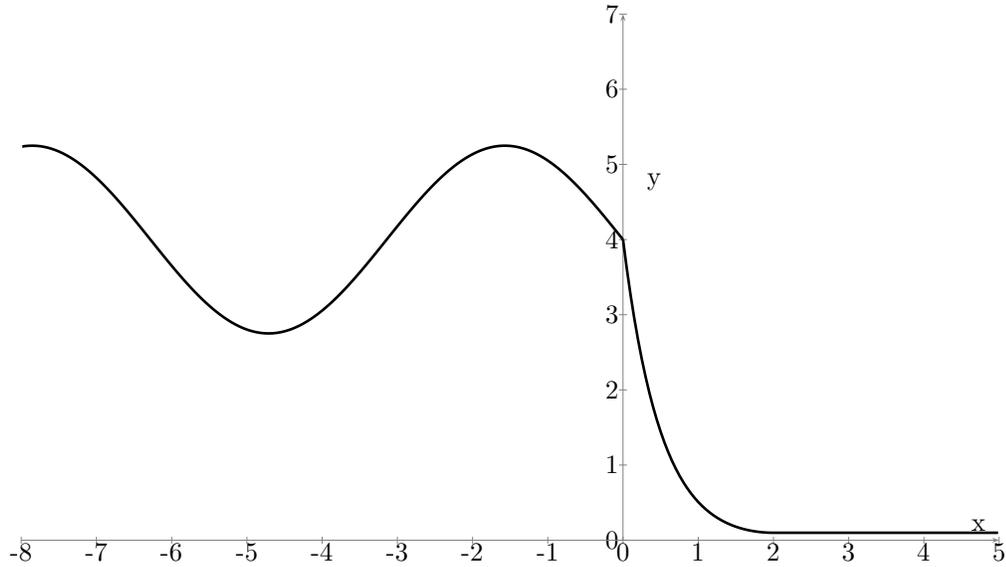


FIGURE 6.5 – Exemple de l’effet tunnel avec une barrière de potentiel entre $x = 0$ et $x = 2$. La fonction représentée est $|\psi(x)|^2$.

états « classiques ». Dans le cas de H_2^+ , les états propres de la position de l’électron ne sont pas des états propres du hamiltonien, et ne sont donc pas des états stationnaires. En termes classiques, on peut dire que la molécule d’ammoniac s’inverse périodiquement. C’est sur ce principe qu’est basé le maser à ammoniac (cf [3], Leçon 6).

Une première modélisation de la radioactivité α utilisant l’effet tunnel a été donnée par Gamow en 1928. Dans ce modèle, la particule α est piégée au sein du noyau par un puits de potentiel, mais peut s’en échapper par effet tunnel. Ce modèle a permis de prédire la demi-vie d’un noyau soumis à ce type de désintégration.

Le microscope à effet tunnel est bien sûr une application directe de cet effet. Dans ce cas, on dispose d’une pointe très fine et d’une surface à examiner. On établit entre les deux une différence de potentiel. En approchant très près la pointe de la surface, les électrons de la surface pourront sauter la barrière par effet tunnel et engendrer un courant, faible mais mesurable, dans la pointe. Un système d’asservissement permet de lever ou d’abaisser la pointe de façon à maintenir le courant constant. De cette façon on réalise une topographie extrêmement précise de la surface.

6.7 Moment cinétique et spin

6.7.1 Moment cinétique

En mécanique classique, le moment cinétique (aussi appelé moment angulaire, en anglais « angular momentum ») d’une particule par rapport à l’origine est :

$$\vec{L} = \vec{r} \times \vec{p}$$

Soit

$$L_x = yp_z - zp_y; \quad L_y = zp_x - xp_z; \quad L_z = xp_y - yp_x$$

On peut associer à ces quantités classiques des observables définies de la façon la plus simple possible :

$$\hat{L}_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y; \quad \hat{L}_y = \hat{z}\hat{p}_x - \hat{x}\hat{p}_z; \quad \hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x \quad (6.35)$$

où l'on a mis des chapeaux sur les q -nombres (observables, opérateurs hermitiens) pour les distinguer des c -nombres. À l'aide des relations de commutation canoniques entre les observables de position et d'impulsion, on montre que

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z; \quad [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x; \quad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y \quad (6.36)$$

Exercice 6.7.1 Démontrer (6.36) et montrer que les crochets de Poisson entre les composantes du moment cinétique classique vérifient des relations semblables.

On peut résumer les relations (6.36) sous la forme

$$\hat{\mathbf{L}} \times \hat{\mathbf{L}} = i\hbar\hat{\mathbf{L}} \quad (6.37)$$

où $\hat{\mathbf{L}}$ est le vecteur de q -nombres $\begin{pmatrix} \hat{L}_x \\ \hat{L}_y \\ \hat{L}_z \end{pmatrix}$. Définissons maintenant trois observables, J_x , J_y et J_z , composantes du vecteur \vec{J} , vérifiant :

$$\vec{J} \times \vec{J} = i\hbar\vec{J} \quad (6.38)$$

On dit alors que \vec{J} est une observable vectorielle de moment cinétique. C'est bien sûr le cas si $\vec{J} = \hat{\mathbf{L}}$, mais il y a un autre cas que nous rencontrerons plus bas, à savoir le spin. Répétons-nous pour bien nous faire comprendre : les relations (6.38) constituent une définition d'une observable de moment cinétique en mécanique quantique. Il existe deux cas : le moment cinétique *orbital* $\hat{\mathbf{L}}$, obtenu par quantification directe du moment cinétique classique, et le spin, que nous verrons plus bas et qui n'a pas d'analogie classique.

Supposons que pour un certain système physique, on ne s'intéresse qu'à son moment cinétique. Manifestement, J_x , J_y et J_z ne commutent pas, on ne pourra pas en mesurer simultanément plus d'une composante. Choisissons arbitrairement de nous intéresser à la composante selon l'axe (Oz). Comme J_z est un opérateur hermitien, on peut le diagonaliser¹⁰. Soit m une valeur propre de J_z . Si l'espace propre associé à cette valeur propre est de dimension $d > 1$, alors on peut montrer que l'espace de Hilbert \mathcal{H} sur lequel agissent J_x, J_y et J_z est une somme directe $\mathcal{H} = \mathcal{H}_1 \oplus \dots \oplus \mathcal{H}_d$, et que les observables de moment cinétique agissent séparément sur chaque terme. On dit que \mathcal{H} est une représentation *réductible* de l'algèbre engendrée par les opérateurs J_x, J_y, J_z . Cette situation signale qu'il existe une dégénérescence due à l'existence d'une autre observable commutant avec J_x, J_y, J_z . Or nous avons décidé de nous intéresser seulement à ces observables, autrement dit nous allons considérer ce qui se passe dans un seul des termes de la somme d'espaces. On suppose donc dans la suite que les valeurs propres de J_z n'ont pas de dégénérescence, autrement dit, nous supposons que \mathcal{H} est une représentation *irréductible*. Définissons

$$J^2 = J_x^2 + J_y^2 + J_z^2$$

On montre facilement que $[J^2, J_x] = [J^2, J_y] = [J^2, J_z] = 0$. Ceci, ajouté au fait que les valeurs propres de J_z (resp. J_x et J_y) n'ont pas de dégénérescence, implique que J^2 est un opérateur scalaire, c'est-à-dire de la forme λId .

À l'aide des relations de commutation et d'un peu d'algèbre, on montre de façon très similaire à ce qu'on a fait pour l'oscillateur harmonique le théorème suivant :

Théorème 6.7.1 *Il existe un nombre j de la forme $j = \frac{n}{2}$, avec $n \in \mathbb{N}$, tel que :*

¹⁰. On a vu que c'était vrai sous certaines conditions, notamment si l'espace des états est de dimension finie : ce sera toujours le cas.

1. $J^2 = j(j+1)\hbar^2 \text{Id}$,

2. Les valeurs propres de $\frac{1}{\hbar}J_z$ sont les $2j+1$ nombres : $-j, -j+1, \dots, j-1, j$.

En particulier $\dim(\mathcal{H}) = 2j+1$.

La démonstration de ce théorème se trouve dans tous les ouvrages de mécanique quantique, bien qu'il ne soit pas toujours énoncé de cette manière. En fait, ce théorème a été énoncé et démontré dès 1914 par Élie Cartan, bien avant la mécanique quantique ! La raison en est que l'algèbre des composantes du moment cinétique est aussi ce qu'on appelle en mathématique l'algèbre de Lie du groupe des rotations $SO(3)$. Nous n'aborderons pas cet aspect des choses, mais nous invitons le lecteur à consulter par exemple [7] chapitre VI, complément B.

6.7.2 Moment cinétique orbital. Application à l'atome d'hydrogène.

Dans le cas d'un moment cinétique orbital, j se note généralement ℓ , et est nécessairement entier. En effet, dans ce cas, $J_z = \hat{L}_z = \hat{x}\hat{p}_y - \hat{y}\hat{p}_x$. En passant en coordonnées sphériques, l'action de \hat{L}_z sur $\psi(x, y, z) = \psi(r, \theta, \varphi)$ s'exprime très simplement :

$$\hat{L}_z = \frac{\hbar}{i} \frac{\partial}{\partial \varphi}$$

Donc les fonctions d'onde propres ψ_m de \hat{L}_z vérifient

$$\frac{\hbar}{i} \frac{\partial}{\partial \varphi} \psi = m\hbar \psi$$

et sont donc de la forme

$$\psi_m(r, \theta, \varphi) = \phi_m(r, \theta) e^{im\varphi}$$

Or si M est un point de coordonnées cartésiennes x, y, z et sphériques r, θ, φ , on a évidemment $\psi(r, \theta, \varphi) = \psi(x, y, z) = \psi(r, \theta, \varphi + 2\pi)$. Donc on doit avoir

$$e^{2im\pi} = 1$$

Ce qui implique que $m \in \mathbb{N}$. Or $m \in \{-\ell; -\ell+1; \dots; \ell-1; \ell\}$, donc ℓ est forcément un entier.

Considérons maintenant un électron dans un potentiel central coulombien (cas de l'atome d'hydrogène ou des ions à un seul électron). Son hamiltonien est donné par l'opérateur

$$\hat{H} = -\frac{\hbar^2}{2m} \Delta + V(r), \text{ où } V(r) = -\frac{e^2}{4\pi\epsilon_0 r} \quad (6.39)$$

Or il est évident que \hat{L}_z commute avec la multiplication par $V(r)$. De plus, nous laissons le lecteur vérifier à titre d'exercice que \hat{L}_z commute aussi avec le Laplacien (c'est très simple sur l'expression du Laplacien en coordonnées sphériques). On a donc montré que

$$[\hat{L}_z, \hat{H}] = 0$$

Autrement dit, on peut diagonaliser simultanément ces deux observables. Il en résulte de façon évidente que \hat{H} commute également avec \hat{L}_y et \hat{L}_x , donc avec \hat{L}^2 . On a donc trois observables, \hat{H} , \hat{L}_z et \hat{L}^2 , représentant respectivement l'énergie, la composante du moment cinétique sur un axe donné, et le carré du moment cinétique total, qui commutent entre elles. Ainsi, une base d'états propres pour l'électron ¹¹ pourra s'écrire sous la forme

$$|n, \ell, m\rangle$$

11. Rappelons que pour l'instant, nous n'avons pas encore parlé de son spin...

où n étiquette les valeurs propres E_n du hamiltonien, ℓ de \hat{L}^2 et m de \hat{L}_z . Rappelons pour mémoire que $\hat{H}|n, \ell, m\rangle = E_n|n, \ell, m\rangle$, $\hat{L}^2|n, \ell, m\rangle = \ell(\ell + 1)\hbar^2|n, \ell, m\rangle$ et $\hat{L}_z|n, \ell, m\rangle = m\hbar|n, \ell, m\rangle$. Pour les états liés, la condition de normalisabilité $\int |\psi|^2 < \infty$ impose aux valeurs de l'énergie de former un ensemble discret, qu'on peut indexer par un entier $n \in \mathbb{N}^*$. Plus précisément, on a :

$$E_n = -\frac{E_I}{n^2}, \text{ où } E_I \approx 13,6 \text{ eV}$$

On peut montrer par ailleurs que $0 \leq \ell < n$. Il en résulte que l'espace propre du hamiltonien associé à la valeur propre E_n est de dimension $\sum_{\ell=0}^{n-1} (2\ell + 1) = n^2$. Pour la résolution de l'équation de Schrödinger correspondant au hamiltonien (6.39), voir [4] p. 231, ou [7] p. 801.

Si on plonge l'atome dans un champ magnétique constant dirigé selon l'axe des z , on brise l'invariance par rotation, et la valeur propre $m\hbar$ qui s'interprète classiquement comme la composante du moment cinétique selon cet axe prend son importance. En effet, au hamiltonien classique vient s'ajouter un terme d'interaction entre l'électron et le champ, donné par $-\vec{\mu} \cdot \vec{B}$, où $\mu = \gamma_0 \vec{L}$ est le moment magnétique. On va « quantifier » cette expression en postulant un terme $-\hat{\mu} \cdot \vec{B} = -\gamma_0 \hat{L}_z B$ dans le hamiltonien quantique. L'action de \hat{H} sur le vecteur propre $|n, \ell, m\rangle$ devient donc :

$$\hat{H}|n, \ell, m\rangle = (E_n - \gamma_0 B m \hbar)|n, \ell, m\rangle$$

On voit donc que le champ magnétique clive le niveau d'énergie E_n en $2\ell + 1$ sous-niveaux. La raie spectrale correspondante se scinde donc en $2\ell + 1$ sous-raies : c'est l'effet Zeeman.

D'après ce qu'on a dit sur le moment cinétique orbital, ℓ est un nombre entier : le nombre de sous-raies observées doit donc être impair¹²...

6.7.3 Existence du spin

...sauf que ce n'est pas le cas ! L'effet Zeeman a été observé par ce dernier dès 1896. Il trouva d'abord un nombre impair de sous-raies, ce à quoi on s'attendait déjà à partir de considérations de physique classique. Mais il trouva aussi un peu plus tard des cas où l'on observe un nombre pair de sous-raies. Cet effet Zeeman « anormal » demeura longtemps une énigme. On raconte d'ailleurs que Pauli était tellement préoccupé par ce problème, qu'un collègue lui demanda un jour pourquoi il avait l'air si maussade. « Comment peut-on aller bien quand on réfléchit à l'effet Zeeman anormal ? », répondit Pauli...

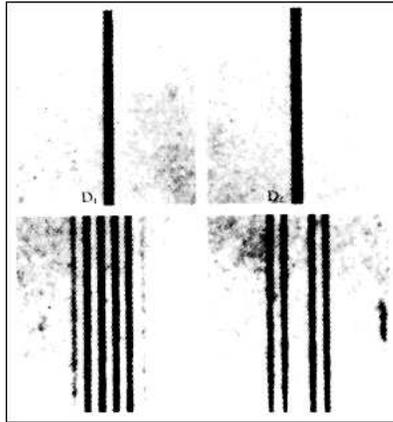


FIGURE 6.6 – Le clivage des raies spectrales sous l'effet d'un champ magnétique. Photo prise par Zeeman lui-même.

¹². L'étude est un peu plus complexe dans le cas où il y a plusieurs électrons, mais le résultat est qualitativement le même.

Nous pouvons cependant rendre compte facilement de cette énigme en supposant qu'il existe pour l'électron une observable de moment cinétique qui n'est pas de nature orbitale. On l'appelle *spin*. Ses valeurs propres, comme nous l'avons vu plus haut, sont nécessairement de la forme $m\hbar$, où m est entier ou demi-entier. Cette hypothèse semble naturelle si on s'en tient à l'analyse algébrique que nous avons faite des observables de type « moment cinétique », mais s'il est aisé d'écrire « un moment cinétique qui n'est pas de nature orbitale », c'est plus difficile de comprendre à quoi cela correspond physiquement. On dit souvent qu'on peut imaginer que l'électron¹³ tourne sur lui-même (c'est d'ailleurs l'origine du mot « spin »), mais si on peut effectivement l'imaginer, cela ne correspond pas à la réalité! D'abord parce que l'électron n'est pas une petite bille (on peut calculer qu'alors un point de la surface de l'électron devrait aller plus de cent fois plus vite que la lumière...), ensuite parce que même si c'était le cas, son moment cinétique serait toujours de nature orbitale. Le spin est le premier exemple d'une observable sans équivalent classique.

Le spin intervient aussi dans l'explication d'une expérience réalisée en 1922 par Otto Stern et Walther Gerlach. Dans cette expérience, un faisceau horizontal de particules neutres¹⁴ pourvues d'un moment magnétique¹⁵ est envoyé dans un champ magnétique inhomogène tel que, classiquement, les atomes subissent une déviation verticale proportionnelle à la projection de leur moment magnétique sur l'axe (Oz). Un écran détecte les atomes à la sortie de l'appareil. Les atomes étant orientés de façon aléatoire, la prédiction de la mécanique classique est qu'ils vont former une tache allongée sur l'écran (idéalement un segment, mais il existe une inévitable dispersion horizontale des vitesses initiales). Le résultat de l'expérience n'est pas du tout celui-là! On voit deux (ou plus, ou une seule, cela dépend de l'atome) taches bien séparées (idéalement deux points).

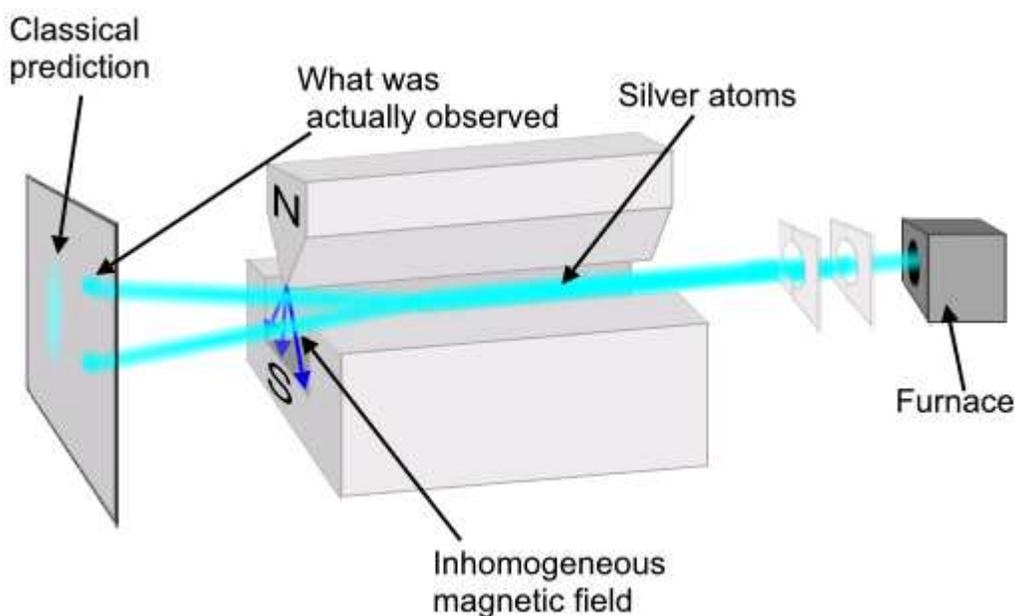


FIGURE 6.7 – L'expérience de Stern et Gerlach.

Cette expérience nous montre que la projection sur l'axe (Oz) du moment magnétique d'un atome est

13. Ou d'autres particules, comme le proton, les quarks, etc.

14. Si les particules sont chargées, elles subissent une force de Lorentz dans le champ magnétique de l'appareil, et cet effet masque celui qu'on veut observer.

15. Historiquement il s'agissait d'atomes d'argent. L'expérience a été refaite un peu plus tard avec des atomes d'hydrogène dans leur état fondamental, avec le même résultat.

quantifié, ce qui n'est pas pour nous surprendre puisqu'elle est proportionnelle à la projection sur le même axe du moment cinétique, dont on sait qu'elle est quantifiée. Mais ce qui est plus intéressant, c'est que dans le cas des atomes d'argent ou d'hydrogène on constate que cette projection ne peut prendre que deux valeurs. Cela montre que le moment cinétique total de l'atome¹⁶ fait intervenir un espace de Hilbert de dimension $2j + 1 = 2$, donc $j = 1/2$ (voir théorème 6.7.1). Ces résultats s'expliquent si on admet que l'électron a un spin $1/2$, ce qui signifie qu'il possède un moment cinétique intrinsèque de valeur $\sqrt{\frac{1}{2}(\frac{1}{2} + 1)\hbar^2} = \frac{\sqrt{3}}{2}\hbar$, dont la projection sur un axe quelconque ne peut valoir que $\pm\frac{\hbar}{2}$. Le moment cinétique d'un atome est ainsi la somme du moment cinétique orbital de ses électrons (qui est nulle dans l'état fondamental) et de leur spin¹⁷ de ses électrons¹⁸.

Nous avons donc deux arguments convaincants, d'origines expérimentales, pour attribuer à l'électron un spin $1/2$. On pourrait objecter que, dans les deux cas, l'expérience ne donne accès qu'à un moment magnétique, et pas directement au moment cinétique intrinsèque de l'électron. L'existence de ce dernier pourrait donc apparaître comme une hypothèse superflue. Néanmoins il existe un argument théorique, développé dans [2], chap. 20, qui montre la nécessité d'introduire le spin : sans celui-ci le moment cinétique total d'un atome isolé ne serait pas conservé. Par ailleurs, la théorie relativiste de l'électron, due à Dirac, introduit aussi le spin comme une nécessité théorique.

6.7.4 Algèbre du spin $1/2$

L'état d'un électron, ou de toute autre particule de spin $1/2$, n'est donc pas complètement déterminé par les observables de position (si l'on fait le choix de la représentation r) ou d'impulsion (en représentation p). Il existe une autre observable, dite de spin, agissant sur un espace « interne » de dimension 2. Nous allons nous intéresser pour l'instant uniquement à cet espace des états de spin, et nous dirons un mot plus loin sur la manière dont on peut représenter le vecteur d'état complet d'une particule de spin $1/2$. En reprenant les notations du paragraphe 6.7.1, on a donc trois observables S_x, S_y et S_z , dont les valeurs propres sont $\pm\frac{\hbar}{2}$. Dans une base de vecteurs propres de S_z , les opérateurs S_x, S_y et S_z auront donc des matrices de la forme $\frac{\hbar}{2}\sigma_x, \frac{\hbar}{2}\sigma_y$ et $\frac{\hbar}{2}\sigma_z$, avec

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Les relations de commutation entre les observables de spin donnent :

$$[\sigma_x, \sigma_y] = 2i\sigma_z \tag{6.40}$$

$$[\sigma_y, \sigma_z] = 2i\sigma_x \tag{6.41}$$

$$[\sigma_z, \sigma_x] = 2i\sigma_y \tag{6.42}$$

De plus, on a $S^2 = \frac{3}{4}\hbar^2\text{Id}$, donc

16. La composition des moments cinétiques est une affaire assez délicate. L'espace de Hilbert du moment cinétique total d'un atome n'est pas irréductible : c'est une somme directe de représentations irréductibles de l'algèbre engendrée par les opérateurs de moment cinétique, correspondant à différentes valeurs de j . L'expérience sur les atomes d'argent montre qu'au moins une de ces valeurs de j vaut $1/2$. La relation avec le spin de l'électron n'est donc pas directe. Cependant, l'état fondamental d'un atome d'hydrogène est à symétrie sphérique, sa contribution orbitale au moment magnétique est donc nulle. Le moment magnétique de l'atome d'hydrogène est alors uniquement relié au spin de l'électron, et le résultat est identique à celui qu'on obtiendrait avec des électrons, en imaginant qu'on puisse annuler la charge électrique de ceux-ci !

17. Le mot spin sert à la fois à désigner le coefficient j qui vaut $1/2$ pour l'électron, l'observable de moment cinétique intrinsèque, et une mesure de cette observable !

18. Il faut également admettre la proportionnalité entre spin et moment magnétique. Le coefficient de proportionnalité s'appelle « rapport gyromagnétique ». Le rôle des noyaux peut être négligé, à cause de la petitesse de leurs rapports gyromagnétiques devant celui de l'électron.

$$\sigma_x^2 + \sigma_y^2 + \sigma_z^2 = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \quad (6.43)$$

Enfin, σ_x, σ_y et σ_z sont des matrices hermitiennes.

Or on sait que $\text{Tr}(AB) = \text{Tr}(BA)$. Il en résulte que la trace d'un commutateur est toujours nulle. D'après (6.41) et (6.42) on en déduit immédiatement que $\text{Tr}(\sigma_x) = \text{Tr}(\sigma_y) = 0$. Or σ_x étant hermitienne, elle est de la forme $\begin{pmatrix} a & \xi \\ \bar{\xi} & b \end{pmatrix}$, avec $a, b \in \mathbb{R}$ et $\xi \in \mathbb{C}$. Mais comme $\text{Tr}(\sigma_x) = 0$, on a $b = -a$. De plus, en multipliant la relation (6.41) à droite par σ_z , et en prenant la trace de chaque côté, on obtient

$$\text{Tr}(\sigma_y \sigma_z^2) - \text{Tr}(\sigma_z \sigma_y \sigma_z) = 2i \text{Tr}(\sigma_x \sigma_z)$$

Or

$$\text{Tr}(\sigma_z \sigma_y \sigma_z) = \text{Tr}(\sigma_z (\sigma_y \sigma_z)) = \text{Tr}((\sigma_y \sigma_z) \sigma_z)$$

D'où $\text{Tr}(\sigma_x \sigma_z) = 0$. Mais comme $\sigma_x \sigma_z = \begin{pmatrix} a & -\bar{\xi} \\ \bar{\xi} & a \end{pmatrix}$, on en déduit que $a = 0$. D'où

$$\sigma_x = \begin{pmatrix} 0 & \xi \\ \bar{\xi} & 0 \end{pmatrix}$$

Par une méthode similaire on trouve que

$$\sigma_y = \begin{pmatrix} 0 & \eta \\ \bar{\eta} & 0 \end{pmatrix}$$

On en déduit par un calcul matriciel direct que

$$\sigma_x \sigma_z = -\sigma_z \sigma_x$$

et

$$\sigma_y \sigma_z = -\sigma_z \sigma_y$$

En utilisant à nouveau (6.41), on trouve

$$\sigma_y \sigma_z = i \sigma_x \Leftrightarrow \begin{pmatrix} 0 & -\eta \\ \bar{\eta} & 0 \end{pmatrix} = i \begin{pmatrix} 0 & \xi \\ \bar{\xi} & 0 \end{pmatrix}$$

d'où $\eta = -i\xi$. La relation (6.43) nous donne finalement $|\xi| = 1$. En posant $\xi = e^{i\theta}$, on a donc

$$\sigma_x = \begin{pmatrix} 0 & e^{i\theta} \\ e^{-i\theta} & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -ie^{i\theta} \\ ie^{-i\theta} & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (6.44)$$

Réciproquement, les relations (6.40), (6.41), (6.42) et (6.43) sont vérifiées par les matrices hermitiennes définies par (6.44). Ce sont donc les matrices les plus générales exprimant les observables des composantes du spin (multipliées par $2/\hbar$) dans une base (e_1, e_2) de vecteurs propres associés respectivement aux valeurs propres $\hbar/2$ et $-\hbar/2$ de S_z . Le reste est affaire de conventions. En passant de la base (e_1, e_2) à la base (u, d) définie par $u = e^{i\theta/2}e_1$ et $d = e^{-i\theta/2}e_2$, on obtient *les matrices de spin de Pauli* :

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (6.45)$$

On peut vérifier directement que des vecteurs propres normalisés associés aux valeurs propres ± 1 des matrices σ_x et σ_z sont donnés, respectivement par

$$|\pm, x\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}, \quad \text{et} \quad |\pm, z\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i \end{pmatrix}$$

Avec des notations similaires, on a bien sûr

$$|+, z\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ et } |-, z\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Exercice 6.7.2 Vérifier que si un électron est dans l'état $|+, z\rangle$, une mesure de son spin selon l'axe (Ox) donne $\hbar/2$ avec la probabilité $1/2$ et $-\hbar/2$ avec la probabilité $1/2$.

Le résultat de cet exercice peut être vérifié expérimentalement à l'aide de deux appareils de Stern-Gerlach disposés l'un derrière l'autre.

Passons maintenant au problème de la représentation du vecteur d'état complet de l'électron. Si on ignore le spin, le vecteur d'état peut être représenté par une fonction d'onde de carré sommable¹⁹, il s'agit donc d'un vecteur de l'espace de Hilbert $\mathcal{H}_{\text{pos}} = L^2(\mathbb{R}^3, \mathbb{C})$. D'autre part, si on ne s'occupe que du spin, on a l'espace des états $\mathcal{H}_{\text{spin}}$ qui est un espace vectoriel de dimension 2 sur \mathbb{C} . Mathématiquement, l'espace des états complet de l'électron est le produit tensoriel de ces deux espaces. Heureusement, dans ce cas, il est inutile de connaître la théorie des produits tensoriel : celui-ci s'identifie à l'espace des fonctions de carré sommable à valeur dans $\mathcal{H}_{\text{spin}}$. Ainsi, l'espace des états de l'électron est :

$$\mathcal{H} = L^2(\mathbb{R}^3, \mathcal{H}_{\text{spin}})$$

En choisissant une base de $\mathcal{H}_{\text{spin}}$, par exemple la base $(|+, z\rangle, |-, z\rangle)$, un élément Ψ de cet espace peut-être représenté par un vecteur de fonctions :

$$\Psi(x, y, z) = \begin{pmatrix} \Psi_1(x, y, z) \\ \Psi_2(x, y, z) \end{pmatrix}$$

Si Ψ et Φ sont deux tels vecteurs, leur produit scalaire est défini par

$$\langle \Psi | \Phi \rangle = \int_{\mathbb{R}^3} \Psi_1 \bar{\Phi}_1 + \int_{\mathbb{R}^3} \Psi_2 \bar{\Phi}_2$$

De sorte que $\|\Psi\|^2 = \|\Psi_1\|^2 + \|\Psi_2\|^2$.

Si une base de \mathcal{H}_{pos} est donnée par $|n, \ell, m\rangle$, comme c'est le cas pour un électron unique dans un potentiel central coulombien, une base de \mathcal{H} sera donnée par $|n, \ell, m, \sigma\rangle$, où $\sigma = \pm$, avec

$$|n, \ell, m, +\rangle = \begin{pmatrix} |n, \ell, m\rangle \\ 0 \end{pmatrix}, \quad |n, \ell, m, -\rangle = \begin{pmatrix} 0 \\ |n, \ell, m\rangle \end{pmatrix}$$

Dans ce cas on a $\dim(\mathcal{H}) = 2 \dim(\mathcal{H}_{\text{pos}}) = 2n^2$. Nous verrons plus loin comment utiliser ce résultat pour donner une explication de la structure du tableau périodique des éléments.

6.7.5 Spin et statistique

En physique classique, il est impossible que deux objets soient absolument identiques par tous leurs aspects. Tout d'abord, lorsqu'on dit une phrase comme « soient deux boules de billard identiques », il s'agit d'une façon approximative de présenter les choses. En regardant ces boules avec un microscope, on verra qu'elles se distinguent par certaines rayures invisibles à l'oeil nu. Cependant, avec des particules élémentaires, comme des électrons, on a vraiment affaire à des objets qui ne se distinguent ni par leur masse, ni par leur charge, ni par leur structure interne. . . puisqu'ils n'en ont pas. Si on peut mesurer une quantité qui diffère pour ces deux électrons, alors on pourra les distinguer. En physique classique, il est impossible que deux objets se trouvent en même temps au même endroit : cela signifie qu'il y a toujours une quantité qui pourra les distinguer, à savoir leur position. Cela nous permet d'en nommer un A et

19. On a vu que c'était un peu plus compliqué que ça, d'une part à cause des états de diffusion. D'autre part, on doit pouvoir dériver une fonction d'onde, ce qui n'est pas toujours possible dans L^2 . On va ignorer pour l'instant ces problèmes mathématiques, qui ne modifient pas notablement la discussion qui suit.

l'autre B , et comme on peut suivre leur trajectoire avec exactitude, et que cette trajectoire ne peut à aucun moment les amener à coïncider, on sait à tout instant lequel est A et lequel est B .

Il en va tout autrement en physique quantique. Même si les postulats que nous avons énoncés pour l'instant ne disent rien sur la manière dont il faut représenter l'état d'un système constitué de deux particules identiques, rien ne nous autorise à penser que la probabilité pour que les deux particules soient détectées au même instant au même endroit est forcément nulle. Il semble bien que l'on puisse envisager par exemple deux particules dans un état complètement délocalisé : l'idée de pouvoir les distinguer par leur position se révèle alors dépourvue de sens²⁰. De plus, même si l'on admet que l'on a pu, à un certain instant t_0 , leur affecter deux positions distinctes, la disparition de la notion de trajectoire bien définie nous oblige à considérer le fait que les particules puissent échanger subrepticement leurs positions à un instant $t > t_0$!

La position ne jouant plus le rôle privilégié qu'elle jouait dans cette question en physique classique, la bonne question à se poser dans le cadre quantique, c'est : deux particules peuvent-elles être exactement dans le même état au même moment ?

Nous allons voir que la réponse dépend, de façon très étonnante, du spin de la particule ! Mais avant toute chose, il nous faut dire un mot de la façon dont on représente les états d'un système composé de deux particules identiques. Disons que ce sont deux « quantons ». Soit \mathcal{H} l'espace des états d'un de ces quantons. Les états d'un système de deux quantons appartiennent au produit tensoriel $\mathcal{H} \otimes \mathcal{H}$. Il est inutile de connaître les détails de la théorie du produit tensoriel pour comprendre ce qui est à l'œuvre ici, tout ce que nous avons besoin de savoir, c'est que si $(|k\rangle)_{k \in K}$ est une base de \mathcal{H} , alors une base de $\mathcal{H} \otimes \mathcal{H}$ est $(|k\rangle \otimes |l\rangle)_{k,l \in K}$. Tout vecteur de $\mathcal{H} \otimes \mathcal{H}$ s'appelle un tenseur (d'ordre 2) et peut se développer de façon unique sous la forme

$$|\psi\rangle = \sum_{k,l} \lambda_{k,l} |k\rangle \otimes |l\rangle$$

Introduisons l'opérateur d'échange P , qui au vecteur de base $|k\rangle \otimes |l\rangle$ associe $|l\rangle \otimes |k\rangle$. Cet opérateur échange, comme son nom l'indique, le rôle joué par les deux quantons. Or les deux quantons étant indiscernables, si le vecteur $|\psi\rangle$ représente un état du système, le vecteur $P|\psi\rangle$ doit représenter le même état, autrement dit, il doit exister un complexe non nul λ , dépendant éventuellement de l'état $|\psi\rangle$, tel que $P|\psi\rangle = \lambda|\psi\rangle$. Or il est clair que $P^2 = \text{Id}$, d'où il découle immédiatement que $\lambda = \pm 1$.

Ainsi, un vecteur de \mathcal{H} ne représente un état du système que s'il vérifie $P|\psi\rangle = |\psi\rangle$, auquel cas on dit qu'il est symétrique, ou $P|\psi\rangle = -|\psi\rangle$, on dit alors qu'il est antisymétrique. Un exemple de vecteur d'état symétrique est $|k\rangle \otimes |k\rangle$, ou encore $|k\rangle \otimes |l\rangle + |l\rangle \otimes |k\rangle$. Le vecteur d'état $|k\rangle \otimes |l\rangle - |l\rangle \otimes |k\rangle$ est quant à lui antisymétrique.

Il est de plus facile de voir qu'on doit toujours avoir le même cas pour un même système de deux mêmes quantons. En effet, si ce système pouvait être dans un état symétrique, représenté par $|\psi\rangle$, et dans un état antisymétrique représenté par $|\phi\rangle$, il pourrait être dans l'état $|\psi\rangle + |\phi\rangle$, qui n'est ni l'un ni l'autre. Il se trouve que (et c'est facilement vérifiable), les tenseurs symétriques forment un sous-espace vectoriel de $\mathcal{H} \otimes \mathcal{H}$, qu'on note parfois $\mathcal{S}^2(\mathcal{H})$, tandis que les tenseurs antisymétriques forment un sous-espace vectoriel $\mathcal{A}^2(\mathcal{H})$.

Ce qu'on a fait jusqu'à présent montre qu'un système de deux quantons identiques a pour espace des états $\mathcal{S}^2(\mathcal{H})$ ou $\mathcal{A}^2(\mathcal{H})$. On ne peut pas aller plus loin à l'aide des principes de la mécanique quantique déjà énoncés.

Exercice 6.7.3 Dans le cas où \mathcal{H} est de dimension finie n , montrer que $\mathcal{S}^2(\mathcal{H})$ et $\mathcal{A}^2(\mathcal{H})$ sont deux s.e.v supplémentaires de dimensions respectives $n(n+1)/2$ et $n(n-1)/2$.

Nous allons maintenant énoncer le principe de Pauli, qui précise la situation pour chaque type de particules, en fonction de leur spin.

20. En mécanique classique, il n'est jamais venu à l'idée de personne de contester que deux particules puissent avoir exactement la même impulsion. Or la position et l'impulsion jouant des rôles duaux en mécanique quantique, il est bien naturel d'autoriser a priori que deux particules occupent la même position.

Principe de Pauli Il existe deux types de particules : les bosons et les fermions. Le vecteur d'état d'un système de deux bosons identiques est toujours symétrique. Le vecteur d'état d'un système de deux fermions identiques est toujours antisymétrique. De plus, les particules de spin entier sont des bosons, et les particules de spin demi-entier (i.e. du type $n/2$ avec n impair) sont des fermions.

Parmi les particules élémentaires, les électrons et les quarks sont des fermions, tandis que le photon est un exemple de boson. Quand un système est composé de plusieurs particules élémentaires, on peut faire abstraction de ses constituants tant que les énergies en jeu sont suffisamment faible pour considérer que le système garde toujours son intégrité. Comme les spins s'ajoutent, on voit que, dans les conditions précédemment énoncées, les neutrons et les protons sont des fermions, et les noyaux d'hélium sont des bosons.

Le principe de Pauli traite en fait d'un nombre quelconque de particules identiques, mais nous nous sommes limités à deux particules pour plus de simplicité.

En fait ce principe n'en est pas vraiment un, car il peut se démontrer dans le cadre de la théorie quantique des champs, qui incorpore mécanique quantique et relativité restreinte. C'est d'ailleurs un des grands succès de cette théorie que d'avoir pu démontrer ce principe, qui prend dans ce cadre le nom de « théorème spin-statistique ».

Nous sommes maintenant en mesure de répondre à la question initiale : deux particules peuvent-elles se trouver dans le même état ? En réalité, ce n'est pas une question très bien posée : pour pouvoir déterminer l'état de l'une des deux particules, il faut pouvoir opérer des mesures sur une particule indépendamment de l'autre, ce qui n'est possible que si les particules sont indépendantes, c'est-à-dire n'interagissent pas. Sous cette restriction, les états du système qui représentent une particule dans l'état $|\psi_1\rangle$ et l'autre dans l'état $|\psi_2\rangle$ sont représentés par des vecteurs de la forme

$$|\psi_1\rangle \otimes |\psi_2\rangle \pm |\psi_2\rangle \otimes |\psi_1\rangle$$

où le signe dépend du type de particule considéré. On voit que pour deux fermions, on obtient le vecteur nul si $|\psi_1\rangle = |\psi_2\rangle$, or le vecteur nul ne représente pas un état. Autrement dit, deux fermions indépendants ne peuvent être dans le même état. Ce phénomène s'appelle « principe d'exclusion de Pauli », même si, comme nous l'avons vu, ce n'est qu'une conséquence du principe de Pauli. Il faut noter que, dans la réalité, deux fermions interagissent presque toujours, aussi ce principe ne s'applique-t-il que lorsque l'on peut négliger cette interaction.

Généralisons maintenant notre propos à un système de n particules identiques. Le principe de Pauli, sous sa forme générale, indique que le vecteur d'état d'un système de n bosons appartient à l'espace $\mathcal{S}^n(\mathcal{H})$ des tenseurs symétriques, et que le vecteur d'état d'un système de n fermions appartient à l'espace $\mathcal{A}^n(\mathcal{H})$ des tenseurs antisymétriques.

Un tenseur symétrique est un tenseur invariant lorsqu'on applique l'opérateur P_{ij} qui échange les i -èmes et j -èmes facteurs du produit tensoriel, et qui généralise l'opérateur d'échange introduit plus haut. Un tenseur antisymétrique est un tenseur qui change de signe quand on lui applique P_{ij} .

Si \mathcal{H} est de dimension finie k , et si $|1\rangle, \dots, |k\rangle$ est une base de cet espace, alors un tenseur T d'ordre n est une combinaison linéaire de tenseurs du type

$$|l_1\rangle \otimes \dots \otimes |l_i\rangle \otimes \dots \otimes |l_j\rangle \otimes \dots \otimes |l_n\rangle \tag{6.46}$$

où $l_i \in \{1, \dots, k\}$. Pour que T soit antisymétrique, il est nécessaire que le coefficient affecté au « monôme » (6.46) soit l'opposé de celui que est affecté au monôme obtenu en échangeant les places de l_i et l_j . Donc si $l_i = l_j$, ce coefficient doit être nul, et le monôme en question ne doit pas intervenir dans la décomposition de T . On voit donc que si $n > k$, on a forcément $T = 0$. Autrement dit, nous venons de montrer que la dimension de $\mathcal{A}^n(\mathcal{H})$ est égale à 0 dès que $n > \dim(\mathcal{H})$. Ce fait mathématique très simple a une conséquence physique remarquable : il ne peut y avoir plus de n fermions dans un système physique, où n est la dimension de l'espace des états du système avec un seul fermion.

Exercice 6.7.4 Montrer que si $n = \dim(\mathcal{H})$, alors $\dim(\mathcal{A}^n(\mathcal{H})) = 1$. (S'inspirer de la démonstration de l'unicité du déterminant.)

Les fermions sont donc contraints de « s'empiler » les uns par dessus les autres, car une fois que le ou les (s'il y a dégénérescence) états de plus basse énergie sont occupés par des fermions, les autres doivent occuper des états d'énergie supérieure. Ceci nous donne une explication qualitative de la structure du tableau périodique des éléments!

En effet, même si le hamiltonien \hat{H} d'un atome à plusieurs électrons est trop complexe pour que l'équation de Schrödinger associée puisse être résolue, on peut en première approximation, le remplacer par un hamiltonien \hat{H}_0 , de la forme

$$\hat{H}_0 = \sum_{i=1}^Z \frac{\hat{p}_i^2}{2m_e} + U(\hat{r}_i)$$

où Z est le nombre d'électrons, et U est un potentiel moyen qui rend compte de façon approchée de la répulsion coulombienne entre les électrons : c'est la méthode de Hartree-Fock ([2] p. 1044). Dans cette approximation, les électrons se comportent comme des fermions indépendants, on peut donc leur appliquer le « principe » d'exclusion de Pauli, énoncé plus haut. Or on a vu précédemment qu'une base de l'espace de Hilbert \mathcal{H} d'un électron est donnée par $|n, \ell, m, \sigma\rangle$. Ceci permet d'expliquer la méthode de remplissage en couches pouvant contenir $\dim(\mathcal{H}) = 2n^2$ électrons, bien connu en chimie. Répétons que cette méthode provient d'une approximation assez grossière. Elle ne fonctionne d'ailleurs correctement que jusqu'à $n = 4$.

Venons-en maintenant aux bosons. Rien ne leur interdit d'occuper le même état, et quand celui-ci est l'état d'énergie la plus basse, c'est ce qu'ils ont tendance à faire : les bosons « s'aggrègent » tous dans l'état de plus basse énergie. Ce fait est à l'origine de la condensation de Bose-Einstein, qui est à l'œuvre dans plusieurs phénomènes spectaculaires comme la supraconductivité ou la superfluidité.

Chapitre 7

Paradoxes et Interprétations

7.1 Chat de Schrödinger

Un des aspects les plus étranges de la mécanique quantique est sans doute la coexistence de deux types d'évolution pour le vecteur d'état : l'évolution **U** et l'évolution **R**, ou réduction du paquet d'onde, dont nous avons déjà parlé.

Schrödinger n'a jamais été convaincu par la réduction du paquet d'onde, où réside le caractère essentiellement aléatoire de la mécanique quantique. Il est resté toute sa vie opposé à l'interprétation dite « de Copenhague » de la mécanique quantique. C'est précisément pour combattre cette interprétation qu'il a imaginé la célèbre expérience de pensée dont nous allons parler maintenant.

On imagine que l'on enferme un chat dans une boîte où se trouve une substance radioactive, et un détecteur de particules relié à une capsule de poison mortel. Lorsque la substance émet une particule, le détecteur la reçoit et répand le poison, qui tue le pauvre animal. Imaginons pour fixer les idées, que la probabilité d'émission d'une particule soit de 1/2 au bout d'une heure. Cela signifie qu'au bout d'une heure, le système qui est dans la boîte peut être décrit par le vecteur d'état suivant :

$$|\psi(1)\rangle = \frac{1}{\sqrt{2}}(|\text{particule émise}\rangle + |\text{particule non émise}\rangle)$$

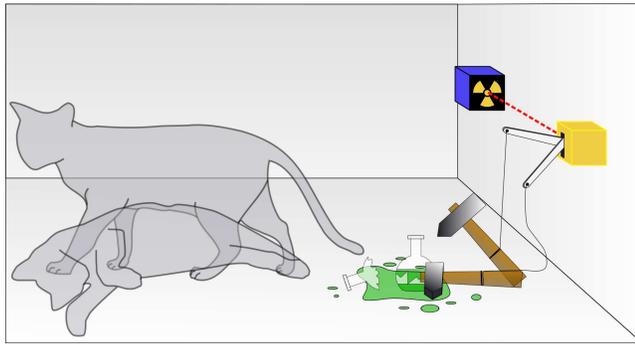


FIGURE 7.1 – Un chat superposé.

Ceci n'est peut-être pas particulièrement troublant, mais cela le devient lorsqu'on décrit ce vecteur, sous la forme équivalente :

$$|\psi(1)\rangle = \frac{1}{\sqrt{2}}(|\text{chat mort}\rangle + |\text{chat vivant}\rangle) \tag{7.1}$$

Le problème est que l'on se trouve avec un état superposé pour un objet macroscopique, et on peut se poser alors beaucoup de questions : pourquoi ne « voit-on » pas de telles superpositions dans la vie de tous les jours ? Que ressent le chat quand il est dans un état superposé mi-mort mi-vivant ?

Selon l'interprétation de Copenhague, l'évolution du système est parfaitement déterministe de l'état $|\psi(0)\rangle = |\text{chat vivant}\rangle$ à l'état $|\psi(1)\rangle$, c'est l'évolution U , et le vecteur $|\psi(1)\rangle$ décrit bien l'état du système au temps $t = 1$ h, et pas seulement la connaissance que nous en avons. Ce n'est que si l'on ouvre la boîte au bout d'une heure que l'on projète le chat soit dans l'état $|\text{mort}\rangle$, soit dans l'état $|\text{vivant}\rangle$, parce qu'à ce moment-là, on effectue une mesure de l'état du système, et il y a réduction du paquet d'onde.

Pour Schrödinger, cette conclusion était clairement absurde : elle signalait un défaut de conception dans l'interprétation de Copenhague. Il faut reconnaître que la « réduction du paquet d'onde » a fait dire beaucoup de sottises, y compris à de grands esprits comme Eugene Wigner, qui pensait que la conscience de l'observateur jouait un rôle dans la réduction du paquet d'onde. À partir de là, on pourrait se demander si la conscience d'un singe ou d'un chien peut suffire à réduire le paquet d'onde. Et que penser si l'ouverture de la boîte est commandée par un système automatique qui note le résultat de l'expérience sans aucune intervention humaine, comme c'est le cas dans la plupart des expériences de physique moderne ?

Mais il y a encore plus amusant : prenons la boîte contenant le chat et l'observateur dans une plus grande boîte, le tout sous le contrôle d'un second observateur. Imaginons en outre que le premier observateur, celui qui est dans la grande boîte, soit un ami des animaux, et qu'à la découverte du cadavre d'un chat il éclate en sanglot, tandis que la vision du chat bien vivant le fasse sauter de joie. En appliquant le même raisonnement à ce super-système que celui que l'on applique au système chat+détecteur, on voit que l'état de l'observateur numéro 1 est, au bout d'une heure, une superposition de l'état joyeux et de l'état triste. Cette superposition disparaît instantanément pour se fixer sur l'une des deux possibilités dès que le deuxième observateur ouvre la grande boîte. Mais supposons que ce dernier décide de ne jamais l'ouvrir. On constate alors que l'évolution du système qui est dans la grande boîte est toujours de type **U**. Or à l'intérieur de la grande boîte, on a au bout d'une heure la réduction du paquet d'onde du chat, ce qui pose une question troublante : comment peut-il y avoir une évolution de type **R** dans un sous-système alors que le vecteur d'état du système complet évolue de façon unitaire, c'est-à-dire déterministe ? Et si on prend comme super-système l'univers entier, on en arrive à se demander comment l'évolution **R** peut simplement exister.

Il y a de nombreuses attitudes possibles face à cette question. La première est de ne pas chercher à comprendre comment c'est possible, mais de se contenter de faire des calculs et de prédire les résultats des expériences (*Shut up and calculate !*). Cette attitude dite *positiviste* est cohérente dans la mesure où ces résultats ont toujours été ceux attendus. Une autre attitude, qui a été celle de Schrödinger et d'Einstein, est de considérer que la mécanique quantique possède une faille, et d'attendre qu'une meilleure théorie, c'est-à-dire une théorie ne possédant pas ce qui à leurs yeux constitue cette faille, vienne la combler. Cette attente s'est focalisée sur ce qu'on appelle les théories de variables cachées, dont nous aurons à reparler. Pendant un certain temps, les deux camps sont restés sur ces positions de principe. Puis, à partir des années 50, sont apparues de nouvelles interprétations de la mécanique quantique, dont la plus célèbre est l'interprétation multiverselle d'Everett. Ces interprétations sont différentes visions de la même théorie physico-mathématique, ce qui signifie en clair qu'elles font exactement les mêmes prédictions, et ne sauraient donc être départagées sur le plan expérimental. En adopter une plutôt qu'une autre est donc (du moins pour le moment) une question philosophique, et non scientifique.

L'interprétation multiverselle, pour fantaisiste qu'elle puisse paraître, est la plus répandue dans le cercle des physiciens théoriciens s'occupant des questions de gravitation quantique et de cosmologie, suivie par l'interprétation de Copenhague, cette dernière étant très majoritairement admise chez les physiciens qui s'occupent de développer des applications de la mécanique quantique.

Disons quelques mots de l'interprétation d'Everett. Pour bien comprendre ce qu'elle dit, et ce qu'elle ne dit pas, distinguons trois questions qu'on peut légitimement se poser à propos de la présentation standard de la mécanique quantique, et qui apparaissent de façon entremêlée dans le paradoxe du chat ¹ :

1. Si bien que cela entraîne parfois certaines confusions, y compris pour l'auteur de ces lignes dans une version précédente de ce texte. . .

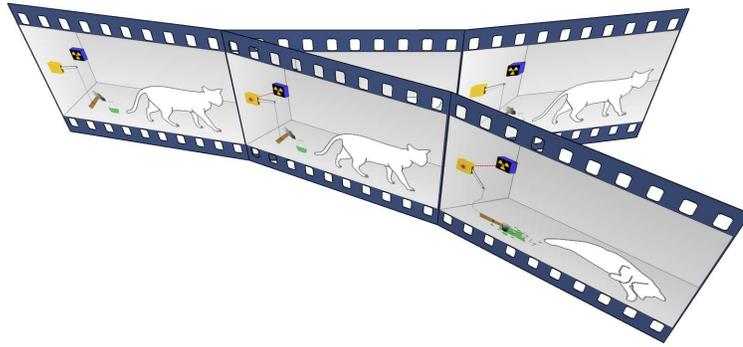


FIGURE 7.2 – L’interprétation multiverselle.

1. Pourquoi y a-t-il une évolution R ?
2. Pourquoi ne voit-on jamais de superpositions ?
3. Comment apparaissent les probabilités ? (Dieu joue-t-il aux dés, si oui comment ?)

Dans l’interprétation de Copenhague, on ne peut pas se poser ces questions : c’est comme ça, ce sont des postulats et ce sont ces postulats qui distinguent cette interprétation des autres. Néanmoins il faut bien admettre que le processus de mesure y joue un rôle un peu mystérieux, en tout cas très particulier.

L’interprétation d’Everett répond de façon radicale à la première question : il n’y a pas d’évolution R ! Le vecteur d’état n’y est plus perçu comme représentant l’état d’un système, mais comme représentant la distribution des états de copies de ce système se trouvant dans des univers différents, tous ces univers ayant leur existence propre. Par exemple, le vecteur (7.1) nous dit² que dans la moitié des univers, le chat est mort au bout d’une heure, et dans l’autre moitié, il est vivant au bout d’une heure. Lorsque l’observateur ouvre la boîte, ou plutôt lorsque les copies (en nombre infini) de l’observateur, se trouvant dans les différentes branches du multivers, ouvrent les copies des boîtes, chacune d’entre elles est informée qu’elle se trouve dans une branche où le chat est mort ou dans une branche où il est vivant.

Les probabilités quantiques n’ont donc pas d’existence objective dans cette interprétation : elles ne sont que subjectives. Objectivement il n’y a que des statistiques. Ceci répond en partie à la question 3 : Dieu ne joue pas aux dés.

Enfin l’interprétation d’Everett ne peut pas à elle seule répondre à la question 2. À cette fin il faut invoquer la théorie de la décohérence. Cette théorie montre que les superpositions d’états quantiques purs sont très fragiles. Un couplage, même très faible, avec l’environnement tend à les remplacer par un mélange statistique d’états classiques. Même si l’état de l’univers tout entier (système étudié + environnement), reste un état quantique pur qui évolue de façon unitaire, lorsque l’on se restreint au système étudié on obtient une apparence d’évolution R , avec un état du système qui saute vers un état classique. Mais vers lequel ? Même si l’on modifie quelque peu l’interprétation de Copenhague pour incorporer la décohérence, la grande différence entre celle-ci et l’interprétation d’Everett se situe précisément dans la réponse à cette question : vers l’un pris au hasard avec une certaine probabilité, pour l’interprétation de Copenhague, vers tous en même temps avec un certain poids statistique, pour l’interprétation d’Everett.

Terminons en signalant que Serge Haroche a obtenu le prix Nobel de physique 2012 pour avoir réussi à mesurer la durée de certains phénomènes de décohérence.

2. En gros. On oublie ici la question des phases.

7.2 Intrication. Paradoxe EPR.

On a vu que parmi les voix discordantes les plus puissantes dans le concert de la mécanique quantique se trouvaient celles de deux de ses pères fondateurs, à savoir Schrödinger et Einstein. Ce dernier n'a eu de cesse d'imaginer des expériences de pensée pour tenter de « coincer » la mécanique quantique, en montrant soit qu'elle était incohérente, soit qu'elle était incomplète³. Subir les assauts de celui qui a été sans aucun doute le plus grand physicien de tous les temps a été une belle épreuve du feu pour la cohérence interne de la mécanique quantique. Elle s'en est toujours bien tiré, et, en outre, les réponses données aux objections du père de la relativité ont permis d'avancer sur le plan théorique, d'éclairer tel ou tel aspect qui avait été négligé.

C'est bien le cas du phénomène *d'intrication quantique* dont nous allons parler maintenant. Ce phénomène a été mis en évidence sur le plan théorique par Einstein, Podolsky et Rosen dans un célèbre article de 1935. Nous allons décrire ici une expérience qui n'est pas exactement celle de l'article original, mais qui permet d'arriver plus simplement aux mêmes conclusions.

On dispose d'un système de deux particules A et B , par exemple un positon et un électron, dont on s'intéresse seulement à l'état de spin. Ces deux particules étant de spin $1/2$, l'état de spin de chacune d'entre elle est représenté par un vecteur de $\mathcal{H} \simeq \mathbb{C}^2$, comme nous l'avons vu plus haut. Un vecteur d'état du système des deux particules sera donc un élément non nul du produit tensoriel $\mathcal{H} \otimes \mathcal{H}$. On peut s'arranger pour que cet état soit l'état dit « singulet » suivant :

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|+, z\rangle \otimes |-, z\rangle - |-, z\rangle \otimes |+, z\rangle) \quad (7.2)$$

On suppose que les deux particules se trouvent initialement au voisinage du point O , et que la particule A se dirige vers Alice tandis que l'autre va vers Bob, qui sont tous les deux munis d'un appareil de Stern-Gerlach leur permettant de mesurer la projection du spin sur un axe quelconque.

L'état $|\Psi\rangle$ possède les propriétés suivantes :

1. Si on mesure la projection du spin de la particule A sur un axe quelconque, on trouve $\hbar/2$ avec la probabilité $1/2$ et $-\hbar/2$ avec la probabilité $1/2$, idem pour la particule B .
2. Les mesures des spins des deux particules ne sont en général pas indépendantes. Par exemple, si on mesure le spin des deux particules selon le même axe (Oz), alors les résultats pour A et B sont opposés. En revanche, si on mesure le spin de A selon (Oz) et le spin de B selon un axe orthogonal (Ox), alors il n'y a aucune corrélation entre les résultats.

On peut démontrer ces propriétés simplement, bien qu'un peu de familiarité avec le produit tensoriel soit nécessaire. C'est l'objet de l'exercice suivant.

Exercice 7.2.1 1. Montrer que si l'on considère une base $(|u\rangle, |v\rangle)$ quelconque de l'espace \mathcal{H} , l'état $|\Psi\rangle$ se réécrit :

$$|\Psi\rangle = \frac{\det P}{\sqrt{2}}(|u\rangle \otimes |v\rangle - |v\rangle \otimes |u\rangle)$$

où P est la matrice de passage de la base $(|u\rangle, |v\rangle)$ vers la base $(|+, z\rangle, |-, z\rangle)$. On voit en particulier qu'un changement de base par une matrice de $SU(2)$ n'affecte pas les composantes de $|\Psi\rangle$. Elles sont donc les mêmes dans la base $(|+, x\rangle, |-, x\rangle)$ par exemple.

2. L'observable « composante du spin selon (Oz) de la particule A » se note $S_z \otimes \text{Id}$. Cette observable agit sur un tenseur décomposé de la façon suivante :

$$S_z \otimes \text{Id}(|u\rangle \otimes |v\rangle) = |S_z u\rangle \otimes |v\rangle$$

Montrer que les vecteurs propres de $S_z \otimes \text{Id}$ sont les vecteurs de la forme $|+, z\rangle \otimes |v\rangle$ (associés à la valeur propre $+\hbar/2$) et les vecteurs de la forme $|-, z\rangle \otimes |v\rangle$ (associés à la valeur propre $-\hbar/2$), où $|v\rangle \neq 0$.

³. Historiquement, Schrödinger a imaginé son expérience du chat un peu après la publication du paradoxe EPR. Elle lui a été inspiré par une remarque d'Einstein à propos des superpositions quantiques.

3. En utilisant la base orthonormée⁴ ($|+, z\rangle \otimes |+, z\rangle, |+, z\rangle \otimes |-, z\rangle, |-, z\rangle \otimes |+, z\rangle, |-, z\rangle \otimes |-, z\rangle$) de $\mathcal{H} \otimes \mathcal{H}$, les questions précédentes, et les postulats de la mécanique quantique, démontrer les propriétés de l'état $|\Psi\rangle$ énoncées ci-dessus.

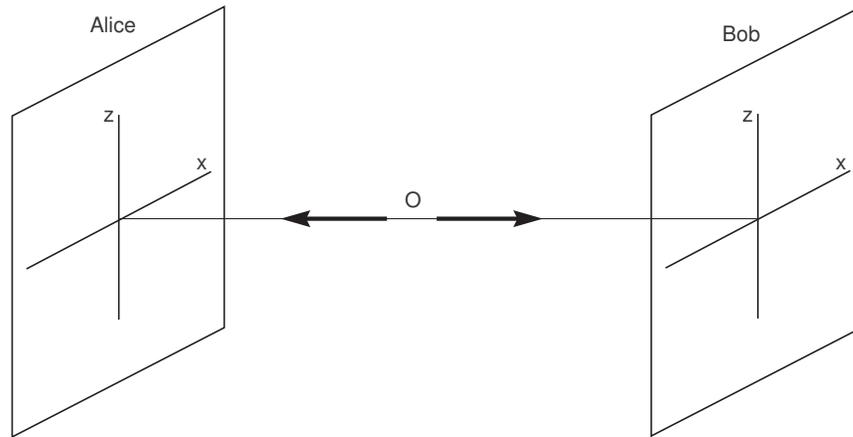


FIGURE 7.3 – L'expérience EPR

Imaginons maintenant que les détecteurs d'Alice et de Bob soient à une très grande distance l'un de l'autre : disons l'un sur la Terre et l'autre sur la Lune. Si Alice et Bob alignent leurs détecteurs, et qu'Alice détecte un spin vers le haut, alors Bob détectera avec certitude un spin vers le bas. De plus, cette propriété ne dépend en rien de l'ordre dans lequel est effectué la mesure : Alice peut effectuer sa mesure avant celle de Bob, ou inversement, ou les deux peuvent l'effectuer simultanément (selon la théorie de la relativité, ces notions d'ordre chronologique sont de toute façon relatives : si les mesures d'Alice et Bob ne sont pas causalement reliées, c'est-à-dire si la lumière n'a pas le temps de passer de l'un à l'autre entre les deux mesures, alors il existe des référentiels dans lesquels Alice effectue sa mesure la première et d'autres dans lesquels c'est Bob, voir [6]). On peut avoir l'impression à ce simple énoncé que la particule sur laquelle Bob va effectuer sa mesure est instantanément informée du résultat obtenu par Alice. Pourtant il n'y a pour l'instant aucune étrangeté : on pourrait parfaitement imaginer que la projection du spin selon (Oz) a été déterminée pour les deux particules au moment de leur départ de O . Il est clair que si on envoie deux messagers depuis O vers Alice et Bob, chacun porteur d'une carte à jouer tirée au sort parmi deux, l'une étant rouge et l'autre noire, il n'y aura pas lieu de s'étonner que la réception d'une carte rouge par Alice implique logiquement la réception d'une carte noire par Bob : ce type de corrélation est parfaitement banal. Mais il y a une très grosse différence dans la situation décrite avec les particules. En effet, rien n'empêche Alice et Bob d'attendre le dernier moment pour décider selon quel axe ils vont mesurer le spin. Imaginons qu'ils tirent à pile ou face au dernier moment pour décider s'ils vont tourner leur appareil pour mesurer la composante selon (Oz) ou selon (Ox) . Dans le cas où ils ont tous les deux tiré la même face de la pièce, leurs résultats de spin seront parfaitement anticorrélés. Comme dans le cas de Schrödinger avec son expérience du chat, il s'agissait pour Einstein, Podolsky et Rosen de montrer que la mécanique quantique révélait une faille. En l'occurrence, ils concluaient que la mécanique quantique était soit incomplète, c'est-à-dire qu'il devait exister certaines variables « cachées », ignorées par la mécanique quantique, dont la connaissance permettrait de prédire avec certitude le résultat des mesures d'Alice et Bob en fonction de l'orientation relative de leurs appareils de mesure (l'équivalent des cartes à jouer, mais en nombre infini), soit que la mécanique quantique était une théorie non locale, c'est-à-dire qu'il pouvait exister un lien à distance entre deux parties d'un système. Pour ces auteurs, il était bien évident qu'une théorie physique sensée se devait d'être locale. Ils pensaient donc disposer

4. Pour la définition du produit scalaire sur $\mathcal{H} \otimes \mathcal{H}$, voir l'appendice C.

d'un argument solide montrant que la mécanique quantique était incomplète. La situation est restée floue pendant assez longtemps. En 1964, John Bell a démontré que quelle que soit la théorie locale à variables cachées qu'on puisse imaginer, les corrélations entre les résultats d'Alice et Bob devaient satisfaire certaines inégalités, tandis que les prédictions de la mécanique quantique violent ces inégalités. Les vérifications expérimentales de ces inégalités sont très délicates, et ce n'est qu'au début des années 1980 que l'équipe d'Alain Aspect de l'institut d'optique d'Orsay a pu montrer de façon claire que les inégalités de Bell sont violées. Autrement dit, il est impossible de décrire les corrélations observées par des corrélations classiques de types « carte à jouer ». Ceci n'exclut pas l'existence de variables cachées, et ne prouve pas que la mécanique quantique soit une théorie complète (comment pourrait-on le prouver ?), mais exclut les théories à variables cachées locales. Ainsi, si des variables cachées existent, elles doivent s'influencer de façon non locale, ce qui n'est pas moins étrange que la non localité quantique.

Remarquons toutefois qu'il est abusif de parler d'influence du résultat de la mesure d'Alice sur celle de Bob, ou réciproquement. Il est plus correct de dire que le système dans l'état $|\Psi\rangle$ ne peut pas être décomposé entre deux sous-systèmes : on dit qu'il est dans un état *intriqué*.

On ne peut pas non plus utiliser cet état intriqué pour communiquer à distance en violant la causalité (c'est-à-dire en transmettant une information plus vite que la lumière). En effet, imaginons qu'Alice essaie d'utiliser cette expérience (en la répétant un certain nombre de fois) pour déterminer la position de l'appareil de mesure de Bob (soit (Ox) , soit (Oz)). Elle peut tourner son appareil selon (Ox) , faire une série de mesures, puis le tourner selon (Oz) et faire une autre série de mesure. Ce qu'elle obtient est une suite de résultats $+\hbar/2$ ou $-\hbar/2$ avec une probabilité $1/2-1/2$, et ce quel que soit ce que fait Bob de son côté. Il n'y a aucun moyen d'en retirer une information sur la position de l'appareil de Bob, ni sur les résultats qu'il a obtenu. Ce n'est qu'en comparant leurs résultats qu'Alice et Bob pourront vérifier que leurs résultats étaient anticorrélés lorsque leurs appareils étaient orientés selon le même axe, mais pour comparer leurs résultats, il faut qu'ils communiquent par un moyen classique, allant au maximum à la vitesse de la lumière !

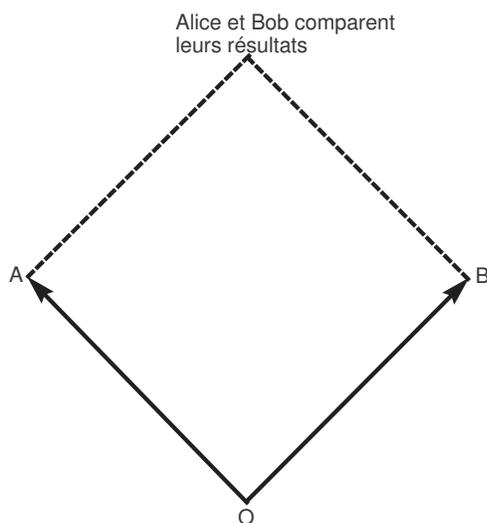


FIGURE 7.4 – Alice et Bob doivent voyager l'un vers l'autre (ligne pointillée), ou s'envoyer un message, pour comparer leurs résultats.

On peut trouver une application très intéressante aux états intriqués : il s'agit de la cryptographie quantique. Le but de la cryptographie quantique n'est pas de coder un message de façon sûre, mais de permettre à deux interlocuteurs de s'envoyer des messages en étant sûr que personne ne les intercepte. Le principe générale est le suivant. Alice envoie une suite aléatoire de particules dans des états de spin

$|\pm, x\rangle$ ou $|\pm, z\rangle$. Bob mesure le spin des particules selon un axe (Ox) ou (Oz), également au hasard. Dans un cas sur deux environ, Bob choisit le même axe qu'Alice et mesure la composante de spin qu'Alice a préparé : ces cas là serviront ultérieurement à transmettre un message. Dans les autres cas, la mesure de Bob donne un résultat aléatoire et n'a pas d'intérêt. Ensuite, Bob transmet publiquement ses choix d'axes et une partie de ses résultats. Si un espion a intercepté les particules d'Alice pour en mesurer le spin, cette dernière pourra s'en apercevoir, car la mesure de l'espion va perturber l'état de certaines particules.

Prenons un exemple. Supposons qu'Alice envoie des particules dans les états

$$|\psi_1\rangle = |+, x\rangle, |\psi_2\rangle = |+, z\rangle, |\psi_3\rangle = |-, z\rangle, |\psi_4\rangle = |+, z\rangle$$

Bob mesure les spins selon les axes (Ox), (Ox), (Oz) et (Oz) successivement. Les mesures de Bob des spins des 1^{ère}, 3^e et 4^e particules sont significatives, mais le résultat de sa seconde mesure est aléatoire et donc inutilisable.

Maintenant supposons qu'un espion ait mesuré les états de spins avant Bob. Comme il ne connaît pas plus que lui quels axes a choisi Alice, il effectue un choix aléatoire d'axes, par exemple (Ox), (Oz), (Ox), (Oz). Après les mesures de l'espion, la première particule est toujours dans l'état $|\psi_1\rangle$, puisque l'espion a mesuré son spin selon l'axe (Ox) et que $|\psi_1\rangle$ est un vecteur propre de S_x . De même, l'état de la 2^e et la 4^e particule n'est pas perturbé. En revanche, l'état de la 3^e particule devient $|\psi'_3\rangle = |+, x\rangle$ ou $|-, x\rangle$ avec probabilité $1/2-1/2$. Supposons que l'on ait $|\psi'_3\rangle = |+, x\rangle$.

Quand Bob va mesurer l'état de la 3^e particule, il a une chance sur deux de trouver $+\hbar/2$. Supposons que ce soit le cas, et supposons alors que Bob publie tous ses choix d'axes et ses résultats pour les 3^e et 4^e particules. Alice constate alors que le résultat de Bob diffère du sien alors qu'il a mesuré le spin selon la même composante : c'est anormal, cela signifie que la transmission a été écoutée.

Dans le cas qui vient d'être décrit, Alice n'a qu'une faible chance de détecter l'écoute de l'espion. Mais lorsque le nombre de particules émises augmente, la probabilité de détection tend rapidement vers 1. Si Alice envoie par exemple 100 particules, et que Bob publie tous ses choix d'axes et 50 de ses résultats, la probabilité qu'une écoute ne soit pas détectée est négligeable. Une fois que la sûreté de la transmission a été ainsi assurée, Alice va indiquer publiquement à Bob quels sont les choix d'axes qu'il doit prendre en compte. Le message d'Alice est constitué par la suite de bits ($+\hbar/2$ correspond par exemple à 1 et $-\hbar/2$ à 0) correspondant à ces choix : cette suite de résultats de mesures de spin est connue d'Alice et Bob et d'eux seuls.

Mais ce qui précède ne fait pas encore intervenir d'états intriqués. Ceux-ci permettent en fait de démontrer le théorème de non-clonage quantique qui assure la sécurité complète de la procédure. Ce théorème affirme qu'il est impossible de fabriquer un double d'un état quantique. Si c'était possible, l'espion pourrait fabriquer des doublons des particules envoyées par Alice, mesurer les états des doublons, et réémettre vers Bob les particules d'Alice non perturbées. Mais cela est impossible par le théorème de non clonage. Voici l'idée de la démonstration (d'après [4], chap. 14). Supposons que l'on désire copier un état $|\psi_1\rangle$. Cet état n'est pas du tout connu. Nous disposons d'un état connu $|\phi\rangle$ qui nous sert de « feuille blanche » sur laquelle on va « photocopier » $|\psi_1\rangle$. L'opération de clonage, si elle est possible, doit donc être une opération du type :

$$C : |\psi_1\rangle \otimes |\phi\rangle \longmapsto |\psi_1\rangle \otimes |\psi_1\rangle \quad (7.3)$$

Supposons maintenant que l'état à copier soit $|\psi_2\rangle$, avec $|\psi_2\rangle \perp |\psi_1\rangle$. L'opération de clonage s'écrit

$$C : |\psi_2\rangle \otimes |\phi\rangle \longmapsto |\psi_2\rangle \otimes |\psi_2\rangle \quad (7.4)$$

Soit maintenant $|\psi_3\rangle = |\psi_1\rangle + |\psi_2\rangle$. Le clonage doit encore s'écrire

$$C : |\psi_3\rangle \otimes |\phi\rangle \longmapsto |\psi_3\rangle \otimes |\psi_3\rangle \quad (7.5)$$

Mais si on ne veut pas que le clonage perturbe l'état original, il ne faut faire aucune opération de mesure. L'opération de clonage doit donc être linéaire. Or la somme de (7.3) et (7.4) est différente de (7.5) :

$$|\psi_1\rangle \otimes |\psi_1\rangle + |\psi_2\rangle \otimes |\psi_2\rangle \neq (|\psi_1\rangle + |\psi_2\rangle) \otimes (|\psi_1\rangle + |\psi_2\rangle)$$

Chapitre 8

Conclusion

Nous arrivons maintenant au terme de ce rapide survol des différents concepts clefs de la mécanique quantique. Nous avons dû, à regret, laisser bien des sujets de côté. Nous espérons que cette introduction donnera au lecteur l'envie de poursuivre plus avant son étude de cette théorie fascinante. Il existe pour cela de très nombreux livres de référence, dont certains sont incontournables par leur exhaustivité ([7]). Pour prendre la suite de cette introduction, nous conseillons vivement la lecture de [3] pour le style vivant et l'enthousiasme communicatif, et [4] qui le complète parfaitement. Nous avons d'ailleurs souvent suivi [4] presque à la lettre, car cet ouvrage nous semble réaliser un très bon compromis entre l'exposition des idées physiques et le formalisme mathématique. Pour les nombreux exercices corrigés et l'accent mis sur le développement historique de la mécanique quantique, l'ouvrage [1] est également chaudement recommandé. Enfin, ceux qui désirent s'initier à l'approche particulière de Feynman, sur laquelle nous avons fait l'impasse, pourront lire [8] avec profit.

Terminons en disant quelques mots des rapports entre les deux grandes théories qui ont révolutionné la physique au XX^e siècle à côté de la mécanique quantique : la relativité restreinte et la relativité générale. Comme nous l'avons signalé, l'équation de Schrödinger est incompatible avec la relativité restreinte. On peut facilement trouver une équation similaire qui possède une forme compatible avec les transformations de Lorentz : il s'agit de l'équation de Klein-Gordon

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

Cette équation avait d'ailleurs été envisagée initialement par Schrödinger, mais il l'a rejetée car elle ne donnait pas les bons résultats pour le spectre de l'hydrogène. En fait, l'équation de Klein-Gordon pose des problèmes d'interprétation car ses solutions ne peuvent pas être utilisées pour construire des densités de probabilité. Une autre équation relativiste, qui est une sorte de racine carrée de l'équation de Klein-Gordon a été trouvée par Dirac. Dans ce cas la fonction d'onde est à valeur dans l'espace de Hilbert d'une particule de spin 1/2. Elle a permis de rendre compte théoriquement de l'existence de l'antimatière, dont le premier exemplaire, le positon, a été découvert deux ans après l'équation de Dirac. Néanmoins, elle conduit elle aussi à certaines incohérences, qui ont été résolues par la construction de la théorie quantique du champ électromagnétique dans les années 1940 par Feynman, Dyson, Schwinger et Tomonaga : l'électrodynamique quantique. Cette première *théorie de champ quantique* fut suivi par d'autres, incorporant les forces nucléaires. Les théories de champ quantiques réalisent la synthèse entre la relativité restreinte et la mécanique quantique. On peut démontrer à l'aide de ces théories la relation spin-statistique, et l'invariance des lois physiques par la symétrie PCT, composée de la symétrie par rapport à un plan, de la conjugaison de charge et du renversement du temps. Le modèle standard des particules élémentaires est une théorie quantique de champs qui décrit avec précision les phénomènes mettant en jeu les forces électromagnétiques et nucléaires.

Si la relativité restreinte a pu être incorporée avec succès dans le cadre quantique, il n'en va pas de même de la relativité générale. Cette dernière théorie est à la fois la meilleure description que l'on connaisse

du champ de gravitation, et un cadre conceptuel mettant en jeu l'espace-temps, et qui a vocation à s'appliquer à toute la physique. Or ce cadre est en conflit violent avec la théorie quantique. La meilleure façon d'unifier relativité générale et physique quantique est aujourd'hui un sujet de recherche actif, qui recèle de sévères disputes et d'âpres débats entre les partisans des différentes approches existantes¹ (théorie des cordes, gravitation quantique à boucles, géométrie non commutative,...). C'est le « nuage noir » que nous a légué la physique du XX^e siècle, et il engendrera, n'en doutons pas, une révolution d'une amplitude comparable à celles provoquées par les problèmes du corps noir et de l'éther. Une révolution qui nous obligera à reconsidérer tous nos concepts, aussi bien relativistes que quantiques. Mais ceci est une autre histoire, qui n'est pas encore écrite...

1. Pour une bonne introduction à ces questions, on pourra consulter [19] ou [20].

Annexe A

Solutions des exercices

Exercice 5.2.1 Soient deux complexes λ, λ' et deux vecteurs v et v' . Pour tout $w \in V$, on a $\epsilon(\lambda v + \lambda' v')(w) = \langle \lambda v + \lambda' v', w \rangle = \bar{\lambda} \langle v, w \rangle + \bar{\lambda}' \langle v', w \rangle = \bar{\lambda} \epsilon(v)(w) + \bar{\lambda}' \epsilon(v')(w) = (\bar{\lambda} \epsilon(v) + \bar{\lambda}' \epsilon(v'))(w)$. Ceci étant vrai pour tout w , on en déduit que les applications $\epsilon(\lambda v + \lambda' v')$ et $\bar{\lambda} \epsilon(v) + \bar{\lambda}' \epsilon(v')$ coïncident. Ainsi ϵ est antilinéaire.

Supposons qu'on ait $\epsilon(v) = \epsilon(v')$. Cela signifie que pour tout $w \in V$ on a $\langle v, w \rangle = \langle v', w \rangle$, ce qui implique que $\langle v - v', w \rangle = 0$ pour tout $w \in V$. Ainsi $v - v'$ est orthogonal à tous les vecteurs de l'espace, ce qui montre que $v - v' = 0$, donc $v = v'$. Par conséquent l'application ϵ est injective.

Exercice 5.2.2 On a pour tout $|v\rangle, |e_i\rangle \langle e_i|v\rangle = |e_i\rangle \times v_i = p_i(|v\rangle)$. Donc les applications p_i et $|e_i\rangle \langle e_i|$ sont égales.

Par ailleurs, il est clair que pour tout $|v\rangle$ on a $v = \sum_i p_i(|v\rangle)$ donc on peut écrire $\text{Id}_V = \sum_i p_i$. Le résultat suit.

Exercice 5.2.3 Pour tous $|v\rangle, |w\rangle$, on a :

$$\begin{aligned} \langle v | (\lambda_1 O_1 + \lambda_2 O_2) | w \rangle &= \lambda_1 \langle v | O_1 | w \rangle + \lambda_2 \langle v | O_2 | w \rangle \\ &= \lambda_1 \langle O_1^* | v \rangle, | w \rangle + \lambda_2 \langle O_2^* | v \rangle, | w \rangle \\ &= \langle (\bar{\lambda}_1 O_1^* + \bar{\lambda}_2 O_2^*) | v \rangle, | w \rangle \end{aligned}$$

D'où le résultat.

Exercice 5.2.4 Pour tous $|v\rangle, |w\rangle$, on a :

$$\begin{aligned} \langle |v\rangle, (O^*)^* |w\rangle \rangle &= \overline{\langle (O^*)^* |w\rangle, |v\rangle} \\ &= \overline{\langle |w\rangle, O^* |v\rangle} \\ &= \langle O^* |v\rangle, |w\rangle \\ &= \langle |v\rangle, O |w\rangle \end{aligned}$$

D'où le résultat.

Par ailleurs on a :

$$\begin{aligned} \langle |v\rangle, O_1 O_2 |w\rangle \rangle &= \langle O_1^* |v\rangle, O_2 |w\rangle \\ &= \langle O_2^* O_1^* |v\rangle, |w\rangle \end{aligned}$$

Exercice 5.2.5

$$\begin{aligned} \overline{\langle u | O | v \rangle} &= \overline{\langle O^* | u \rangle, | v \rangle} \\ &= \langle | v \rangle, O^* | u \rangle \\ &= \langle v | O^* | u \rangle \end{aligned}$$

Exercice 5.2.6 $O = O^* \Leftrightarrow iO = iO^* = -(iO)^*$

Exercice 5.2.7 Soit O un opérateur hermitien, et $|v\rangle$ un vecteur propre de O associé à la valeur propre λ . Alors $\langle Ov|v\rangle = \langle v|Ov\rangle \Rightarrow \langle \lambda v|v\rangle = \langle v|\lambda v\rangle \Rightarrow \bar{\lambda}\langle v|v\rangle = \lambda\langle v|v\rangle \Rightarrow \bar{\lambda} = \lambda$, puisque $|v\rangle$ est non nul par définition d'un vecteur propre. Le cas d'un opérateur antihermitien se traite de la même façon.

Si U est unitaire, et si v est un vecteur propre de U pour la valeur propre λ , alors on a $\langle Uv|Uv\rangle = \langle v|v\rangle \Rightarrow \bar{\lambda}\lambda\langle v|v\rangle = \langle v|v\rangle \Rightarrow |\lambda|^2 = 1$. D'où le résultat.

Exercice 5.2.8 Soit $(a_n)_{n \in \mathbb{N}}$ une suite d'opérateurs convergeant vers l'opérateur a . Il suffit de montrer que la suite $(a_n^*)_{n \in \mathbb{N}}$ converge vers a^* pour montrer la continuité du passage à l'adjoint. Or on a $\|a_n - a\| = \|(a_n - a)^*\| = \|a_n^* - a^*\|$. Comme a_n tend vers a , on a $\|a_n - a\|$ qui tend vers 0, donc $\|a_n^* - a^*\|$ également, ce qui prouve que a_n^* tend vers a^* .

Exercice 5.2.9

1. Notons $(s_n(a))$ la suite définissant l'exponentielle (voir le théorème 5.2.4). Pour tout $n \in \mathbb{N}$, il est clair que $s_n(a)^* = s_n(a^*)$. En passant à la limite de chaque côté, on obtient $\exp(a)^* = \exp(a^*)$ (l'interversion de la limite et du passage à l'adjoint est justifiée puisque cette dernière opération est continue).
2. Il suffit d'appliquer le point précédent avec $a = a^*$.
3. Si $a^* = -a$ on obtient $\exp(a)^* = \exp(a^*) = \exp(-a) = \exp(a)^{-1}$. En conclusion, $\exp(a)$ est bien unitaire.

Exercice 5.3.1 Écrivons pour simplifier $U = U(t, t_0)$. On a par définition $\frac{dU}{dt} = -\frac{i}{\hbar}UH = \frac{1}{i\hbar}UH$, d'où $\frac{dU^*}{dt} = -\frac{1}{i\hbar}HU^*$. D'où :

$$\frac{dO_H}{dt} = \frac{dU^*}{dt}O_S U + U^*O_S \frac{dU}{dt} = -\frac{1}{i\hbar}HU^*O_S U + \frac{1}{i\hbar}U^*O_S UH = \frac{1}{i\hbar}[O_H, H]$$

Exercice 6.1.1

$$\begin{aligned} \langle A \rangle_\psi &= \sum_{i=1}^k \alpha_i P(a \rightarrow \alpha_i | \psi) \\ &= \frac{1}{\|\psi\|^2} \sum_{i=1}^k \alpha_i \|\pi_i |\psi\rangle\|^2 \\ &= \frac{1}{\|\psi\|^2} \sum_{i=1}^k \alpha_i \langle \psi | \pi_i^* \pi_i | \psi \rangle \\ &= \frac{1}{\|\psi\|^2} \sum_{i=1}^k \alpha_i \langle \psi | \pi_i | \psi \rangle \\ &= \frac{1}{\|\psi\|^2} \langle \psi | \sum_{i=1}^k \alpha_i \pi_i | \psi \rangle \\ &= \frac{1}{\|\psi\|^2} \langle \psi | A | \psi \rangle \end{aligned}$$

Exercice 6.2.1 On a $\langle e_1 | X | e_1 \rangle = \frac{1}{2} (e^{-i\theta} \quad -1) \begin{pmatrix} -d/2 & 0 \\ 0 & d/2 \end{pmatrix} \begin{pmatrix} e^{i\theta} \\ -1 \end{pmatrix} = 0$. On trouve de même $\langle e_2 | X | e_2 \rangle = 0$.

Pour l'écart-type, on peut remarquer que $X^2 = \frac{d^2}{4} \text{Id}$, d'où $(\Delta X)_{e_1} = (\langle e_1 | X^2 | e_1 \rangle - \langle e_1 | X | e_1 \rangle^2)^{1/2} = (\frac{d^2}{4} - 0)^{1/2} = d/2$.

Exercice 6.5.1 Rappelons la valeur de l'intégrale gaussienne :

$$\int_{-\infty}^{+\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}$$

En appliquant ceci à (6.26), on trouve $\|\psi_0\|^2 = \int_{-\infty}^{+\infty} |\psi_0(x)|^2 dx = C_0^2 \sqrt{\frac{\hbar}{\omega m}}$. Pour obtenir $\|\psi_0\| = 1$, il suffit donc de prendre $C_0 = \left(\frac{\omega m}{\hbar}\right)^{1/4}$.

Exercice 6.5.2 Tout d'abord calculons $\langle \hat{q} \rangle_{|0\rangle}$. On a $Q = \frac{1}{\sqrt{2}}(a + a^*)$, donc $\sqrt{\frac{m\omega}{\hbar}} \langle 0 | \hat{q} | 0 \rangle = \langle 0 | Q | 0 \rangle = \frac{1}{\sqrt{2}}(\langle 0 | a | 0 \rangle + \langle 0 | a^* | 0 \rangle) = 0$. Pour des raisons similaires, on a aussi $\langle 0 | \hat{p} | 0 \rangle = 0$. Remarquons qu'on peut conduire les calculs sans problème directement en représentation r :

$$\langle 0 | \hat{q} | 0 \rangle = \int_{-\infty}^{+\infty} \bar{\psi}_0(x) x \psi_0(x) dx$$

Or l'intégrande étant une fonction impaire, son intégrale sur \mathbb{R} est nulle, d'où le résultat. On a donc

$$\begin{aligned} (\Delta \hat{q})_{|0\rangle} &= (\langle 0 | \hat{q}^2 | 0 \rangle)^{1/2} \\ &= \sqrt{\frac{\hbar}{\omega m}} (\langle 0 | Q^2 | 0 \rangle)^{1/2} \\ &= \sqrt{\frac{\hbar}{\omega m}} (\langle 0 | \frac{1}{2}(a^2 + (a^*)^2 + aa^* + a^*a) | 0 \rangle)^{1/2} \\ &= \sqrt{\frac{\hbar}{2\omega m}} (\langle 0 | aa^* | 0 \rangle)^{1/2}, \text{ car } a|0\rangle = \langle 0 | a^* = 0 \\ &= \sqrt{\frac{\hbar}{2\omega m}} (\langle 0 | 1 + a^*a | 0 \rangle)^{1/2} \\ &= \sqrt{\frac{\hbar}{2\omega m}} \end{aligned}$$

Comme $\hat{p}^2 = \hbar\omega m P^2$ et $P^2 = \frac{1}{2}(aa^* + a^*a - a^2 - (a^*)^2)$, on trouve de même $\langle \hat{p}^2 \rangle_{|0\rangle} = \sqrt{\frac{\hbar\omega m}{2}}$, d'où le résultat.

Exercice 6.5.3 Il s'agit respectivement de $|\psi_0(x)|^2$, $|\psi_1(x)|^2$ et $|\psi_2(x)|^2$, par définition. En utilisant le postulat 5, on peut écrire par exemple :

$$P(q \rightarrow x_0, |0\rangle) = |\langle x_0 | 0 \rangle|^2$$

En représentation r , le bra $\langle x_0 |$ appliqué à une fonction ψ retourne la valeur $\psi(x_0)$. C'est bien une forme linéaire sur l'espace des fonctions que l'on considère¹. En utilisant la définition du produit scalaire pour les fonctions, si le bra $\langle x_0 |$ était le dual d'un ket $|\delta_{x_0}\rangle$, on pourrait écrire

$$\langle x_0 | \psi \rangle = \int_{\mathbb{R}} \delta_{x_0}(x) \psi(x) dx = \psi(x_0) \quad (\text{A.1})$$

On a omis la barre de conjugaison, car δ_{x_0} est réelle. On l'appelle la « distribution de Dirac ». Il faut bien comprendre que l'expression (A.1) est purement symbolique : il n'existe aucune fonction ayant cette propriété. Une telle fonction serait partout nulle sauf en x_0 où elle aurait une valeur infinie. De plus son intégrale sur \mathbb{R} vaudrait 1. On peut trouver des fonctions qui approchent δ_{x_0} , mais son véritable statut est d'être un bra, c'est-à-dire une forme linéaire sur l'espace des fonctions, et dans ce cas, ce bra n'est le dual d'aucun ket.

Exercice 6.7.1 Il se trouve que l'on peut répondre aux deux questions en même temps. On commence par vérifier de façon immédiate que, dans toute algèbre, le commutateur vérifie :

$$[a, bc] = [a, b]c + b[a, c] \quad (\text{A.2})$$

Comme on a de façon évidente $[x, y] = -[y, x]$, on en déduit que

$$[ab, c] = -[c, ab] = [a, c]b + a[b, c] \quad (\text{A.3})$$

1. Sur lequel je ne m'étendrai pas...

En utilisant successivement ces deux relations, on trouve que

$$[ab, cd] = [a, c]bd + a[b, c]d + c[a, d]b + ca[b, d] \quad (\text{A.4})$$

Or on vérifie à l'aide de la formule de Leibniz que le crochet de Poisson vérifie

$$\{f, gh\} = \{f, g\}h + g\{f, h\} \quad (\text{A.5})$$

Comme on a également $\{f, g\} = -\{g, f\}$, la relation (A.4) vaut aussi bien avec des crochets de Poisson qu'avec des commutateurs. Or, à l'aide de cette relation on peut ramener $[\hat{L}_x, \hat{L}_y]$ aussi bien que $\{L_x, L_y\}$ aux relations de commutations canoniques. Faisons-le pour les commutateurs :

$$\begin{aligned} [\hat{L}_x, \hat{L}_y] &= [\hat{y}\hat{p}_z - \hat{z}\hat{p}_y, \hat{z}\hat{p}_x] \\ &= [\hat{y}\hat{p}_z, \hat{z}\hat{p}_x] - [\hat{z}\hat{p}_y, \hat{z}\hat{p}_x] - [\hat{y}\hat{p}_z, \hat{x}\hat{p}_z] + [\hat{z}\hat{p}_y, \hat{x}\hat{p}_z] \\ &= \hat{y}[\hat{p}_z, \hat{z}]\hat{p}_x + \hat{x}[\hat{z}, \hat{p}_z]\hat{p}_y, \text{ d'après (A.4) et les relations de commutation canoniques} \\ &= i\hbar(-\hat{y}\hat{p}_x + \hat{x}\hat{p}_y) = i\hbar\hat{L}_z \end{aligned}$$

En vertu de ce qu'on a dit plus haut, le calcul est exactement le même avec les crochets de Poisson, sauf que le facteur $i\hbar$ n'apparaît pas.

Exercice 6.7.2 D'après le postulat 4, la valeur s_x d'une mesure du spin selon l'axe (Ox) ne peut que donner $\pm\hbar/2$, puisque ce sont les valeurs propre de l'observable associée σ_x . Les états étant normalisés, on a, suivant le postulat 5 :

$$P(s_x \rightarrow +\hbar/2; |+, z\rangle) = |\langle +, x|+, z\rangle|^2 = \left(\frac{1}{\sqrt{2}} \times (1 \times 1 + 1 \times 0)\right)^2 = 1/2$$

L'autre probabilité cherchée est égale à $1 - 1/2 = 1/2$.

Exercice 7.2.1

1. Soit $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ la matrice de passage de la base $(|u\rangle, |v\rangle)$ vers la base $(|+, z\rangle, |-, z\rangle)$. On a donc $|+, z\rangle = a|u\rangle + b|v\rangle$, $|-, z\rangle = c|u\rangle + d|v\rangle$. En introduisant ces formules dans la définition de $|\Psi\rangle$, et en utilisant la bilinéarité du produit tensoriel, qui permet de développer celui-ci comme un produit, le résultat vient.
2. On a

$$S_z \otimes \text{Id}(|\pm, z\rangle \otimes |v\rangle) = (S_z|\pm, z\rangle) \otimes |v\rangle = \pm\frac{\hbar}{2}|\pm, z\rangle \otimes |v\rangle$$

Les vecteurs de la forme $|\pm, z\rangle \otimes |v\rangle$ sont bien des vecteurs propres de $S_z \otimes \text{Id}$ associés à la valeur propre $\pm\hbar/2$. Maintenant, on vérifie immédiatement que l'ensemble $V_+ = \{|+, z\rangle \otimes |v\rangle, |v\rangle \in \mathcal{H}\}$, et l'ensemble $V_- = \{|-, z\rangle \otimes |v\rangle, |v\rangle \in \mathcal{H}\}$, sont des sous-espaces vectoriels de $\mathcal{H} \otimes \mathcal{H}$ de dimension 2. Comme ils sont associés à des valeurs propres différentes, ils sont en somme directe, ainsi $\mathcal{H} \otimes \mathcal{H} = V_+ \oplus V_-$, ce qui montre que l'on a bien trouvé toutes les valeurs propres et tous les vecteurs propres de $S_z \otimes \text{Id}$

3. (a) Une mesure de spin de la particule A selon l'axe (Oz) produit le résultat $s_z^A = \pm\hbar/2$, puisque ce sont les valeurs propres de l'observable $S_z \otimes \text{Id}$. Pour calculer les probabilités de ces résultats, on applique le postulat 5. Pour éviter des notations trop lourdes, notons $|+, z\rangle = |+\rangle$, $|+, z\rangle \otimes |+, z\rangle = |++\rangle$, et ainsi de suite. La probabilité de trouver la valeur $+\hbar/2$ lorsque le système est dans l'état $|\Psi\rangle = \frac{1}{\sqrt{2}}(|+-\rangle - |-+\rangle)$ est :

$$\begin{aligned} P(s_z^A \rightarrow +\hbar/2, |\Psi\rangle) &= |\langle ++|\Psi\rangle|^2 + |\langle +-|\Psi\rangle|^2 \\ &= \frac{1}{2}(|\langle ++\rangle\langle +-| - \langle +-|\langle ++\rangle|^2 + |\langle ++\rangle\langle -|- \rangle - \langle +-|\langle -|+\rangle|^2) = 1/2 \end{aligned}$$

Maintenant, soit un axe quelconque dirigée par le vecteur \vec{k} . Le spin d'une particule selon l'axe dirigé par \vec{k} ne peut prendre que les valeurs $\hbar/2$ et $-\hbar/2$. En utilisant la première question, on peut remplacer la base $(|+, z\rangle, |-, z\rangle)$ par la base $(|+, \vec{k}\rangle, |-, \vec{k}\rangle)$ où $|\pm, k\rangle$ est un vecteur propre associé à la valeur propre $\pm\hbar/2$. Le changement de base étant orthonormé, il se fait par une matrice unitaire. L'expression de l'état $|\Psi\rangle$ ne change donc que par une phase sans conséquence. Par ailleurs, les vecteurs propres de l'observable $S_{\vec{k}} \otimes \text{Id}$ sont de la forme $|\pm, k\rangle \otimes |v\rangle$. Donc on peut conduire exactement le même calcul en remplaçant partout z par \vec{k} .

Pour ce qui est de la particule B, il faut s'intéresser à l'observable $\text{Id} \otimes S_z$, ou plus généralement $\text{Id} \otimes S_{\vec{k}}$. Les calculs sont symétriques des précédents.

- (b) Supposons que la mesure de A soit faite avant² celle de B. Alors si on trouve $+\hbar/2$ pour le spin selon (Oz) , l'état $|\Psi\rangle$ se trouve projeté sur l'état (on sous-entend toujours les z pour simplifier) :

$$\langle ++ | \Psi \rangle | ++ \rangle + \langle +- | \Psi \rangle | +- \rangle = | +- \rangle$$

Or l'état $| +- \rangle$ est un état propre associé à la valeur propre $-\hbar/2$ de l'observable $\text{Id} \otimes S_z$. Ainsi, une mesure ultérieure du spin de la particule B selon l'axe (Oz) donne forcément le résultat $-\hbar/2$. On vérifie de même que si la mesure de A est $-\hbar/2$, une mesure ultérieure du spin selon (Oz) de la particule B donne $+\hbar/2$.

Supposons maintenant que la mesure de B soit faite avant celle de A. Observons qu'une base de vecteur propre associé à la valeur propre $+\hbar/2$ de l'observable $\text{Id} \otimes S_z$ est $(| ++ \rangle, | - + \rangle)$. Ainsi, si le résultat de B est $+\hbar/2$, l'état $|\Psi\rangle$ se trouve projeté sur

$$\langle ++ | \Psi \rangle | ++ \rangle + \langle - + | \Psi \rangle | - + \rangle = - | - + \rangle$$

Or $| - + \rangle$ est un vecteur propre de $S_z \otimes \text{Id}$ associé à la valeur propre $-\hbar/2$. Donc le résultat est le même que précédemment : les mesures de A et B sont toujours opposées.

Supposons maintenant que A mesure la composante du spin de sa particule selon (Oz) , et qu'ensuite B fasse sa mesure selon (Ox) . Comme précédemment, après la mesure de A l'état se trouve projeté soit sur $| + -, z \rangle := | +, z \rangle \otimes | -, z \rangle$, dans le cas d'un résultat $+\hbar/2$, soit dans l'état $| - +, z \rangle$ en cas de résultat $-\hbar/2$. Plaçons-nous dans le premier cas. La probabilité qu'a B de trouver $+\hbar/2$ en mesurant le spin selon (Ox) vaut, avec des notations que nous espérons claires :

$$\begin{aligned} P(s_B^x \rightarrow +\hbar/2) &= |\langle ++, x | + -, z \rangle|^2 + |\langle - +, x | + -, z \rangle|^2 \\ &= |\langle +, x | -, z \rangle \langle +, x | +, z \rangle|^2 + |\langle -, x | -, z \rangle \langle +, x | +, z \rangle|^2 \\ &= \left| \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} \right|^2 + \left| -\frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}} \right|^2 = \frac{1}{2} \end{aligned}$$

d'où le résultat.

2. Si A et B ne sont pas causalement reliées, cette supposition n'a pas de sens absolu selon la théorie de la relativité, mais dépend du référentiel galiléen dans lequel on s'exprime. Néanmoins nous allons voir que le résultat est le même dans tous les cas, il est donc indépendant du référentiel.

Annexe B

Condensé de mécanique analytique

Dans la présentation traditionnelle de la mécanique classique, on recherche la trajectoire $t \mapsto \gamma(t)$ d'un point matériel connaissant sa position initiale $\gamma(t_i)$, sa vitesse initiale $\dot{\gamma}(t_i)$, et la force (ou la résultante des forces) $\vec{F}(t)$ agissant sur lui. Le tout fonctionne à l'aide du principe fondamental de la dynamique. La mécanique analytique est une reformulation de la mécanique classique dans laquelle on cherche quelle est la trajectoire suivie par un point matériel passant du point M_i au point M_f entre les instants t_i et t_f . Autrement dit on cherche $t \mapsto \gamma(t)$ vérifiant

$$\gamma(t_i) = M_i \text{ et } \gamma(t_f) = M_f \quad (\text{B.1})$$

Pour chaque courbe vérifiant (B.1), on définit l'action :

$$S(\gamma) = \int_{t_i}^{t_f} \mathcal{L}(\gamma(t), \dot{\gamma}(t), t) dt \quad (\text{B.2})$$

où l'intégrande s'appelle le *lagrangien*. Tout système mécanique est caractérisé par un lagrangien. Dans le cas d'une particule de masse m dans \mathbb{R}^3 soumise à un potentiel¹ V dépendant du point et éventuellement du temps, le lagrangien est égal à

$$\mathcal{L}(q, \dot{q}, t) = \frac{1}{2} m \|\dot{q}\|^2 - V(q, t) \quad (\text{B.3})$$

où q désigne collectivement les trois variables de position (par exemple $q = (x_1, x_2, x_3)$ en coordonnées cartésiennes), et \dot{q} désigne les trois variables de vitesse. Notons que la fonction \mathcal{L} est définie pour toutes valeurs de ses 6 variables (plus le temps), indépendamment du fait que ces variables représentent la position ou la vitesse d'une trajectoire effectivement suivie. Le lagrangien n'est rien d'autre qu'une fonction de plusieurs variables, tandis que l'action a pour argument une trajectoire : on dit parfois que c'est une fonctionnelle.

Pour savoir quelle est la trajectoire effectivement suivie, on applique le principe de moindre action.

Principe de moindre action : La trajectoire suivie rend l'action $S(\gamma)$ extrémale.

Notons que ce principe porte mal son nom puisque l'action peut être maximale et pas forcément minimale. En fait, ce principe serait mieux formulé en ces termes : la trajectoire suivie annule la variation δS de l'action, lorsque les extrémités M_i et M_f sont fixées. La quantité $\delta S(\gamma)$ est la variation au premier ordre de l'action lorsqu'on passe de la trajectoire γ à une trajectoire très voisine, tout en gardant les extrémités fixées. Le principe de moindre action réclame donc en fait qu'une sorte de dérivée généralisée, qu'on pourrait noter $\frac{\delta S}{\delta \gamma}$ s'annule. On peut montrer que l'annulation de δS équivaut aux équations d'Euler-Lagrange :

1. Il s'agit là encore de l'énergie potentielle.

$$\frac{\partial \mathcal{L}}{\partial q_i} = \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \quad (\text{B.4})$$

où $i = 1, 2, 3$ pour une particule dans \mathbb{R}^3 , mais ce formalisme se généralise à un nombre plus grand de degrés de liberté (par exemple pour un système de particules).

La quantité

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \quad (\text{B.5})$$

s'appelle *moment conjugué* de la variable q_i . Les moments conjugués s'appellent aussi des impulsions. Dans le cas d'une particule dans \mathbb{R}^3 soumise à un potentiel dépendant seulement du point de l'espace et éventuellement du temps, les impulsions et les quantités de mouvement $m\dot{x}_i$ coïncident (dans des coordonnées cartésiennes). Mais ce n'est pas le cas en général, par exemple en présence d'un champ magnétique (car la force de Lorentz dépend de la vitesse).

Dans le formalisme de hamilton, on décrit l'état d'un système à l'aide des variables q_i et p_i . L'espace dans lequel elles varient s'appelle *espace des phases*. Pour décrire la dynamique dans l'espace des phases, il faut d'abord construire le hamiltonien. Celui s'obtient en inversant, si c'est possible, la relation (B.5) afin d'écrire \dot{q}_i en fonction des p_j et des q_j . On peut alors poser

$$H(q, p, t) := \sum_i p_i \dot{q}_i(q, p) - \mathcal{L}(q, p, t) \quad (\text{B.6})$$

On montre alors, en calculant la différentielle dH , que l'on a, pour tout i :

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \text{ et } \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad (\text{B.7})$$

Ce sont les équations du mouvement, dites équations canoniques de Hamilton-Jacobi.

Soient maintenant deux fonctions f, g des variables p, q, t . Ces fonctions sont des « observables classiques ». Le crochet de Poisson de f et g est défini (lorsque les dérivées partielles existent), par :

$$\{f, g\} := \sum_i \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} \right) \quad (\text{B.8})$$

Rappelons que i parcourt $1..n$, où n est le nombre de degrés de liberté du système (par exemple $n = 3$ dans le cas d'une particule dans l'espace).

On en déduit immédiatement les relations

$$\{q_i, q_j\} = \{p_i, p_j\} = 0; \quad \{q_i, p_j\} = \delta_{ij} \quad (\text{B.9})$$

En utilisant les équations de Hamilton-Jacobi, on trouve, pour toute fonction f comme ci-dessus :

$$\frac{df}{dt} = \{f, H\} + \frac{\partial f}{\partial t} \quad (\text{B.10})$$

Si f ne dépend pas explicitement du temps, et si son crochet de Poisson avec le hamiltonien est nul, alors on voit que f est une constante du mouvement.

Annexe C

Produit tensoriel

Commençons par le plus simple. Soient V et W deux K -espaces vectoriels de dimension finie de bases respectives $(v_i)_{1 \leq i \leq n}$ et $(w_i)_{1 \leq i \leq p}$. Alors $V \otimes W$ est un K -espace vectoriel de dimension np , dont une base se note $(v_i \otimes w_j)_{1 \leq i \leq n, 1 \leq j \leq p}$. La construction de $V \otimes W$ est indépendante des bases et s'étend à la dimension infinie. Si $v = \sum_i \lambda_i v_i \in V$ et $w = \sum_j \mu_j w_j \in W$, le produit tensoriel $v \otimes w \in V \otimes W$ est :

$$v \otimes w = \sum_{i,j} \lambda_i \mu_j v_i \otimes w_j$$

Le produit tensoriel définit donc une application bilinéaire $V \times W \rightarrow V \otimes W$, qui à (v, w) associe $v \otimes w$. Un élément de $V \otimes W$ s'appelle un tenseur. Un tenseur de la forme $v \otimes w$ s'appelle un tenseur décomposé. Un tenseur général est une combinaison linéaire de tenseurs décomposés.

Le produit tensoriel possède la propriété suivante : pour toute application bilinéaire $\phi : V \times W \rightarrow X$, où X est un K -espace vectoriel, il existe une unique application linéaire T telle que $T(v \otimes w) = \phi(v, w)$, pour tout $v \in V$ et $w \in W$. Cela signifie que l'on pose, pour tout tenseur décomposé $v \otimes w$:

$$T(v \otimes w) = \phi(v, w)$$

où ϕ est bilinéaire, alors T est bien définie et s'étend de façon unique en une application linéaire de $V \otimes W$ dans X .

Le produit tensoriel est associatif. Cela signifie que l'on peut identifier les K -e.v. $U \otimes (V \otimes W)$ et $(U \otimes V) \otimes W$.

Lorsqu'on a un espace de fonctions, du type $L^2(\mathbb{R}^3, \mathbb{C})$, et un \mathbb{C} -espace vectoriel de dimension finie V , alors le produit tensoriel $L^2(\mathbb{R}^3, \mathbb{C}) \otimes V$ s'identifie à l'espace de fonctions $L^2(\mathbb{R}^3, V)$ via l'application définie sur les tenseurs décomposés par

$$f \otimes v \mapsto (x \mapsto f(x)v)$$

Cela signifie que le tenseur $f \otimes v$ s'identifie à l'application de $\mathbb{R}^3 \rightarrow V$ qui à x associe $f(x)v$.

Si maintenant on a deux espaces de fonctions, comme par exemple $L^2(X, \mathbb{C})$ et $L^2(Y, \mathbb{C})$, où X et Y sont par exemple des parties de \mathbb{R}^3 , alors le produit tensoriel $L^2(X, \mathbb{C}) \otimes L^2(Y, \mathbb{C})$ peut être vu comme un sous-espace de $L^2(X \times Y, \mathbb{C})$ via l'identification du tenseur $f \otimes g$ et de la fonction de deux variables $(x; y) \mapsto f(x)g(y)$. Les fonctions de deux variables définies sur $X \times Y$ et à valeur dans \mathbb{C} peuvent ne pas être des combinaisons linéaires (rappelons qu'une combinaison linéaire est toujours une somme finie) du type $(x; y) \mapsto \sum_i \lambda_i f_i(x)g_i(y)$. Néanmoins, l'espace des fonctions de deux variables $L^2(X \times Y, \mathbb{C})$ est une certaine complétion de $L^2(X, \mathbb{C}) \otimes L^2(Y, \mathbb{C})$. Cela signifie que toute fonction de deux variables peut être suffisamment bien approchée par une combinaison linéaire du type précédent. Il en résulte qu'en mécanique quantique, l'espace de Hilbert de deux particules dans l'espace (sans autre degré de liberté que leur position, pour simplifier) est $L^2(\mathbb{R}^3 \times \mathbb{R}^3, \mathbb{C})$. Un vecteur d'état d'un tel système de particule est une fonction de deux variables (vectorielles) $\Psi(x, y)$, ou encore une fonction de 6 variables

réelles $\Psi(x_1, x_2, x_3, y_1, y_2, y_3)$, où x_1, x_2, x_3 sont les coordonnées de la première particule et y_1, y_2, y_3 les coordonnées de la seconde.

Enfin si $\langle \cdot, \cdot \rangle$ est un produit scalaire sur \mathcal{H} , et (\cdot, \cdot) un produit scalaire sur \mathcal{K} , on vérifie qu'en posant sur les tenseurs décomposés

$$\langle u \otimes v, u' \otimes v' \rangle := \langle u, v \rangle (u', v')$$

et en étendant cette définition par bilinéarité, on définit un produit scalaire sur $\mathcal{H} \otimes \mathcal{K}$.

Annexe D

Constantes

La notation 1,234567(89) signifie $1,234567 \pm 0,000089$.

Constante de Planck : $h = 6,62606896(33) \times 10^{-34} \text{J.s}$

Constante de Planck réduite $\hbar = \frac{h}{2\pi} = 1,054571628(53) \text{J.s}$

Constante de Boltzmann $k = 1,3806504(24) \times 10^{-23} \text{J.K}^{-1}$

Vitesse de la lumière : $c = 2,99792458 \times 10^8 \text{m.s}^{-1}$ (valeur exacte par définition)

masse de l'électron : $m_e = 9,10938215(45) \times 10^{-31} \text{kg}$

masse du proton : $m_p = 1,672621637(83) \times 10^{-27} \text{kg}$

charge de l'électron : $e = -1,602176487(40) \times 10^{-19} \text{C}$.

permittivité du vide : $\epsilon_0 = \frac{1}{\mu_0 c^2} \approx 8,854187817 \times 10^{-12} \text{A}^2 \text{s}^4 \text{kg}^{-1} \text{m}^{-3}$ (valeur définie)

Bibliographie

- [1] C. Aslangul, *Mécanique quantique*, tome 1, de boeck
- [2] C. Aslangul, *Mécanique quantique*, tome 2, de boeck
- [3] J.L. Basdevant, *12 leçons de mécanique quantique*, Vuibert, 2006
- [4] J.L. Basdevant, J. Dalibard, *Mécanique quantique*, Les éditions de l'école Polytechnique, 2008
- [5] P. Bergmann, *Introduction to the theory of relativity*, Dover, 1976
- [6] F. Besnard, *Introduction à la relativité restreinte et générale*, poly disponible sur <http://pagesperso-orange.fr/fabien.besnard/cours/EPF/relativity.pdf>
- [7] C. Cohen-Tannoudji, B. Diu, F. Laloë, *Mécanique quantique*, Hermann, 1973.
- [8] R. Feynman, *Les cours de physique de Feynman, Tome 3 : Mécanique Quantique*, Dunod, 2003
- [9] W. Greiner, L. Neise, H. Stöcker, *Thermodynamics and statistical mechanics*, Springer, 1995
- [10] L. Marchildon, *Quantum Mechanics*, Springer, 2002
- [11] H. Kragh, *Max Planck, the reluctant revolutionary*, Physicsworld, <http://physics-world.com/cws/article/print/373>
- [12] A. Messiah, *Quantum Mechanics*, Dover, 1999
- [13] R. Omnès, *Comprendre la mécanique quantique*, EDP Sciences, 2000
- [14] A. Pais, *Subtle is the Lord*, Oxford University Press, 2005
- [15] R. Penrose, *The Road to Reality, A complete guide to the laws of the universe*, Jonathan Cape, 2004 (trad. en français : À la découverte de l'univers, Odile Jacob, 2007)
- [16] Planck, La révolution quantique, *Les génies de la science*, No 27, mai-juillet 2006
- [17] J. Schwartz, M. McGuinness, *Introducing Einstein*, Icon Books
- [18] T. Coudreau, P. Milman, *Domestiquer l'intrication quantique*, Pour la Science No 361, Nov. 2007
- [19] L. Smolin, *Three roads to quantum gravity*, Basic Books, 2001
- [20] M. Lachièze-Rey, *Au-delà de l'Espace et du temps, la nouvelle physique*, Le Pommier, 2003