SORBONNE UNIVERSITÉ

Master de physique – Sorbonne Université

Mécanique quantique

Concepts fondamentaux

Kasimir Malevitch, « Supremus n° 56 »

Ces notes de cours sont fondées sur le cours de *concepts fondamentaux de la mécanique quantique* dispensé à Sorbonne Université en première année de master. Ce cours présuppose acquis les bases de la mécanique quantique enseignées en licence, qui ne seront rappelées que brièvement, ainsi que des bases d'analyse et d'algèbre.

Note : les section marquées d'un étoile $\star$ sont d'un niveau mathématique plus élevé et leur étude approfondie n'est pas exigée.

Bibliographie

La lecture d'ouvrages de référence est vivement encouragée. En particulier :

- *Mécanique quantique*, Claude Cohen-Tannoudji, Bernard Diu et Franck Laloë, éditions `Hermann` (tomes I et II) et `EDP sciences` (tome III) ;
- *Modern Quantum mechanics*, Jun John Sakurai, `Cambridge University Press` ;
- *Mécanique quantique*, Jean-Louis Basdevant et Jean Dalibard, `Éditions de l'École polytechnique` ;
- *Mécanique quantique*, Claude Aslangul, `De Boeck`.

Pour les aspects mathématiques, on pourra consulter l'ouvrage suivant :

- *Mathematical Methods in Quantum Mechanics,* Gerald Teschl, Graduate Studies in Mathematics, Volume 157, Amer. Math. Soc., Providence, 2014,

relativement accessible, mis en ligne gracieusement par l'auteur à l'adresse suivante
`https://www.mat.univie.ac.at/ gerald/ftp/book-schroe/`

Dernière mise à jour

Le 6 décembre 2024

Table des matières

1 Espace des états et observables **7**
1.1 Espace des états et formalisme de Dirac 7
1.2 Opérateurs et observables 14
1.3 Valeurs moyennes et mesure 26

Appendices **33**
1.A Séries de Fourier et espace de Hilbert des fonctions de carré sommable 33
1.B Expérience de Stern et Gerlach 34

2 Fonctions d'onde et paquets d'ondes **37**
2.1 L'espace de Hilbert $\mathcal{L}^2(\mathbb{R}^3)$ 37
2.2 Représentation de position 39
2.3 Représentation d'impulsion 44
2.4 Particule dans une boîte, électrons libres dans les solides 49
2.5 Extensions auto-adjointes de l'opérateur impulsion $\star$ 52

Appendices **56**
2.A Transformation de Fourier 56
2.B L'oscillateur harmonique 59

3 Symétries en mécanique quantique **64**
3.1 Mécanique lagrangienne : rappels 65
3.2 Symétries en physique classique 66
3.3 Quelques éléments de théorie des groupes 73
3.4 Symétries en mécanique quantique 83

4 Symétries II : symétries spatiales **88**
4.1 Rotations dans l'espace 88
4.2 Groupes de translations 96
4.3 Parité 98
4.4 Le groupe euclidien 102

5 Systèmes périodiques **104**
5.1 Réseau unidimensionnel . . . 105
5.2 Aspects mathématiques des réseaux . . . 107
5.3 Structures cristallines . . . 111
5.4 Théorème de Bloch . . . 117

6 Produit tensoriel et particules identiques **120**
6.1 Produit tensoriel . . . 120
6.2 Particules identiques . . . 128
6.3 Théorème spin-statistique . . . 136
6.4 Espaces de Fock . . . 136

7 Électrons dans les solides **145**
7.1 Gaz de fermions libres, surface de Fermi . . . 145
7.2 Potentiel périodique et structure de bandes . . . 149

8 Évolution temporelle **156**
8.1 Opérateur d'évolution . . . 156
8.2 Représentations de Schrödinger et Heisenberg . . . 165
8.3 Renversement du temps . . . 170

9 Opérateur densité et entropies **176**
9.1 Opérateur densité et mélange statistique . . . 176
9.2 Entropie d'un système quantique, entropie d'intrication . . . 183

10 Introduction à l'intégrale de chemin **188**
10.1 L'intégrale de chemin . . . 188
10.2 Propagateur d'une particule libre : intégrale de chemin . . . 190
10.3 Intégrale de chemin et limite classique . . . 193
10.4 De l'intégrale de chemin à l'équation de Schrödinger . . . 194

Appendices **196**
10.A Régularisation zêta . . . 196

11 Théorie des perturbations stationnaires **199**
11.1 Position du problème . . . 199
11.2 Perturbations dans le cas non-dégénéré . . . 205
11.3 Perturbations dans le cas dégénéré . . . 213

12 Limite semi-classique : approximation WKB **218**
12.1 Équation de Hamilton–Jacobi . . . 219
12.2 Fonction d'onde WKB . . . 220
12.3 Approximation WKB et intégrale de chemin . . . 223
12.4 Points de rebroussement et effet tunnel . . . 223

13 Moment cinétique I : rotations et spin **228**
13.1 Représentations du groupe des rotations . . . 229
13.2 Groupe $\mathsf{SU}(2)$ et spineurs . . . 232
13.3 Représentations unitaires irreductibles de $\mathsf{SU}(2)$ . . . 238

14 Moment cinétique II : harmoniques sphériques **245**
14.1 Fonctions sur la sphère . . . 245
14.2 Harmoniques sphériques . . . 250
14.3 Propriétés des harmoniques sphériques . . . 255
14.4 Exemples d'harmoniques sphériques . . . 258

15 Potentiel central et atome d'hydrogène **261**
15.1 Hamiltonien à potentiel central . . . 261
15.2 Potentiel coulombien et atome d'hydrogène . . . 266
15.3 Analyse physique des solutions . . . 271

16 Moment cinétique III : composition **276**
16.1 Composition de moments cinétiques : exemples . . . 276
16.2 Théorie générale : coefficients de Clebsch–Gordan . . . 284

Appendices **290**
16.A Coefficients de Clebsch–Gordan et matrices de rotation . . . 290

17 Structure fine et effet Zeeman **292**
17.1 Structure fine de l'atome d'hydrogène . . . 292
17.2 Le théorème de Wigner–Eckart . . . 299
17.3 L'effet Zeeman . . . 306

Appendices **310**
17.A Calcul d'intégrales radiales . . . 310

18 Perturbations dépendantes du temps **312**
18.1 Représentation d'interaction . . . 312
18.2 Théorie des perturbations dépendantes du temps . . . 316
18.3 Règle d'or de Fermi . . . 319

19 Interaction matière-rayonnement **323**
19.1 Hamiltonien d'une particule chargée . . . 323
19.2 Invariance de jauge . . . 325
19.3 Absorption et émission spontanée . . . 330

20 Quantification du champ électromagnétique **336**
20.1 Développement en modes . . . 337
20.2 Quantification canonique . . . 339

20.3 Hélicité du photon 341
20.4 Énergie du vide 342
20.5 Le modèle de Jaynes–Cummings 346

Cours n°1

Espace des états et observables

En mécanique classique hamiltonienne, l'état d'un système à un instant donné correspond à un point de l'*espace des phases* de coordonnées généralisées $\{q^i, p_i\}$; les grandeurs physiques caractérisant le système sont des fonctions $f(q^i, p_i; t)$ de ces coordonnées et éventuellement du temps. Via les équations de Hamilton on peut alors déduire – au moins en principe – l'évolution du système et prédire sont état à tout instant ultérieur.

En mécanique quantique, dans le formalisme de Dirac [1], un système est caractérisé par un *vecteur d'état* duquel on peut tirer les valeurs moyennes d'opérateurs appelés *observables*, liées aux grandeurs physiques caractérisant le système. L'objectif de ce cours est de reformuler les principaux postulats de la mécanique quantique, déjà vus en licence, en développant d'avantage les structures mathématiques associées, reliées aux notions d'espace de Hilbert et d'opérateur auto-adjoint.

1.1 Espace des états et formalisme de Dirac

La structure mathématique de l'espace des états dépend du type de système considéré. On pourrait à première vue distinguer trois situations qualitativement differentes :

a) le système n'a accès qu'à un nombre fini d'états. L'espace des états est de dimension finie ;

b) le système a accès à un nombre infini dénombrable d'états, c.-à.-d. que l'on peut les distinguer par un indice $n \in \mathbb{N}$;

c) le système a accès à un nombre infini indénombrable d'états, typiquement distingués par la valeur d'un nombre réel $\alpha \in \mathbb{R}$.

Le premier cas de figure se rencontre souvent lorsqu'on isole quelques états d'un système pour en obtenir une description simplifiée, que les autres états d'énergie soient situés à des échelles d'énergie très différentes ou que certains degrés de liberté ne jouent pas de rôle important. On peut citer l'*expérience de Stern et Gerlach*, que nous décrirons plus loin, où seuls les degrés de liberté de spin (système à deux états) sont importants, et non les degrés de liberté de translation (continuum d'états).

Le deuxième cas de figure se rencontre pour une particule confinée par un potentiel, comme l'oscillateur harmonique, l'atome hydrogénoïde [1] ou, plus simplement, une particule dans une boîte avec des conditions de bord de Dirichlet ; dans tous les cas le système a accès uniquement à des niveaux d'énergie discrets.

Le troisième cas apparaît par exemple pour une particule libre dans un espace infini, ou en présence d'un potentiel si l'énergie de la particule est telle qu'elle n'est pas, classiquement, confinée par celui-ci, car son énergie peut prendre des valeurs arbitraires dans un intervalle de la forme $[a, +\infty[$. Nous verrons cependant que la différence entre ces deux derniers cas de figure est plus subtile qu'on pourrait le croire.

1.1.1 Espace des états de dimension finie

Si le système n'a accès qu'à un nombre fini d'états, l'espace des états $\mathcal{H}$ possède la structure d'un *espace vectoriel complexe de dimension finie*. Les vecteurs de cet espace sont commodément notés, dans la notation introduite par Paul Dirac, par des « **kets** » de la forme $|\alpha\rangle$, où α est une étiquette désignant un certain état. La structure d'espace vectoriel implique en particulier que

$$\forall\, |\alpha\rangle\,, |\beta\rangle \in \mathcal{H}\ , \quad \forall \lambda,\, \mu \in \mathbb{C}\ , \quad \lambda\,|\alpha\rangle + \mu\,|\beta\rangle \in \mathcal{H}\,. \tag{1.1}$$

L'espace des états possède en outre une structure d'*espace hermitien*, étant muni d'un *produit hermitien*, c.-à.-d. d'une application

$$\iota : \left\{ \begin{array}{lcl} \mathcal{H} \times \mathcal{H} & \to & \mathbb{C} \\ (|\alpha\rangle\,, |\beta\rangle) & \mapsto & \iota(|\alpha\rangle\,, |\beta\rangle) \in \mathbb{C} \end{array} \right. \tag{1.2}$$

qui possède comme propriétés importantes la sesquilinéarité [2]

$$\iota\left(|\alpha\rangle\,, \mu_1\,|\beta_1\rangle + \mu_2\,|\beta_2\rangle\right) = \mu_1\,\iota\left(|\alpha\rangle\,, |\beta_1\rangle\right) + \mu_2\,\iota\left(|\alpha\rangle\,, |\beta_2\rangle\right) \tag{1.3a}$$

$$\iota\left(\lambda_1\,|\alpha_1\rangle + \lambda_2\,|\alpha_2\rangle\,, |\beta\rangle\right) = \bar{\lambda}_1\,\iota\left(|\alpha_1\rangle\,, |\beta\rangle\right) + \bar{\lambda}_2\,\iota\left(|\alpha_2\rangle\,, |\beta\rangle\right) \tag{1.3b}$$

ainsi que

$$\iota(|\alpha\rangle\,, |\beta\rangle) = \overline{\iota(|\beta\rangle\,, |\alpha\rangle)}\,, \tag{1.4a}$$

$$\forall\, |\alpha\rangle \in \mathcal{H}\ , \quad \iota(|\alpha\rangle\,, |\alpha\rangle) \geqslant 0\,. \tag{1.4b}$$

$$\forall\, |\alpha\rangle \in \mathcal{H}\ , \quad \iota(|\alpha\rangle\,, |\alpha\rangle) = 0 \implies |\alpha\rangle = 0\,. \tag{1.4c}$$

Le produit hermitien (1.2) étant défini positif d'après les équations (1.4b,1.4c), il permet de définir une norme sur l'espace des états

$$\|\,|\alpha\rangle\,\| \overset{\text{déf.}}{=} \sqrt{\iota(|\alpha\rangle\,, |\alpha\rangle)} \geqslant 0\,. \tag{1.5}$$

1. Le cas de l'atome est complexe : même si le nombre de niveaux discrets (états liés) est infini, au-delà d'une certaine énergie il existe un continuum d'états asymptotiques.

2. Attention à ne pas confondre avec la bilinéarité du produit scalaire réel ! On notera en particulier, dans l'équation (1.3b), des complexes conjugués des coefficients de la combinaison linéraire $\lambda_1\,|\alpha_1\rangle + \lambda_2\,|\alpha_2\rangle$.

On peut munir cet espace de dimension finie d'une base, et on peut en particulier définir une base orthonormée

$$\left\{ |\ell\rangle \ , \ \ell = 1, \dots, n \,\middle|\, \iota\big(|\ell\rangle \, , |\ell'\rangle \big) = \delta_{\ell,\ell'} \right\} \ , \tag{1.6}$$

où n est la dimension de l'espace. On peut ainsi développer tout vecteur de $\mathcal{H}$ sur cette base :

$$|\alpha\rangle = \sum_{\ell=1}^{n} \alpha_\ell \, |\ell\rangle \ , \quad \alpha_\ell \in \mathbb{C} \, . \tag{1.7}$$

Tout espace hermitien de dimension n est isomorphe à $\mathbb{C}^n$.

Il est très utile de considérer l'*espace dual* $\mathcal{H}^\star$ de l'espace vectoriel $\mathcal{H}$, c.-à.-d. l'espace des formes linéaires φ sur $\mathcal{H}$,

$$\varphi : \left\{ \begin{array}{ccc} \mathcal{H} & \to & \mathbb{C} \\ |\alpha\rangle & \mapsto & \varphi(|\alpha\rangle) \in \mathbb{C} \end{array} \right. \qquad \text{t.q.} \quad \varphi(\lambda \, |\alpha\rangle + \mu \, |\beta\rangle) = \lambda \varphi(|\alpha\rangle) + \mu \varphi(|\beta\rangle) \, . \tag{1.8}$$

Cet espace possède lui-même une structure d'espace vectoriel.

Pour un espace hermitien, il existe un isomorphisme canonique entre $\mathcal{H}$ et son dual $\mathcal{H}^\star$. À tout vecteur $|\alpha\rangle$ on associe une forme linéaire, appelée « **bra** » et notée $\langle\alpha|$ dans le formalisme de Dirac, de la manière suivante :

$$\langle\alpha| : \left\{ \begin{array}{ccc} \mathcal{H} & \to & \mathbb{C} \\ |\beta\rangle & \mapsto & \langle\alpha|\beta\rangle \overset{\text{déf.}}{=} \iota(|\alpha\rangle \, , \, |\beta\rangle) \in \mathbb{C} \end{array} \right. \tag{1.9}$$

Si $\mathcal{H}$ est muni d'une base orthonormée $\{|\ell\rangle \ , \ell = 1, \dots, n\}$, nous avons :

$$\langle\ell| : \left\{ \begin{array}{ccc} \mathcal{H} & \to & \mathbb{C} \\ |\alpha\rangle & \mapsto & \langle\ell|\alpha\rangle \overset{\text{déf.}}{=} \iota(|\ell\rangle \, , \, |\alpha\rangle) \in \mathbb{C} \end{array} \right. \tag{1.10}$$

et l'ensemble $\{\langle\ell| \, , \, \ell = 1, \dots, n\}$ forme une base orthonormée de $\mathcal{H}^\star$, appelée *base duale*, qui vérifie la propriété évidente :

$$\langle\ell|r\rangle = \delta_{\ell,r} \, . \tag{1.11}$$

Cette notion permet en particulier de noter le produit hermitien de manière compacte (notation « bra-ket ») :

$$\langle\alpha|\beta\rangle \overset{\text{déf.}}{=} \iota(|\alpha\rangle \, , \, |\beta\rangle) \, . \tag{1.12}$$

Redonnons dans cette notation l'importante propriété (1.4a) :

$$\boxed{\langle\beta|\alpha\rangle = \overline{\langle\alpha|\beta\rangle}} \tag{1.13}$$

La propriété de sesquilinéarité (1.3) du produit hermitien implique finalement la correspondance suivante entre une combinaisons linéaires de vecteurs et son dual, soit entre un *ket* et le *bra* associé :

$$\lambda_1 \, |\alpha_1\rangle + \lambda_2 \, |\alpha_2\rangle \ \longleftrightarrow \ \bar{\lambda}_1 \, \langle\alpha_1| + \bar{\lambda}_2 \, \langle\alpha_2| \, . \tag{1.14}$$

On notera, de manière cruciale, la conjugaison complexe des coefficients λ_1 et λ_2. Plus généralement, la décomposition sur une base orthonormée (1.7) donne la décomposition suivante du produit hermitien :

$$\langle\beta|\alpha\rangle = \iota\left(\sum_{\ell=1}^{n}\beta_\ell\,|\ell\rangle\,,\sum_{r=1}^{n}\alpha_r\,|r\rangle\right) = \sum_{\ell=1}^{n}\sum_{r=1}^{n}\bar{\beta}_\ell\alpha_r\,\langle\ell|r\rangle = \sum_{\ell=1}^{n}\bar{\beta}_\ell\alpha_\ell\,. \tag{1.15}$$

1.1.2 Espace des états de dimension infinie

Lorsque l'espace des états n'est plus de dimension finie, c.-à.-d. lorsque le système a accès à un nombre infini d'états distincts, la description de l'espace des états devient plus complexe ; nous nous bornerons ici à mettre en avant les principes importants et leurs conséquences physiques.

La structure mathématique pertinente pour traiter le problème est celle d'*espace de Hilbert*, d'après les travaux du mathématicien allemand David Hilbert (1862-1943). Un espace de Hilbert est une généralisation en dimension infinie d'un espace hermitien, lui-même généralisation complexe d'un espace euclidien.

Définition 1 (espace de Hilbert). *Un espace de Hilbert $\mathcal{H}$ est un espace vectoriel complexe muni d'un produit hermitien $\iota : \mathcal{H}\times\mathcal{H}\to\mathbb{C}$, satisfaisant aux propriétés (1.3) et (1.4). Cet espace est en outre un espace métrique complet.*

La notion d'espace métrique complet est un peu plus technique. Premièrement, l'espace de Hilbert est un espace métrique car on peut y définir une notion de distance entre deux éléments,

$$d(|\alpha\rangle\,,|\beta\rangle) \overset{\text{déf.}}{=} \|\,|\alpha\rangle - |\beta\rangle\,\| = \sqrt{\iota(|\alpha\rangle - |\beta\rangle\,,|\alpha\rangle - |\beta\rangle)}. \tag{1.16}$$

Un espace métrique $\mathcal{H}$ est dit *complet* si toute suite de Cauchy [3] converge dans $\mathcal{H}$. Comme son nom l'indique, un espace complet est un espace dans lequel il n'y a pas de « vides ». Un contre-exemple est fourni par les nombres rationnels, car on peut construire une suite de Cauchy de nombres rationnels convergeant vers $\sqrt{2}$ qui n'est pas rationnel.

Cette notion mathématique de complétude est essentielle pour que les principes de la mécanique quantique s'appliquent correctement. Sans cela, il n'y aurait aucune garantie que l'évolution dans le temps d'un état physique correspondant à un vecteur d'état $|\alpha(0)\rangle\in\mathcal{H}$ au début de l'expérience donne au bout d'un temps t un autre vecteur d'état $|\alpha(t)\rangle$ appartenant au même espace des états $\mathcal{H}$, suivant les règles d'évolution temporelle que nous étudierons au chapitre 8.

Sur l'espace de Hilbert muni de la norme $\|\ \|$ issue du produit hermitien on peut démontrer l'*inégalité de Cauchy–Schwartz* :

3. Rappel d'analyse : une suite $\{u_n; n\in\mathbb{N}\}$ dans un espace métrique est dite de Cauchy si

$$\forall\,\epsilon>0,\ \exists N\in\mathbb{N}\ \text{t.q.}\ \forall r,s>N,\ \|\,u_r-u_s\,\|<\epsilon.$$

Intuitivement, cela signifie que pour n grand les termes de la suite se rapprochent les uns des autres.

Théorème 1 (inégalité de Cauchy–Schwartz).

$$\forall\, |\alpha\rangle\,, |\beta\rangle \in \mathcal{H}\ , \quad |\iota(|\alpha\rangle\,, |\beta\rangle)| \leqslant \| \, |\alpha\rangle \, \| \, \| \, |\beta\rangle \, \|\ , \tag{1.17}$$

et l'égalité est satisfaite si et seulement si $|\alpha\rangle$ *et* $|\beta\rangle$ *sont colinéaires.*

Espace dual $\star$

Comme dans le cas d'un espace des états de dimension finie, on peut définir un *espace dual* $\mathcal{H}^\star$ à un espace de Hilbert $\mathcal{H}$, comme l'espace des applications linéaires $\mathsf{f} : \mathcal{H} \to \mathbb{C}$ bornées, c.-à.-d. telles que

$$|\mathsf{f}| \overset{\text{déf.}}{=} \sup_{|\psi\rangle \in \mathcal{H},\, \| \, |\psi\rangle \, \| = 1} |\mathsf{f}(|\psi\rangle)| < \infty\,. \tag{1.18}$$

On peut montrer qu'une application linéaire est bornée si et seulement si elle est continue.

La relation entre les éléments de $\mathcal{H}$ et de $\mathcal{H}^\star$ est cependant plus délicate à établir. On peut démontrer le *thèorème de représentation de Riesz*, qui s'énonce comme suit :

Théorème 2 (théorème de Riesz). *Soit* $\mathcal{H}$ *un espace de Hilbert muni d'un produit scalaire* ι *et une forme linéaire* $\mathsf{f} : \mathcal{H} \to \mathbb{C}$ *bornée. Il existe un unique vecteur* $|\beta\rangle \in \mathcal{H}$ *tel que, pour tout vecteur* $|\alpha\rangle \in \mathcal{H}$*,* $\mathsf{f}(|\alpha\rangle)$ *est donnée par le produit scalaire de* $|\beta\rangle$ *avec* $|\alpha\rangle$ *:*

$$\mathsf{f}(|\alpha\rangle) = \iota(|\beta\rangle\,, |\alpha\rangle)\,.$$

Suivant la notation de Dirac, on associera alors au vecteur ou « ket » $|\beta\rangle$ définissant la forme linéaire $\mathsf{f} \in \mathcal{H}^\star$ un « bra » $\langle\beta|$ tel que :

$$\langle\beta|\alpha\rangle \overset{\text{déf.}}{=} \iota(|\beta\rangle\,, |\alpha\rangle) = \mathsf{f}(|\alpha\rangle)\,. \tag{1.19}$$

Tout comme pour un espace des états de dimension finie, on peut donc associer à chaque *ket* (resp. à chaque combinaison linéaire de kets), un unique *bra* (resp. une combinaison linéaire de bras) par un isomorphisme anti-linéaire :

$$\lambda_1\, |\alpha_1\rangle + \lambda_2\, |\alpha_2\rangle \ \longleftrightarrow\ \bar\lambda_1\, \langle\alpha_1| + \bar\lambda_2\, \langle\alpha_2|\ . \tag{1.20}$$

Cette correspondance entre kets et bras peut être également notée en utilisant le *conjugaison hermitienne* $\dagger$, soit :

$$\big(\, |\alpha\rangle \,\big)^\dagger \overset{\text{déf.}}{=} \langle\alpha|\ , \quad \big(\, \langle\alpha| \,\big)^\dagger \overset{\text{déf.}}{=} |\alpha\rangle\ . \tag{1.21}$$

L'espace $\mathcal{H}^\star$ engendré par les bras est lui-même un espace de Hilbert, dont le produit scalaire $\iota^\star$ est défini par :

$$\iota^\star(\langle\alpha|\,, \langle\beta|) = \overline{\iota(|\alpha\rangle\,, |\beta\rangle)} = \overline{\langle\alpha|\beta\rangle} = \langle\beta|\alpha\rangle\ , \tag{1.22}$$

conformément à la propriété (1.20).

Bases de l'espace des états

La notion de base est là encore un peu plus complexe pour un espace de Hilbert que pour un espace vectoriel ordinaire en raison de leur dimension possiblement infinie. Un exemple typique est fourni par le développement en série de Fourier des fonctions de carré sommable sur un intervalle ; les principes et résultats principaux sont résumés à la fin de ce cours dans l'appendice 1.A.

De manière générale, on peut définir dans un espace de Hilbert $\mathcal{H}$ une famille orthonormée $\{|\ell\rangle \in \mathcal{H},\, \ell \in \mathcal{I}\}$ où $\mathcal{I}$ est un ensemble, telle que

$$\forall \ell, \ell' \in \mathcal{I}\,,\ \langle \ell | \ell' \rangle = 0 \quad \text{si} \quad \ell \neq \ell'\,, \tag{1.23a}$$

$$\forall \ell \in \mathcal{I}\,,\ \| \, |\ell\rangle \, \| = 1\,. \tag{1.23b}$$

Pour que cette famille orthonormée forme une base orthonormée de $\mathcal{H}$, il faut de plus que l'ensemble des combinaisons linéaires des vecteurs de la famille $\{|\ell\rangle \in \mathcal{H},\, \ell \in \mathcal{I}\}$ soit *dense* dans $\mathcal{H}$, c.-à.-d. que tout voisinage d'un élément quelconque de $\mathcal{H}$ contienne au moins une combinaison linéaire des vecteurs $|\ell\rangle$. On peut montrer la densité de cette famille dans $\mathcal{H}$ est équivalente à la proposition suivante :

$$\forall \ell \in \mathcal{I}\,,\ \langle \ell | \nu \rangle = 0 \implies |\nu\rangle = 0_{\mathcal{H}}\,, \tag{1.24}$$

soit que le complément orthogonal du sous-espace engendré par les vecteurs de base $|\ell\rangle$ ne contient que le vecteur nul $0_{\mathcal{H}}$.

Pour un espace de Hilbert quelconque, on peut toujours définir une base orthonormée, appelée base de Hilbert ; lorsque cette base est dénombrable, c.-à.-d. lorsque l'ensemble $\mathcal{I}$ est dénombrable, l'espace de Hilbert est dit *séparable*. [4]

Dans ce cas, on peut indexer la base comme $\{|n\rangle\,, n \in \mathcal{I} = \{1, 2, \ldots\} \subset \mathbb{N}\}$ et, pour tout $|\nu\rangle \in \mathcal{H}$, l'égalité entre $|\nu\rangle$ et son développement sur la base orthonormée signifie que

$$|\nu\rangle = \sum_{n \in \mathcal{I}} \langle n | \nu \rangle \, |n\rangle \ \Leftrightarrow \lim_{N \to \infty} \left\| \, |\nu\rangle - \sum_{n=1}^{N} \langle n | \nu \rangle \, |n\rangle \, \right\| = 0\,, \tag{1.25}$$

soit que $|\nu\rangle$ est « infiniment proche » de son développement sur la base.

C'est dans ce sens qu'il faut comprendre le développement sur une base dans un espace de Hilbert ; le membre de gauche de (1.25) définissant une série, il s'agit d'une question de convergence du même type que celle concernant les séries de fonctions en analyse. [5]

Pour toute famille orthonormée d'un espace de Hilbert $\mathcal{H}$ on peut également démontrer l'inégalité de Bessel :

$$\forall\, |\nu\rangle \in \mathcal{H}\ , \sum_{\ell \in \mathcal{I}} |\langle \ell | \nu \rangle|^2 \leqslant \langle \nu | \nu \rangle\ , \tag{1.26}$$

4. Cette propriété est indépendante du choix de la base ; deux bases orthonormées quelconques d'un même espace de Hilbert ont la même cardinalité, c.-à.-d. qu'elles sont toutes les deux soit finies (et de même dimension), soit dénombrables, soit non dénombrables.

5. Lorsque la base est finie, l'égalité entre $|\nu\rangle$ et son développement peut évidemment s'exprimer sans prendre de limite.

qui devient, lorsque cette famille forme une base de Hilbert l'égalité de Parceval

$$\forall\, |\nu\rangle \in \mathcal{H}\ , \quad \langle\nu|\nu\rangle = \sum_{\ell\in\mathcal{I}} |\langle\ell|\nu\rangle|^2\ , \tag{1.27}$$

qui joue un rôle important en mécanique quantique.

D'après le théorème 2 on peut associer à tout vecteur $|\ell\rangle$ d'une base orthonormée de l'espace de Hilbert $\mathcal{H}$ un vecteur de base $\langle\ell| \in \mathcal{H}^\star$ de l'espace dual. La définition (1.22) nous indique que ces vecteurs sont orthonormés. On montre qu'ils forment une base de $\mathcal{H}^\star$, et donc que l'espace de Hilbert et son dual ont la même cardinalité (en particulier la même dimension si celle-ci est finie).

1.1.3 Premier postulat de la mécanique quantique

Ces propriétés mathématiques étant établies, le rôle central de l'espace des états en mécanique quantique est du au postulat suivant :

Postulat 1 (espace des états). *L'état d'un système quantique donné est entièrement déterminé par un vecteur d'état $|\alpha\rangle$ appartenant à l'espace des états $\mathcal{H}$ du système, possédant la structure mathématique d'un espace de Hilbert séparable.*

L'hypothèse de séparabilité de $\mathcal{H}$ assure qu'un ensemble dénombrable d'expériences est suffisant pour déterminer l'état d'un système physique.

Ce postulat donne la structure mathématique de l'espace des états mais n'en donne pas une description explicite. Quelle est l'information nécessaire pour caractériser l'état d'un système quantique ? Intuitivement, deux états physiques peuvent être considérés comme distincts s'il existe des grandeurs physiques associées prenant des valeurs distinctes ; une prescription précise sera donnée au chapitre 1.2.

Espace projectif des états physiques

La correspondance entre l'état d'un système quantique et un vecteur d'un espace de Hilbert, l'espace des états $\mathcal{H}$, n'est pas bijective ; deux vecteurs de $\mathcal{H}$ correspondent au même état physique du système s'ils sont proportionnels entre eux (dans le corps des complexes).

Pour décrire sans redondance les états d'un système on peut ainsi définir une relation d'équivalence sur $\mathcal{H}$:

$$|\nu\rangle \sim |\nu'\rangle \ \text{ si } \quad \exists\lambda \in \mathbb{C}^\star\ , \quad |\nu'\rangle = \lambda\,|\nu\rangle\ , \tag{1.28}$$

Le quotient de l'espace des états par cette relation d'équivalence définit un *espace projectif*,[6] noté $\mathbb{P}[\mathcal{H}]$, dont les éléments sont appelés *rayons*. Chaque rayon est associé à un état du système quantique considéré.

6. Lorsque l'espace des états est de dimension finie l'espace associé est isomorphe à l'espace projectif complexe $\mathbb{CP}^n$. Par exemple, l'espace projectif $\mathbb{CP}^1$, sous-espace de $\mathbb{C}^2$ défini par la relation $(z_1, z_2) \sim \lambda(z_1, z_2)$ avec $\lambda \in \mathbb{C}^\star$, n'est rien d'autre que la sphère ordinaire S^2 dans un espace à trois dimensions.

Pour simplifier les notations on choisit généralement pour chaque classe d'équivalence un représentant de norme unité par rapport à la norme (1.5) :

$$|\alpha\rangle \longrightarrow |\alpha\rangle_{\text{N}} \overset{\text{déf.}}{=} \frac{1}{\sqrt{\langle\alpha|\alpha\rangle}}\,|\alpha\rangle \;, \quad \forall\, |\alpha\rangle \neq 0_{\mathcal{H}}\,. \tag{1.29}$$

Il subsiste cependant une ambiguïté, liée à la phase globale arbitraire du vecteur normé ainsi défini. Cette phase n'a pas d'impact, comme il se doit, sur la valeur des grandeurs physiques et peut être choisie arbitrairement pour simplifier les expressions.

Remarque 1 *En raison de l'hypothèse de séparabilité de l'espace de Hilbert en mécanique quantique, il n'existe pas de notion de « base continue ». De tels « vecteurs de base » peuvent être employés par commodité comme intermédiaires de calcul, voir le chapitre 2, mais ne peuvent en aucun cas être associés à des états physiques du système.*

1.2 Opérateurs et observables

Aux grandeurs physiques caractérisant un système, telles la quantité de mouvement, le moment cinétique ou l'énergie, sont associées en mécanique quantique des *opérateurs*, qui sont des applications linéaires de l'espace des états dans lui-même :

$$\mathcal{O} : \left\{ \begin{array}{ccc} \mathsf{D}(\mathcal{O}) \subset \mathcal{H} & \to & \mathcal{H} \\ |\alpha\rangle & \mapsto & \mathcal{O}\,|\alpha\rangle \end{array} \right. , \quad \mathcal{O}\left(\lambda_1\,|\alpha_1\rangle + \lambda_2\,|\alpha_2\rangle\right) = \lambda_1 \mathcal{O}\,|\alpha_1\rangle + \lambda_2 \mathcal{O}\,|\alpha_2\rangle \;, \tag{1.30}$$

où $\mathsf{D}(\mathcal{O})$ est le domaine de définition de l'opérateur $\mathcal{O}$. On considère généralement des opérateurs pour lesquels $\mathsf{D}(\mathcal{O})$ est dense dans $\mathcal{H}$, afin qu'ils soient définis pour « la plupart » des états du système. On montre alors le théorème suivant :

Théorème 3 *Soit une application linéaire* $\mathcal{O} : \mathsf{D}(\mathcal{O}) \subset \mathcal{H} \to \mathcal{H}$. *Si* $\mathsf{D}(\mathcal{O})$ *est dense dans* $\mathcal{H}$ *et si* $\mathcal{O}$ *est bornée, c.-à.-d.*

$$\|\mathcal{O}\| \overset{\text{déf.}}{=} \sup_{|\psi\rangle\in\mathcal{H},\, \||\psi\rangle\|=1} \|\,\mathcal{O}\,|\psi\rangle)\,\| < \infty\,, \tag{1.31}$$

alors il existe une unique extension (continue) de $\mathcal{O}$ *en une application* $\mathcal{H} \to \mathcal{H}$. *L'ensemble de ces applications bornées et linéaires sur* $\mathcal{H}$ *sera noté* $\mathfrak{L}(\mathcal{H})$.

Essentiellement, les opérateurs bornés sur un espace de Hilbert séparable se comportent comme des opérateurs agissant sur un espace de dimension finie ; les difficultés viennent avec les opérateurs non bornés, qui sont monnaie courante en mécanique quantique.

La composition d'actions successives d'opérateurs conduit naturellement à la notion de produit d'opérateurs :

$$\forall\,|\alpha\rangle \in \mathsf{D}(\mathcal{P}) \;\text{ t.q. }\; \mathcal{P}\,|\alpha\rangle \in \mathsf{D}(\mathcal{O}) \;, \quad \big(\mathcal{O}\,\mathcal{P}\big)\,|\alpha\rangle \overset{\text{déf.}}{=} \mathcal{O}\big(\mathcal{P}\,|\alpha\rangle\,\big) \tag{1.32}$$

Ce produit est associatif et distributif, c.-à.-d.

$$\forall |\alpha\rangle \in \mathcal{H} \ , \quad \mathcal{O}\mathcal{P}\mathcal{Q}\,|\alpha\rangle = \big(\mathcal{O}\mathcal{P}\big)\mathcal{Q}\,|\alpha\rangle = \mathcal{O}\big(\mathcal{P}\mathcal{Q}\big)\,|\alpha\rangle \ , \tag{1.33a}$$

$$\mathcal{O}\left(\mathcal{P}+\mathcal{Q}\right)|\alpha\rangle = \mathcal{O}\mathcal{P}\,|\alpha\rangle + \mathcal{O}\mathcal{Q}\,|\alpha\rangle \ . \tag{1.33b}$$

En revanche, la définition (1.32) indique que le produit de deux opérateurs quelconque n'est pas commutatif en général. L'absence de commutativité entre deux opérateurs donnés est mesurée par le *commutateur*, defini comme :

$$\boxed{[\,\mathcal{O},\,\mathcal{P}\,] \overset{\text{déf.}}{=} \mathcal{O}\mathcal{P} - \mathcal{P}\mathcal{O}} \tag{1.34}$$

Le commutateur peut être vu comme un produit intérieur sur l'espace des opérateurs qui possède plusieurs propriétés importantes [7] :

$$\text{Antisymétrie :} \quad [\,\mathcal{P},\,\mathcal{O}\,] = -\,[\,\mathcal{O},\,\mathcal{P}\,] \ , \tag{1.35a}$$

$$\text{Bilinéarité :} \quad [\,\lambda_1\mathcal{O}_1 + \lambda_2\mathcal{O}_2,\,\mathcal{P}\,] = \lambda_1\,[\,\mathcal{O}_1,\,\mathcal{P}\,] + \lambda_2\,[\,\mathcal{O}_2,\,\mathcal{P}\,] \ , \tag{1.35b}$$

$$\text{Identité de Jacobi :} \quad [\,\mathcal{O},\,[\,\mathcal{P},\,\mathcal{Q}\,]\,] + [\,\mathcal{Q},\,[\,\mathcal{O},\,\mathcal{P}\,]\,] + [\,\mathcal{P},\,[\,\mathcal{Q},\,\mathcal{O}\,]\,] = 0 \ . \tag{1.35c}$$

D'autres identités sont utiles pour effectuer des calculs. On peut citer par exemple

$$[\,\mathcal{O},\,\mathcal{P}\mathcal{Q}\,] = [\,\mathcal{O},\,\mathcal{P}\,]\,\mathcal{Q} + \mathcal{P}\,[\,\mathcal{O},\,\mathcal{Q}\,] \ . \tag{1.36}$$

1.2.1 Opérateurs hermitiens et auto-adjoints ⋆

Étant donné qu'à tout vecteur ou *ket* $|\alpha\rangle$, par le théorème 2, on peut associer une forme linéaire ou *bra* $\langle\alpha|$, on peut définir le *conjugué hermitien* ou *adjoint* $\mathcal{O}^\dagger$ d'un opérateur $\mathcal{O} \in \mathfrak{L}(\mathcal{H})$, donc borné, en utilisant la conjugaison hermitienne d'un vecteur, voir équation (1.21) :

$$\mathcal{O}\,|\alpha\rangle \ \in \mathcal{H} \ \longleftrightarrow \ \langle\alpha|\,\mathcal{O}^\dagger \overset{\text{déf.}}{=} (\mathcal{O}\,|\alpha\rangle)^\dagger \ \in \mathcal{H}^\star \ . \tag{1.37}$$

Une autre manière équivalente de définir l'adjoint d'un opérateur est via le produit scalaire hermitien entre deux vecteurs :

$$\forall |\alpha\rangle\,, |\beta\rangle \in \mathcal{H}\,, \ \iota\left(|\alpha\rangle\,, \mathcal{O}\,|\beta\rangle\right) = \iota\left(\mathcal{O}^\dagger\,|\alpha\rangle\,, |\beta\rangle\right) \ . \tag{1.38}$$

D'après la définition précédente on montre aisément les propriétés suivantes :

$$\left(\mathcal{O}^\dagger\right)^\dagger = \mathcal{O}\,, \tag{1.39a}$$

$$(\lambda\mathcal{O} + \mu\mathcal{P})^\dagger = \bar{\lambda}\mathcal{O}^\dagger + \bar{\mu}\mathcal{P}^\dagger\,, \tag{1.39b}$$

$$(\mathcal{O}\mathcal{P})^\dagger = \mathcal{P}^\dagger\mathcal{O}^\dagger\,. \tag{1.39c}$$

7. Ces propriétés sont partagées par les crochets de Poisson, défini par l'éqn. (3.38), dont il sont le pendant quantique. Les lecteurs portés sur les mathématiques feront également le parallèle avec le crochet de Lie.

Un opérateur $\mathcal{O} \in \mathfrak{L}(\mathcal{H})$ est dit *auto-adjoint* si

$$\mathcal{O}^\dagger = \mathcal{O}\,. \tag{1.40}$$

La définition d'un opérateur auto-adjoint non borné demande plus de soin ; il ne s'agit nullement d'une question académique, par exemple l'opérateur hamiltonien est en général non borné ! On définit tout d'abord un opérateur *hermitien* ou *symétrique* :

Définition 2 (opérateur hermitien). *Soit $\mathcal{O}$ un opérateur dont le domaine de définition $\mathsf{D}(\mathcal{O})$ est dense dans $\mathcal{H}$. Cet opérateur est dit hermitien ou symétrique si*

$$\forall\,|\alpha\rangle\,,\,|\beta\rangle \in \mathsf{D}(\mathcal{O})\;,\quad \iota(|\alpha\rangle\,,\mathcal{O}\,|\beta\rangle) = \iota(\mathcal{O}\,|\alpha\rangle\,,|\beta\rangle)\,. \tag{1.41}$$

On montre qu'un opérateur est hermitien si et seulement si la forme quadratique associée

$$\mathsf{q}_\mathcal{O} : \left\{ \begin{array}{lcl} \mathsf{D}(\mathcal{O}) & \to & \mathbb{C} \\ |\alpha\rangle & \mapsto & \mathsf{q}_\mathcal{O}(|\alpha\rangle) = \iota\,(|\alpha\rangle\,,\mathcal{O}\,|\alpha\rangle) \end{array} \right. \tag{1.42}$$

est à valeurs dans $\mathbb{R}$ (la réciproque est évidente !). Définissons maintenant l'adjoint $\mathcal{O}^\dagger$ d'un opérateur $\mathcal{O}$ dont le domaine de définition $\mathsf{D}(\mathcal{O})$ est dense dans $\mathcal{H}$. Le domaine de définition de $\mathcal{O}^\dagger$ constituant le point le plus délicat, nous commencerons par déterminer celui-ci.

Définition 3 (opérateur auto-adjoint). *Soit un opérateur $\mathcal{O}$ dont le domaine de définition $\mathsf{D}(\mathcal{O})$ est dense dans $\mathcal{H}$. Le domaine de définition l'opérateur adjoint $\mathcal{O}^\dagger$ est donné par*

$$\mathsf{D}(\mathcal{O}^\dagger) = \big\{\, |\alpha\rangle \in \mathcal{H} \;\big|\; \exists\,|\beta\rangle \in \mathcal{H}\,,\ \forall\,|\gamma\rangle \in \mathsf{D}(\mathcal{O})\,,\ \iota(|\alpha\rangle\,,\mathcal{O}\,|\gamma\rangle) = \iota(|\beta\rangle\,,|\gamma\rangle)\,,\ \big\}\,. \tag{1.43}$$

On définit alors l'adjoint de $\mathcal{O}$ par

$$\forall\,|\alpha\rangle \in \mathsf{D}(\mathcal{O}^\dagger)\;,\quad \mathcal{O}^\dagger\,|\alpha\rangle = |\beta\rangle\;, \tag{1.44}$$

où le vecteur $|\beta\rangle \in \mathcal{H}$ est défini comme ci-dessus, éqn. (1.43). Un opérateur $\mathcal{O}$ est dit auto-adjoint s'il est hermitien et si $\mathsf{D}(\mathcal{O}) = \mathsf{D}(\mathcal{O}^\dagger)$.

La différence entre un opérateur auto-adjoint et un opérateur hermitien est la suivante. Premièrement, si un opérateur $\mathcal{O}$ est hermitien, comme

$$\forall\,|\alpha\rangle\,,\,|\gamma\rangle \in \mathsf{D}(\mathcal{O})\,,\ \iota(|\alpha\rangle\,,\mathcal{O}\,|\gamma\rangle) = \iota(\mathcal{O}\,|\alpha\rangle\,,|\gamma\rangle)\,, \tag{1.45}$$

on a $\mathsf{D}(\mathcal{O}) \subseteq \mathsf{D}(\mathcal{O}^\dagger)$ – en identifiant $\mathcal{O}\,|\alpha\rangle$ avec $|\beta\rangle$ dans (1.43) – et deuxièmement, d'après (1.44),

$$\forall\,|\alpha\rangle \in \mathsf{D}(\mathcal{O})\,,\ \mathcal{O}^\dagger\,|\alpha\rangle = \mathcal{O}\,|\alpha\rangle\;. \tag{1.46}$$

Cet opérateur est *auto-adjoint* si et seulement si $\mathsf{D}(\mathcal{O}^\dagger) = \mathsf{D}(\mathcal{O})$.

Remarque 2 *Pour qu'un opérateur $\mathcal{O}$ hermitien soit auto-adjoint il faut et il suffit qu'il ait le même domaine de définition que son adjoint $\mathcal{O}^\dagger$.*

Propriété 1 *Soient $\mathcal{O}$ et $\mathcal{P}$ deux opérateurs linéraires sur $\mathcal{H}$, tels que $\mathsf{D}(\mathcal{O}) \subseteq \mathsf{D}(\mathcal{P})$ et $\forall\, |\alpha\rangle \in \mathsf{D}(\mathcal{O})$, $\mathcal{O}\,|\alpha\rangle = \mathcal{P}\,|\alpha\rangle$. On a alors*

$$\mathsf{D}(\mathcal{P}^\dagger) \subseteq \mathsf{D}(\mathcal{O}^\dagger) \quad \textit{et} \quad \forall\, |\alpha\rangle \in \mathsf{D}(\mathcal{P}^\dagger)\,,\ \mathcal{O}^\dagger\,|\alpha\rangle = \mathcal{P}^\dagger\,|\alpha\rangle\ . \tag{1.47}$$

La démonstration est aisée en revenant à la définition (1.43). Soit $|\alpha\rangle \in \mathsf{D}(\mathcal{P}^\dagger)$. On peut trouver $|\beta\rangle \in \mathcal{H}$ tel que, pour tout $|\gamma\rangle \in \mathsf{D}(\mathcal{P})$, $\iota(|\alpha\rangle\,, \mathcal{P}\,|\gamma\rangle) = \iota(|\beta\rangle\,, |\gamma\rangle)$. Cela est vrai en particulier si $|\gamma\rangle \in \mathsf{D}(\mathcal{O})$ et on a $\mathcal{P}\,|\gamma\rangle = \mathcal{O}\,|\gamma\rangle$ donc $|\alpha\rangle$ appartient également à $\mathsf{D}(\mathcal{O}^\dagger)$.

En conséquence, étendre le domaine de définition d'un opérateur $\mathcal{O}$ aura pour conséquence de restreindre celui de son adjoint $\mathcal{O}^\dagger$. Un opérateur auto-adjoint, pour lequel $\mathsf{D}(\mathcal{O}) = \mathsf{D}(\mathcal{O}^\dagger)$, est un cas « optimum » et il n'y a pas d'intérêt à essayer d'étendre son domaine de définition.

On peut définir la *fermeture* d'un opérateur par $\mathcal{O}^\mathsf{f} \overset{\text{déf.}}{=} \left(\mathcal{O}^\dagger\right)^\dagger$. Le résultat précédent nous indique que

$$\mathsf{D}(\mathcal{O}) \subseteq \mathsf{D}(\mathcal{O}^\mathsf{f}) \subseteq \mathsf{D}(\mathcal{O}^\dagger)\,, \tag{1.48}$$

et $\mathcal{O}^\mathsf{f}$, qui coïncide avec $\mathcal{O}$ dans $\mathsf{D}(\mathcal{O})$, est une extension de $\mathcal{O}$.

Extension auto-adjointe d'un opérateur hermitien

Soit $\mathcal{O}$ un opérateur hermitien, qui ne serait pas auto-adjoint car $\mathsf{D}(\mathcal{O}) \subsetneqq \mathsf{D}(\mathcal{O}^\dagger)$. Afin de définir une observable de la mécanique quantique (voir 1.2.4 plus loin), on peut chercher à élargir son domaine de définition afin d'obtenir un opérateur auto-adjoint ; on obtient alors une *extension auto-adjointe* de l'opérateur $\mathcal{O}$. Deux cas de figure peuvent se présenter :

— la fermeture de $\mathcal{O}$ est un opérateur auto-adjoint (ce qui veut dire que $\mathcal{O}^\mathsf{f}$ et $\mathcal{O}^\dagger$ ont même domaine de définition). Dans ce cas $\mathcal{O}^\mathsf{f}$ est l'unique extension auto-adjointe de $\mathcal{O}$;

— si $\mathcal{O}^\mathsf{f}$ n'est pas auto-adjoint, se pose la question de l'existence d'une extension auto-adjointe et de son unicité.

Le deuxième cas de figure est naturellement le plus délicat à traiter. Contentons-nous de mentionner un cas particulier, très important en pratique. Soit $\mathcal{O}$ un opérateur hermitien borné inférieurement, c.-à.-d. tel qu'il existe $\mu \in \mathbb{R}$ vérifiant :

$$\forall\, |\alpha\rangle \in \mathsf{D}(\mathcal{O})\,,\ \iota(|\alpha\rangle\,, \mathcal{O}\,|\alpha\rangle) \geqslant \mu \,\||\alpha\rangle\|^2\ . \tag{1.49}$$

Alors $\mathcal{O}^\dagger$ est une extension auto-adjointe de $\mathcal{O}$, qui est essentiellement unique. Dans le cadre de la mécanique quantique, ce résultat s'applique au hamiltonien d'un système dont le spectre en énergie est borné inférieurement.

Ces questions d'extensions auto-adjointes ne sont pas réservées à des systèmes ou opérateurs « exotiques » mais se rencontrent dans les cas les plus courants, du moment que l'espace des états n'est pas de dimension finie. Des exemples seront présentés dans le chapitre 2 et illustreront ces notions de manière moins abstraite.

Exemple 1 (opérateur multiplication). On considère $\mathcal{L}^2(\mathbb{R})$, l'espace de fonctions de carré sommable sur $\mathbb{R}$, sur lequel agit l'opérateur dit *opérateur multiplication* :

$$\mathsf{M}_u\ :\ \mathsf{f} \in \mathcal{L}^2(\mathbb{R})\ \mapsto\ u\,\mathsf{f}\,, \tag{1.50}$$

où u est une fonction donnée $x \in \mathbb{R} \mapsto u(x) \in \mathbb{C}$. L'ensemble de définition de cet opérateur est

$$D(M_u) = \{f \in \mathcal{L}^2(\mathbb{R}) | u\, f \in \mathcal{L}^2(\mathbb{R})\}\,. \tag{1.51}$$

Pour obtenir l'opérateur adjoint on écrit, pour $f, g \in D(M_u)$,

$$\iota(g, M_u f) = \int dx\, \overline{g(x)} u(x) f(x) = \int dx\, \overline{\overline{u(x)} g(x)} f(x) = \iota(M_u^\dagger g, f)\,. \tag{1.52}$$

Ainsi l'opérateur adjoint est défini par [8]

$$M_u^\dagger :\ f \ \mapsto\ \overline{u} f\,,\ D(M_u^\dagger) = \{f \in \mathcal{L}^2(\mathbb{R}) | \overline{u}\, f \in \mathcal{L}^2(\mathbb{R})\} \tag{1.53}$$

et $D(M_u^\dagger) = D(M_u)$. En particulier, l'opérateur M_u est auto-adjoint si la fonction associée $x \mapsto u(x)$ est à valeurs dans $\mathbb{R}$.

1.2.2 Notation de Dirac et opérateurs

Dans le formalisme de Dirac, les composantes d'un opérateur $\mathcal{O}$ agissant sur l'espace des états s'expriment comme :

$$\mathcal{O}_{\alpha\beta} \overset{\text{déf.}}{=} \langle\alpha|\, \mathcal{O}\, |\beta\rangle\ . \tag{1.54}$$

Ils peuvent être calculés de deux manières équivalentes d'après l'équation (1.38) :

— action du *bra* $\langle\alpha|$ sur le *ket* $\mathcal{O}\,|\beta\rangle$ (action « à droite » de l'opérateur $\mathcal{O}$) ;

— action du *bra* $\langle\alpha|\,\mathcal{O} = \left(\mathcal{O}^\dagger\,|\alpha\rangle\right)^\dagger$ sur le *ket* $|\beta\rangle$ (action « à gauche » de l'opérateur $\mathcal{O}$).

Soit une base orthonormée de l'espace des états $\{|n\rangle\,,\, n \in \mathcal{I}\}$. Un opérateur $\mathcal{O}$ borné (donc défini sur $\mathcal{H}$) est caractérisé complètement par ses « éléments de matrice » $\mathcal{O}_{mn}$ sur cette base [9]. En notation de Dirac on peut représenter cet opérateur de la manière suivante :

$$\boxed{\mathcal{O} = \sum_{r,s\in\mathcal{I}} \mathcal{O}_{rs}\, |r\rangle\, \langle s|} \tag{1.55}$$

On prêtera attention à l'ordre des termes dans cet expression. Le *ket* apparaît avant le bra, ce n'est donc pas le produit hermitien de deux vecteurs ! On vérifie que

$$\begin{aligned}\langle m|\, \mathcal{O}\, |n\rangle &= \sum_{r,s\in\mathcal{I}} \mathcal{O}_{rs}\, \langle m|r\rangle\, \langle s|n\rangle \\ &= \sum_{r,s\in\mathcal{I}} \mathcal{O}_{rs} \delta_{m,r} \delta_{s,n} \\ &= \mathcal{O}_{mn}\,,\end{aligned} \tag{1.56}$$

8. En toute rigueur dans l'équation (1.52) nous supposons que g appartient à $D(M_u)$ mais il se peut que $D(M_u^\dagger)$ contienne des fonctions qui ne le vérifient pas !

9. Le concept de matrice est utilisé ici de manière un peu abusive, car la base n'est pas nécessairement de dimension finie ; comme $\mathcal{H}$ est supposé séparable, elle est néanmoins dénombrable et on obtient une matrice de taille infinie.

en utilisant l'orthonormalité de la base, éqn. (1.23).

De même on peut exprimer les éléments de matrice de l'adjoint d'un opérateur $\mathcal{O}$. On a :

$$\begin{aligned}\mathcal{O}^\dagger{}_{mn} &= \langle m|\,\mathcal{O}^\dagger\,|n\rangle = \iota\left((\mathcal{O}^\dagger)^\dagger\right)|m\rangle\,,|n\rangle\big) = \iota\left(\mathcal{O}\,|m\rangle\,,|n\rangle\right)\\ &= \overline{\iota\left(|n\rangle\,,\mathcal{O}\,|m\rangle\right)} = \overline{\mathcal{O}_{nm}}\,,\end{aligned} \tag{1.57}$$

où on a utilisé la propriété (1.4a) du produit hilbertien. On en déduit que la conjugaison hermitienne d'un opérateur $\mathcal{O}$ sur un espace hermitien revient à prendre la transposée et la conjuguée complexe de la matrice associée à l'opérateur :

$$\boxed{\mathcal{O}^\dagger \overset{\text{déf.}}{=} \overline{\mathcal{O}^{\mathrm{T}}}} \tag{1.58}$$

En travaillant directement avec la notation de Dirac, le conjugué hermitien de l'expression (1.55) donne :

$$\mathcal{O}^\dagger = \sum_{r,s\in\mathcal{I}} \left(\mathcal{O}_{rs}\,|r\rangle\,\langle s|\right)^\dagger = \sum_{r,s\in\mathcal{I}} \overline{\mathcal{O}_{rs}}\left(\langle s|\right)^\dagger\left(|r\rangle\right)^\dagger = \sum_{r,s\in\mathcal{I}} \overline{\mathcal{O}_{rs}}\,|s\rangle\,\langle r| = \sum_{r,s\in\mathcal{I}} \left(\overline{\mathcal{O}^{\mathrm{T}}}\right)_{sr}|s\rangle\,\langle r|\,, \tag{1.59}$$

où on a utilisé à la deuxième étape le fait que le conjugué hermitien d'un produit inverse l'ordre des termes (comme pour la transposée d'un produit de matrices).

Un cas particulièrement important en pratique de la notation (1.55) est l'expression de l'opérateur identité dans la notation de Dirac. Cet opérateur est défini naturellement comme $\mathcal{I} : |\alpha\rangle \mapsto |\alpha\rangle$ pour tout $|\alpha\rangle \in \mathcal{H}$. Ses éléments de matrice dans toute base orthonormale sont donc :

$$\mathcal{I}_{mn} = \langle m|\,\mathcal{I}\,|n\rangle = \delta_{m,n}\,, \tag{1.60}$$

d'où on tire, d'après (1.55), la *relation de fermeture*, ou résolution de l'identité, sur l'espace des états pour la base choisie :

$$\boxed{\mathcal{I} = \sum_{m\in\mathcal{I}} |m\rangle\,\langle m|} \tag{1.61}$$

Cette identité étant valable pour toute base orthonormée, on peut en déduire en des formules de changement de bases pour un opérateur quelconque. Soit donc $\{|n\rangle\,,\, n \in \mathcal{I}\}$ et $\{|\tilde{n}\rangle\,,\, \tilde{n} \in \mathcal{I}\}$ deux bases orthonormées de l'espace des états. On a

$$\mathcal{O}_{\tilde{m}\tilde{n}} = \langle\tilde{m}|\,\mathcal{O}\,|\tilde{n}\rangle = \langle\tilde{m}|\,\mathcal{I}\mathcal{O}\mathcal{I}\,|\tilde{n}\rangle = \sum_{m,n} \langle\tilde{m}|m\rangle\,\langle m|\,\mathcal{O}\,|n\rangle\,\langle n|\tilde{n}\rangle = \sum_{m,n} \langle\tilde{m}|m\rangle\,\mathcal{O}_{mn}\,\langle n|\tilde{n}\rangle \tag{1.62}$$

Le *bra-ket* $\langle n|\tilde{n}\rangle$ correspond donc aux éléments de matrice de la matrice de changement de base $\mathcal{U}$ entre la base $\{|n\rangle\}$ à la base $\{|\tilde{n}\rangle\}$, $\mathcal{U}_{n\tilde{n}} = \langle n|\tilde{n}\rangle$. Si on applique maintenant la relation (1.62) à l'opérateur identité, on trouve

$$\mathcal{I}_{\tilde{m},\tilde{n}} = \delta_{\tilde{m},\tilde{n}} = \sum_{m,n} \langle\tilde{m}|m\rangle\,\delta_{mn}\,\langle n|\tilde{n}\rangle = \sum_{m} \langle\tilde{m}|m\rangle\,\langle m|\tilde{n}\rangle\,. \tag{1.63}$$

En identifiant $\langle m|\tilde{n}\rangle = \mathcal{U}_{m\tilde{n}}$ et $\langle\tilde{m}|m\rangle = \overline{\langle m|\tilde{m}\rangle} = \overline{\mathcal{U}}_{m\tilde{m}}$ on a montré la propriété :

$$\boxed{\mathcal{U}^\dagger\mathcal{U} = \mathcal{I}} \tag{1.64}$$

ou de manière équivalente $U^{-1} = U^\dagger$. Un opérateur vérifiant cette propriété est dit *unitaire*; il préserve le produit hermitien car

$$\forall |\alpha\rangle, |\beta\rangle \in \mathcal{H}, \ \iota(U|\alpha\rangle, U|\beta\rangle) = \langle\alpha| U^\dagger U |\beta\rangle = \langle\alpha|\beta\rangle = \iota(|\alpha\rangle, |\beta\rangle). \tag{1.65}$$

Nous avons ainsi montré qu'*un changement de base entre deux bases orthonormées de l'espace des états correspond à un opérateur unitaire.*

1.2.3 Projection orthogonale

Soit $\mathcal{H}$ un espace de Hilbert et $\mathcal{F}$ un sous-espace vectoriel de $\mathcal{H}$, c.-à.-d. stable par combinaison linéaire. Le *complément orthogonal* de $\mathcal{F}$ dans $\mathcal{H}$ est défini de la manière suivante :

$$\mathcal{F}^\perp \overset{\text{déf.}}{=} \left\{ |\alpha\rangle \in \mathcal{H}, \forall |\beta\rangle \in \mathcal{F}, \langle\alpha|\beta\rangle = 0 \right\}. \tag{1.66}$$

$\mathcal{F}^\perp$ est un sous-espace vectoriel fermé de $\mathcal{H}$. On peut démontrer le théorème de projection :

Théorème 4 (théorème de projection). *Soit $\mathcal{F}$ un sous-espace vectoriel complet de $\mathcal{H}$. Pour tout $|\alpha\rangle \in \mathcal{H}$, il existe un unique vecteur $|\phi\rangle \in \mathcal{F}$ tel que la norme de $\| |\alpha\rangle - |\phi\rangle \|$ soit minimale. En outre $|\phi\rangle$ est l'unique élément de $\mathcal{F}$ tel que $|\alpha\rangle - |\phi\rangle \in \mathcal{F}^\perp$.*

En conséquence, pour tout sous-espace $\mathcal{F} \subset \mathcal{H}$, on peut définir une décomposition orthogonale

$$\mathcal{H} = \mathcal{F} \oplus \mathcal{F}^\perp, \tag{1.67}$$

obtenue en décomposant tout $|\alpha\rangle \in \mathcal{H}$ comme $|\alpha\rangle = |\phi\rangle + \big(|\alpha\rangle - |\phi\rangle\big)$. L'application linéaire

$$\Pi_\mathcal{F} \ : \quad \left\{ \begin{array}{ccc} \mathcal{H} & \to & \mathcal{F} \\ |\alpha\rangle & \mapsto & |\phi\rangle \end{array} \right. \tag{1.68}$$

est le *projecteur orthogonal* sur le sous-espace $\mathcal{F}$. On vérifie facilement que cet opérateur est bien un projecteur :

$$(\Pi_\mathcal{F})^2 = \Pi_\mathcal{F} \tag{1.69a}$$

$$\iota(\Pi_\mathcal{F}|\alpha\rangle, |\beta\rangle) = \iota(|\alpha\rangle, \Pi_\mathcal{F}|\beta\rangle) \quad \forall |\alpha\rangle, |\beta\rangle \in \mathcal{H}. \tag{1.69b}$$

Dans une base adaptée on peut écrire le projecteur $\Pi_\mathcal{F}$ sous une forme particulièrement simple. On commence par choisir une base orthonormée de $\mathcal{F}$, $\{|\imath\rangle \in \mathcal{F}, \imath \in \mathcal{I}\}$, qu'on complète par une base de $\mathcal{F}^\perp$, $\{|\jmath\rangle \in \mathcal{F}^\perp, \jmath \in \mathcal{J}\}$, pour former une base de $\mathcal{H}$; cette base est naturellement orthonormée car $\langle\imath|\jmath\rangle = 0$ en raison de la décomposition (1.67). On a alors :

$$\Pi_\mathcal{F} = \sum_{\imath \in \mathcal{I}} |\imath\rangle \langle\imath| \ . \tag{1.70}$$

1.2.4 Opérateur auto-adjoints et observables

En mécanique quantique une classe particulière d'opérateurs sur l'espace des états joue un rôle primordial, les opérateurs *auto-adjoints* qui sont leur propre adjoint :

$$\mathcal{O}^\dagger = \mathcal{O}\,, \tag{1.71}$$

dont les aspects mathématiques ont été discutés dans la sous-section 1.2.1. En reprenant la discussion après l'éqn. (1.54), on peut faire agir les opérateurs auto-adjoints indistinctement à gauche ou à droite pour le calcul d'un éléments de matrice.

Un exemple simple d'opérateur auto-adjoint est fourni par le projecteur orthogonal (1.70) sur un sous-espace $\mathcal{F} \subset \mathcal{H}$. En utilisant la formulation présentée à l'éqn. (1.59) nous avons :

$$(\Pi_{\mathcal{F}})^\dagger = \left(\sum_{\imath,\imath' \in \mathcal{I}} \delta_{\imath\imath'} \, |\imath\rangle \, \langle \imath'| \right)^\dagger = \sum_{\imath,\imath' \in \mathcal{I}} \overline{\delta_{\imath\imath'}} \, |\imath'\rangle \, \langle \imath| = \sum_{\imath',\imath \in \mathcal{I}} \delta_{\imath'\imath} \, |\imath'\rangle \, \langle \imath| = \Pi_{\mathcal{F}}\,. \tag{1.72}$$

Naturellement cela s'applique au cas particulier de l'opérateur identité sur $\mathcal{H}$.

Spectre d'un opérateur auto-adjoint $\star$

Dans le cas d'un espace hermitien (donc de dimension finie), où les opérateurs sont associés à des matrices carrées complexes, *un opérateur hermitien, qui est automatiquement auto-adjoint, est diagonalisable et ses valeurs propres sont toutes réelles ; il existe en outre une base orthonormée de vecteurs propres.* La diagonalisation de l'opérateur revient à un changement de base entre une base orthonormée et une autre. D'après la discussion après l'éqn. (1.63) la matrice de changement de base U associée est une matrice *unitaire*, c.-à.-d. satisfaisant la relation (1.64).

Dans un espace de Hilbert de dimension infinie, une classe d'opérateurs, les opérateurs auto-adjoints compacts, se comportent essentiellement de la même manière que les opérateurs sur les espaces hermitiens. On démontre le théorème suivant :

Théorème 5 *Soit $\mathcal{O}$ un opérateur linéaire agissant sur un espace de Hilbert $\mathcal{H}$. On suppose $\mathcal{O}$ borné, voir l'éqn. (1.31) et compact, c.-à.-d. tel que l'image d'une partie bornée de $\mathcal{H}$ a pour adhérence une partie compacte de $\mathcal{H}$.*[10]*. Cet opérateur est auto-adjoint si et seulement s'il existe une base orthonormée de $\mathcal{H}$ formée de vecteurs propres de $\mathcal{O}$ et les valeurs propres sont toutes des nombres réels :*

$$\mathcal{O}^\dagger = \mathcal{O} \;\Leftrightarrow\; \exists\, \{|\mathfrak{r}\rangle\,, \mathfrak{r} \in \mathcal{I}\} \;\; \textit{base orthonormée de} \;\; \mathcal{H} \;\; \textit{t.q.} \;\; \forall \mathfrak{r} \in \mathcal{I}\,,\; \mathcal{O}\,|\mathfrak{r}\rangle = \lambda_{\mathfrak{r}}\,|\mathfrak{r}\rangle \;\;,\quad \lambda_{\mathfrak{r}} \in \mathbb{R}\,. \tag{1.73}$$

Les valeurs propres $\lambda_{\mathfrak{r}}$ ne sont pas nécessairement toutes distinctes ; l'ensemble des vecteurs propres de $\mathcal{O}$ associés à une valeur propre $\lambda_{\mathfrak{r}}$ forme un sous-espace de Hilbert $\mathcal{H}_{\lambda_{\mathfrak{r}}}$, de dimension complexe supérieure ou égale à un. Si elle est strictement supérieure à un la dite valeur propre est dite *dégénérée*.

10. $\mathcal{A}$ est une partie compacte de $\mathcal{H}$ si de toute suite d'éléments de $\mathcal{A}$ on peut extraire une sous-suite convergente.

En mécanique quantique, un exemple d'opérateur compact (mis à par le cas trivial d'opérateurs agissant sur un espace de dimension finie) est fourni par l'*opérateur densité* qui sera discuté plus loin. La grande majorité des opérateurs ne sont ni compacts ni même bornés (par exemple l'impulsion, le moment cinétique orbital, etc...) et il est nécessaire d'utiliser un formalisme plus général, celui de la *théorie spectrale* dont nous nous bornerons à survoler certains aspects utiles pour la physique.

De manière tout à fait générale, on définit tout d'abord le *spectre* d'un opérateur, notion plus générale que celle de valeur propre.

Définition 4 (spectre d'un opérateur). *Soit* $\mathcal{O} : \mathrm{D}(\mathcal{O}) \subseteq \mathcal{H} \to \mathcal{H}$ *un opérateur. Le spectre* $\sigma(\mathcal{O})$ *de* $\mathcal{O}$ *est la partie de* $\mathbb{C}$ *définie par :*

$$\sigma(\mathcal{O}) = \mathbb{C} \setminus \{z \in \mathbb{C} \,|\, (\mathcal{O} - z\mathcal{I})^{-1} \textit{existe et est borné}\} \tag{1.74}$$

En particulier, si $\mathcal{O} - z\mathcal{I}$ possède un noyau non trivial, alors $z \in \sigma(\mathcal{O})$, et z est appelée *valeur propre* de $\mathcal{O}$: un vecteur $|\alpha\rangle \in \ker(\mathcal{O} - z)$ est un *vecteur propre* associé à cette valeur propre. Il faut veiller à distinguer les notions de spectre et de valeur propre dans le cas général. Les valeurs propres forment *le spectre ponctuel* de l'opérateur ; le spectre peut posséder d'autres composantes, en particulier continues.

Exemple 2 (opérateur position). Considérons l'opérateur agissant sur l'espace de fonctions $\mathcal{L}^2(\mathbb{R})$ comme $\mathcal{X} : \psi(x) \mapsto x\psi(x)$ (il s'agit de l'opérateur position qui sera traité dans le cours n° 2, exemple d'un opérateur de multiplication, cf. exemple 1).

On cherche à déterminer à quelle condition l'application $\psi(x) \mapsto \frac{1}{z-x}\psi(x)$ est bornée dans $\mathcal{L}^2(\mathbb{R})$; cela revient à chercher $M > 0$ tel que $|\frac{1}{z-x}| \leqslant M$, $\forall x \in \mathbb{R}$. On montre alors sans difficulté que le spectre est donné par

$$\sigma(\mathcal{X}) = \{z \in \mathbb{C} \,|\, \nexists \epsilon > 0\,,\ \forall x \in \mathbb{R}\,,\ |z - x| \geqslant \epsilon\} = \mathbb{R}\,. \tag{1.75}$$

Ainsi le spectre de l'opérateur position correspond – comme on pourrait s'y attendre – à tout l'axe réel. Cependant cet opérateur n'admet ni valeurs propres ni vecteurs propres ! Nous expliquerons dans le cours n° 2 que les « vecteurs propres généralisés » associés ne font pas partie de l'espace de Hilbert $\mathcal{L}^2(\mathbb{R})$.

La notion de spectre d'un opérateur étant établie, on peut caractériser les opérateurs auto-adjoints de la manière suivante :

Théorème 6 *Soit* $\mathcal{O}$ *un opérateur hermitien sur un espace de Hilbert.* $\mathcal{O}$ *est un opérateur auto-adjoint si et seulement si* $\sigma(\mathcal{O}) \subseteq \mathbb{R}$.

Concernant les éventuels vecteurs propres associés à un opérateur hermitien, on peut montrer premièrement :

Lemme 1 *Soit* $\mathcal{O}$ *un opérateur hermitien possédant des vecteurs propres* $|\psi_r\rangle$ *et des valeurs propres* λ_r *associées (c.-à.-d. telles que le noyau de* $\mathcal{O} - \lambda_r\mathcal{I}$ *soit non-trivial). Alors les vecteurs propres associés à des valeurs propres distinctes sont orthogonaux :*

$$\mathcal{O}\,|\psi_1\rangle = \lambda_1\,|\psi_1\rangle\ ,\ \mathcal{O}\,|\psi_2\rangle = \lambda_2\,|\psi_2\rangle\ ,\ \langle\psi_1|\psi_2\rangle = 0 \ \textit{si}\ \lambda_1 \neq \lambda_2\,. \tag{1.76}$$

Si on peut construire une base orthonormée de vecteurs propres, on a le théorème suivant :

Théorème 7 *Soit $\mathcal{O}$ un opérateur hermitien sur $\mathcal{H}$ tel qu'il existe une base orthonormée de $\mathcal{H}$ formée de vecteurs propres de $\mathcal{O}$. Alors $\mathcal{O}$ est un opérateur auto-adjoint (ou plus précisément sa fermeture $\mathcal{O}^{\mathsf{f}}$ l'est).*

Remarquons que la réciproque n'est pas nécessairement vraie – voir l'exemple de l'opérateur position ci-dessus.

Finalement, l'analogue de la diagonalisation pour des opérateurs auto-adjoints sur un espace de Hilbert constitue le fameux *théorème spectral* dont une version – légèrement édulcorée – est la suivante [11] :

Théorème 8 (théorème spectral). *Soit $\mathcal{O}$ un opérateur borné et auto-adjoint sur un espace de Hilbert $\mathcal{H}$. Il existe un opérateur unitaire $\mathcal{U}$ et une fonction réelle et bornée f telle que*

$$\mathcal{U}^{\dagger}\mathcal{O}\mathcal{U} = \mathcal{M}_{\mathsf{f}}\,, \tag{1.77}$$

où $\mathcal{M}_{\mathsf{f}}$ est l'opérateur multiplication, considéré dans l'exemple 1, associé à la fonction f.

Les valeurs prises par la fonction f sont ici l'analogue des valeurs propres. Moyennant quelques raffinements supplémentaires, le théorème spectral se généralise aux opérateurs auto-adjoints non bornés en des termes proches.

Une autre version du théorème spectral, d'une plus grande utilité, fait appel à la théorie de la mesure et ne sera pas développée dans ces notes pour cette raison. Schématiquement, nous avons

$$\mathcal{O} = \int_{\sigma(\mathcal{O})} \lambda \, \mathrm{d}\Pi_{\lambda}(\mathcal{O})\,, \tag{1.78}$$

où $\mathrm{d}\Pi_{\lambda}(\mathcal{O})$ est une *mesure spectrale*, équivalent dans le cas continu de la projection sur le sous-espace propre associé à λ dans le cas discret.

Postulat sur les observables

En mécanique hamiltonienne, l'état d'un système physique est caractérisé d'une part par les coordonnées généralisées q^{i} et les impulsions conjuguées p_{i}, formant un système de coordonnées sur l'espace des phases du système, et d'autre part par des fonctions $\mathsf{f}(\mathsf{q}^{\mathsf{i}}, \mathsf{p}_{\mathsf{i}}; \mathsf{t})$ de l'espace des phases et éventuellement du temps associées à d'autres grandeurs comme le moment cinétique ou l'énergie.

Lors du passage à la mécanique quantique via la quantification canonique, les grandeurs physiques deviennent des opérateurs appelés *observables*, applications linéaires de l'espace des états $\mathcal{H}$ dans lui-même. Les grandeurs physiques classiques prenant uniquement des valeurs réelles, on souhaite imposer une condition analogue aux observables. Demander aux observables de correspondre à une matrice réelle est trop restrictif et dépend naturellement de la base choisie ; pour imposer que les quantités physiques associées aux observables soient réelles, il suffit d'imposer que le spectre de ces opérateurs soit réel. Nous sommes donc amenés à formuler le postulat suivant :

11. Le théorème complet affirme aussi l'existence d'un espace mesurable E sur lequel est défini f.

Postulat 2 (observables). *Les grandeurs mesurables caractérisant un système quantique sont associées à des opérateurs auto-adjoints agissant sur l'espace de Hilbert $\mathcal{H}$ correspondant à l'espace des états du système. Ces opérateurs sont appelés observables.*

L'observable la plus importante en mécanique quantique est l'*opérateur hamiltonien* $\mathcal{H}$, version quantique du hamiltonien (3.34) de la mécanique classique. Comme nous le verrons plus tard, dans les systèmes pour lesquels l'énergie est conservée, le spectre du hamiltonien, voir l'éqn. (1.75) pour la définition, donne les énergies accessibles au système.

1.2.5 Opérateurs commutant deux à deux

Deux états physiques associés à des vecteurs de l'espace de Hilbert $\mathcal{H}$ – ou plutôt des rayons de l'espace projectif $\mathbb{P}[\mathcal{H}]$ associé – sont distincts si au moins une grandeur physique du système permet de les différencier. Il semble donc naturel de considérer une base de l'espace des états $\mathcal{H}$ constituée de vecteurs propres des différentes observables associées à ces grandeurs physiques, pour autant qu'une telle base existe – voir la sous-section 1.2.4 pour une discussion approfondie.

Cette intuition se heurte à une difficulté importante. Les observables considérées constituent un ensemble d'opérateurs auto-adjoints qui n'ont aucune raison, a priori, de commuter entre eux. Physiquement, deux opérateurs qui ne commutent pas correspondent à des observables *incompatibles*. En suivant les postulats qui seront énoncés dans la sous-section 1.3, il n'est pas possible de mesurer simultanément les grandeurs associées à ces deux observables, car l'ordre dans lequel les deux mesures sont effectuées va changer le résultat.

Pour obtenir une base de l'espace des états il faut ainsi considérer des observables *compatibles*, c.-à.-d. qui commutent entre elles. Nous pouvons premièrement énoncer le théorème suivant :

Théorème 9 *Soient $\mathcal{O}$ et $\mathcal{P}$ deux opérateurs, hermitiens sur un espace hermitien $\mathcal{H}$, ou auto-adjoints bornés et compacts sur un espace de Hilbert $\mathcal{H}$. Ces opérateurs commutent si et seulement si ils admettent une base commune de vecteurs propres :*

$$[\mathcal{O},\mathcal{P}]=0 \;\Leftrightarrow\; \exists\,\{|r\rangle, r\in\mathcal{I}\} \;\textit{base orthonormée de}\; \mathcal{H} \;\textit{t.q.}\; \forall r\in\mathcal{I}, \begin{cases} \mathcal{O}|r\rangle &= \lambda_r|r\rangle\,, \\ \mathcal{P}|r\rangle &= \mu_r|r\rangle\,. \end{cases} \tag{1.79}$$

De manière équivalente, on peut écrire que

$$\boxed{[\mathcal{O},\mathcal{P}]=0 \;\Leftrightarrow\; \exists U \textit{ unitaire t.q. } U\mathcal{O}U^\dagger=\mathrm{DIAG}(\lambda_1,\lambda_2,\ldots) \textit{ et } U\mathcal{P}U^\dagger=\mathrm{DIAG}(\mu_1,\mu_2,\ldots)} \tag{1.80}$$

Une généralisation évidente de ce théorème nous indique qu'un ensemble d'opérateurs commutant deux à deux admet une base commune de vecteurs propres.

Dans le cas d'opérateurs auto-adjoints non compacts, donc n'admettant pas de bases de vecteurs propres a priori, l'analogue du thèorème 9 est

Théorème 10 (opérateurs auto-adjoints commutants). *Soit* $\{\mathcal{O}_i, i = 1, \ldots, n\}$ *une famille d'opérateurs auto-adjoints sur un espace de Hilbert* $\mathcal{H}$ *commutant deux à deux. Il existe alors un opérateur unitaire* U *et un jeu de fonctions* $\{f_i, i = 1, \ldots, n\}$ *à valeurs dans* $\mathbb{R}$ *tels que*

$$U\mathcal{O}_i U^\dagger = M_{f_i} \ , \quad \forall i = 1, \ldots, n \, , \tag{1.81}$$

où M_{f_i} *est l'opérateur multiplication associé à la fonction* f_i*.*

1.2.6 Ensemble complet d'observables commutantes

Deux vecteurs de l'espace des états $\mathcal{H}$ non colinéaires (c.-à.-d. n'appartenant pas au même rayon dans $\mathbb{P}[\mathcal{H}]$) doivent correspondre à deux états physiques distincts, et doivent donc pouvoir être distingués l'un de l'autre par au moins un mesure physique.

Dans le cas de figure le plus simple, on peut trouver une observable $\mathcal{O}$, par exemple le hamiltonien du système, telle que sont spectre ponctuel soit non-dégénéré et qu'il existe une base orthonormée de $\mathcal{H}$ constituée de vecteurs propres de $\mathcal{O}$ associés à des valeurs propres toutes distinctes :

$$\langle \ell | \ell' \rangle = \delta_{\ell,\ell'} \, , \ \mathcal{O} \, |\ell\rangle = \lambda_\ell \, |\ell\rangle \ , \ \lambda_\ell \neq \lambda_{\ell'} \text{ si } \ell \neq \ell' \, , \tag{1.82}$$

soit une base orthonormée constituée de vecteurs propres de $\mathcal{O}$. Ainsi, préparer le système dans un état propre de $\mathcal{O}$ détermine complètement son état quantique car il existe un unique rayon correspondant dans $\mathbb{P}[\mathcal{H}]$.

Un cas de figure fréquent est le suivant. Plusieurs vecteurs d'états linéairement indépendants correspondent à des vecteurs propres de même valeur propre λ pour une certaine observable $\mathcal{O}$; les vecteurs propres engendrent un sous-espace propre $\mathcal{H}_\lambda$ de dimension supérieure à un. On considère alors une autre observable $\mathcal{P}$ commutant avec $\mathcal{O}$, et il existe une base de $\mathcal{H}_\lambda$ constituée de vecteurs propres de $\mathcal{P}$ avec des valeurs propres deux à deux toutes distinctes et l'observable $\mathcal{P}$ permet ainsi de « préciser » l'état du système.

De manière générale, il faut considérer un ensemble de plusieurs observables pour spécifier entièrement l'état du système et on peut définir :

Définition 5 (ensemble complet d'observables commutantes). *Soit un système quantique dont l'espace des états correspond à un espace de Hilbert séparable* $\mathcal{H}$*. On considère une famille finie d'opérateurs autoadjoints* $\{\mathcal{O}_\ell, \ell = 1, 2, \ldots\}$*, commutant deux à deux :*

$$\forall \ell, \ell' \ , \quad [\, \mathcal{O}_\ell, \, \mathcal{O}_{\ell'} \,] = 0 \, . \tag{1.83}$$

On suppose également que $\mathcal{H}$ *admette une base orthonormée constituée de vecteurs propres communs à tous les opérateurs* $\mathcal{O}_\ell$ [12]*, telle que, pour tout vecteur de la base, il existe au moins une observable de la famille* $\{\mathcal{O}_\ell\}$ *pour laquelle la valeur propre associée soit non dégénérée. En d'autres termes, la donnée des valeurs propres* $\{\lambda_\ell\}$ *associées aux observables de la famille* $\{\mathcal{O}_\ell\}$ *caractérise complètement le vecteur d'état associé à une constante multiplicative près.*

$$\forall \ell, \left\{ \begin{array}{l} \mathcal{O}_\ell \, |\alpha\rangle = \lambda_\ell \, |\alpha\rangle \\ \mathcal{O}_\ell \, |\beta\rangle = \lambda_\ell \, |\beta\rangle \end{array} \right. \Rightarrow \ \exists a \in \mathbb{C}^\star \, , \ |\beta\rangle = a \, |\alpha\rangle \ . \tag{1.84}$$

12. Si les opérateurs sont tous compacts, ou si $\mathcal{H}$ est de dimension finie, cela est garanti par le théorème 9.

Un tel ensemble d'observables est appelé ensemble complet d'observables commutantes, ou ECOC.

Pour des opérateurs dont le spectre contient des composantes continues, comme l'observable de position qui sera discutée en détail dans le cours n° 2, une généralisation de ce concept existe, même s'il ne s'agit pas de vecteurs propres – ni même de vecteurs de l'espace de Hilbert – à proprement parler.

1.3 Valeurs moyennes et mesure

Les observables d'un système quantique sont des endormorphismes auto-adjoints agissant sur l'espace des états $\mathcal{H}$ du système, et correspondant à des grandeurs physiques caractérisant les différents états possibles du système.

En mécanique classique hamiltonienne, connaissant la position du système $(\mathfrak{q}^i, \mathfrak{p}_i)$ dans l'espace des phases, ont peut sans ambiguïté en déduire la valeur de toute grandeur physique correspondant à une certaine fonction $\mathfrak{f}(\mathfrak{q}^i, \mathfrak{p}_i; \mathfrak{t})$ de l'espace des phases et du temps. En mécanique quantique, à un vecteur d'état quelconque $|\alpha\rangle$, on ne peut associer de manière déterministe la valeur que prendra une grandeur physique associée à une observable – cela n'aurait mathématiquement pas de sens dans le formalisme d'espaces de Hilbert employé.

1.3.1 Valeurs moyennes d'observables

Les quantités pertinentes en mécanique quantique sont les *valeurs moyennes* d'observables, qui donnent la moyenne des résultats obtenus pour la mesure de la grandeur physique associée au sens statistique ; il s'agit d'un postulat fondamental de la mécanique quantique.

Postulat 3 (valeurs moyennes d'observables). *Soit un système quantique dans un état caractérisé, à l'instant de la mesure, par un vecteur $|\alpha\rangle \in \mathcal{H}$ de norme unité – plus précisément par une classe d'équivalence dans $\mathbb{P}[\mathcal{H}]$ contenant $|\alpha\rangle$ – et $\mathcal{O}$ une observable associée à une grandeur physique mesurable. La valeur moyenne des mesures de cette grandeur, pour un ensemble statistique de copies du système préparées dans le même état, est donnée par*

$$\boxed{\langle \mathcal{O} \rangle_\alpha \overset{\text{déf.}}{=} \langle \alpha | \, \mathcal{O} \, | \alpha \rangle} \tag{1.85}$$

Ce postulat s'applique à toute observable du système, faisant partie ou non de l'ensemble complet d'observables commutantes.

1.3.2 Premier postulat de la mesure

La prédiction de la valeur moyenne d'une observable ne nous renseigne pas sur les résultats possibles d'une mesure individuelle, ni sur la probabilité d'obtenir telle ou telle valeur. Lorsque le vecteur d'état $|\alpha\rangle$ se trouve correspondre à un vecteur propre de l'observable $\mathcal{O}$ de valeur propre λ le résultat de la mesure donnera de manière certaine la valeur λ.

Dans la suite de cette sous-section, on considérera uniquement le cas des opérateurs sur des espaces hermitiens ou des opérateurs compacts sur des espaces de Hilbert. La généralisation aux opérateurs non-compacts nécessiterait d'introduire la théorie mathématique de la mesure que seuls ceux d'entre vous ayant étudié l'intégrale de Lebesgue connaissent.

Pour traiter du cas général, on décompose tout d'abord l'espace des états en somme directe des sous-espaces propres de $\mathcal{O}$:

$$\boxed{\mathcal{H} = \bigoplus_{\mathrm{r}} \mathcal{H}_{\lambda_\mathrm{r}} \ , \quad \forall \mathrm{r} \neq \mathrm{r}' \ , \quad \lambda_\mathrm{r} \neq \lambda_{\mathrm{r}'} \, .} \tag{1.86}$$

L'orthogonalité entre les différents sous-espaces propres découle de l'hermiticité de $\mathcal{O}$. En effet, on a

$$\forall \, |\alpha\rangle \in \mathcal{H}_{\lambda_\mathrm{r}}, \, |\beta\rangle \in \mathcal{H}_{\lambda_{\mathrm{r}'}} \qquad \langle\beta| \, \mathcal{O} \, |\alpha\rangle = \lambda_\mathrm{r} \, \langle\beta|\alpha\rangle = \lambda_{\mathrm{r}'} \, \langle\beta|\alpha\rangle \implies \lambda_\mathrm{r} = \lambda_{\mathrm{r}'} \text{ ou } \langle\beta|\alpha\rangle = 0 \, , \tag{1.87}$$

où nous avons fait agir $\mathcal{O}$ tantôt à droite sur $|\alpha\rangle$, tantôt à gauche sur $\langle\beta|$.

On peut alors décomposer tout vecteur $|\alpha\rangle$ comme somme de ses projetés orthogonaux sur les différents sous-espaces propres :

$$|\alpha\rangle = \sum_{\lambda_\mathrm{r}} \Pi_{\lambda_\mathrm{r}} \, |\alpha\rangle \ , \quad \Pi_{\lambda_\mathrm{r}} \, |\alpha\rangle \in \mathcal{H}_{\lambda_\mathrm{r}} \, . \tag{1.88}$$

Pour simplifier les notations, on travaillera avec des vecteurs d'états normés, voir éqn. (1.5).

On munit chacun des sous espaces propres d'une base orthonormée $\{|e^\mathrm{r}{}_\ell\rangle \ , \ \ell \in \mathcal{I}_\mathrm{r}\}$, c.-à.-d. une famille de vecteurs dense dans $\mathcal{H}_{\lambda_\mathrm{r}}$ telle que

$$\forall \ell \in \mathcal{I}_\mathrm{r} \ , \quad \mathcal{O} \, |e^\mathrm{r}{}_\ell\rangle = \lambda_\mathrm{r} \, |e^\mathrm{r}{}_\ell\rangle \ , \quad \langle e^\mathrm{r}{}_\ell | e^\mathrm{r}{}_{\ell'}\rangle = \delta_{\ell\ell'} \, , \tag{1.89}$$

et on peut ainsi décomposer le projeté orthogonal du vecteur *normé* $|\alpha\rangle$ sur $\mathcal{H}_{\lambda_\mathrm{r}}$:

$$\Pi_{\lambda_\mathrm{r}} \, |\alpha\rangle = \sum_{\ell \in \mathcal{I}_\mathrm{r}} \alpha_\mathrm{r}{}^\ell \, |e^\mathrm{r}{}_\ell\rangle \ . \tag{1.90}$$

Il est important de remarquer que le vecteur $\Pi_{\lambda_\mathrm{r}} \, |\alpha\rangle$ n'est pas de norme unité ; sa norme contient une information cruciale, qui constitue le postulat suivant :

Postulat 4 (premier postulat de la mesure). *Soit un système quantique dans un état caractérisé par un vecteur $|\alpha\rangle \in \mathcal{H}$ de norme unité à l'instant de la mesure. Soit $\mathcal{O}$ une observable associée à une grandeur physique mesurable dont le spectre, c.-à.-d. l'ensemble des valeurs propres distinctes, est donné par une famille $\{\lambda_\mathrm{r}\}$. Ces valeurs propres fournissent les résultats possibles de la mesure.*

— *si $|\alpha\rangle$ est un vecteur propre de $\mathcal{O}$, c.-à.-d. si $\mathcal{O} \, |\alpha\rangle = \lambda_\mathrm{r} \, |\alpha\rangle$, le résultat de la mesure donnera la valeur λ_r de manière certaine ;*

— *si* $|\alpha\rangle$ *n'est pas un vecteur propre de* $\mathcal{O}$*, on ne peut déterminer le résultat de la mesure avec certitude. La probabilité d'obtenir la valeur* λ_r *comme résultat de la mesure est donnée par la norme carrée du projeté orthogonal* $|\alpha\rangle$ *sur le sous-espace propre* $\mathcal{H}_{\lambda_r}$ *:*

$$\boxed{p(\lambda_r) = \| \Pi_{\lambda_r} |\alpha\rangle \|^2 = \sum_{\ell \in \mathcal{I}_r} \left| \alpha_r{}^{\ell} \right|^2} \tag{1.91}$$

Le calcul de $p(\lambda_r)$ pour toutes les valeurs propres λ_r distinctes associées à l'observable $\mathcal{O}$ donne une distribution de probabilité pour les résultats possibles de la mesure.

Vérifions que $p(\lambda_r)$ possède les bonnes propriétés pour constituer une distribution de probabilité. On vérifie premièrement, en utilisant l'inégalité de Bessel (1.26) pour une base orthonormée de $\mathcal{H}_\lambda$ considérée comme famille orthonormée dans $\mathcal{H}$, que $p(\lambda) \leqslant 1$. On remarque ensuite qu'en raison de la décomposition (1.86) de l'espace de Hilbert, la somme des projecteurs sur tous les sous-espaces propres donne l'opérateur identité :

$$\sum_{\lambda_r} \Pi_{\lambda_r} = \mathfrak{I}\,, \tag{1.92}$$

d'où on déduit que

$$\sum_{\lambda_r} p(\lambda_r) = \sum_{\lambda_r} \| \Pi_{\lambda_r} |\alpha\rangle \|^2 = \left\| \sum_{\lambda_r} \Pi_{\lambda_r} |\alpha\rangle \right\|^2 = \| \, |\alpha\rangle \, \|^2 = 1\,, \tag{1.93}$$

en utilisant l'orthogonalité entre les sous-espace propres.

On peut considérer que ce premier postulat de la mesure découle du postulat 3 portant sur les valeurs moyennes d'opérateurs. En effet l'équation (1.91) correspond à la valeur moyenne dans l'état $|\alpha\rangle$ de la projection orthogonale Π_λ, qui est un opérateur auto-adjoint sur $\mathcal{H}$. Dans cette perspective, les valeurs propres associées à Π_λ sont 1 et 0, correspondant respectivement aux vecteurs de base appartenant au sous-espace propre $\mathcal{H}_\lambda$ et n'y appartenant pas.

D'un autre point de vue la valeur moyenne (1.85) d'un opérateur $\mathcal{O}$ peut être reliée à la distribution de probabilité associée à la mesure de ses valeurs propres. À l'aide de la résolution de l'identité (1.92), nous avons, comme $\Pi_{\lambda_r} |\alpha\rangle \in \mathcal{H}_{\lambda_r}$ pour tout $|\alpha\rangle$:

$$\begin{aligned} \langle\alpha| \mathcal{O} |\alpha\rangle &= \langle\alpha| \mathcal{O} \sum_r \Pi_{\lambda_r} |\alpha\rangle = \sum_r \langle\alpha| \mathcal{O} \left(\Pi_{\lambda_r} |\alpha\rangle\right) \\ &= \sum_r \lambda_r \langle\alpha| \Pi_{\lambda_r} |\alpha\rangle = \sum_r \lambda_r \langle\alpha| \Pi^2_{\lambda_r} |\alpha\rangle \end{aligned} \tag{1.94}$$

où nous avons utilisé la propriété (1.69a) d'un projecteur, soit finalement, un projecteur étant auto-adjoint :

$$\boxed{\langle\mathcal{O}\rangle_\alpha = \sum_r p(\lambda_r)\, \lambda_r} \tag{1.95}$$

1.3.3 Second postulat de la mesure

Le premier postulat de la mesure fournit une probabilité d'obtenir une certaine valeur λ_r comme résultat d'une mesure physique associée à une certaine observable $\mathcal{O}$. Que peut-on dire de l'état du système après cette mesure ?

Supposons qu'une deuxième mesure associée à la même observable $\mathcal{O}$ soit effectuée immédiatement après sur le même système, sans laisser au système le temps d'évoluer et d'interagir avec l'extérieur après la première mesure.

Il semble pertinent de supposer – ce qui est validé par l'expérience – que cette deuxième mesure donne le même résultat que la première de manière certaine. Cela n'est compatible avec le postulat 4 que si la première mesure projette de manière « instantanée » le vecteur d'état du système sur un vecteur appartenant au sous-espace propre associé à la valeur λ_r. Ce postulat, souvent nommé « réduction du paquet d'ondes » se formule de la manière suivante :

Postulat 5 (second postulat de la mesure : réduction du paquet d'ondes). *Soit un système quantique dans un état caractérisé par un vecteur $|\alpha\rangle \in \mathcal{H}$ de norme unité à l'instant de la mesure et $\mathcal{O}$ une observable associée à une grandeur physique mesurable. Supposons que la mesure effectuée sur le système donne la valeur λ_r, appartenant au spectre de $\mathcal{O}$. L'état du système juste après la mesure est alors donné par la projection orthogonale de $|\alpha\rangle$ sur le sous-espace propre $\mathcal{H}_{\lambda_r}$. En normant le vecteur obtenu on obtient*

$$\boxed{|\alpha\rangle \xrightarrow{\text{valeur propre } \lambda_r \text{ mesurée}} \frac{1}{\| \Pi_{\lambda_r} |\alpha\rangle \|} \Pi_{\lambda_r} |\alpha\rangle} \tag{1.96}$$

Une deuxième mesure de la *même observable* juste après la première ne change rien à l'état du système, en utilisant la propriété (1.69a) de l'opérateur de projection.

Nous présentons dans l'appendice 1.B un exemple célèbre d'expérience de mesure en mécanique quantique, l'expérience de Stern et Gerlach.

Mesures successives

Considérons deux observables $\mathcal{O}$ (resp. $\tilde{\mathcal{O}}$) de spectre $\{\lambda_r\}$ (resp. de spectre $\{\tilde{\lambda}_r\}$) commutant entre elles, c.-à.-d. vérifiant $[\mathcal{O}, \tilde{\mathcal{O}}] = 0$. Considérons alors l'expérience de pensée suivante :

- une mesure de la grandeur associée à $\mathcal{O}$ donne le résultat correspondant à la valeur propre λ_r ;
- immédiatement après, une mesure de la grandeur associée à $\tilde{\mathcal{O}}$ donne le résultat correspondant à la valeur propre $\tilde{\lambda}_{\tilde{r}}$.

L'état du système, après les mesures successives, peut être déduit du postulat 5. Il est obtenu par projections successives sur les sous-espaces propres correspondant respectivement aux deux valeurs propres des deux observables (en normant le résultat final) :

$$|\alpha\rangle \longrightarrow \frac{1}{\| \Pi_{\tilde{\lambda}_{\tilde{r}}} \Pi_{\lambda_r} |\alpha\rangle \|} \Pi_{\tilde{\lambda}_{\tilde{r}}} \Pi_{\lambda_r} |\alpha\rangle \ . \tag{1.97}$$

Les deux observables étant commutantes, il existe une base orthonormée $\{|e^{r}{}_{\ell}\rangle\ ,\ \ell \in \mathcal{I}_r\}$ du sous-espace propre $\mathcal{H}_{\lambda_r}$ pour la valeur propre λ_r de $\mathcal{O}$ dont les vecteurs de base sont également vecteurs propres de $\tilde{\mathcal{O}}$:

$$\forall \ell \in \mathcal{I}_r\ ,\ \exists \tilde{r}\ ,\ \tilde{\mathcal{O}}\,|e^{r}{}_{\ell}\rangle = \tilde{\lambda}_{\tilde{r}}\,|e^{r}{}_{\ell}\rangle\ . \tag{1.98}$$

En conséquence, la deuxième mesure ne peut donner n'importe quel résultat, mais uniquement des valeurs $\tilde{\lambda}_{\tilde{r}}$ compatibles avec l'éqn. (1.98), c.-à.-d. associées à des vecteurs de base dans l'intersection des sous-espaces propres $\mathcal{H}_{\tilde{\lambda}_{\tilde{r}}}$ de la deuxième observable et $\mathcal{H}_{\lambda_r}$ de la première observable.

Suivant la manière dont les sous-espaces propres de $\mathcal{O}$ et $\tilde{\mathcal{O}}$ sont imbriqués, cela peut donner le résultat de la deuxième mesure de manière certaine connaissant le résultat de la première mesure.

On peut considérer la même expérience de pensée, mais avec deux observables qui ne commutent pas entre elles : $[\mathcal{O}, \tilde{\mathcal{O}}] \neq 0$. De telles observables sont dites *incompatibles*. Dans ce cas, même si la règle (1.97) reste vraie, on ne peut donner de résultat général analogue à (1.98) concernant le résultat de deux mesures successives. La manière dont la connaissance d'une observable affecte celle de l'autre est contenue dans le principe d'incertitude.

1.3.4 Principe d'incertitude

Lorsque deux observables $\mathcal{O}$ et $\mathcal{P}$ ne commutent pas, il n'est pas possible de trouver une base commune de vecteurs propres [13]

Du point de vue de la mesure quantique, la non-commutativité de deux observables peut s'interpréter dans l'expérience de pensée suivante. Supposons qu'on effectue une mesure de l'observable $\mathcal{P}$, et que le résultat de la mesure λ_r indique que le vecteur d'état est projeté sur le sous-espace propre $\mathcal{H}_{\lambda_r}$. On effectue ensuite une mesure de l'observable $\mathcal{O}$, et que le résultat de la mesure μ_ℓ indique que le vecteur d'état est projeté sur le sous-espace propre $\mathcal{H}_{\mu_\ell}$. Si on effectue ensuite une nouvelle mesure de $\mathcal{P}$, nous n'avons génériquement aucune garantie de trouver la même valeur λ_r.

Cela signifie que la distribution de probabilité associée aux mesures successives de $\mathcal{O}$ puis de $\mathcal{P}$ sera différente de la distribution de probabilité associée aux mesures successives de $\mathcal{P}$ puis de $\mathcal{O}$; on ne peut donner alors un sens, dans la limite où l'intervalle de temps entre les mesures tend vers zéro, à la mesure simultanée de ces deux grandeurs.

Le principe d'incertitude exprime le fait, qualitativement, que l'acte de mesure de la grandeur associée à l'observable $\mathcal{O}$ va projeter le vecteur d'état sur le sous-espace propre correspondant et de fait va, génériquement, augmenter notre incertitude sur le résultat de la mesure de l'autre observable $\mathcal{P}$.

Pour formuler mathématiquement ce principe, introduisons la variance d'une observable

13. Cela n'interdit certaintement pas $\mathcal{O}$ et $\mathcal{P}$ d'avoir des vecteurs propres communs, si leur commutateur $[\mathcal{O}, \mathcal{P}]$ possède un noyau non trivial. Par exemple, les matrices $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ et $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ ne commutent pas mais admettent $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ comme vecteur propre commun.

dans un certain état $|\alpha\rangle$:

$$\mathrm{Var}_\alpha(\mathcal{O}) \overset{\text{déf.}}{=} \langle(\mathcal{O} - \langle\mathcal{O}\rangle_\alpha)^2\rangle_\alpha = \langle\alpha|\,(\mathcal{O} - \langle\mathcal{O}\rangle_\alpha)^2\,|\alpha\rangle\ , \tag{1.99}$$

où la valeur moyenne $\langle\cdot\rangle_\alpha$ est définie par l'équation (1.85) ; la variance sera également notée $\Delta_\alpha^2(\mathcal{O})$, cette quantité étant positive.

En toute généralité, en utilisant les propriétés des espaces de Hilbert on peut démontrer le théorème suivant :

Théorème 11 (principe d'incertitude). *Soient $\mathcal{O}$ et $\mathcal{P}$ deux opérateurs hermitiens sur un espace de Hilbert $\mathcal{H}$. Pour tout vecteur $|\alpha\rangle \in \mathsf{D}(\mathcal{O}\mathcal{P}) \cap \mathsf{D}(\mathcal{P}\mathcal{O})$,*

$$\boxed{\Delta_\alpha(\mathcal{O})\Delta_\alpha(\mathcal{P}) \geqslant \frac{1}{2}\big|\langle[\,\mathcal{O},\,\mathcal{P}\,]\rangle_\alpha\big|} \tag{1.100}$$

Une condition suffisante pour obtenir l'égalité entre le membre de gauche et de droite est que $|\alpha\rangle$ soit un vecteur propre de $\mathcal{O}$ ou $\mathcal{P}$, ou bien que

$$\exists\lambda \in \mathbb{R}^\star\,,\ (\mathcal{P} - \langle\mathcal{P}\rangle_\alpha)\,|\alpha\rangle = \mathrm{i}\lambda(\mathcal{O} - \langle\mathcal{O}\rangle_\alpha)\,|\alpha\rangle\ . \tag{1.101}$$

Preuve. Pour démontrer cette relation, on introduit tout d'abord les opérateurs « centrés » associés à $\mathcal{O}$ et $\mathcal{P}$:

$$\mathcal{O}_\alpha \overset{\text{déf.}}{=} \mathcal{O} - \langle\mathcal{O}\rangle_\alpha\ , \quad \mathcal{P}_\alpha \overset{\text{déf.}}{=} \mathcal{P} - \langle\mathcal{P}\rangle_\alpha\,, \tag{1.102}$$

On peut alors écrire

$$\Delta_\alpha(\mathcal{O}) = \|\,\mathcal{O}_\alpha\,|\alpha\rangle\,\|\ , \quad \Delta_\alpha(\mathcal{P}) = \|\,\mathcal{P}_\alpha\,|\alpha\rangle\,\|\ . \tag{1.103}$$

L'inégalité de Cauchy–Schwartz, éqn. (1.17) donne :

$$\big|\iota\,(\mathcal{O}_\alpha\,|\alpha\rangle\,,\,\mathcal{P}_\alpha\,|\alpha\rangle)\,\big| \leqslant \|\,\mathcal{O}_\alpha\,|\alpha\rangle\,\|\,\|\,\mathcal{P}_\alpha\,|\alpha\rangle\,\|\ , \tag{1.104}$$

et permet d'obtenir grâce à l'éqn. (1.103)

$$\big|\iota\,(\mathcal{O}_\alpha\,|\alpha\rangle\,,\,\mathcal{P}_\alpha\,|\alpha\rangle)\,\big| \leqslant \Delta_\alpha(\mathcal{O})\Delta_\alpha(\mathcal{P})\,. \tag{1.105}$$

On décompose ensuite le produit $\mathcal{O}_\alpha\mathcal{P}_\alpha$ comme

$$\mathcal{O}_\alpha\mathcal{P}_\alpha = \frac{1}{2}\,[\,\mathcal{O}_\alpha,\,\mathcal{P}_\alpha\,] + \frac{1}{2}\,\{\,\mathcal{O}_\alpha,\,\mathcal{P}_\alpha\,\}\,, \tag{1.106}$$

où nous avons introduit l'*anti-commutateur*

$$\{\,\mathcal{O}_\alpha,\,\mathcal{P}_\alpha\,\} \overset{\text{déf.}}{=} \mathcal{O}_\alpha\mathcal{P}_\alpha + \mathcal{P}_\alpha\mathcal{O}_\alpha\,. \tag{1.107}$$

Les opérateurs $\mathcal{O}_\alpha$ et $\mathcal{P}_\alpha$ étant hermitiens, $\{\,\mathcal{O}_\alpha,\,\mathcal{P}_\alpha\,\}$ et $\mathrm{i}\,[\,\mathcal{O}_\alpha,\,\mathcal{P}_\alpha\,]$ le sont également. On peut donc écrire

$$\big|\iota\,(\mathcal{O}_\alpha\,|\alpha\rangle\,,\,\mathcal{P}_\alpha\,|\alpha\rangle)\,\big|^2 = \big|\,\langle\alpha|\,\mathcal{O}_\alpha\mathcal{P}_\alpha\,|\alpha\rangle\,\big|^2 = \frac{1}{4}\big|\,\langle\alpha|\,[\,\mathcal{O}_\alpha,\,\mathcal{P}_\alpha\,]\,|\alpha\rangle\,\big|^2 + \frac{1}{4}\big|\,\langle\alpha|\{\,\mathcal{O}_\alpha,\,\mathcal{P}_\alpha\,\}|\alpha\rangle\,\big|^2\,, \tag{1.108}$$

où on a utilisé le fait que la forme quadratique associée à un opérateur hermitien est réelle, voir éqn. (1.42), et décomposé la norme carrée dans l'équation (1.108) en contributions de la partie réelle et imaginaire.

En remarquant que $[\mathcal{O}_\alpha, \mathcal{P}_\alpha] = [\mathcal{O}, \mathcal{P}]$ et en combinant cette dernière relation avec l'inégalité (1.105), on obtient la relation d'incertitude (1.100).

Pour prouver la seconde partie du théorème, on remarque premièrement que si $|\alpha\rangle$ est un vecteur propre de $\mathcal{O}$ (resp. de $\mathcal{P}$) on a $\mathcal{O}_\alpha|\alpha\rangle = 0$ (resp. $\mathcal{P}_\alpha|\alpha\rangle = 0$) et dans les deux cas de figure $\langle\alpha|[\mathcal{O}, \mathcal{P}]|\alpha\rangle = 0$, les opérateurs étant hermitiens. Les membre de gauche et de droite de (1.100) sont donc tous les deux nuls.

Dans le cas où $|\alpha\rangle$ n'est ni un vecteur propre de $\mathcal{O}$ ni un vecteur propre de $\mathcal{P}$, pour que l'inégalité de Cauchy–Schwarz (1.104) soit une égalité il faut que les deux vecteurs soient colinéaires, soit

$$\exists z \in \mathbb{C}^\star\,,\ \mathcal{P}_\alpha|\alpha\rangle = z\mathcal{O}_\alpha|\alpha\rangle\ . \tag{1.109}$$

Il faut ensuite que le dernier terme dans l'éqn. (1.108) soit nul, soit

$$\langle\alpha|\{\mathcal{O}_\alpha, \mathcal{P}_\alpha\}|\alpha\rangle = 0\,. \tag{1.110}$$

Si on injecte (1.109) dans cette dernière égalité, on obtient $(z + \bar{z})\,\|\,\mathcal{O}_\alpha|\alpha\rangle\,\| = 0$, et donc finalement que $\mathrm{Re}\,(z) = 0$.

□

Dans cette formulation générale du principe d'incertitude, la borne inférieure de l'inégalité (1.100) dépend explicitement de l'état $|\alpha\rangle$ dans lequel se trouve le système. Il se peut donc très bien, si $|\alpha\rangle$ est dans le noyau de l'opérateur $[\mathcal{O}, \mathcal{P}]$, que l'inégalité prenne la forme $\Delta_\alpha(\mathcal{O})\Delta_\alpha(\mathcal{P}) \geqslant 0$ même si les opérateurs ne commutent pas entre eux, et n'induise donc aucune contrainte. [14]

À retenir

— L'espace des états de la mécanique quantique est associé à un espace de Hilbert séparable qui, s'il est de dimension finie, se réduit à un espace hermitien ;

— La norme des vecteurs d'états n'ayant pas de signification physique, l'espace des états a la structure d'un espace projectif composé de rayons ; On choisit en général pour chaque rayon un représentant de norme unité ;

— à chaque vecteur de cet espace on peut associer une forme linéaire bornée, et vice-versa. Cette correspondance est à l'origine du formalisme « bra-kets » de Dirac ;

— Les grandeurs physiques sont associées à des *observables*, opérateurs linéraires auto-adjoints sur l'espace des états. Différence entre opérateurs hermitiens et auto-adjoints. De tels opérateurs possèdent un spectre réel ;

14. En revanche, si le commutateur est proportionnel à l'identité, le membre de droite de l'éqn. (1.100) est indépendant de l'état du système en considérant que le vecteur $|\alpha\rangle$ correspondant est de norme unité.

— Pour des opérateurs auto-adjoints sur un espace hermitien, le spectre se compose de valeurs propres réelles associées à des vecteurs propres et il est possible de construire une base orthonormée de vecteurs propres sur l'espace des états ;

— Premier postulat de la mesure : la probabilité de que le résultat de la mesure d'une observables corresponde à une certaine valeur propre est le module carré de la projection orthogonale du vecteur d'état sur le sous-espace propre correspondant ;

— Second postulat de la mesure : après une mesure ayant donné une certaine valeur propre, le vecteur d'état est le projeté orthogonal du vecteur d'état d'avant la mesure sur le sous-espace propre associé ;

— Principe d'incertitude.

Appendices

1.A Séries de Fourier et espace de Hilbert des fonctions de carré sommable

Considérons, pour illustrer la notion de base d'un espace de Hilbert, l'espace de Hilbert $\mathcal{L}^2([0,1])$ des fonctions $|f\rangle$ d'une variable réelle de carré sommable sur l'intervalle $[0,1]$, muni du produit hermitien

$$\langle f|g\rangle = \int_0^1 dx\, \overline{f(x)} g(x)\,, \tag{1.111}$$

d'où on tire la norme $\|\quad\|_{\mathcal{L}^2}$ définie par

$$\| f \|_{\mathcal{L}^2} = \langle f|f\rangle^{1/2} = \left(\int_{[0,1]} |f|^2 \right)^{1/2}. \tag{1.112}$$

Un développement en série de Fourier de $f \in \mathcal{L}^2([0,1])$ consiste à exprimer toute fonction de cet espace sur une base d'exponentielles complexes :

$$\{ |n\rangle \leftrightarrow e_n : x \mapsto e^{2i\pi n x}\,,\ n \in \mathbb{Z} \}\,, \quad \langle n|m\rangle = \int_0^1 dx\, e^{2i\pi(m-n)x} = \delta_{n,m}\,. \tag{1.113}$$

L'orthonormalité de la famille $\{|n\rangle\,, n \in \mathbb{Z}\}$ est évidente, pour montrer qu'elle forme une base orthonormée (et dénombrable) de $\mathcal{L}^2([0,1])$ il faudrait en plus montrer que l'ensemble des combinaisons linéaires des $|n\rangle$ est dense dans $\mathcal{L}^2([0,1])$, c.-à.-d. que tout voisinage d'une fonction de carré sommable, aussi petit soit-il, contient au moins une série de Fourier.

Les coefficients de Fourier $\{\alpha_n,\, n \in \mathbb{Z}\}$ associés à une fonction $|f\rangle \in \mathcal{L}^2([0,1])$ sont donnés par la projection de f sur les les vecteurs de base, soit

$$\alpha_n = \langle n|f\rangle = \int_0^1 dx\, \overline{e_n(x)} f(x)\,. \tag{1.114}$$

Naïvement, on souhaiterait identifier toute fonction $f \in \mathcal{L}^2([0,1])$ avec son développement en série de Fourier :

$$\forall x \in [0,1]\,,\ f(x) \stackrel{?}{=} \sum_{n\in\mathbb{Z}} \alpha_n e_n(x)\,. \tag{1.115}$$

Cette égalité « ponctuelle » n'est pas vraie pour toute fonction $f \in \mathcal{L}^2([0,1])$, mais seulement avec des hypothèses supplémentaires de continuité. En revanche on peut toujours affirmer que la série de Fourier converge vers f au sens de la norme de $\mathcal{L}^2([0,1])$, c.-à.-d.

$$\forall \epsilon > 0\,,\quad \exists M \in \mathbb{N}\,,\quad \forall N > M\,,\quad \left\| f - \sum_{n=-N}^{N} \alpha_n e_n \right\|_{\mathcal{L}^2} < \epsilon\,. \tag{1.116}$$

C'est dans ce sens qu'il faut comprendre la notion de développement d'un vecteur de l'espace de Hilbert $\mathcal{L}^2([0,1])$ sur la base (1.113) lorsqu'on écrit

$$|f\rangle = \sum_{n\in\mathbb{Z}} |n\rangle\,\langle n|f\rangle\ . \tag{1.117}$$

1.B Expérience de Stern et Gerlach

L'expérience de Stern et Gerlach, pensée par Stern et réalisée avec Gerlach en 1922 [2] a joué un rôle important dans la compréhension de la mécanique quantique, et constitue un point de repère très utile pour comprendre la notion de mesure en mécanique quantique.

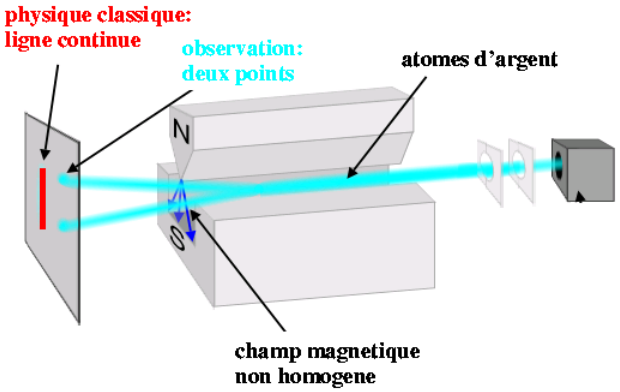

FIGURE 1.1 – Schéma de l'expérience de Stern et Gerlach (source : Wikipedia Commons)

Commençons par une description du dispositif expérimental, voir fig. 1.1. Des atomes d'argent sont chauffés dans un four, à droite sur le schéma, et une petite ouverture laisse sortir un faisceau d'atomes qui est collimaté par une série de diaphragmes. Dans l'état électronique fondamental, l'atome d'argent porte un moment magnétique du au spin 1/2 du $47^{\text{ième}}$ des électrons que comporte l'atome (les 46 autres n'apportant pas de contribution nette au spin total). Comme nous le verrons, l'opérateur moment magnétique $\vec{\mu}$ est proportionnel à l'opérateur de spin $\vec{S}$:

$$\vec{\mu} = -\frac{\mu_B}{\hbar}\vec{S}\,, \tag{1.118}$$

où la constante positive μ_B est appelée magnéton de Bohr. L'atome dans l'état fondamental possède donc, outre ses degrés de liberté associés à sa trajectoire dans l'espace, des degrés de liberté associés au spin de son électron externe.

Le faisceau d'atomes passe ensuite dans un champ magnétique $\vec{B}$ inhomogène, orienté selon l'axe Oz d'un repère orthonormé de l'espace et variant selon Oz également. Classiquement, un moment magnétique $\vec{\mu}$ dans un champ magnétique inhomogène subit une force $F = \vec{\nabla}\left(\vec{\mu}\cdot\vec{B}\right) = \mu_z\,(\partial_z B)\,\vec{e}_z$.

Si on place un écran perpendiculaire au faisceau, après son passage par la cavité contenant le champ magnétique, on va donc mesurer une déviation du faisceau proportionnelle au moment magnétique selon l'axe Oz des atomes d'argent : cette déviation s'effectuera vers le haut ou vers le bas suivant que μ_z soit respectivement négatif ou positif. La mesure de la position des atomes d'argent frappant l'écran est une mesure de leur moment magnétique, et donc de leur spin électronique.

Dans un cadre de physique classique, on assimilerait le spin $\vec{S}$ de l'électron à un moment cinétique ordinaire, et S_z, projection de $\vec{S}$ sur l'axe O_z, pourrait prendre toutes les valeurs réelles entre une valeur maximale $\|\vec{S}\|$ et une valeur minimale $-\|\vec{S}\|$, ce qui donnerait un continuum de points sur l'écran après la détection de beaucoup d'atomes, voir la figure.

En mécanique quantique, l'espace des états réduit associé aux seuls degrés de liberté de spin pour une particule de spin 1/2 est de dimension 2, et il existe une base dans laquelle l'opérateur $\mathcal{S}_z$ prend la forme matricielle diagonale suivante :

$$\mathcal{S}_z \mathrel{\widehat{=}} \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.119}$$

Cette matrice possède donc deux vecteurs propres normés, respectivement, en notation matricielle :

$$\begin{aligned} |+\rangle_z &\mathrel{\widehat{=}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathcal{S}_z|+\rangle_z = \frac{\hbar}{2}|+\rangle_z, \\ |-\rangle_z &\mathrel{\widehat{=}} \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \mathcal{S}_z|-\rangle_z = -\frac{\hbar}{2}|-\rangle_z, \end{aligned} \tag{1.120}$$

À chacun de ces états propres va correspondre une certaine déviation dans le champ magnétique ; ainsi on n'observera que deux taches localisées sur l'écran au lieu d'un continuum de points. Ce résultat est une preuve expérimentale de la validité de la mécanique quantique.

En ne prenant en compte que les degrés de liberté de spin, un vecteur d'état normé quelconque, correspondant à un atome d'argent sortant du four, est de la forme (en choisissant la phase globale de manière appropriée) :

$$|\psi\rangle = \cos\theta\,|+\rangle_z + e^{i\phi}\sin\theta\,|-\rangle_z,\ \theta \in [0,\pi/2[,\ \phi \in [0,2\pi[. \tag{1.121}$$

Suivant le formalisme développé dans ce chapitre, nous pouvons premièrement calculer la

valeur moyenne du spin selon l'axe Oz dans cet état :

$$\begin{aligned}\langle \mathcal{S}_z \rangle_\psi &= \langle\psi|\, \mathcal{S}_z\, |\psi\rangle = \left(\cos\theta\ {}_z\langle +| + e^{-i\phi}\sin\theta\ {}_z\langle -|\right) \mathcal{S}_z \left(\cos\theta\, |+\rangle_z + e^{i\phi}\sin\theta\, |+\rangle_z\right) \\ &= \frac{\hbar}{2}\left(\cos\theta\ {}_z\langle +| + e^{-i\phi}\sin\theta\ {}_z\langle -|\right)\left(\cos\theta\, |+\rangle_z - e^{i\phi}\sin\theta\, |+\rangle_z\right) \\ &= \frac{\hbar}{2}\left(\cos^2\theta - \sin^2\theta\right) = \frac{\hbar}{2}\cos 2\theta\,. \end{aligned} \tag{1.122}$$

Comme on pouvait s'y attendre, si un atome est préparé dans l'état $|+\rangle_z$ (resp. $|-\rangle_z$) en choisissant $\theta = 0$ (resp. $\theta = \pi/2$) la valeur moyenne obtenue est $+\hbar/2$ (resp. $-\hbar/2$).

Deuxièmement, les probabilités de mesurer les valeurs $\pm\hbar/2$ pour le spin sont données respectivement par :

$$p(+) = \|\, \mathcal{P}_+ |\psi\rangle \,\|^2 = \|\, |+\rangle_z\ {}_z\langle +|\psi\rangle \,\|^2 = |\, {}_z\langle +|\psi\rangle\,|^2 = \cos^2\theta\,, \tag{1.123a}$$

$$p(-) = \|\, \mathcal{P}_- |\psi\rangle \,\|^2 = \|\, |-\rangle_z\ {}_z\langle -|\psi\rangle \,\|^2 = |\, {}_z\langle -|\psi\rangle\,|^2 = \sin^2\theta\,. \tag{1.123b}$$

Le dispositif de Stern et Gerlach se prête également à des expériences mettant en jeu des mesures successives, comme présenté vers l'éqn. (1.97). Il suffit pour cela d'imaginer qu'on perce une ouverture dans l'écran de la figure 1.1, au niveau de la tache correpondant à un des états possibles de spin. Si par exemple l'ouverture se situe dans la partie haute de l'écran, tous les atomes sortant du dispositif seront dans l'état de spin $|-\rangle_z$, suivant le postulat 5 (réduction du paquet d'ondes).

Bibliographie

[1] P. A. M. Dirac, "The Principles of Quantum Mechanics," Oxford University Press, 1930.

[2] W. Gerlach and O. Stern, "Der experimentelle Nachweis der Richtungsquantelung im Magnetfeld," Zeitschrift für Physik **9** (1922) 349. doi :10.1007/BF01326983

Cours n°2

Fonctions d'onde et paquets d'ondes

Dans ce chapitre le formalisme mis en place au cours n°1 sera appliqué aux opérateurs position et impulsion, faisant le lien entre l'approche abstraite fondée sur les espace de Hilbert et l'approche plus concrète à la mécanique quantique en termes de fonctions d'onde. Nous étudierons ensuite en détail deux exemples, la particule libre et l'oscillateur harmonique.

2.1 L'espace de Hilbert $\mathcal{L}^2(\mathbb{R}^3)$

Dans l'approche de Schrödinger une particule massive sans spin est décrite par une fonction d'onde :

$$\psi \; : \; \left\{ \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{C} \\ \vec{x} & \mapsto & \psi(\vec{x}) \end{array} \right. \tag{2.1}$$

Afin que cette description soit compatible avec le formalisme général décrit au cours n°1, l'espace de fonctions auquel appartient ψ doit avoir la structure d'un espace de Hilbert séparable. Pour simplifier les notations, nous considérerons dans la plupart des développements qui suivent des fonctions d'onde à une dimension, c.-à.-d. des fonctions $\mathbb{R} \to \mathbb{C}$; la généralisation ne comporte aucune difficulté particulière.

Par hypothèse, les fonctions d'ondes sont des éléments de $\mathcal{L}^2(\mathbb{R})$, espace des fonctions de $\mathbb{R}$ dans $\mathbb{C}$ de norme finie par rapport à la norme $\mathcal{L}^2$

$$\| \, |\psi\rangle \, \| \stackrel{\text{déf.}}{=} \sqrt{\int \mathrm{d}x \; |\psi \, (x)|^2} \, , \tag{2.2}$$

découlant du produit hermitien

$$\langle \phi | \psi \rangle \stackrel{\text{déf.}}{=} \int \mathrm{d}x \, \overline{\phi \, (x)} \psi \, (x) \; . \tag{2.3}$$

On peut montrer que cet espace satisfait bien à toutes le propriétés d'un espace de Hilbert.

Rappelons qu'un espace de Hilbert est séparable si et seulement si toute base orthonormée de cet espace est dénombrable. Une base dénombrable de $\mathcal{L}^2(\mathbb{R})$ est obtenue de la manière

suivante. On définit tout d'abord les *polynômes de Hermite* par la relation :

$$H_n(x) \overset{\text{déf.}}{=} (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \ , \quad n \in \mathbb{N} \, . \tag{2.4}$$

Les trois premiers de ces polynômes sont donnés par $H_0(x) = 1$, $H_1(x) = 2x$ et $H_2(x) = 4x^2 - 2$. Ils satisfont à la relation d'orthogonalité [1] :

$$\forall n, n' \in \mathbb{N} \ , \quad \int dx \, e^{-x^2} \ H_n(x) H_{n'}(x) = \sqrt{\pi} 2^n n! \delta_{n,n'} \, . \tag{2.5}$$

Cette relation nous indique comment construire une base de $\mathcal{L}^2(\mathbb{R})$. On définit les *fonctions de Hermite* comme :

$$\psi_n(x) = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} \, e^{-\frac{x^2}{2}} H_n(x) \ , \quad n \in \mathbb{N} \, . \tag{2.6}$$

À l'aide de l'identité (2.5) on obtient :

$$\langle \psi_n | \psi_{n'} \rangle = \int dx \, , \overline{\psi}_n(x) \psi_{n'}(x) = \delta_{n,n'} \, , \tag{2.7}$$

et les fonctions de Hermite forment une famille orthonormée de l'espace de Hilbert $\mathcal{L}^2(\mathbb{R})$. On montre que cette famille est dense dans $\mathcal{L}^2(\mathbb{R})$, dont elle constitue donc une base. Cette base étant dénombrable, l'espace de Hilbert est séparable.

On peut remarquer que les fonctions de Hermite (2.6) sont solutions de l'équation différentielle :

$$\frac{d^2}{dx^2} \psi_n(x) + \left(2n + 1 - x^2 \right) \psi_n(x) = 0 \, , \tag{2.8}$$

qui apparaît dans le contexte de l'oscillateur harmonique, voir l'appendice 2.B.

Revenons pour finir au cas de l'espace des fonctions d'onde dans l'espace à trois dimensions $\mathcal{L}^2(\mathbb{R}^3)$. Cet espace est également un espace de Hilbert séparable, muni du produit hermitien

$$\langle \phi | \psi \rangle \overset{\text{déf.}}{=} \int d^3x \, \overline{\phi \left(\vec{x} \right)} \psi \left(\vec{x} \right) \ , \tag{2.9}$$

et une base orthonormée de cet espace est simplement fournie par le produit de fonctions de Hermite pour les coordonnées cartésiennes :

$$\Psi_{n_1 n_2 n_3}(\vec{x}) = \psi_{n_1}(x) \psi_{n_2}(y) \psi_{n_3}(z) \, . \tag{2.10}$$

Du point de vue mathématique, il serait naturel d'utiliser cette base de fonctions pour étudier les fonctions d'onde d'une particule massive. Néanmoins comme nous le verrons à la section 2.3.2, cette base n'est pas adaptée à tous les systèmes physiques, n'étant pas en particulier constituée de vecteurs propres de l'opérateur hamiltonien d'une particule libre.

1. Remarquons que la mesure d'intégration sur $\mathbb{R}$ est ici donnée la mesure gaussiennne $e^{-x^2} dx$ et non la mesure habituelle dx.

2.2 Représentation de position

Une fonction d'onde ψ est associée à un vecteur $|\psi\rangle$ de l'espace de Hilbert $\mathcal{L}^2(\mathbb{R}^3)$, muni du produit hermitien (2.3). Pour simplifier les notations, la discussion sera restreinte aux fonctions d'onde à une dimension, éléments de $\mathcal{L}^2(\mathbb{R})$, la généralisation ne posant pas de difficulté conceptuelle.

2.2.1 Distribution de Dirac

On aimerait définir une forme linéaire sur $\mathcal{L}^2(\mathbb{R})$ qui permette d'obtenir comme résultat la valeur de la fonction d'onde en un point x donné ; l'objectif est de pouvoir donner un sens, in fine, à la probabilité de présence en mécanique quantique. On pose

$$\boxed{\langle x| \ : \quad \left\{ \begin{array}{ccc} \mathcal{L}^2(\mathbb{R}) & \to & \mathbb{C} \\ |\psi\rangle & \mapsto & \langle x|\psi\rangle \overset{\text{déf.}}{=} \psi(x)\,. \end{array} \right.} \tag{2.11}$$

La forme linéaire satisfaisant cette relation pour toute fonction est la *distribution de Dirac*. Au lieu de la notation (2.11) on utilise également $\delta_x(\psi) = \psi(x)$, ou en notation de Dirac $\langle \delta_x|\psi\rangle = \psi(x)$.

Remarquons que cette forme linéaire n'est pas bornée, au sens de l'équation (1.18), car même si $|\psi\rangle$ est de norme unité par rapport à la norme $\mathcal{L}^2$ la fonction $\psi(x)$ peut prendre des valeurs arbitrairement grandes. Elle n'appartient donc pas à $\mathcal{L}^2(\mathbb{R})^\star$, espace dual de $\mathcal{L}^2(\mathbb{R})$ comme défini dans la sous-section 1.1.2. Il faut donc travailler avec un espace de formes linéaires plus grand que $\mathcal{L}^2(\mathbb{R})^\star$ (au sens de l'inclusion) ce qui nous conduit à considérer un espace de fonction plus petit que $\mathcal{L}^2(\mathbb{R})$ sur lequel les formes linéaires agissent.

La définition des distributions [2] (plus précisément des *distributions tempérées*) fait appel à l'espace de Schwartz $\mathcal{S}(\mathbb{R})$ des fonctions infiniment dérivables à décroissance rapide :

$$\mathcal{S}(\mathbb{R}) \overset{\text{déf.}}{=} \left\{ f \in \mathcal{C}^\infty(\mathbb{R}),\, \forall \alpha, \beta \in \mathbb{N}^*,\, \sup_{x\in\mathbb{R}} \left| x^\alpha \frac{\mathrm{d}^\beta f(x)}{\mathrm{d}x^\beta} \right| < \infty \right\}, \tag{2.12}$$

dont font partie les fonctions de Hermite (2.6). Ce sous-espace de $\mathcal{L}^2(\mathbb{R})$ n'est pas un espace de Hilbert – sinon il existerait une bijection avec son dual $\mathcal{S}(\mathbb{R})^\star$, l'espace des distributions dont fait partie la distribution de Dirac – néanmoins il est dense dans $\mathcal{L}^2(\mathbb{R})$.

Pour de nombreux calculs de mécanique quantique, laissant de côté la rigueur mathématique, il est utile d'interpréter cette relation comme

$$\psi(x) = \langle x|\psi\rangle \overset{?}{=} \int \mathrm{d}u\, \delta_x(u)\psi(u)\,, \tag{2.13}$$

pour une certaine « fonction » δ_x (le complexe conjugué peut être omis car la distributions de Dirac est réelle). Cependant, l'expression (2.117) n'a pas de sens du point de vue de la théorie de l'intégration, que ce soit l'intégrale de Riemann ou l'intégrale de Lebesgue ; il faudrait une fonction identiquement nulle pour tout $u \neq x$ donnant cependant une intégrale non-nulle ;

d'une certaine manière la valeur en x devrait être infinie.[2] On utilise également la notation légèrement différente :

$$\int du\, \delta(x-u)\psi(u) = \psi(x)\,. \tag{2.14}$$

Ce point de vue peu rigoureux permet également d'obtenir simplement plusieurs propriétés (correctes !) de la distribution de Dirac. Par exemple,

$$\int dy\, \delta'(y)\, \psi(y) = -\psi'(0)\,, \tag{2.15}$$

évoquant une intégration par parties, ou bien

$$\forall a \in \mathbb{R}^*\,, \quad \int du\, \delta(au)\psi(u) = \frac{1}{|a|}\psi(0)\,, \tag{2.16}$$

évoquant un changement de variables.

2.2.2 Opérateur position

Au vecteur de coordonnées $\vec{x}$ de la physique classique, on fait correspondre l'opérateur position, un triplet d'observables

$$\vec{X} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}\,, \tag{2.17}$$

qui sont des applications linéaires auto-adjointes sur l'espace de Hilbert $\mathcal{H}$. Ces opérateurs commutent naturellement entre eux, suivant le processus de quantification canonique, comme leurs crochets de Poisson sont nuls :

$$[X, Y] = [X, Z] = [Y, Z] = 0\,. \tag{2.18}$$

Pour donner un sens à la mesure de la position de la particule associée à cet espace des états, on souhaiterait naïvement construire une base de vecteurs propres communs à ces opérateurs, grâce aux théorèmes 5 et 9 du chapitre 1. Cependant, cet opérateur n'est pas borné et les hypothèses du théorème 5 ne sont pas satisfaites, on ne peut a priori construire une base orthonormée de $\mathcal{L}^2(\mathbb{R}^3)$ constituée de vecteurs propres de l'opérateur position $\vec{X}$.[3]

2. On peut cependant construire une famille de fonctions à un paramètre, de plus en plus « piquées » lorsque ce paramètre ϵ tend vers zéro, approchant la distribution de Dirac ; par exemple, la famille de fonctions gaussiennes $f_\epsilon : x \mapsto \frac{1}{\epsilon\sqrt{\pi}} e^{-x^2/\epsilon^2}$. Une telle famille de fonctions appartient à $\mathcal{L}^2(\mathbb{R})$ et, $\mathcal{L}^2(\mathbb{R})$ étant un espace de Hilbert et donc isomorphe à son dual $\mathcal{L}^2(\mathbb{R})^\star$, on peut associer à ces fonctions des formes linéaires $\psi \mapsto \int dx\, f_\epsilon(x)\psi(x)$.

3. On peut objecter que sur un intervalle fermé de $\mathbb{R}$ (correspondant à une particule 1d dans une boîte) l'opérateur position est borné ; cependant l'opérateur position n'est de toute manière pas compact et la conclusion est inchangée.

Pour simplifier les notations nous nous concentrons sur un opérateur de position à une dimension. Dans le cours n° 1, example 2 nous avons étudié le spectre de l'opérateur position et montré que

$$\sigma(\mathcal{X}) = \mathbb{R}\,, \tag{2.19}$$

en accord avec l'intuition physique qui suggère que les valeurs propres de $\mathcal{X}$ correspondent aux positions possibles de la particule. Bien que le spectre de l'opérateur position soit non vide, il n'admet aucun vecteur propre ! Il faudrait en effet considérer une équations aux vecteurs propres de la forme $x\psi(x) = \lambda\psi(x)$ avec $\psi \in \mathcal{L}^2(\mathbb{R})$.

Il est licite de considérer des « vecteurs propres » de $\mathcal{X}$, non pas dans l'espace de fonctions $\mathcal{L}^2(\mathbb{R})$, mais dans l'espace des distributions $\mathcal{S}(\mathbb{R})^\star$ précédemment présenté, c.-à.-d. l'espace des formes linéaires sur l'espace de fonctions de Schwartz $\mathcal{S}(\mathbb{R})$. Ces vecteurs propres ne sont rien d'autre que les distributions de Dirac définies par l'éqn. (2.11) :

$$\langle x|\,\mathcal{X} = x\,\langle x|\ , \tag{2.20}$$

comme l'indique la relation suivante :

$$\boxed{\langle x|\,\mathcal{X}\,|\psi\rangle = x\psi(x)} \tag{2.21}$$

c.-à.-d. que l'action de l'opérateur position sur la fonction d'onde $\psi(x)$ revient [4] à la multiplication par x.

Reprenant le formalisme mathématiquement peu rigoureux – mais fort pratique – utilisé pour écrire la relation (2.117), nous définirons une famille d'« états propres de position » $\{|x\rangle\,, x \in \mathbb{R}\}$ par la propriété : [5]

$$\mathcal{X}\,|x\rangle = x\,|x\rangle\ \ , \quad x \in \mathbb{R}\,. \tag{2.22}$$

Ces kets ne sont pas des éléments de l'espace des états $\mathcal{L}^2(\mathbb{R})$ – ce ne sont même pas des fonctions – et ne peuvent donc représenter un état physique d'un système quantique. L'observable position étant auto-adjointe,

$$\langle x'|\,\mathcal{X}\,|x\rangle = x\,\langle x'|x\rangle = x'\,\langle x'|x\rangle\ , \tag{2.23}$$

suivant que l'on choisisse l'action à gauche ou à droite. Cette relation n'a un sens (du moins l'approche peu rigoureuse employée) que si

$$\boxed{\langle x'|x\rangle = \delta(x' - x)} \tag{2.24}$$

4. Notons que la fonction $x \mapsto x\psi(x)$ n'est pas nécessairement dans $\mathcal{L}^2(\mathbb{R})$ mais que l'espace de Schwartz $\mathcal{S}(\mathbb{R}) \subset \mathcal{L}^2(\mathbb{R})$ est stable par cette opération.

5. Il existe un formalisme mathématique permettant de définir de tels objets, mais sortant du cadre des espaces de Hilbert.

C'est dans ce sens que l'on peut considérer $\{|x\rangle\ x \in \mathbb{R}\}$ comme une « base orthonormée continue » de l'espace de Hilbert $\mathcal{L}^2(\mathbb{R})$ des fonctions d'onde, bien que ces fonctions n'en fassent pas partie et que toute base orthonormée de $\mathcal{L}^2(\mathbb{R})$ soit dénombrable. [6]

On peut alors réinterpréter l'expression intégrale (2.14) pour la distribution de Dirac comme

$$\psi(x) = \langle x|\psi\rangle = \int du\, \delta(x-u)\, \psi(u) = \int du\, \langle x|u\rangle\, \langle u|\psi\rangle = \langle x| \left(\int du\, |u\rangle\, \langle u|\psi\rangle \right) , \qquad (2.26)$$

l'expression entre parenthèses du dernier terme ressemblant au développement sur une base orthonormée en notation de Dirac, voir éqn. (1.25). Cela conduit naturellement à la version continue de la relation de fermeture (1.61), donnée par

$$\boxed{\mathfrak{I} = \int dx\ |x\rangle\, \langle x|} \qquad (2.27)$$

Ces relations, bien qu'incorrectes mathématiquement, peuvent être vues comme des « raccourcis » permettant d'obtenir simplement des identités correctes, alors que l'utilisation d'un formalisme rigoureux serait souvent plus fastidieuse.

Étendant le postulat 3 du cours n°1 aux observables à spectre continu, on peut par exemple dériver la valeur moyenne de la position d'une particule dans un état $|\psi\rangle$ comme :

$$\begin{aligned} \langle \vec{X} \rangle_\psi &\overset{\text{déf.}}{=} \langle\psi|\, \vec{X}\, |\psi\rangle = \langle\psi| \left(\int d^3x\ |\vec{x}\rangle\, \langle\vec{x}| \right) \vec{X}\, |\psi\rangle = \int d^3x\, \langle\psi|\vec{x}\rangle\, \vec{x}\, \langle\vec{x}|\psi\rangle \\ &= \int d^3x\ |\psi(\vec{x})|^2\, \vec{x}\,, \end{aligned} \qquad (2.28)$$

si l'intégrale correspondante converge. On peut de même définir les moments d'ordre supérieur de la distribution de probabilité associée, par exemple la variance (1.99).

2.2.3 Densité de probabilité de présence

Suivant le postulat 4 du cours n°1, la probabilité de mesurer la valeur propre λ_n pour une observable est donnée par le carré du module du projeté du vecteur d'état $|\psi\rangle$ sur le sous-espace propre correspondant ; si le sous-espace propre est de dimension 1, notant $|n\rangle$ un vecteur propre normé correspondant, cette probabilité est $p(\lambda_n) = |\langle n|\psi\rangle|^2$. Plus généralement, on a $p(\lambda_n) = \| \Pi_{\lambda_n} |\psi\rangle \|^2$ où Π_{λ_n} est le projecteur orthogonal sur le sous-espace propre correspondant.

6. On pourrait imaginer une autre manière d'introduire les « vecteurs de base » de la représentation de position. On définirait la famille $\{|x\rangle\ x \in \mathbb{R}\}$ par la propriété :

$$\forall x \in \mathbb{R}\ , \quad \forall\, |\psi\rangle \in \mathcal{L}^2(\mathbb{R})\ , \quad \langle\psi|x\rangle = \overline{\langle x|\psi\rangle} = \overline{\psi(x)}\,, \qquad (2.25)$$

ce qui induirait cependant les mêmes difficultés mathématiques (de ce point de vue, on remarque déjà que $\mathcal{S}(\mathbb{R})$ n'est pas un espace de Hilbert, et qu'il n'est pas isomorphe à son dual).

Étant donné que le bra $\langle\vec{x}|$ associé à la distribution de Dirac centrée en $\vec{x}$ correspond à un vecteur propre de l'opérateur position $\vec{\mathcal{X}}$ (strictement parlant, dans l'espace $\mathcal{S}(\mathbb{R}^3)^\star$ des distributions), il est naturel d'étendre ce postulat de la mesure à la position d'une particule.

Cependant, de même que, considérant une variable aléatoire continue $X \in [0,1]$, la probabilité $p(X = 1/2)$ de tirer exactement la valeur $1/2$ est nulle, la probabilité de trouver la particule « exactement » au point de coordonnées $\vec{x}$ est nulle. On peut en revanche associer au module carré de la fonction d'onde $|\psi(\vec{x})|^2$ une densité de probabilité, ce qui mène au postulat suivant :

Postulat 1 (densité de probabilité de présence).
Soit une particule décrite par le vecteur d'état $|\psi\rangle \in \mathcal{L}^2(\mathbb{R}^3)$ *normé (c.-à.-d. tel que* $\langle\psi|\psi\rangle = 1$*) à un instant donné. La probabilité que la particule soit localisée à cet instant dans un domaine* D *de* $\mathbb{R}^3$ *est donnée par :*

$$\boxed{p(D) = \int_D |\langle\vec{x}|\psi\rangle|^2 \, d^3x = \int_D |\psi(\vec{x})|^2 \, d^3x} \tag{2.29}$$

On vérifie que (en « notation de physiciens ») :

$$p(\mathbb{R}^3) = \int_{\mathbb{R}^3} d^3x \, \langle\psi|\vec{x}\rangle \, \langle\vec{x}|\psi\rangle = \langle\psi| \left(\int_{\mathbb{R}^3} d^3x \, |\vec{x}\rangle \, \langle\vec{x}| \right) |\psi\rangle = \langle\psi|\psi\rangle = 1 \, . \tag{2.30}$$

Il existe également un analogue du postulat 5, qui mérite ici son nom de *réduction du paquet d'ondes* :

Postulat 2 (réduction du paquet d'ondes). *Soit une particule décrite par le vecteur d'état normé* $|\psi\rangle \in \mathcal{L}^2(\mathbb{R}^3)$ *à un instant donné. Si le résultat de la mesure de sa position indique qu'elle se trouve dans un domaine* $D \subset \mathbb{R}^3$*, alors le vecteur d'état de la particule est modifié au moment de la mesure suivant :*

$$|\psi\rangle \xrightarrow{\text{détection dans le domaine } D} \frac{1}{\left\| \int_D d^3x \, |x\rangle \, \langle x|\psi\rangle \right\|} \int_D |x\rangle \, \langle x|\psi\rangle \, . \tag{2.31}$$

En termes de fonction d'onde, nous avons

$$\psi(\vec{x}) \longrightarrow \frac{1}{\sqrt{\int_D d^3x \, |\psi(\vec{x})|^2}} \, \delta_D(\vec{x}) \, \psi(\vec{x}) \, , \tag{2.32}$$

où δ_D est la fonction indicatrice de D :

$$\delta_D(\vec{x}) = \begin{cases} 1 & \text{si} \quad \vec{x} \in D \, , \\ 0 & \text{si} \quad \vec{x} \notin D \, . \end{cases} \tag{2.33}$$

Notons que l'expression éqn. (2.32) est mathématiquement correcte. [7]

Naturellement, la fonction d'onde du système va évoluer après l'instant de la mesure, suivant le phénomène d'*étalement du paquet d'ondes* qui sera étudié plus loin.

7. Nous supposerons naturellement que D est mesurable et de mesure non-nulle.

2.3 Représentation d'impulsion

L'espace des phases de la mécanique hamiltonienne est paramétré par les coordonnées généralisées et leurs moments conjugués, les impulsions. La formulation de la mécanique quantique présentée ici étant essentiellement hamiltonienne, il est naturel de considérer, après l'observable position $\vec{\mathcal{X}}$, l'observable $\vec{\mathcal{P}}$ associée à l'impulsion.

Pour une particule dans un volume infini, cette observable possède également, comme nous le verrons, un spectre continu, et son analyse se heurte au même difficultés. Dans une première approche, examinons l'action de cette observables sur la fonction d'onde $\psi(x) = \langle x|\psi\rangle$ associée à un vecteur d'état $|\psi\rangle \in \mathcal{L}^2(\mathbb{R})$.

D'après la procédure de quantification canonique, voir le postulat 1 du cours n°10, l'opérateur d'impulsion $\mathcal{P}$ est caractérisé par son commutateur avec l'opérateur de position $\mathcal{X}$, appelé *relation de commutation canonique* :

$$\boxed{[\,\mathcal{X},\,\mathcal{P}\,] = i\hbar} \tag{2.34}$$

qui devient à trois dimensions

$$\boxed{[\,\mathcal{X}_i,\,\mathcal{P}_j\,] = i\hbar\,\delta_{i,j}} \tag{2.35}$$

On déduit d'une part du commutateur (2.34) la relation :

$$\langle x|\,[\,\mathcal{X},\,\mathcal{P}\,]\,|\psi\rangle = i\hbar\,\langle x|\psi\rangle = i\hbar\psi(x)\,, \tag{2.36}$$

et d'autre part, en utilisant la « relation de fermeture » (2.27) :

$$\begin{aligned}\langle x|\,[\,\mathcal{X},\,\mathcal{P}\,]\,|\psi\rangle &= \langle x|\,\mathcal{X}\mathcal{P}\,|\psi\rangle - \langle x|\,\mathcal{P}\mathcal{X}\,|\psi\rangle \\ &= x\int dx'\,\langle x|\,\mathcal{P}\,|x'\rangle\,\langle x'|\psi\rangle - \int dx'\,\langle x|\,\mathcal{P}\,|x'\rangle\,\langle x'|\,\mathcal{X}\,|\psi\rangle \\ &= \int dx'\,\langle x|\,\mathcal{P}\,|x'\rangle\,\big(x-x'\big)\psi(x')\,. \end{aligned} \tag{2.37}$$

Les « éléments de matrice » de $\mathcal{P}$ dans la base $|x\rangle$ doivent donc satisfaire l'équation intégrale

$$\forall\psi \in \mathcal{L}^2(\mathbb{R})\;,\quad \int dx'\,\langle x|\,\mathcal{P}\,|x'\rangle\,\big(x-x'\big)\psi(x') = i\hbar\psi(x)\,, \tag{2.38}$$

pour tout vecteur d'état $|\psi\rangle$ (avec les réserves mathématiques habituelles). La solution de cette équation est donnée par

$$\boxed{\langle x|\,\mathcal{P}\,|x'\rangle = i\hbar\delta'(x-x')} \tag{2.39}$$

En effet on a bien, à l'aide de l'éqn. (2.15) :

$$\int dx'\,\big(x-x'\big)\psi(x')\delta'(x-x') = -\frac{d}{dx'}\big(x-x'\big)\psi(x')\Big|_{x=x'} = \psi(x)\,. \tag{2.40}$$

On peut ensuite déduire de l'éqn. (2.39) l'action de l'opérateur $\mathcal{P}$ sur les fonctions d'ondes :

$$\boxed{\langle x|\,\mathcal{P}\,|\psi\rangle = -i\hbar\frac{d}{dx}\psi(x)} \tag{2.41}$$

et à trois dimensions

$$\boxed{\langle \vec{x}|\,\vec{\mathcal{P}}\,|\psi\rangle = -i\hbar\vec{\nabla}\psi(\vec{x})} \tag{2.42}$$

L'opérateur d'impulsion est donc proportionnel à l'opérateur de dérivation.

L'étude des vecteurs propres de l'opérateur impulsion peut se faire selon la même logique que pour l'opérateur position ; nous verrons ensuite comment relier ces deux représentations. On considère la famille $\{\langle p|\ ,\ p\in\mathbb{R}\}$ de formes linéaires par

$$\forall p\in\mathbb{R}\ ,\quad \langle p|\,\mathcal{P} = p\,\langle p|\ . \tag{2.43}$$

L'action de ces formes linéaires sur un vecteur d'état $|\psi\rangle$ de l'espace de Hilbert [8] $\mathcal{H}\cong\mathcal{L}^2(\mathbb{R})$ définit une fonction $p\mapsto\hat{\psi}(p)$, via :

$$\langle p|\ :\quad \left\{\begin{array}{lcl}\mathcal{L}^2(\mathbb{R}) & \to & \mathbb{C}\\ |\psi\rangle & \mapsto & \langle p|\psi\rangle \overset{\text{déf.}}{=} \hat{\psi}(p)\,.\end{array}\right. \tag{2.44}$$

La fonction $p\mapsto\hat{\psi}(p)$ est la transformée de Fourier de la fonction $x\mapsto\psi(x)$, voir l'appendice 2.A. L'action de l'opérateur impulsion sur ces fonctions d'ondes en impulsion est donnée par la multiplication par p :

$$\langle p|\,\mathcal{P}\,|\psi\rangle = p\hat{\psi}(p)\,. \tag{2.45}$$

Comme pour l'opérateur position définissons une famille d'« états propres d'impulsion », $\{|p\rangle\,,p\in\mathbb{R}\}$:

$$\mathcal{P}\,|p\rangle = p\,|p\rangle\ \ ,\quad p\in\mathbb{R}\,, \tag{2.46}$$

qui satisfont, suivant un raisonnement analogue à celui conduisant à l'éqn. (2.24), la condition d'orthonormalité généralisée :

$$\langle p|p'\rangle = \delta(p-p')\,, \tag{2.47}$$

et permettent d'obtenir la relation de fermeture

$$\mathfrak{I} = \int dp\ |p\rangle\,\langle p|\ . \tag{2.48}$$

Le passage de la représentation d'impulsion à la représentation de position s'effectue de la manière suivante. Nous cherchons à obtenir l'expression des « éléments de matrice » $\langle x|p\rangle$, à l'aide de l'équation (2.41). Nous avons d'une part :

$$\langle x|\,\mathcal{P}\,|p\rangle = p\,\langle x|p\rangle\ , \tag{2.49}$$

8. Comme expliqué précédemment en toute rigueur il faudrait considérer l'espace de Schwartz $\mathcal{S}(\mathbb{R})$ dense dans $\mathcal{L}^2(\mathbb{R})$.

et d'autre part

$$\langle x|\,\mathcal{P}\,|p\rangle = -i\hbar\frac{d}{dx}\,\langle x|p\rangle\ . \tag{2.50}$$

L'équation différentielle $-i\hbar\frac{d}{dx}\,\langle x|p\rangle = p\,\langle x|p\rangle$ a pour solution :

$$\langle x|p\rangle = \mathcal{N}e^{\frac{ipx}{\hbar}}\,, \tag{2.51}$$

et la constante de normalisation s'obtient en écrivant

$$\delta(p-p') = \langle p'|p\rangle = \int dx\ \langle p'|x\rangle\,\langle x|p\rangle = |\mathcal{N}|^2\int dx\, e^{\frac{i(p-p')x}{\hbar}}\,. \tag{2.52}$$

En utilisant l'équation (2.117) de l'appendice 2.A puis la propriété (2.16) nous avons

$$\int dx\, e^{\frac{i(p-p')x}{\hbar}} = 2\pi\,\delta\left(\frac{p-p'}{\hbar}\right) = 2\pi\hbar\delta(p-p')\,. \tag{2.53}$$

Avec un choix de phase approprié nous obtenons donc :

$$\boxed{\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{\frac{ipx}{\hbar}}} \tag{2.54}$$

et son conjugué

$$\langle p|x\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{-\frac{ipx}{\hbar}}\,, \tag{2.55}$$

qui se généralisent immédiatement à trois dimensions comme :

$$\langle\vec{x}|\vec{p}\rangle = \frac{1}{(2\pi\hbar)^{3/2}}e^{\frac{i\vec{p}\cdot\vec{x}}{\hbar}} \tag{2.56a}$$

$$\langle\vec{p}|\vec{x}\rangle = \frac{1}{(2\pi\hbar)^{3/2}}e^{\frac{-i\vec{p}\cdot\vec{x}}{\hbar}} \tag{2.56b}$$

Sans surprise, le passage de la représentation position à la représentation impulsion correspond à la transformation de Fourier, voir app. 2.A. Nous avons en effet,

$$\langle p|\psi\rangle = \langle p|\int dx\,|x\rangle\,\langle x|\psi\rangle\ \leftrightarrow \hat{\psi}(p) = \int dx\,\langle p|x\rangle\,\psi(x) = \frac{1}{\sqrt{2\pi\hbar}}\int dx e^{-\frac{ipx}{\hbar}}\psi(x)\,. \tag{2.57}$$

Les fonctions d'onde $\langle\vec{x}|\vec{p}\rangle$ données par l'équation (2.56a) sont naturellement des *ondes planes* ; n'étant pas des éléments de $\mathcal{L}^2(\mathbb{R})$, elles ne peuvent correspondre à des états physiques du système. Cependant, la transformation de Fourier inverse d'une fonction d'onde de carré sommable – et donc physique – peut être vue comme une décomposition sur une « base » d'ondes planes, bien que ces dernières ne soient pas dans $\mathcal{L}^2(\mathbb{R})$:

$$\psi(x) = \langle x|\psi\rangle = \int dp\ \langle x|p\rangle\,\langle p|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}}\int dp\,\hat{\psi}(p)\,e^{\frac{ipx}{\hbar}}\,. \tag{2.58}$$

Lorsqu'une particule est confinée dans une boîte de volume fini la situation est différente car l'onde plane est alors évidemment de carré sommable ; vous noterez également la normalisation différente des fonctions d'onde entre le cas du volume infini, éqn. (2.56a), et du volume fini.

2.3.1 Relation d'incertitude d'Heisenberg

Historiquement, le principe d'incertitude donné par l'équation (1.100) du théorème 11 a été donné par Heisenberg pour les opérateurs positions et impulsion [3].

Les opérateurs $\mathcal{X}$ et $\mathcal{P}$ obéissent à la relation de commutation (2.34) qui a pour particularité que le commutateur est proportionnel à l'opérateur identité. Ainsi, en utilisant l'équation (1.100) pour des vecteurs de norme unité on trouve la relation

$$\forall\, |\psi\rangle \in \mathcal{H}\,,\ \| \,|\psi\rangle\, \| = 1\,,\ \boxed{\Delta_\psi(\mathcal{X})\Delta_\psi(\mathcal{P}) \geqslant \frac{1}{2}\hbar}\,, \tag{2.59}$$

dont le membre de droite est indépendant de l'état $|\psi\rangle$ du système. Cette borne nous indique, qualitativement, que plus la distribution de probabilité de position possède une variance faible – donc plus la position est connue avec précision – plus la distribution de probabilité d'impulsion possède une variance importante - donc moins l'impulsion est connue avec précision – et inversement. Mathématiquement cette relation reflète une propriété des transformations de Fourier comme nous allons le voir.

La deuxième partie du théorème 11 donnait une condition suffisante sur le vecteur d'état pour que l'égalité entre les membres de gauche et de droite dans (2.59) soit satisfaite ; cela correspond au meilleur compromis possible entre l'incertitude sur la position et sur l'impulsion. On cherche $|\psi\rangle \in \mathcal{L}^2(\mathbb{R})$ tel que :

$$\exists \lambda \in \mathbb{R}\,,\quad (\mathcal{P} - \langle\mathcal{P}\rangle_\psi)\,|\psi\rangle = i\lambda(\mathcal{X} - \langle\mathcal{X}\rangle_\psi)\,|\psi\rangle\ . \tag{2.60}$$

En multipliant cette équation par le bra $\langle x|$ on trouve

$$\langle x|\,\mathcal{P}\,|\psi\rangle - \langle\mathcal{P}\rangle_\psi \psi(x) = i\lambda(x - \langle\mathcal{X}\rangle_\psi)\psi(x)\,, \tag{2.61}$$

et, en utilisant (2.41) l'équation différentielle

$$\frac{d\psi(x)}{dx} = \frac{i}{\hbar}\big(\langle\mathcal{P}\rangle_\psi + i\lambda(x - \langle\mathcal{X}\rangle_\psi)\big)\psi(x)\,, \tag{2.62}$$

dont la solution normalisée est :

$$\boxed{\psi(x) = \left(\frac{\lambda}{\pi\hbar}\right)^{1/4} e^{-\frac{\lambda}{2\hbar}\left(x - \langle\mathcal{X}\rangle_\psi\right)^2 + \frac{i}{\hbar}\langle\mathcal{P}\rangle_\psi x}} \tag{2.63}$$

On remarquera que la normalisation de cette fonction est par rapport à la norme de $\mathcal{L}^2(\mathbb{R})$, c.-à.-d. $\int |\psi|^2 = 1$. La fonction d'onde donnée par l'équation (2.63), qui appartient à $\mathcal{S}(\mathbb{R}) \subset \mathcal{L}^2(\mathbb{R})$, est appelée *paquet d'ondes gaussien*.

La densité de probabilité de présence associée est une distribution gaussienne centrée en $\langle\mathcal{X}\rangle_\psi$. On a alors naturellement

$$\Delta_\psi(\mathcal{X}) = \sqrt{\frac{\hbar}{\lambda}}\,, \tag{2.64}$$

et la relation d'incertitude nous donne sans calcul explicite

$$\Delta_\psi(X)\Delta_\psi(P) = \frac{1}{2}\hbar \implies \Delta_\psi(P) = \frac{1}{2}\sqrt{\hbar\lambda}\,. \tag{2.65}$$

La transformée de Fourier d'une gaussienne étant une gaussienne, la distribution de probabilité en impulsion est aussi gaussienne. Les équations (2.64) et (2.65) nous indiquent que la variance de la distribution de position est inversement proportionnelle à celle de la distribution d'impulsion. Contraindre la position de la particule par un choix approprié de paquet d'onde va donner une plus grande dispersion d'impulsion, et inversement.

2.3.2 Espace des états d'une particule libre

Dans le cours n°1, nous avons développé, de manière semi-rigoureuse, le formalisme permettant de décrire l'espace des états d'un système quantique en tant qu'espace de Hilbert. Ce formalisme va être utilisé ici pour étudier un système physique simple, celui d'une particule libre massive sans spin. En complément nous rappelerons dans l'appendice 2.B l'espace des états de l'oscillateur harmonique qui, paradoxalement, est plus simple du point de vue mathématique.

Considérons donc le hamiltonien décrivant une particule libre, massive et sans spin dans un espace infini à trois dimensions :

$$\mathcal{H}_0 = \frac{1}{2m}\vec{P}^2 = \frac{1}{2m}\left(P_x^2 + P_y^2 + P_z^2\right)\,. \tag{2.66}$$

cet opérateur est naturellement hermitien ; pour définir un opérateur auto-adjoint associé il faudra préciser un domaine de définition approprié.

Dans la représentation de position, les états propres de ce Hamiltonien satisfont à l'équation différentielle :

$$\langle\vec{x}|\,\mathcal{H}_0\,|\psi\rangle = E\,\langle\vec{x}|\psi\rangle \;\Leftrightarrow\; -\Delta\,\psi(\vec{x}) = \frac{2mE}{\hbar^2}\,\psi(\vec{x})\,. \tag{2.67}$$

Si on choisit $|\psi\rangle \in S(\mathbb{R}^3)$ on a

$$-\Delta\psi(\vec{x}) = \mathcal{F}^{-1}\left(\vec{k}^2\hat{\psi}(\vec{k})\right)(\vec{x})\,, \tag{2.68}$$

où $\mathcal{F}^{-1}$ est la transformée de Fourier inverse, voir appendice 2.A, équation (2.105).

Le théorème spectral 8 est vérifié car on peut transformer le hamiltonien (2.66) en opérateur de multiplication $\hat{\psi}(k) \mapsto k^2\hat{\psi}(k)$ par transformée de Fourier qui est une transformation unitaire :

$$\boxed{\left(\mathcal{F}\,\mathcal{H}_0\,\mathcal{F}^{-1}\right)\hat{\psi}(k) = \frac{\hbar^2k^2}{2m}\,\hat{\psi}(k)} \tag{2.69}$$

Cette relation suggère un domaine de définition naturel pour ce hamiltonien, à partir du domaine de l'opérateur multiplication apparaissant ici :

$$D(\mathcal{H}_0) = \{\psi \in \mathcal{L}^2(\mathbb{R}),\, \vec{k}^2\hat{\psi}(\vec{k}) \in \mathcal{L}^2(\mathbb{R})\}\,. \tag{2.70}$$

Sur ce domaine, exemple d'*espace de Sobolev*, $\mathcal{H}_0$ est un opérateur auto-adjoint. Son spectre est purement continu, et donné par

$$\boxed{\sigma(\mathcal{H}_0) = \left\{ \frac{\hbar^2 \vec{k}^2}{2m},\ \vec{k} \in \mathbb{R}^3 \right\} \cong \mathbb{R}^+} \tag{2.71}$$

Le spectre ponctuel de l'opérateur $\mathcal{H}_0$ est vide ; il n'admet ni valeurs propres ni donc vecteurs propres dans $D(\mathcal{H}_0) \subset \mathcal{H}$. Il est plus naturel en physique de travailler plutôt avec la variable d'impulsion

$$\vec{p} = \hbar \vec{k}. \tag{2.72}$$

Si on accepte de considérer des vecteurs ne faisant pas partie de l'espace de Hilbert, en reprenant le formalisme développé à la section 2.3, il est évident que les « vecteurs propres » généralisés d'impulsion,

$$\vec{\mathcal{P}} \, |\vec{p}\rangle = \vec{p} \, |\vec{p}\rangle \ , \tag{2.73}$$

sont également « vecteurs propres » du hamiltonien $\mathcal{H}_0$:

$$\boxed{\mathcal{H}_0 \, |\vec{p}\rangle = \frac{1}{2m} \vec{\mathcal{P}}^2 \, |\vec{p}\rangle = \frac{1}{2m} \vec{p}^{\,2} \, |\vec{p}\rangle} \tag{2.74}$$

À trois dimensions, il existe une dégénerescence infinie des états d'énergie, car tous les kets $|\vec{p}\rangle$ tels que $\vec{p}$ appartient à la sphère $\vec{p}^2 = p$, $p \in \mathbb{R}_+^\star$ correspondent à la même énergie pour la particule. En représentation de position, ces « vecteurs propres » sont naturellement des ondes planes, voir l'éqn. (2.54) :

$$\boxed{\langle \vec{x} | \vec{p} \rangle = \left(\frac{1}{2\pi\hbar} \right)^{3/2} e^{\frac{i \vec{p} \cdot \vec{x}}{\hbar}}} \tag{2.75}$$

Insistons pour finir sur le fait que ces kets ne peuvent être associés physiquement à des états physiques. Un état physique correspond à un *paquet d'ondes*, une fonction de $\mathcal{L}^2(\mathbb{R}^3)$, pouvant être vue, via une transformation de Fourier, comme une superposition d'ondes planes. Un exemple est le paquet d'ondes gaussien donné par l'équation (2.63). Ces vecteurs d'états n'étant pas états propres du hamiltonien $\mathcal{H}_0$, il doivent nécessairement évoluer avec le temps.

2.4 Particule dans une boîte, électrons libres dans les solides

L'espace des états étudié dans la sous-section 2.3.2 est associé à une particule libre dans un espace infini ; il existe cependant des situations où, pour des raisons physiques ou techniques, il est utile de considérer une particule confinée à l'intérieur d'un volume, par exemple dans une boîte cubique.

Un tel modèle est particulièrement utile pour décrire la physique des électrons dans les solides, plus précisément celle des électrons de valence des métaux alcalins comme le sodium.

Ces électrons ne sont évidemment pas libres, mais soumis à un potentiel dont l'allure typique est présentée sur la Fig. 2.1. Le potentiel possède des valeurs élevées à l'extérieur du cristal ; l'énergie du vide E_{vac} est l'énergie à fournir à un électron pour l'extraire du cristal (de l'ordre de 5 à 10 eV). Les symboles (+) indiquent la position des coeurs ioniques, chargés positivement.

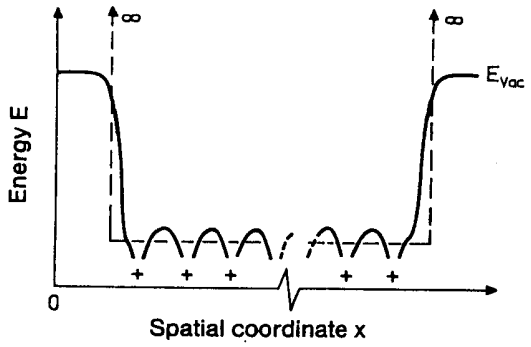

FIGURE 2.1 – Allure du potentiel pour un électron dans un cristal. Source Ibach et Luth [1]

Dans le *modèle des électrons libres*, on suppose que le potentiel V prend une valeur constante V_0 sur tout le volume du cristal, et que le potentiel est infini à l'extérieur, c.-à.-d. que l'on considère un puits bordé de barrières infinies, représenté par la courbe en traits tiretés sur la figure. Un électron est ainsi décrit par le hamiltonien suivant

$$\mathcal{H} = \frac{1}{2m}\vec{\mathcal{P}}^2 + V_0\,, \tag{2.76}$$

sur un domaine

$$R = [0, L_x] \times [0, L_y] \times [0, L_z] \subset \mathbb{R}^3\,, \tag{2.77}$$

où L_x, L_y et L_z correspondent aux dimensions du cristal (ayant la forme d'un parallélépipède rectangle) le long des axes correspondants. Mathématiquement, la présence de barrières infinies aux bords du cristal revient à imposer des conditions aux limites *de Dirichlet* à la fonction d'onde de la particule :

$$\Psi(0,y,z) = \Psi(L_x,y,z) = 0\ , \quad \Psi(x,0,z) = \Psi(x,L_y,z) = 0\ , \quad \Psi(x,y,0) = \Psi(x,y,L_z) = 0\,. \tag{2.78}$$

Il serait en effet impossible sinon d'obtenir des états propres de $\mathcal{H}$ d'énergie finie. On considère donc le domaine de définition suivant pour le hamiltonien (2.76) [9] :

$$D(\mathcal{H}) = \left\{ |\Psi\rangle \in \mathcal{L}^2(R) \text{ t.q. } \quad \mathcal{H}\,|\Psi\rangle \in \mathcal{L}^2(R) \text{ et } \quad \Psi(0,y,z) = \Psi(L_x,y,z) = \cdots = 0 \right\}. \tag{2.79}$$

Dans la suite, on choisira $V_0 = 0$ en redéfinissant l'origine des énergies. L'opérateur hamiltonien $\mathcal{H}$ ainsi défini possède un spectre discret :

$$\sigma\big(\mathcal{H}\big) = \left\{ \frac{\hbar^2\pi^2}{2m}\left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2}\right)\ ,\ (n_x,n_y,n_z) \in \big(\mathbb{N}^*\big)^3 \right\}. \tag{2.80}$$

9. La question du caractère auto-adjoint du hamiltonien pour ce domaine de définition sera évoquée dans la section 2.5.

avec les vecteurs propres normés $|\vec{n}\rangle$ associés correspondant aux fonctions d'onde :

$$\Psi_{\vec{n}}(\vec{x}) = \langle \vec{x}|\vec{n}\rangle = \sqrt{\frac{2^3}{L_x L_y L_z}} \sin\left(\frac{\pi n_x x}{L_x}\right) \sin\left(\frac{\pi n_y y}{L_y}\right) \sin\left(\frac{\pi n_z z}{L_z}\right) ,$$
$$0 \leqslant x \leqslant L_x\,,\ 0 \leqslant y \leqslant L_y\,,\ 0 \leqslant z \leqslant L_z\,. \quad (2.81)$$

On vérifie que :

$$\| \, |\vec{n}\rangle \, \|^2 = \frac{8}{L_x L_y L_z} \int_0^{L_x} dx \, \sin^2\left(\frac{\pi n_x x}{L_x}\right) \int_0^{L_y} dy \, \sin^2\left(\frac{\pi n_y y}{L_y}\right) \int_0^{L_z} dz \, \sin^2\left(\frac{\pi n_z z}{L_z}\right) = 1\,. \quad (2.82)$$

2.4.1 Conditions aux limites périodiques

Les conditions aux limites de Dirichlet, motivées physiquement dans le modèle des électrons libres, voir fig. 2.1, ou pour décrire des modèles de matière condensée en général, ne sont néanmoins pas les plus couramment utilisées. Les vecteurs propres du hamiltonien donnés par l'éqn. (2.81) ne sont pas des vecteurs propres d'impulsion ; cette propriété est liée au fait que les conditions aux limites de Dirichlet ne respectent pas la symétrie de translation, aspect qui sera discuté en détail au chapitre 7.

Parmi les conditions aux limites préservant le caractère auto-adjoint du hamiltonien (2.76), voir la section 2.5, nous choisirons des *conditions aux limites périodiques* :

$$\boxed{\Psi(0,y,z) = \Psi(L_x,y,z)\ , \quad \Psi(x,0,z) = \Psi(x,L_y,z)\ , \quad \Psi(x,y,0) = \Psi(x,y,L_z)} \quad (2.83)$$

Géométriquement, cela revient à identifier les faces opposées du parallélépipède rectangle associé au solide – de la même manière qu'identifier les deux extrémités d'un segment donne un cercle. Ces conditions aux limites ne sont pas physiques, mais l'erreur commise devient négligeable dans la limite où le solide devient macroscopique, car il s'agit d'« effets de bord », proportionnels à la surface du solide, alors que les propriétés extensives comme l'énergie interne sont proportionnels à son volume. [10]. Du point de vue des symétries, que nous adopterons dans les chapitres suivants, on peut remarquer que la symétrie de translation, exacte avec des conditions aux limites périodique, n'est qu'une symétrie approchée avec des conditions aux limites de Dirichlet. [11] Le spectre de l'opérateur hamiltonien $\mathcal{H}$ est également discret pour ce choix de conditions aux limites et donné par :

$$\boxed{\sigma\big(\mathcal{H}\big) = \left\{ \frac{2\pi^2\hbar^2}{m} \left(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2} \right)\ ,\ (n_x,n_y,n_z) \in \mathbb{Z}^3 \right\}} \quad (2.84)$$

10. À noter cependant que dans certains matériaux, comme ceux dit » isolant topologiques », les effets physiques importants se déroulent à la surface et non dans le volume !

11. D'un point de vue mathématique, on pourra imposer de manière équivalente les conditions aux limites périodiques en considérant la fonction d'onde non pas comme un élément de $\mathcal{L}^2(R)$ mais comme un élément de $\mathcal{L}^2(\mathbb{R}^3)$ soumis à des conditions de périodicité :

$$\forall (x,y,z) \in \mathbb{R}^3\ , \quad \forall (k,\ell,m) \in \mathbb{Z}^3\ , \quad \Psi\big(x + kL_x, y + \ell L_y, z + mL_z\big) = \Psi\big(x,y,z\big)\,.$$

On remarquera en comparant l'éqn. (2.80) et l'éqn. (2.84) que le choix de conditions aux limites – ou, dit autrement, d'extension auto-adjointe du hamiltonien – donne un spectre en énergie différent, donc des prévisions physiques distinctes. Les vecteurs propres associés à ce spectre discret sont donnés par :

$$\boxed{\Psi_{\vec{n}}(x,y,z) = \langle \vec{x}|\vec{n}\rangle = \frac{1}{\sqrt{L_x L_y L_z}}\, e^{i\vec{k}\cdot\vec{x}}\,, \quad \vec{k} = \left(\tfrac{2\pi n_x}{L_x}, \tfrac{2\pi n_y}{L_y}, \tfrac{2\pi n_z}{L_z}\right), \quad (n_x, n_y, n_z) \in \mathbb{Z}^3} \tag{2.85}$$

Ces vecteurs forment une famille orthonormée de l'espace de Hilbert $\mathcal{L}^2(R)$. On vérifie aisément que :

$$\langle \vec{m}|\vec{n}\rangle = \frac{1}{L_x}\int_0^{L_x} dx\, e^{\frac{2i\pi}{L_x}(n_x-m_x)x} \frac{1}{L_y}\int_0^{L_y} dy\, e^{\frac{2i\pi}{L_y}(n_y-m_y)y} \frac{1}{L_x}\int_0^{L_z} dz\, e^{\frac{2i\pi}{L_z}(n_z-m_z)z} = \delta_{\vec{m},\vec{n}}\,. \tag{2.86}$$

On peut en outre montrer (cf. étude des séries de Fourier) que l'espace engendré par les combinaisons linéaires de ces vecteurs est dense dans $\mathcal{L}^2(R)$, et forme donc une base orthonormée de cet espace de Hilbert composée d'états propres du hamiltonien $\mathcal{H}$.

2.5 Extensions auto-adjointes de l'opérateur impulsion ⋆

Dans la représentation de position, l'opérateur impulsion est associé à l'opérateur de dérivation spatiale, ou gradient dans le cas d'un système à trois dimensions :

$$\vec{\mathcal{P}} = -i\hbar\vec{\nabla}\,, \tag{2.87}$$

agissant sur l'espace des fonctions de carré sommable, $\mathcal{L}^2(\mathbb{R}^3)$ dans le cas d'un espace infini, ou $\mathcal{L}^2(R \subset \mathbb{R}^3)$ dans le cas où le système est confiné dans une région R de l'espace ; ce choix n'est pas sans conséquences sur les propriété de l'opérateur.

Pour simplifier la discussion, considérons un système confiné dans un intervalle $[0, L] \in \mathbb{R}$, c.-à.-d. une particule à une dimension dans un puits de potentiel infini de largeur L, étudié en détail dans la section 2.4 (pour le cas tridimensionnel). Choisissons premièrement comme domaine de définition de l'opérateur $\mathcal{P} = -i\hbar d/dx$ l'ensemble des fonctions sur $[0, L]$ de carré sommable, de dérivée première de carré sommable, et s'annulant aux extrémités de l'intervalle (conditions de Dirichlet) :

$$D(\mathcal{P}) = \left\{\phi \in \mathcal{L}^2([0,L]) \text{ t.q. } \quad \phi' \in \mathcal{L}^2([0,L]) \text{ et } \quad \phi(0) = \phi(L) = 0\right\}. \tag{2.88}$$

Afin de déterminer si l'opérateur $\mathcal{P}$ est auto-adjoint, et sur quel domaine, calculons

$$\begin{aligned}\iota\big(\psi, \mathcal{P}\phi\big) - \iota\big(\mathcal{P}\psi, \phi\big) &= \int_0^L dx \left(\bar{\psi}(x)\big(-i\hbar\phi'(x)\big) - \big(-i\hbar\bar{\psi}'(x)\big)\phi(x)\big)\right) \\ &= -i\hbar\int_0^L dx\, \frac{d}{dx}\big(\bar{\psi}(x)\phi(x)\big) = -i\hbar\big(\bar{\psi}(L)\phi(L) - \bar{\psi}(0)\phi(0)\big)\,. \end{aligned} \tag{2.89}$$

Ainsi, l'opérateur $\mathcal{P}$ est tout d'abord, comme attendu, un opérateur hermitien sur son domaine de définition car, si $\phi, \psi \in \mathrm{D}(\mathcal{P})$, nous constatons que $\imath\big(\psi,\, \mathcal{P}\phi\big) - \imath\big(\mathcal{P}\psi,\, \phi\big) = 0$.

Crucialement, remarquons que l'annulation de $\bar{\psi}(L)\phi(L) - \bar{\psi}(0)\phi(0)$ est une conséquence des conditions aux limites pour ϕ, c.-à.-d. $\phi(0) = \phi(L) = 0$, et ceci indépendamment des conditions aux limites pour ψ ! Cela signifie que l'opérateur adjoint de $\mathcal{P}$ (bien que, littéralement, $\mathcal{P}^\dagger$ soit égal à $-i\hbar \mathrm{d}/\mathrm{d}x$) possède un domaine de définition plus grand que celui de l'opérateur $\mathcal{P}$:

$$\mathrm{D}(\mathcal{P}^\dagger) = \big\{ \phi \in \mathcal{L}^2([0, L]) \text{ t.q.} \quad \phi' \in \mathcal{L}^2([0, L]) \big\} \supsetneq \mathrm{D}(\mathcal{P})\,. \tag{2.90}$$

Ainsi l'opérateur impulsion n'est pas auto-adjoint pour une particule dans un puits de potentiel avec des conditions de Dirichlet, pourtant physiquement naturelles ! Cela peut s'expliquer en évoquant le concept de symétrie, qui sera développé en détail au chapitre 3 et suivants ; en effet, les conditions de Dirichlet brisent la symétrie de translation dont le générateur infinitésimal est l'opérateur impulsion.

Cherchons des conditions aux limites moins restrictives sur les fonctions d'onde de la particule telles que l'opérateur impulsion soit auto-adjoint, en suivant la logique déjà évoquée au chapitre 1 ; plus le domaine de définition d'un opérateur $\mathcal{O}$ est restreint plus celui de son adjoint $\mathcal{O}^\dagger$ est élargi, et réciproquement. On pose ainsi :

$$\phi(L) = \lambda\phi(0)\ , \quad \lambda \in \mathbb{C}^\star\,. \tag{2.91}$$

En reportant ces conditions aux limites dans le calcul effectué à l'équation (2.89), nous trouvons la contrainte :

$$\imath\big(\psi,\, \mathcal{P}\phi\big) - \imath\big(\mathcal{P}\psi,\, \phi\big) = -i\hbar\big(\bar{\psi}(L)\phi(L) - \bar{\psi}(0)\phi(0)\big) = -i\hbar\phi(0)\big(\lambda\bar{\psi}(L) - \bar{\psi}(0)\big) \stackrel{!}{=} 0\,. \tag{2.92}$$

Pour que cette égalité soit vérifiée, il faut donc que $\bar{\psi}(L) = (1/\lambda)\bar{\psi}(0)$, soit

$$\psi(L) = \frac{1}{\bar{\lambda}}\psi(0)\,. \tag{2.93}$$

L'opérateur $\mathcal{P}$ est donc auto-adjoint sur le domaine tel que les deux conditions aux limites (2.91,2.93) sont compatibles, soit pour

$$\lambda = 1/\bar{\lambda} \quad \Longrightarrow \quad \exists \theta \in [0,\, \pi[\text{ t.q.} \quad \lambda = e^{i\theta}\,. \tag{2.94}$$

Il existe ainsi une infinité d'extensions auto-adjointes $\mathcal{P}_\theta$ de l'opérateur impulsion $\mathcal{P}$ sur un intervalle, paramétrées par un angle $\theta \in [0, 2\pi[$, telles que :

$$\mathrm{D}(\mathcal{P}_\theta) = \big\{ \phi \in \mathcal{L}^2([0, L]) \text{ t.q.} \quad \phi' \in \mathcal{L}^2([0, L]) \text{ et} \quad \phi(L) = e^{i\theta}\phi(0) \big\}\,. \tag{2.95}$$

Cette famille d'opérateurs auto-adjoints contient comme cas particuliers les conditions aux limites périodiques couramment utilisées, $\phi(0) = \phi(L)$, mais également les conditions aux limites anti-périodiques, $\phi(0) = -\phi(L)$.

Le choix de l'extension auto-adjointe de $\mathcal{P}$ a des conséquences physiquement observables. Premièrement toute fonction d'onde de carré sommable sur $[0, L]$ peut se développer en modes de Fourier correspondant aux fonctions propres orthonormées de $\mathcal{P}_\theta$:

$$\phi(x) = \sum_{n\in\mathbb{Z}} c_n(\theta)\phi_n(x)\,, \quad \phi_n(x) = \tfrac{1}{\sqrt{L}}\, e^{\frac{2i\pi\nu}{L}x}\,, \quad \nu = n + \frac{\theta}{2\pi}\,, \tag{2.96}$$

avec

$$c_n(\theta) = \tfrac{1}{\sqrt{L}} \int_0^L e^{-\frac{2i\pi}{L}\nu x}\phi(x)\,. \tag{2.97}$$

Ces fonctions propres orthonormées ont des valeurs propres données par :

$$-i\hbar\frac{d}{dx}\phi_n(x) = \frac{2\pi\hbar}{L}\big(n + \tfrac{\theta}{2\pi}\big)\phi_n(x)\,, \tag{2.98}$$

et la probabilité de mesurer l'une de ces valeurs pour une particule dont la fonction d'onde est donnée par $\psi(x)$ est égale à $|c_n(\theta)|^2$, où le coefficient de Fourier est donné par l'équation (2.97).

Le cas d'une particule soumise à une barrière de potentiel, c.-à.-d. restreinte à un intervalle de la forme $[0, +\infty[$, est encore plus surprenant, car *il n'existe aucune extension auto-adjointe de l'opérateur impulsion $\mathcal{P}$!* Sans en donner ici la démonstration mathématique rigoureuse, donnons un argument heuristique, fondé sur les calculs précédents. Si on considère que les fonctions d'onde, en plus d'être de carré sommable, sont continues, elles ont nécessairement une limite nulle en l'infini. La limite $L \to \infty$ de l'éqn. (2.89) donne alors la condition d'hermiticité $\bar{\psi}(0)\phi(0) = 0$. Avec la condition au bord $\phi(0) = 0$, cette contrainte est satisfaite mais en contrepartie la valeur de $\psi(a)$ est complètent arbitraire. Au contraire du cas précédent, il n'existe pas de condition au bord pour ϕ plus générale, qui assurerait que $D(\mathcal{P}) = D(\mathcal{P}^\dagger)$.

Finalement, pour le système associé à une particule non-confinée, c.-à.-d. dont les états physiques sont associés à des fonctions d'onde de $\mathcal{L}^2(\mathbb{R})$, étudié à la section 2.3, l'opérateur impulsion est d'emblée auto-adjoint, ou, dit autrement, il n'admet qu'une unique extension auto-adjointe. Il n'admet aucune valeur propre, et son spectre est donné par $\sigma(\mathcal{P}) = \mathbb{R}$.

Une analyse similaire peut s'appliquer au hamiltonien d'une particule libre massive (2.76) sur un intervalle $[0, L]$, étudié dans la section 2.4 ; nous en donnons ici les grandes lignes. Considérons un domaine de définition $D(\mathcal{H}) \subset \mathcal{L}^2([0, L])$, et demandons que le hamiltonien soit auto-adjoint sur ce domaine par un choix approprié de conditions aux limites satisfaites pour tout $\phi \in D(\mathcal{H})$. Une condition nécessaire est :

$$\forall \phi \in D(\mathcal{H})\,, \quad \iota\big(\phi, \mathcal{H}\phi\big) - \iota\big(\mathcal{H}\phi, \phi\big) \stackrel{!}{=} 0\,, \tag{2.99}$$

qui peut s'exprimer comme :

$$\begin{aligned} 0 = \iota\big(\phi, \mathcal{H}\phi\big) - \iota\big(\mathcal{H}\phi, \phi\big) &= -\frac{\hbar^2}{2m}\int_0^L dx\, \Big(\overline{\phi(x)}\phi''(x) - \overline{\phi''(x)}\phi(x)\Big) \\ &= -\frac{\hbar^2}{2m}\Big[\overline{\phi(x)}\phi'(x) - \overline{\phi'(x)}\phi(x)\Big]_0^L\,. \end{aligned} \tag{2.100}$$

Cette condition peut s'écrire comme :

$$\left|L\phi'(0) - i\phi(0)\right|^2 + \left|L\phi'(L) + i\phi(L)\right|^2 = \left|L\phi'(0) + i\phi(0)\right|^2 + \left|L\phi'(L) - i\phi(L)\right|^2 \quad (2.101)$$

Le membre de gauche et de droite de cette équation peuvent respectivement se comprendre comme

$$\left|\begin{pmatrix} L\phi'(0) - i\phi(0) \\ L\phi'(L) + i\phi(L) \end{pmatrix}\right|^2 = \left|L\phi'(0) - i\phi(0)\right|^2 + \left|L\phi'(L) + i\phi(L)\right|^2,$$

et

$$\left|\begin{pmatrix} L\phi'(0) + i\phi(0) \\ L\phi'(L) - i\phi(L) \end{pmatrix}\right|^2 = \left|L\phi'(0) + i\phi(0)\right|^2 + \left|L\phi'(L) - i\phi(L)\right|^2.$$

On peut ainsi décrire l'espace des solutions de la condition (2.99) comme :

$$\begin{pmatrix} L\phi'(0) - i\phi(0) \\ L\phi'(L) + i\phi(L) \end{pmatrix} = U \begin{pmatrix} L\phi'(0) + i\phi(0) \\ L\phi'(L) - i\phi(L) \end{pmatrix}. \quad (2.102)$$

où U est une matrice complexe 2×2 unitaire. Un choix de U détermine un choix de domaine de définition $\mathcal{D}(\mathcal{H}) \subset \mathcal{L}^2([0, L])$ pour lequel on peut montrer que le hamiltonien est auto-adjoint, choix influençant les propriétés physique du système, en premier lieu le spectre du hamiltonien. Les conditions de Dirichlet sont par exemple associées au choix de la matrice identité, $U = \mathbb{I}_2$.

Les conditions aux limites périodiques correspondent au choix $U = i\sigma_2 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, comme on le voit en prenant la différence entre les deux équations obtenues à partir de (2.102). Vérifions alors explicitement que le hamiltonien $\mathcal{H}$ est bien auto-adjoint. Soit $\psi \in D(\mathcal{H})$, et donc périodique, et $\phi \in \mathcal{L}^2([0, L])$ (plus précisément, avec la condition supplémentaire que sa dérivée seconde existe et est de carré sommable). Nous avons :

$$\begin{aligned} \iota\big(\phi, \mathcal{H}\psi\big) - \iota\big(\mathcal{H}\phi, \psi\big) &= -\frac{\hbar^2}{2m}\Big(\overline{\phi(L)}\psi'(L) - \overline{\phi'(L)}\psi(L) - \overline{\phi(0)}\psi'(0) + \overline{\phi'(0)}\psi(0)\Big) \\ &\Longrightarrow \quad \Big(\overline{\phi(L)} - \overline{\phi(0)}\Big)\psi'(0) = \Big(\overline{\phi'(L)} - \overline{\phi'(0)}\Big)\psi(0) \stackrel{!}{=} 0 \end{aligned} \quad (2.103)$$

Pour que cette relation soit satisfaite indépendamment des valeurs de ψ et ψ' en $x = 0$, il faut nécessairement que ϕ soit également périodique. On en déduit que le domaine de l'adjoint de $\mathcal{H}$ est identique au domaine de $\mathcal{H}$ choisi ; les conditions périodiques définissent donc bien un domaine sur lequel $\mathcal{H}$ est auto-adjoint.

Le message à retenir de cette analyse, qui s'applique à de nombreux exemples, est que la donnée de l'expression littérale du hamiltonien n'est pas nécessairement suffisante pour caractériser le système physique. Le choix de l'extension auto-adjointe du hamiltonien, qui revient ici aux choix des conditions aux limites, est partie intégrante de la définition du problème, ce choix étant *in fine* dicté par des considérations physiques.

À retenir

— les fonctions de Hermite forment une base dénombrable de l'espace $\mathcal{L}^2(\mathbb{R})$ des fonctions d'onde ;

— la définition de la distribution de Dirac dans l'espace dual ;

— le postulat de réduction du paquet d'ondes ;

— les paquets d'ondes gaussiens, meilleur compromis entre l'incertitude sur la position et sur l'impulsion ;

— l'opérateur impulsion correspondant à la dérivation dans la représentation de position ;

— le passage de la représentation position à la représentation impulsion donné par la transformation de Fourier ;

— le hamiltonien d'une particule libre se diagonalise par transformation de Fourier, néanmoins les ondes planes ne sont pas des états physiques ;

— les conditions aux limites de Dirichlet et périodiques pour des particules libres dans une boîte ;

— les extensions auto-adjointes des opérateurs impulsion et hamiltonien pour une particule libre.

Appendices

2.A Transformation de Fourier

Soit f une fonction $\mathbb{R} \to \mathbb{R}$ appartenant à $\mathcal{L}^1(\mathbb{R})$, c.-à.-d. intégrable sur $\mathbb{R}$. On définit la transformée de Fourier de f par

$$\mathcal{F} \; : \quad k \; \mapsto \; \mathcal{F}(f)(k) = \frac{1}{\sqrt{2\pi}} \int dx \, e^{-ikx} \, f(x) \, . \tag{2.104}$$

On utilise également la notation $\hat{f}$ pour la transformée de Fourier de f. Si $\hat{f}$ est également intégrable sur $\mathbb{R}$, on peut définir la transformée de Fourier inverse

$$\mathcal{F}^{-1} \; : \quad \hat{f} \; \mapsto \; \mathcal{F}^{-1}(\hat{f})(x) = \frac{1}{\sqrt{2\pi}} \int dk \, e^{ikx} \, \hat{f}(k) \, , \tag{2.105}$$

qui coïncide avec f presque partout (en effet, $\mathcal{F}^{-1}(\hat{f})$ est continue alors que f ne l'est pas nécessairement).

Il est possible d'étendre la notion de transformation de Fourier des fonctions $f \in \mathcal{L}^1(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$ (c.-à.-d. à la fois sommable et de carré sommable) aux fonctions de carré sommable, en utilisant le théorème suivant :

Théorème 1 (théorème de Plancherel). *Soit* $f \in \mathcal{L}^2(\mathbb{R})$*. On définit la transformation de Fourier tronquée de* f *comme :*

$$\hat{f}_R : k \mapsto \frac{1}{\sqrt{2\pi}} \int_{-R}^{R} dx\, e^{-ikx} f(x) , \quad R > 0 . \tag{2.106}$$

Lorsque $R \to \infty$*,* $\hat{f}_R$ *converge pour la norme* $\| \quad \|_{\mathcal{L}^2}$ *vers une fonction* $\hat{f} \in \mathcal{L}^2(\mathbb{R})$*, la transformée de Fourier–Plancherel de* f*. On peut ainsi étendre la transformation de Fourier d'une application de* $\mathcal{L}^1(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$ *vers* $\mathcal{L}^2(\mathbb{R})$ *une application de* $\mathcal{L}^2(\mathbb{R})$ *dans lui-même. Cette application est une isométrie de* $\mathcal{L}^2(\mathbb{R})$*, c.-à.-d.*

$$\| f \|_{\mathcal{L}^2} = \left\| \hat{f} \right\|_{\mathcal{L}^2} \tag{2.107}$$

qui s'écrit plus explicitement

$$\int dx\, |f(x)|^2 = \int dk\, |\hat{f}(k)|^2 . \tag{2.108}$$

Finalement, on remarque que $\mathcal{F}^2(f)(x) = f(-x)$ ce qui implique que $\mathcal{F}^4$ est l'opérateur identité. En conséquence le spectre de l'opérateur $\mathcal{F}$, vu comme opérateur $\mathcal{L}^2(\mathbb{R}) \to \mathcal{L}^2(\mathbb{R})$, est $\sigma(\mathcal{F}) = \{1, -1, i, -i\}$.

Dans le contexte de la transformée de Fourier, il est avantageux de considérer des fonctions appartenant à l'espace de Schwartz $\mathcal{S}(\mathbb{R})$ défini par l'éqn. (2.12). En effet, la transformée de Fourier d'une fonction décroît d'autant plus vite à l'infini que cette fonction est régulière et, inversement, la transformée de Fourier d'une fonction est d'autant plus régulière que cette fonction décroît vite à l'infini. L'espace de Schwartz étant composé de fonctions infiniment dérivable *et* à décroissance rapide, on montre que :

Proposition 1 *Soit* $f \in \mathcal{S}(\mathbb{R})$*, alors la transformée de Fourier de* f*, définie par*

$$\mathcal{F}(f) \ : \quad k \ \mapsto \ \mathcal{F}(\, f\,)(k) = \frac{1}{\sqrt{2\pi}} \int dx\, e^{-ikx} f(x) , \tag{2.109}$$

appartient également à $\mathcal{S}(\mathbb{R})$ *et l'application* $\mathcal{F}$ *est une bijection de* $\mathcal{S}(\mathbb{R})$ *sur lui-même.* $\mathcal{F}$ *est également une isométrie pour la norme* $\| \ \|_{\mathcal{L}^2}$ *et donc une transformation unitaire sur* $\mathcal{S}(\mathbb{R})$ *:*

$$\mathcal{F}^\dagger \mathcal{F} = \mathcal{I} . \tag{2.110}$$

L'espace $\mathcal{S}(\mathbb{R})$ étant dense dans $\mathcal{L}^2(\mathbb{R})$, cela donne une manière alternative de définir la transformée de Fourier dans ce dernier espace. Cela indique aussi, par continuité, que la transformée de Fourier sur $\mathcal{S}(\mathbb{R})$ s'étend à une transformation unitaire sur $\mathcal{L}^2(\mathbb{R})$.

Propriété 1 *On peut montrer simplement les deux propriétés suivantes de la transformation de Fourier concernant la dérivation,*

$$\forall f \in \mathcal{S}(\mathbb{R}) , \quad \mathcal{F}(\, f' \,) : \ k \mapsto ik\mathcal{F}(\, f\,)(k) , \tag{2.111a}$$

$$\forall f \in \mathcal{S}(\mathbb{R}) , \quad \mathcal{F}(\, f\,)' : \ k \mapsto -i\mathcal{F}(\, xf\,)(k) , \tag{2.111b}$$

respectivement par intégrations par parties et en intervertissant la dérivation et l'intégration.

Passons maintenant à l'espace des distributions tempérées $\mathcal{S}(\mathbb{R})^\star$, c.-à.-d. l'espace des formes linéraires continues sur $\mathcal{S}(\mathbb{R})$. On peut en particulier associer à toute fonction $f \in \mathcal{L}^1(\mathbb{R})$ une distribution $\langle f|$, via [12]

$$\forall |\phi\rangle \in \mathcal{S}(\mathbb{R}) \ , \quad |\phi\rangle \mapsto \langle f|\phi\rangle = \int_{\mathbb{R}} dx\, \overline{f(x)}\phi(x) \in \mathbb{C}\,, \tag{2.112}$$

la convergence de l'intégrale étant garantie par la décroissance rapide de ϕ. La fonction f admettant une transformée de Fourier, on a

$$\langle \mathcal{F}(f)|\phi\rangle = \int_{\mathbb{R}} dy\, \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dx\, e^{ixy}\overline{f(x)}\, \phi(y) = \int_{\mathbb{R}} dx\, \overline{f(x)} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dy\, e^{ixy}\phi(y) = \left\langle f \middle| \mathcal{F}^{-1}(\phi)\right\rangle \ . \tag{2.113}$$

Nous pouvons alors généraliser cette propriété afin de *définir* la transformation de Fourier d'une distribution quelconque – l'espace $\mathcal{S}(\mathbb{R})^\star$ contenant des éléments qui ne sont associés à aucune fonction, comme la distribution de Dirac – de la manière suivante :

Définition 1 *Soit $\langle D| \in \mathcal{S}(\mathbb{R})^\star$ une distribution tempérée, c.-à.-d. une application linéaire continue sur l'espace de Schwartz $\mathcal{S}(\mathbb{R})$. La transformée de Fourier $\mathcal{F}_D$ de $\langle D|$ est définie par :*

$$\forall |\phi\rangle \in \mathcal{S}(\mathbb{R}) \ , \quad \langle \mathcal{F}_D|\phi\rangle = \left\langle D \middle| \mathcal{F}^{-1}(\phi)\right\rangle \ . \tag{2.114}$$

Grâce a cette définition il est possible de calculer des transformées de Fourier qui n'ont pas de sens pour des fonctions, en particulier pour la distribution de Dirac δ_x. D'après les définitions (2.112) et (2.11) nous avons :

$$\forall |\phi\rangle \in \mathcal{S}(\mathbb{R}) \ , \quad \langle \mathcal{F}(\delta_x)|\phi\rangle = \left\langle \delta_x \middle| \mathcal{F}^{-1}(\phi)\right\rangle = \mathcal{F}^{-1}(\phi)(x) = \frac{1}{\sqrt{2\pi}} \int dy\, e^{ixy}\phi(y)\,. \tag{2.115}$$

ainsi la transformée de Fourier de δ_x est la distribution associée à la fonction $e_x : k \mapsto$

12. Nous avons fait une entorse à la définition habituelle des distributions pour rendre les notations compatibles avec celles de la mécanique quantique. La distribution associée à une fonction $f \in \mathcal{L}^1(\mathbb{R})$ est habituellement définie comme la forme linéaire

$$\phi \in \mathcal{S}(\mathbb{R}) \mapsto \int dx\, f(x)\phi(x),$$

et avec cette convention l'équation (2.112) définit la distribution complexe conjuguée, associée à $\bar{f}$. Le lien entre les deux notation est fait en remarquant que

$$\int dx \overline{f(x)}\phi(x) = \overline{\int dx f(x)\overline{\phi(x)}} \ \Leftrightarrow \langle \bar{f}|\phi\rangle = \overline{\langle f|\bar{\phi}\rangle}\,.$$

Un effet de ce changement de convention est de remplacer dans la dernière égalité de (2.113) et dans (2.114) la transformation de Fourier par la transformation inverse.

$\frac{1}{\sqrt{2\pi}}e^{-ikx}$. Pour la transformation inverse nous obtenons

$$\forall|\phi\rangle \in \mathcal{S}(\mathbb{R})\ , \quad \langle \mathcal{F}^{-1}(e_x)|\phi\rangle = \langle e_x|\mathcal{F}(\phi)\rangle = \int du \frac{e^{ixu}}{\sqrt{2\pi}} \left(\int dy \frac{1}{\sqrt{2\pi}} e^{-iuy}\phi(y) \right)$$
$$= \frac{1}{2\pi} \int dy \left(\int du\, e^{iu(x-y)} \right) \phi(y)\,. \tag{2.116}$$

Comme, par ailleurs, $\mathcal{F}^{-1}(e_x) = \delta_x$ par définition de la transformée de Fourier inverse, nous en déduisons, en notation de physiciens, que

$$\int du\, e^{iu(x-y)} = 2\pi\delta_x(y) = 2\pi\delta(x-y)\,. \tag{2.117}$$

La généralisation de cette importante formule à trois dimensions est immédiate :

$$\boxed{\int d^3u\, e^{i\vec{u}\cdot(\vec{x}-\vec{x}')} = (2\pi)^3\delta(\vec{x}-\vec{x}')} \tag{2.118}$$

2.B L'oscillateur harmonique

L'oscillateur harmonique est un modèle simple approprié pour décrire la physique d'une particule massive au voisinage d'un minimum de son potentiel, voir fig. 2.2.

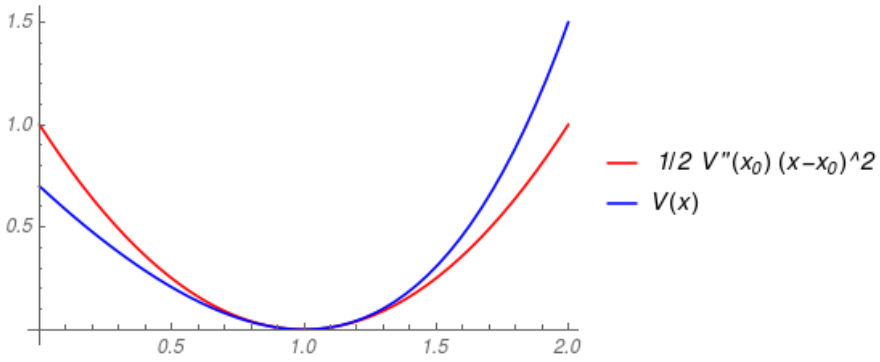

FIGURE 2.2 – Approximation quadratique d'un potentiel au voisinage d'un minimum.

À l'aide des principes de la quantification canonique, l'opérateur hamiltonien associé à ce système est donné par :

$$\boxed{\mathcal{H} = \frac{1}{2m}\mathcal{P}^2 + \frac{1}{2}m\omega^2\mathcal{X}^2} \tag{2.119}$$

Cet opérateur est bien auto-adjoint, car $\mathcal{X}$ et $\mathcal{P}$ le sont.

Nous cherchons à déterminer le spectre de cet opérateur, qui correspond aux énergies accessibles au système. Ce hamiltonien n'est a priori pas un opérateur borné, car l'énergie du

système peut être arbitrairement grande ; nous donnerons plus bas un domaine de définition naturel pour cet opérateur.

En représentation de position l'équation aux valeurs propres pour le hamiltonien prend la forme :

$$\langle x|\,\mathcal{H}\,|\psi\rangle = E\,\langle x|\psi\rangle \tag{2.120}$$

En utilisant les équations (2.20), (2.42) et la relation de fermeture (2.27) nous avons

$$\langle x|\,\mathcal{X}^2\,|\psi\rangle = x^2\,\langle x|\psi\rangle = x^2\psi(x) \tag{2.121a}$$

$$\begin{aligned}\langle x|\,\mathcal{P}^2\,|\psi\rangle &= \int dx'\,\langle x|\,\mathcal{P}\,|x'\rangle\,\langle x'|\,\mathcal{P}\,|\psi\rangle = i\hbar\int dx'\delta'(x-x')\left(-i\hbar\frac{d}{dx'}\psi(x')\right)\\ &= -\hbar^2\frac{d^2\psi(x)}{dx^2}\end{aligned} \tag{2.121b}$$

Nous obtenons ainsi l'équation différentielle

$$\boxed{\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} = \left(\frac{m\omega^2x^2}{2} - E\right)\psi(x)} \tag{2.122}$$

Si on divise par $\frac{\hbar\omega}{2}$ et on introduit la variable sans dimension $u = \sqrt{\frac{m\omega}{\hbar}}x$, l'équation (2.122) devient, avec $f(u) = \psi(x)$:

$$\frac{d^2}{du^2}f(u) + \left(\frac{2E}{\hbar\omega} - u^2\right)f(u) = 0 \tag{2.123}$$

La comparaison avec l'équation (2.8) nous indique que nous avons obtenu l'équation d'Hermite, à la condition que

$$\frac{2E}{\hbar\omega} = 2n+1 \;\Leftrightarrow\; \boxed{E = \hbar\omega\left(n+\frac{1}{2}\right)\;,\quad n\in\mathbb{N}} \tag{2.124}$$

Les fonctions propres correspondantes sont les fonctions de Hermite ψ_n données par l'équation (2.6). Ces dernières formant une base orthonormée de $\mathcal{L}^2(\mathbb{R})$, nous avons bien obtenu une base orthonormée de l'espace des états formée de vecteurs propres du hamiltonien.

Revenons maintenant sur le choix du domaine de définition de $\mathcal{H}$. Le hamiltonien (2.119) comprenant dans le deuxième terme l'opérateur de multiplication par x^2, il n'est pas défini sur $\mathcal{L}^2(\mathbb{R})$. Au vu de l'analyse précédente un domaine de définition naturel est, en termes de la variable u sans dimension,

$$D(\mathcal{H}) = \text{Vect}\left(\left\{u^r e^{-u^2/2},\, r\in\mathbb{N}^\star\right\}\right) \subset \mathcal{L}^2(\mathbb{R})\,. \tag{2.125}$$

ce domaine, qui est dense dans $\mathcal{L}^2(\mathbb{R})$, contient naturellement les fonctions de Hermite.

L'oscillateur harmonique s'étudie plus communément par des méthodes algébriques. Introduisons les *opérateurs d'échelle* :

$$a \overset{\text{déf.}}{=} \sqrt{\frac{m\omega}{2\hbar}}X + \frac{i}{\sqrt{2m\hbar\omega}}P\,, \tag{2.126a}$$

$$a^\dagger \overset{\text{déf.}}{=} \sqrt{\frac{m\omega}{2\hbar}}X - \frac{i}{\sqrt{2m\hbar\omega}}P\,. \tag{2.126b}$$

Comme la notation le suggère, comme X et P sont des opérateurs auto-adjoints a et $a^\dagger$ sont conjugués hermitiens l'un de l'autre. En utilisant la relation de commutation canonique (2.34) on obtient facilement que

$$\boxed{\left[\,a,\,a^\dagger\,\right] = 1} \tag{2.127}$$

En exprimant X et P en termes de a et $a^\dagger$ et en remplaçant dans le hamiltonien (2.119) on a finalement :

$$\boxed{\mathcal{H} = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right)} \tag{2.128}$$

On définit l'opérateur

$$\mathcal{N} \overset{\text{déf.}}{=} a^\dagger a\,, \tag{2.129}$$

hermitien d'après la propriété (1.39c). De manière évidente, les vecteurs propres de $\mathcal{N}$ sont identiques aux vecteurs propres de $\mathcal{H}$ et cet opérateur est auto-adjoint sur le même domaine $D(\mathcal{H})$ donné par l'éqn. (2.125). On calcule ensuite les commutateurs suivants, grâce à la propriété (1.36) :

$$[\mathcal{N},\,a\,] = -a \tag{2.130a}$$

$$\left[\mathcal{N},\,a^\dagger\,\right] = a^\dagger \tag{2.130b}$$

On remarque également que le domaine $D(\mathcal{H})$ est stable sous l'action de a et $a^\dagger$. Supposons maintenant l'existence d'un vecteur propre normé $|n\rangle \in D(\mathcal{H})$ pour l'opérateur $\mathcal{N}$:

$$\mathcal{N}\,|n\rangle = n\,|n\rangle \;\;,\quad \langle n|n\rangle = 1\,, \tag{2.131}$$

ou $n \in \sigma(\mathcal{N}) \subset \mathbb{R}$ car $\mathcal{N}$ est un opérateur auto-adjoint. Les identités (2.133) permettent d'écrire

$$\mathcal{N}\,(a\,|n\rangle) = (a\mathcal{N} - a)\,|n\rangle = (n-1)\,(a\,|n\rangle)\;, \tag{2.132a}$$

$$\mathcal{N}\left(a^\dagger\,|n\rangle\right) = \left(a^\dagger\mathcal{N} + a^\dagger\right)|n\rangle = (n+1)\left(a^\dagger\,|n\rangle\right)\;. \tag{2.132b}$$

$$\tag{2.132c}$$

Ainsi $a\,|n\rangle$ et $a^\dagger\,|n\rangle$ sont également des vecteurs propres de $\mathcal{N}$, s'ils sont non nuls, contenus dans $D(\mathcal{H})$ en raison de la stabilité de $D(\mathcal{H})$ sous l'action de a et $a^\dagger$, associés respectivement aux valeurs propres $n-1$ et $n+1$. De la même manière, on trouve par induction que le

vecteur $a^\ell |n\rangle$ (resp. $(a^\dagger)^\ell |n\rangle$) avec $\ell \in \mathbb{N}^\star$ s'il est non nul, est vecteur propre de $\mathcal{N}$ pour la valeur propre $n - \ell$ (resp. $n + \ell$).

On peut ensuite calculer la norme de ces vecteurs propres en utilisant les identités (2.133) et le commutateur (2.127) :

$$\| a |n\rangle \|^2 = \langle n| a^\dagger a |n\rangle = \langle n| \mathcal{N} |n\rangle = n \langle n|n\rangle = n \,, \tag{2.133a}$$

$$\left\| a^\dagger |n\rangle \right\|^2 = \langle n| a a^\dagger |n\rangle = \langle n| (\mathcal{N} + 1) |n\rangle = (n+1) \langle n|n\rangle = n + 1 \,. \tag{2.133b}$$

Le produit hermitien sur $\mathcal{H}$ étant défini positif, la norme de tout vecteur doit être positive ou nulle.

La solution du problème est de supposer l'existence d'un *état fondamental* de valeur propre nulle,

$$\mathcal{N} |0\rangle = 0 \,, \quad \langle 0|0\rangle = 1 \,, \tag{2.134}$$

à partir duquel les autres états sont obtenus. La relation (2.133a) nous indique que ce vecteur satisfait également à l'équation :

$$\boxed{a |0\rangle = 0} \tag{2.135}$$

L'énergie de cet état fondamental, $E_0 = \frac{\hbar\omega}{2}$, est non nulle, en raison des fluctuations quantiques autour du minimum du potentiel.

D'après la relation le vecteur $a^\dagger |0\rangle$, et par induction les vecteurs de la forme $(a^\dagger)^n |0\rangle$, sont non-nuls et sont donc vecteurs propres de $\mathcal{N}$ avec la valeur propre n. Calculons leur norme à l'aide du commutateur (2.127) et des relations (2.132) :

$$\begin{aligned} \left\| (a^\dagger)^n |0\rangle \right\|^2 &= \langle 0| a^n (a^\dagger)^n |0\rangle = \langle 0| a^{n-1} (a a^\dagger) (a^\dagger)^{n-1} |0\rangle = \langle 0| a^{n-1} (a^\dagger a + 1) (a^\dagger)^{n-1} |0\rangle \\ &= \langle 0| a^{n-1} (\mathcal{N} + 1) (a^\dagger)^{n-1} |0\rangle = \langle 0| a^{n-1} (n - 1 + 1) (a^\dagger)^{n-1} |0\rangle \\ &= n \langle 0| a^{n-1} (a^\dagger)^{n-1} |0\rangle = \cdots = n! \langle 0|0\rangle = n! \end{aligned} \tag{2.136}$$

Finalement montrons qu'une famille orthonormée de vecteurs propres de $\mathcal{N}$, et donc de $\mathcal{H}$, est fournie par

$$\boxed{|n\rangle = \frac{1}{\sqrt{n!}} \left(a^\dagger\right)^n |0\rangle} \tag{2.137}$$

Ces vecteurs sont normés d'après le calcul précédent. Pour montrer qu'ils sont orthogonaux, on choisit $m, n \in \mathbb{N}^\star$ avec $m \neq n$ et on suppose en toute généralité que $m = n + r$, $r \in \mathbb{N}^\star$. On a, d'après (2.136) :

$$\langle m|n\rangle = \frac{1}{\sqrt{m!n!}} \langle 0| a^m \left(a^\dagger\right)^n |0\rangle = \frac{1}{\sqrt{m!n!}} \langle 0| a^r a^n \left(a^\dagger\right)^n |0\rangle = \frac{n!}{\sqrt{m!n!}} \langle 0| a^r |0\rangle = 0 \,. \tag{2.138}$$

Finalement nous pouvons faire le lien entre l'approche fonctionnelle et l'approche algébrique. L'équation (2.135) définissant l'état fondamental devient en représentation position

$$0 = \langle x| a |0\rangle = \langle x| \left(\sqrt{\frac{m\omega}{2\hbar}} X + \frac{i}{\sqrt{2m\hbar\omega}} P \right) |0\rangle = \sqrt{\frac{\hbar}{2m\omega}} \left(\frac{m\omega}{\hbar} x + \frac{d}{dx} \right) \psi_0(x) \,, \tag{2.139}$$

donc la solution normalisée est unique, à une phase près, et est donnée par une gaussienne :

$$\boxed{\phi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega}{2\hbar}x^2}} \tag{2.140}$$

l'unicité de cette solution garantit celle de l'état fondamental $|0\rangle$ et, par induction, celle de tout les vecteurs propres ; le hamiltonien de l'oscillateur harmonique a ainsi un spectre discret et non dégénéré. Les autres états s'obtiennent par action de $a^\dagger$ et on peut écrire

$$\boxed{\phi_n(x) = \frac{1}{\sqrt{n!}} \langle x| \left(a^\dagger\right)^n |0\rangle = \frac{1}{\sqrt{n!}} \left(\frac{\hbar}{2m\omega}\right)^{n/2} \left(\frac{m\omega}{\hbar}x - \frac{d}{dx}\right)^n \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega}{2\hbar}x^2}} \tag{2.141}$$

qui ne sont rien d'autres que les solutions obtenues par l'approche d'analyse, en termes de fonctions de Hermite.

On reprend en effet la variable sans dimension $u = \sqrt{\frac{m\omega}{\hbar}}x$ et on pose $\psi_n(u)du = \phi_n(x)dx$, ce qui donne après simplification des constantes

$$\psi_n(u) = \frac{1}{\sqrt{2^n n!\sqrt{\pi}}} \left(u - \frac{d}{du}\right)^n e^{-u^2/2}, \tag{2.142}$$

et on démontre ensuite par une récurrence immédiate sur n :

$$e^{u^2/2} \left(u - \frac{d}{du}\right)^n e^{-u^2/2} = (-1)^n e^{u^2} \frac{d^n}{du^n} e^{-u^2}, \tag{2.143}$$

qui permet d'identifier les ψ_n avec les fonctions de Hermite, voir l'équation (2.6).

Bibliographie

[1] Harald Ibach et Hans Lüth, "Solid-State Physics An Introduction to Principles of Materials Science," Springer, 2009

[2] L. Schwartz, "Théorie des distributions," 1957, Éditions Hermann.

[3] W. a Heisenberg, "Uber den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik," Z. Phys. **43**, 172 (1927). doi :10.1007/BF01397280

Cours n°3

Symétries en mécanique quantique

Les symétries jouent un rôle absolument crucial en physique en géneral, et en physique quantique en particulier. Connaître les symétries d'un problème physique permet d'organiser les degrés de liberté du système selon cette symétrie et facilite grandement la recherche de solutions du problème posé, qui doivent respecter cette symétrie. En outre il existera des quantités conservées au cours du temps associées à cette symétrie dont le rôle est fondamental dans la classification des solutions.

En physique quantique les symétries jouent un autre rôle, tout aussi important. Dans un système en interaction, les couplages entre les degrés de liberté reçoivent en général des corrections, comme les corrections fines et hyperfines à l'atome d'hydrogène. Généralement, tous les couplages possibles peuvent apparaître, en raison des fluctuations quantiques. Cependant, s'il existe une symétrie dans le problème, et que cette symétrie est exacte, elle va interdire a un grand nombre de couplages d'apparaître ce qui reduira drastiquement la forme et le nombre des corrections quantiques possibles. Pour revenir à l'exemple de l'atome d'hydrogène, toutes les corrections doivent respecter la symétrie de rotation du système.

En théorie quantique des champs, extension relativiste de la mécanique quantique et langage dans lequel s'exprime la physique des particules, le rôle des symétries est encore plus profond. Mis à part quelques principes très généraux (unitarité, renormalisabilité,...) la théorie physique des particules élémentaires et de leurs interactions est *uniquement* déterminée par les symétries du problème et la manière dont les différentes particules se transforment par rapport à ces symétries. Une fois énoncées ces données le modèle standard des particules élémentaires est complètement déterminé mis à part les valeurs d'une trentaine de paramètres libres.

Dans ce cours nous présenterons quelques outils généraux sur l'utilisation des symétries, d'une part en mécanique classique et d'autre part, en plus grand détail, en mécanique quantique. Les cours suivants seront consacrés à l'application de ces outils à diverses symétries de la Nature.

3.1 Mécanique lagrangienne : rappels

On considère un système dynamique classique dont les degrés de liberté sont représentés par des coordonnées généralisées $\{q^i(t);\, i = 1, \ldots, n\}$ évoluant au cours du temps, entre l'instant initial $t = 0$ et l'instant final $t = T$. On associe au système une fonction $\mathcal{L}(q^i(t), \dot{q}^i(t), t)$ appelée *lagrangien*, dépendant des coordonnées généralisées, de leurs dérivées temporelles et éventuellement explicitement du temps. L'intégrale temporelle du lagrangien pendant le temps d'évolution est une fonctionnelle appelée *action* :

$$\mathcal{S}[q] \overset{\text{déf.}}{=} \int_0^T \mathrm{d}t\, \mathcal{L}(q^i(t), \dot{q}^i(t), t)\,. \tag{3.1}$$

Cette action a la dimension du produit d'une énergie par un temps.

Le système va suivre au cours du temps une trajectoire bien définie dans l'espace balayée par les coordonnées généralisées $q^i(t)$. Considérons que le système occupe des positions initiales et finales données $q^i(0) = q^i_0$ et $q^i(T) = q^i_T$. La trajectoire entre ces deux points est alors régie par le *principe de moindre action* : les fonctions $\{q^i(t);\, i = 1, \cdots n\}$ doivent être choisies de telle sorte que l'action du système (3.1) soit minimale (de manière plus générale, un extremum local ou un point col). Pour exprimer ce principe de manière mathématique, on considère une variation des fonctions $q^i(t)$ autour de la trajectoire $Q^i(t)$ effectivement suivie par le système :

$$q^i(t) = Q^i(t) + \delta^i(t)\ , \quad \delta^i(0) = \delta^i(T) = 0\,. \tag{3.2}$$

La variation au premier ordre correspondante de l'action (3.1) est alors donnée par :

$$\begin{aligned}
\delta\mathcal{S} &= \int_0^T \mathrm{d}t\, \Big(\mathcal{L}(Q^i(t) + \delta^i(t), \dot{Q}^i(t) + \dot{\delta}^i(t), t) - \mathcal{L}(Q^i(t), \dot{Q}^i(t), t)\Big) \\
&= \int_0^T \mathrm{d}t\, \left(\frac{\partial \mathcal{L}(q^i, \dot{q}^i, t)}{\partial q^i}\delta^i(t) + \frac{\partial \mathcal{L}(q^i, \dot{q}^i, t)}{\partial \dot{q}^i}\dot{\delta}^i(t) + \mathcal{O}(\delta^2)\right) \\
&= \int_0^T \mathrm{d}t\, \sum_{i=1}^{n} \left(\frac{\partial \mathcal{L}(q^i, \dot{q}^i, t)}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}(q^i, \dot{q}^i, t)}{\partial \dot{q}^i}\right)\delta^i(t) + \left[\sum_{i=1}^{n} \delta^i \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{q}^i}\right]_0^T .
\end{aligned} \tag{3.3}$$

Le deuxième terme de cette dernière expression va s'annuler en vertu des conditions aux limites (3.2). On demande alors que le premier terme s'annule pour une variation arbitraire $\delta^i(t)$ autour de la trajectoire, pour obtenir un point stationnaire de l'action. Cela donne les *équations d'Euler–Lagrange* :

$$\frac{\partial \mathcal{L}(q^i, \dot{q}^i, t)}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}(q^i, \dot{q}^i, t)}{\partial \dot{q}^i} = 0\,. \tag{3.4}$$

Cette formulation de la mécanique a pour avantage de ne faire apparaître que les degrés de liberté réels du système.

Considérons un ensemble de N particules libres de masses m_k et de coordonnées $\vec{x}^k(t)$, avec $k = 1, \ldots, N$. On suppose que ces particules interagissent via un potentiel $V(\vec{x}^\ell)$ dépendant uniquement de leurs positions respectives. Le lagrangien du système est alors :

$$\mathcal{L} = \frac{1}{2} \sum_{k=1}^{N} m_k \|\dot{\vec{x}}^k\|^2 - V(\vec{x}^l)\,, \tag{3.5}$$

ce qui conduit via (3.4) aux équations habituelles de la mécanique du point en présence d'un potentiel :

$$m_k \ddot{\vec{x}}^k + \frac{\partial V(\vec{x}^l)}{\partial \vec{x}^k} = 0\,. \tag{3.6}$$

Dans la dérivation des équations d'Euler–Lagrange (3.4), on considère une variation de la trajectoire autour de celle effectivement suivie par le système. Cependant, il ne s'agit ici que d'une étape intermédiaire de calcul !

Du point de vue physique, le système suit la trajectoire obéissant aux équations du mouvement et aux conditions aux limites en $t = 0$ et $t = T$, correspondant à un jeu de fonctions $\{q^i(t), i = 1, \ldots, n\}$, et n'explore pas les autres chemins qui ne minimisent pas l'action.

3.2 Symétries en physique classique

Avant d'aborder le rôle des symétries en mécanique quantique, nous commencerons par rappeler quelques resultats importants concernant les symétries en mécanique quantique. La question peut être abordée soit du point de vue de la mécanique lagrangienne, soit du point de vue de la mécanique hamiltonienne.

3.2.1 Charges de Noether

Commençons par l'approche lagrangienne. On considère un système mécanique, de coordonnées généralisées $\{q^i(t), i = 1, \ldots, n\}$, décrit par une action

$$\mathcal{S} = \int_{t_0}^{t_1} dt\, \mathcal{L}(q^i, \dot{q}^i, t)\,. \tag{3.7}$$

Soit $q_c(t)$ une solution classique, c.-à.-d. une fonction $q_c(t)$, $t \in [t_0, t_1]$, satisfaisant les équations d'Euler–Lagrange (3.4).

On considère une symétrie continue du système, correspondant à une transformation continue des coordonnées :

$$q^i(t) \mapsto \tilde{q}^i[q^j(t), \dot{q}^j(t)]\,, \tag{3.8}$$

Cette transformation est une *symétrie* de l'action (3.7) si cette dernière est invariante, ou est uniquement modifiée par un terme de dérivée totale qui ne change pas les équations

du mouvement [1] Pour une transformation infinitésimale autour de la solution $q_c(t)$, nous écrivons

$$\delta^i(t) \overset{\text{déf.}}{=} \tilde{q}^i[q^j_c(t), \dot{q}^j_c(t)] - q^i_c(t)\,. \tag{3.9}$$

Contrairement aux variations que nous avions considérées pour dériver les équations d'Euler–Lagrange, éqn. (3.2), nous ne demandons pas à δq^i de s'annuler en $t = t_0$ et $t = t_1$.

La variation de l'action (3.7) au premier ordre est donnée, après intégration par parties, par :

$$\begin{aligned}\delta\mathcal{S} &= \int_{t_0}^{t_1} dt \left\{ \mathcal{L}(q^i + \delta^i, \dot{q}^i + \dot{\delta}^i, t) - \mathcal{L}(q^i, \dot{q}^i, t) \right\} \\ &= \int_{t_0}^{t_1} dt \sum_i \left(\frac{\partial \mathcal{L}}{\partial q^i} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}^i} \right) \Big|_{q^i = q^i_c} \delta q^i + \left[\sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}^i}\, \delta q^i \right]_{t=t_0}^{t=t_1}\,. \end{aligned} \tag{3.10}$$

Le premier terme s'annule car $q_c(t)$ satisfait aux équations d'Euler–Lagrange ; demandant que le deuxième terme soit égal à une dérivée totale nous obtenons :

$$\left[\sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}^i}\, \delta q^i - \Lambda \right]_{t=t_0}^{t=t_1} = 0\,. \tag{3.11}$$

Les instants t_0 et t_1 étant arbitraires, nous arrivons à une version simplifiée du théorème de Noether :

Théorème 1 (théorème de Noether). *Soit un système mécanique décrit par un lagrangien $\mathcal{L}(q^i, \dot{q}^i, t)$ possédant une symétrie continue associée au changement de coordonnées infinitésimal $q^i \mapsto q^i + \delta q^i$. À cette symétrie est associée une charge de Noether,*

$$\boxed{Q \overset{\text{déf.}}{=} \sum_i \frac{\partial \mathcal{L}}{\partial \dot{q}^i}\, \delta q^i - \Lambda} \tag{3.12}$$

où le second terme correspond à la dérivée totale éventuelle obtenue par variation de l'action par la symétrie. Cette charge est conservée, c.-à.-d. qu'elle reste constante au cours du temps.

Énergie

Considérons un système dont le lagrangien ne dépend pas explicitement du temps. Le système est alors invariant par translation dans le temps. Soit la transformation

$$\begin{cases} t & \mapsto & \tilde{t} = t - \tau \\ q^i(t) & \mapsto & \tilde{q}^i(\tilde{t}) = q^i(t)\,. \end{cases} \tag{3.13}$$

1. De manière équivalente, on autorise le lagrangien à se transformer comme $\mathcal{L} \mapsto \mathcal{L} + \frac{d}{dt}\Lambda$ sous l'action de la symétrie.

La variation des coordonnées associée est :

$$\delta q^i(t) = \tilde{q}^i(t) - q^i(t) = \tilde{q}^i(\tilde{t} + \tau) - q^i(t) = \frac{dq^i}{dt}\tau + \mathcal{O}(\tau^2)\,. \tag{3.14}$$

Calculons la variation de l'action.

$$\delta\mathcal{S} = \int_{t_0}^{t_1} dt\, \left\{\mathcal{L}(q^i + \dot{q}^i\tau, \dot{q}^i + \ddot{q}^i\tau) - \mathcal{L}(q^i, \dot{q}^i)\right\} = \tau\int_{t_0}^{t_1} dt\, \frac{d\mathcal{L}}{dt} + \mathcal{O}(\tau^2) = \tau\left[\,\mathcal{L}\,\right]_{t=t_0}^{t=t_1}. \tag{3.15}$$

La charge conservée obtenue n'est rien d'autre que le hamiltonien du système,

$$H = \frac{1}{\tau}\left(\sum_i \frac{\partial\mathcal{L}}{\partial\dot{q}^i}\,\delta q^i - \Lambda\right) = \sum_i \frac{\partial\mathcal{L}}{\partial\dot{q}^i}\dot{q}^i - \mathcal{L}\,. \tag{3.16}$$

Ainsi, si le lagrangien ne dépend pas explicitement du temps, l'énergie du système est constante dans le temps.

Quantité de mouvement

Considérons maintenant un système dont le lagrangien ne dépend pas explicitement de la coordonnée q^j – mais pouvant dépendre de sa dérivée temporelle. Le système est alors invariant par translation spatiale dans cette direction. Soit la transformation

$$\begin{cases} q^j(t) & \mapsto & q^j(t) + \lambda \\ q^i(t) & \mapsto & q^i(t) \qquad i \neq j \end{cases} \tag{3.17}$$

La variation des coordonnées associée est :

$$\delta q^i(t) = \lambda\delta^{i,j}\,. \tag{3.18}$$

Calculons la variation de l'action.

$$\delta\mathcal{S} = \int_{t_0}^{t_1} dt\, \left\{\mathcal{L}(q^i + \lambda\delta^{i,j}, \dot{q}^i, t) - \mathcal{L}(q^i, \dot{q}^i, t)\right\} = \lambda\int_{t_0}^{t_1} dt\, \frac{\mathcal{D}\mathcal{L}}{\mathcal{D}q^j} = 0\,, \tag{3.19}$$

par hypothèse. La quantité conservée n'est rien d'autre que la j-ième composante du vecteur quantité de mouvement :

$$p^j = \frac{\partial\mathcal{L}}{\partial\dot{q}^j}\,. \tag{3.20}$$

Moment cinétique

Considérons finalement un système invariant par rotation, c.-à.-d. invariant par la transformation

$$q^i(t) \;\mapsto\; \sum_j R^i{}_j q^j(t) \tag{3.21}$$

où R est une matrice de rotation, matrice 3×3 satisfaisant

$$\det R = 1 \,, \quad R^T R = \mathbb{I} \,. \tag{3.22}$$

On considère une rotation infinitésimale autour de l'identité. On développe :

$$q^i \;\mapsto\; \sum_j \left(\delta^i{}_j - \theta \mathfrak{T}^i{}_j\right) q^j \,, \tag{3.23}$$

avec $|\theta| \ll 1$. La matrice R étant de déterminant unité on a

$$\det(\mathbb{I} + \theta\mathfrak{T}) = 1 - \theta \mathrm{Tr}\,\mathfrak{T} + \mathcal{O}(\theta^2) \quad \Longrightarrow \quad \mathrm{Tr}\,\mathfrak{T} = 0 \,, \tag{3.24}$$

et l'orthogonalité de R implique que

$$R^T R = (\mathbb{I} - \theta\mathfrak{T}^T)(\mathbb{I} - \theta\mathfrak{T}) = \mathbb{I} - \theta(\mathfrak{T}^T + \mathfrak{T}) + \mathcal{O}(\theta^2) \quad \Longrightarrow \quad \mathfrak{T}^T + \mathfrak{T} = 0 \,. \tag{3.25}$$

La matrice $\mathfrak{T}$ est donc antisymmétrique (ce qui implique évidemment qu'elle soit de trace nulle). Une matrice 3×3 antisymétrique possède 3 composantes indépendantes, étant de la forme

$$\mathfrak{T} = \begin{pmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{pmatrix} \,, \quad a, b, c, \in \mathbb{R} \,. \tag{3.26}$$

Les trois composantes de $\mathfrak{T}$ peuvent être vues comme les trois composantes d'un (pseudo)-vecteur. Introduisons pour cela le symbole totalement antisymétrique de Levi–Civita :

$$\epsilon^{123} = 1 \,, \quad \epsilon^{ijk} = \begin{cases} \epsilon_\sigma & \text{si} \quad \exists \sigma \in S_3,\, \sigma(ijk) = (123) \\ 0 & \text{sinon} \end{cases} \,, \tag{3.27}$$

où σ désigne une permutation de trois éléments appartenant au groupe symétrique S_3, dont ϵ_σ est la signature. [2] On peut alors écrire

$$\mathfrak{T}^i{}_j = \sum_k \epsilon^i{}_{jk} n^k \,, \quad \vec{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \,. \tag{3.28}$$

Si on choisit $\vec{n}$ de norme unité dans l'éqn. (3.23), ce vecteur donne l'axe de rotation alors que θ désigne l'angle de la rotation infinitésimale.

On considère donc une variation infinitésimale des coordonnées dans le lagrangien de la forme :

$$\delta q^i(t) = -\theta \sum_j \mathfrak{T}^i{}_j q^j(t) \,, \tag{3.29}$$

qui peut s'écrire en notation vectorielle comme

$$\delta \vec{q} = \theta\, \vec{n} \wedge \vec{q} \,. \tag{3.30}$$

2. Le symbole de Levi–Civita intervient naturellement dans le calcul du produit mixte ou déterminant à trois dimensions. On vérifie aisément que $\det(\vec{u}, \vec{v}, \vec{w}) = \vec{u} \cdot (\vec{v} \wedge \vec{w}) = \sum_{i,j,k} \epsilon^{ijk} u_i v_j w_k$.

Calculons la variation de l'action. Si le lagrangien est invariant par rotation, il ne dépend que de $q \overset{\text{déf.}}{=} \|\vec{q}\|$ et de $v \overset{\text{déf.}}{=} \|\dot{\vec{q}}\|$. Nous avons alors :

$$\begin{aligned}\delta\mathcal{S} &= \int_{t_0}^{t_1} dt \left\{ \frac{\mathcal{D}\mathcal{L}}{\mathcal{D}q}\frac{\mathcal{D}q}{\mathcal{D}q^i}\delta q^i + \frac{\mathcal{D}\mathcal{L}}{\mathcal{D}v}\frac{\mathcal{D}v}{\mathcal{D}\dot{q}^i}\delta\dot{q}^i \right\} = \int_{t_0}^{t_1} dt \left\{ \frac{\mathcal{D}\mathcal{L}}{\mathcal{D}q}\frac{1}{q}\vec{q}\cdot\delta\vec{q} + \frac{\mathcal{D}\mathcal{L}}{\mathcal{D}v}\frac{1}{v}\dot{\vec{q}}\cdot\delta\dot{\vec{q}} \right\} \\ &= \theta \int_{t_0}^{t_1} dt \left\{ \frac{\mathcal{D}\mathcal{L}}{\mathcal{D}q}\frac{1}{q}\vec{q}\cdot(\vec{n}\wedge\vec{q}) + \frac{\mathcal{D}\mathcal{L}}{\mathcal{D}v}\frac{1}{v}\dot{\vec{q}}\cdot(\vec{n}\wedge\dot{\vec{q}}) \right\} = 0\,, \end{aligned} \tag{3.31}$$

où nous avons utilisé l'antisymétrie du produit mixte.

La quantité conservée obtenue n'est rien d'autre que le moment cinétique projeté le long du vecteur $\vec{n}$ associé à l'axe de rotation :

$$\sum_i \frac{\partial\mathcal{L}}{\partial\dot{q}^i}\delta q^i = \theta\,\vec{p}\cdot(\vec{n}\wedge\vec{q}) = \theta\,\vec{n}\cdot\vec{L}\,, \tag{3.32}$$

où $\vec{p}$ est le vecteur quantité de mouvement et $\vec{L} = \vec{q}\wedge\vec{p}$ le vecteur moment cinétique.

3.2.2 Formalisme hamiltonien

Commençons par quelques rappels de mécanique hamiltonienne. Pour un système décrit par un lagrangien $\mathcal{L}(q^i(t), \dot{q}^i(t), t)$ dépendant de coordonnées généralisées $q^i(t)$, on définit premièrement les impulsions généralisées conjuguées aux q^i par :

$$p_i \overset{\text{déf.}}{=} \frac{\partial\mathcal{L}(q^i, \dot{q}^i, t)}{\partial\dot{q}^i}\,. \tag{3.33}$$

Par une transformation de Legendre on obtient alors le hamiltonien du système, fonction des variables indépendantes $\{q^i, p_i;\ i = 1, \ldots, n\}$ et du temps :

$$\mathcal{H}(p_i, q^i, t) \overset{\text{déf.}}{=} \sum_i p_i\dot{q}^i - \mathcal{L}\,. \tag{3.34}$$

Les variables indépendantes satisfont aux équations de Hamilton, qui sont du premier ordre :

$$\dot{q}^i = \frac{\partial\mathcal{H}}{\partial p_i}\,, \quad \dot{p}_i = -\frac{\partial\mathcal{H}}{\partial q^i}\,. \tag{3.35}$$

Un des avantages de cette formulation de la mécanique est que l'état du système est entièrement déterminé par la donnée des $\{q^i, p_i;\ i = 1, \ldots, n\}$, qui correspondent à un système de coordonnés sur l'*espace des phases* du système.

En mécanique hamiltonienne, l'évolution temporelle d'une fonction $f(q^i, p_i, t)$ des variables $\{q^i, p_i\}$ et du temps s'établit de la manière suivante. Premièrement,

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_i \left(\frac{\partial f}{\partial q^i}\dot{q}^i + \frac{\partial f}{\partial p_i}\dot{p}_i \right)\,. \tag{3.36}$$

En utilisant les équations de Hamilton (3.35) cela s'écrit

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \{ f, \mathcal{H} \} , \tag{3.37}$$

où nous avons introduit les *crochets de Poisson* entre deux fonctions sur l'espace des phases :

$$\{f, g\} \overset{\text{déf.}}{=} \sum_i \left(\frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i} \right) . \tag{3.38}$$

Nous avons en particulier

$$\{q^j, p_k\} = \sum_i \left(\frac{\partial q^j}{\partial q^i} \frac{\partial p_k}{\partial p_i} - \frac{\partial q^j}{\partial p_i} \frac{\partial p_k}{\partial q^i} \right) = \sum_i \delta^{i,j} \delta_{i,k} = \delta^j{}_k , \tag{3.39}$$

en termes du symbole de Kronecker

$$\delta_{i,j} \overset{\text{déf.}}{=} \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases} , \tag{3.40}$$

ce qui permet de donner les équations de Hamilton sous la forme :

$$\frac{dp_i}{dt} = \{p_i, \mathcal{H}\} = -\frac{\partial \mathcal{H}}{\partial q_i} \tag{3.41a}$$

$$\frac{dq_i}{dt} = \{q_i, \mathcal{H}\} = \frac{\partial \mathcal{H}}{\partial p_i} . \tag{3.41b}$$

L'équation d'évolution (3.37) donne une autre propriété importante de la mécanique.

Propriété 1 *une fonction de l'espace des phases sans dépendance temporelle explicite est constante dans le temps si et seulement si son crochet de Poisson avec le hamiltonien est nul.*

$$\frac{df(q^i, p_i)}{dt} = 0 \;\Leftrightarrow\; \{f, \mathcal{H}\} = 0 \tag{3.42}$$

Les conséquences de ce résultat apparaîtront plus loin dans ce chapitre.

*
* *

Revenons maintenant aux considérations de symétrie. Remarquons tout d'abord qu'une charge de Noether, voir éqn. (3.12), s'exprime dans les variables canoniques comme

$$Q = \sum_i p_i \delta q^i(q^j, p_j, t) - \Lambda(q^j, p_j, t) . \tag{3.43}$$

La charge Q étant une constante du mouvement, partant de l'éqn. (3.37) nous trouvons la forme hamiltonienne du théorème de Noether :

$$0 = \frac{dQ}{dt} = \frac{\partial Q}{\partial t} + \{ H, Q \} . \tag{3.44}$$

Dans le cas – le plus fréquent – où Q ne dépend pas explicitement du temps, le crochet de Poisson du hamiltonien avec la charge de Noether est nul : $\{H, Q\} = 0$.

Il existe un lien profond entre les symétries et les transformations canoniques de la mécanique hamiltonienne (cf. cours de pré-rentrée et [1]). Rappelons qu'une transformation canonique est un changement de variables dans l'espace des phases, c.-à.-d. , en se restreignant ici aux transformations ne dépendant pas explicitement du temps,

$$q^i \mapsto \tilde{q}^i(q^j, p_j)\,, \quad p_i \mapsto \tilde{p}_i(q^j, p_j)\,, \tag{3.45}$$

qui préservent les relations de Hamilton (3.35) ou, de manière équivalente, les crochets de Poisson canoniques, c.-à.-d. satisfassent

$$\{\,\tilde{q}^i,\,\tilde{p}_j\,\}_{p,q} = \delta^i_j\,, \quad \{\,\tilde{q}^i,\,\tilde{q}^j\,\}_{p,q} = 0\,, \quad \{\,\tilde{p}_i,\,\tilde{p}_j\,\}_{p,q} = 0\,, \tag{3.46}$$

où $\{\,a,\,b\,\}_{p,q} \overset{\text{déf.}}{=} \sum_i (\partial_{q^i} a \partial_{p_i} b - \partial_{p_i} a \partial_{q^i} b)$. On montre que de telles transformations s'expriment en termes d'une *fonction génératrice* $F(q^i, \tilde{p}_i)$ telle que

$$dF = \sum_i p_i dq^i + \sum_i \tilde{q}^i d\tilde{p}_i \implies p_i = \frac{\partial F}{\partial q^i}\,, \quad \tilde{q}^i = \frac{\partial F}{\partial \tilde{p}_i}\,. \tag{3.47}$$

On peut alors vérifier que, les coordonnées $\{\tilde{q}^i, \tilde{p}_i\}$ conservent les équation de Hamilton et ainsi que cette transformation est bien canonique. [3] On note que la fonction génératrice particulière $F_e(q^i, \tilde{p}_i) = \sum_i q_i \tilde{p}^i$ donne la transformation identité car $p_i = \partial_{q^i} F = \tilde{p}_i$ et $\tilde{q}^i = \partial_{\tilde{p}_i} F = q^i$.

Considérons un développement autour de la transformation identité de la fonction génératrice, correspondant à

$$F = \sum_j q_j \tilde{p}^j + \varepsilon f(q_i, \tilde{p}^i) + \mathcal{O}(\varepsilon^2)\,. \tag{3.48}$$

On a alors

$$\tilde{q}^i = \frac{\partial F}{\partial \tilde{p}_i} = q^i + \varepsilon \frac{\partial f}{\partial \tilde{p}_i} \tag{3.49a}$$

$$p_i = \frac{\partial F}{\partial q^i} = \tilde{p}_i + \varepsilon \frac{\partial f}{\partial q^i} \tag{3.49b}$$

3. Si on exige que les hamiltoniens $H(p, q)$ et $\tilde{H}(\tilde{p}, \tilde{q})$ soient associés au même principe de moindre action, on doit avoir

$$\delta S = \delta \int \big(p \cdot \dot{q} - H(p, q)\big) = \delta \int \big(\tilde{p} \cdot \dot{\tilde{q}} - \tilde{H}(\tilde{p}, \tilde{q})\big) \implies p \cdot \dot{q} - H(p, q) = \tilde{p} \cdot \dot{\tilde{q}} - \tilde{H}(\tilde{p}, \tilde{q}) + \frac{d\Lambda}{dt}$$

Choisissons Λ de la forme $\Lambda = F(q, \tilde{p}) - \tilde{q} \cdot \tilde{p}$. On a alors

$$\frac{d\Lambda}{dt} = \frac{\partial F}{\partial q}\dot{q} + \frac{\partial F}{\partial \tilde{p}}\dot{\tilde{p}} - \tilde{q} \cdot \dot{\tilde{p}} - \dot{\tilde{q}} \cdot \tilde{p}$$

et ainsi

$$\Big(p - \frac{\partial F}{\partial q}\Big) \cdot \dot{q} - H(p, q) = \Big(\frac{\partial F}{\partial \tilde{p}} - \tilde{q}\Big) \cdot \dot{\tilde{p}} - \tilde{H}(\tilde{p}, \tilde{q})$$

donnant le résultat annoncé.

Au premier ordre en ε, on peut assimiler f à une fonction de (q_i, p^i) et écrire la variation infinitésimale des coordonnées comme

$$\delta q^i \overset{\text{déf.}}{=} \tilde{q}^i - q^i = \varepsilon \frac{\partial f}{\partial p_i} \tag{3.50a}$$

$$\delta p_i \overset{\text{déf.}}{=} \tilde{p}_i - p_i = -\varepsilon \frac{\partial f}{\partial q^i} \tag{3.50b}$$

Ces relations peuvent s'exprimer en termes des crochets de Poisson comme

$$\delta q^i = \varepsilon \left\{ q^i, f \right\} \tag{3.51a}$$

$$\delta p_i = \varepsilon \{ p_i, f \} \tag{3.51b}$$

Considérons maintenant la variation du hamiltonien H par cette transformation. Nous avons

$$\begin{aligned} \delta H &= \sum_i \left(\frac{\partial H}{\partial q^i} \delta q^i + \frac{\partial H}{\partial p^i} \delta p^i \right) = \varepsilon \sum_i \left(\frac{\partial H}{\partial q^i} \frac{\partial f}{\partial p_i} - \frac{\partial H}{\partial p^i} \frac{\partial f}{\partial q^i} \right) \\ &= \varepsilon \{ H, f \} . \end{aligned} \tag{3.52}$$

Soit une transformation canonique infinitésimale f ayant un crochet de Poisson nul avec le hamiltonien : $\{ H, f \} = 0$. Comme f ne dépend pas explicitement du temps par hypothèse, f est une quantité conservée ou charge de Noether. Nous pouvons réinterpréter alors les éqns. (3.51) de la manière suivante : *les quantités conservées associées à des symétries sont les générateurs infinitésimaux des transformations canoniques correspondantes.*

Remarquons pour finir que si on choisit $f = H$ et on identifie ε à un intervalle de temps infinitésimal δt, les éqns. (3.51) deviennent

$$\delta q^i = \delta t \left\{ q^i, H \right\} , \quad \delta p^i = \delta t \left\{ p^i, H \right\} , \tag{3.53}$$

qui ne sont rien d'autre que les équations de Hamilton sous la forme (3.41). On peut ainsi considérer le hamiltonien comme le générateur de l'évolution dans le temps du système.

3.3 Quelques éléments de théorie des groupes

Mathématiquement, le opérations de symétrie sur des système physiques forment des groupes, et ainsi l'étude des propriétés des symétries en mécanique quantique s'exprime naturellement dans le langage de la théorie des groupes. Nous commencerons par donner quelques définitions générales.

Définition 1 (groupe). *Un groupe* G *est un ensemble muni d'une loi de composition interne* $G \times G \to G$ *satisfaisant aux propriétés suivantes :*

- *élément neutre* : $$\exists e \in G , \ \forall g \in G , e \cdot g = g \cdot e = g , \tag{3.54a}$$
- *inverse* : $$\forall g \in G , \ \exists g^{-1} \in G , \ g^{-1} \cdot g = g \cdot g^{-1} = e , \tag{3.54b}$$
- *associativité* : $$\forall g_1, g_2, g_3 \in G , \ (g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3) . \tag{3.54c}$$

L'élément $e \in G$ est appelé *élément neutre* du groupe. Un groupe mathématique est une entité abstraite, qui peut être réduite à une table de multiplication donnant la « règle du jeu » pour multiplier deux éléments entre eux pour donner un élément du groupe. On utilisera parfois la notation $(G, \cdot)$ pour le groupe, permettant de préciser de quelle loi de composition interne il s'agit.

Définition 2 (groupe abélien ou commutatif). *Un groupe G est dit abélien ou commutatif s'il satisfait*

$$\forall g_1, g_2 \in G\ , \quad g_1 \cdot g_2 = g_2 \cdot g_1\ . \tag{3.55}$$

Un groupe qui ne vérifie pas cette propriété pour tous ses éléments est dit non-abélien.

Pour un groupe abélien l'ordre dans lequel agissent les éléments du groupe n'a pas d'importance. C'est le cas par exemple pour les translations ou les rotations dans le plan. Au contraire, pour des rotations dans l'espace d'axes différents, l'ordre des rotations importe ; le groupe correspondant est non-abélien. Nous ferons plus loin le lien entre la non-commutativité des groupes et la non-commutativité en mécanique quantique.

Une grande partie des symétries des systèmes quantiques proviennent de transformations de l'espace, telles les rotations et les translations. De ce point de vue les propriété de groupe sont assez naturelles ; par exemple une transformation qui amène un point de l'espace sur un autre point sera toujours réversible, donc il existera une transformation inverse. À proprement parler une transformation de l'espace n'est pas un groupe, mais correspond à une *action de groupe*[4] du groupe abstrait correspondant :

Définition 3 (action de groupe). *Soit G un groupe et E un ensemble. Une action de G sur E est une application $\mathfrak{a}_g : G \times E \to E$ qui satisfait aux propriétés suivantes :*

$$\forall g, h \in G\ ,\ \forall x \in E\ ,\ \mathfrak{a}_g\big(\mathfrak{a}_h(x)\big) = \mathfrak{a}_{g\cdot h}(x)\ , \tag{3.56a}$$

$$\forall g\ ,\ \forall x \in E\ ,\ \mathfrak{a}_e(x) = x\ . \tag{3.56b}$$

Ces propriétés assurent la compatibilité entre l'action sur E et la structure de groupe de G. Si E est un ensemble fini, on peut associer pour tout $g \in E$ l'application $\mathfrak{a}_g$ à une certaine permutation des éléments de E.

Pour revenir à l'exemple ci-dessus, les translations de points de l'espace doit être vues en tant qu'action du groupe des translations sur l'ensemble $\mathbb{R}^3$.

Il arrive fréquemment de considérer des applications d'un groupe dans un autre. De même qu'une application linéaire d'un espace vectoriel dans un autre préserve la structure d'espace vectoriel, on peut demander à ce qu'une telle application préserve la structure de groupe.

Définition 4 (homomorphisme de groupe). *Soit $(G, \cdot)$ et $(\tilde{G}, \circ)$ deux groupes et $\rho : G \to G'$ une application. Cette application est un homomorphisme de groupe si*

$$\forall g_1, g_2 \in G\ , \quad \rho(g_1 \cdot g_2) = \rho(g_1) \circ \rho(g_2)\ . \tag{3.57}$$

Cette définition implique en particulier que $\rho(e_G) = e_{\tilde{G}}$ et que $\rho(g^{-1}) = [\rho(g)]^{-1}$ (*exercice* : le montrer).

4. Naturellement cela n'a rien à voir avec une action de groupe dans le langage judiciaire !

3.3.1 Groupes discrets et groupes continus

Un *groupe discret* est tel que l'ensemble contenant tous les éléments du groupe est dénombrable. Un exemple simple est fourni par $\mathbb{Z}$, l'ensemble des entiers relatifs, muni de la loi de composition interne donnée par l'addition ; il est évident que toutes les propriétés de l'éqn. (3.56) sont satifsaites et que le groupe est commutatif.

Un exemple de système physique admettant une action du groupe $\mathbb{Z}$ est donné par une chaîne d'atomes périodique à une dimension, voir la figure 3.1 où a désigne le pas du réseau. L'action de l'élément $k \in \mathbb{Z}$ sur la chaîne correspond à « sauter » de l'atome n à l'atome $n + k$ de la chaîne.

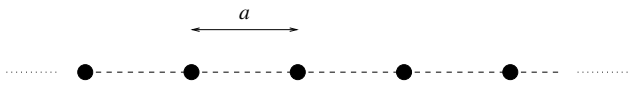

FIGURE 3.1 – Chaîne d'atomes à une dimension.

Parmi les groupes discrets certains groupes ne contiennent qu'un nombre fini d'éléments ; ce sont des *groupes finis* ; le nombre d'éléments d'un groupe G fini, appelé *ordre du groupe*, est noté $\#G$. Un groupe fini est entièrement spécifié par une table $\#G \times \#G$ contenant tous les produits d'éléments de G.

Un classe de groupes finis et abéliens est donnée par les groupes de type $\mathbb{Z}_n$, c.-à.-d. des entiers relatifs identifés modulo n.[5] On peut décrire $\mathbb{Z}_n$ complètement en prenant les entiers $\{0, 1, \ldots, n-1\}$ et en calculant leurs sommes deux à deux modulo n. Par exemple, le groupe $\mathbb{Z}_2$ contient les entiers $\{0, 1\}$ avec les règles de calcul $0 + 0 \equiv 0 \mod 2$, $0 + 1 \equiv 1 \mod 2$ et $1 + 1 \equiv 0 \mod 2$. Dans la notation générale des groupes cela donne la table

$\mathbb{Z}_2$

$\cdot$	e	g
e	$e \cdot e = e$	$e \cdot g = g$
g	$g \cdot e = g$	$g \cdot g = e$

(3.58)

Les groupes dont les éléments ne sont pas dénombrables sont des *groupes continus*. Pour être plus précis, les éléments d'un tel groupe sont décrits localement, c.-à.-d. au voisinage d'un certain élément $g_0 \in G$, par un ensemble de paramètres réels. Par exemple, les matrices des rotations à deux dimensions s'expriment comme

$$g(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad \theta \in [0, 2\pi[\tag{3.59}$$

À trois dimensions, comme nous le verrons au cours n° 4, les éléments du groupe des rotations à trois dimensions se notent $g(\theta, \phi, \psi)$ et dépendent ainsi de trois paramètres continus. Pour un groupe continu quelconque les paramètres décrivant localement le groupe peuvent être associés à un système de coordonnées sur un certain espace, appelé *variété de groupe*.

Reprenons l'exemple (3.59). Il est évident que θ peut être associée à la coordonnée sur un cercle. Le produit des rotations $g(\theta) \cdot g(\theta_0) = g(\theta_0 + \theta)$ est interprété comme une *action du*

5. Les mathématiciens préfèrent utiliser la notation $\mathbb{Z}/n\mathbb{Z}$.

groupe sur lui même qui, dans cet exemple, translate le point de coordonnée θ_0 au point de coordonnée $\theta_0 + \theta$. De même l'inversion $g(\theta) \mapsto g(\theta)^{-1} = g(-\theta)$ correspond à $\theta \mapsto -\theta$ sur le cercle.

Un groupes continu G ou plus précisément un *groupe de Lie*, est défini comme un espace admettant une action de G sur lui-même, ainsi qu'une opération d'inversion. Cet espace doit avoir les caractéristiques d'une variété différentiable. [6]

3.3.2 Sous-groupes

Il est fréquent de considérer uniquement une partie d'un groupe de symétrie. Par exemple, considérons un atome d'hydrogène en mécanique quantique ; ce système est invariant par rotation dans l'espace donc les solutions du problème doivent respecter cette symétrie. Si on plonge cet atome dans un champ magnétique, le système sera uniquement invariant par les rotations dans le plan orthogonal à $\vec{B}$ et seule cette symétrie résiduelle sera pertinente pour la solution du problème. Pour les mêmes raisons que les symétries d'origine forment un groupe, les symétries résiduelles forment également un groupe, inclus dans le premier.

Définition 5 (sous-groupe). *Soit* $(G, \cdot)$ *un groupe et* $H \subset G$ *une partie non-vide de* G. H *est un sous-groupe de* G *si* H *est stable par rapport au produit* $\cdot$ *et par rapport à l'inversion. En d'autres termes,*

$$\forall h_1, h_2 \in H\,,\ h_1 \cdot h_2^{-1} \in H\,. \tag{3.60}$$

En particulier, un sous-groupe contient toujours l'élément neutre. Les sous-groupes discrets contiennent naturellement uniquement des sous-groupes discrets, alors que les groupes continus contiennent à la fois des sous groupes discrets et continus en général.

Exemple 1 (groupe $\mathbb{Z}_4$). Le groupe $\mathbb{Z}_4$ est composé des éléments $\{0, 1, 2, 3\}$ et la loi de composition interne est l'addition modulo 4. Il est aisé de réaliser que $H = \{0, 2\}$ est un sous-groupe de $\mathbb{Z}_4$. Étant donné que $2 + 2 \equiv 0 \mod 4$, on remarque que H est isomorphe au groupe $\mathbb{Z}_2$ présenté dans la table (3.58). De manière générale, $\mathbb{Z}_n$ admet un sous-groupe isomorphe à $\mathbb{Z}_p$ si p divise n

Les groupes de Lie admettent une classe particulièrement importante de sous-groupes continus, les sous-groupes à un paramètre.

Définition 6 (sous-groupe à un paramètre). *Soit* $(G, \cdot)$ *un groupe de Lie. Un sous-groupe à un paramètre est une application continue*

$$\gamma : \begin{cases} \mathbb{R} & \to & G \\ t & \mapsto & \gamma(t) \end{cases} \tag{3.61}$$

qui est également un homomorphisme de groupe $(\mathbb{R}, +) \to (G, \cdot)$*, c.-à.-d.*

$$\forall t_1, t_2 \in \mathbb{R}\,, \quad \gamma(t_1 + t_2) = \gamma(t_1) \cdot \gamma(t_2)\,. \tag{3.62}$$

6. Une variété différentiable et un espace localement équivalent à $\mathbb{R}^n$ tel que les applications permettant de « recoller » les systèmes de coordonnées sur les différentes parties de l'espace sont différentiables.

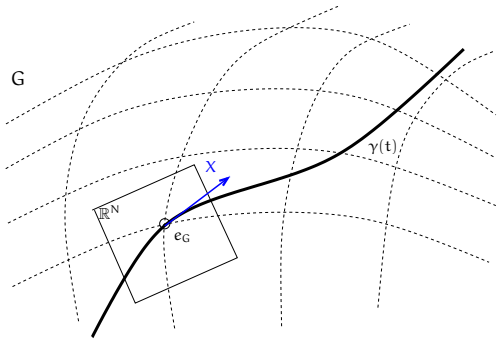

FIGURE 3.2 – Sous-groupe à un paramètre et espace tangent à l'élément neutre

Il est évident que ce sous-groupe (exercice : vérifier qu'il s'agit bien d'un sous-groupe !) est abélien car $(\mathbb{R}, +)$ l'est.

Si on se rappelle qu'un groupe de Lie peut être vu comme une variété différentiable, un sous-groupe à un paramètre est une courbe tracée sur cet espace, voir la fig. 3.2. En outre, étant associé à un homéomorphisme de groupe, nous avons tonjours $\gamma(0) = e$, où e est l'élément neutre de G. Dans la limite $t \to 0$, on se restreint à un voisinage de l'identité où on peut assimiler le groupe à son espace tangent à l'identité, T_eG, isomorphe à $\mathbb{R}^N$ pour un certain N. À tout sous-groupe à un paramètre γ on peut donc associer un élément X de son espace tangent, tel que

$$X = \frac{d}{dt}\gamma(t)\Big|_{t=0} . \tag{3.63}$$

Les vecteurs $X \in T_eG$ de l'espace tangent à l'identité correspondent à l'*algèbre de Lie* $\mathfrak{g}$ du groupe G. Ses éléments représentent les générateurs d'une transformation de groupe infinitésimale. Inversement

Théorème 2 *Soit G un groupe de Lie et $\mathfrak{g}$ l'algèbre de Lie associée. Pour tout $X \in \mathfrak{g}$ il existe un unique sous-groupe à un paramètre γ_X tel que $X = \frac{d}{dt}\gamma(t)\big|_{t=0}$. On définit alors l'application exponentielle par*

$$\exp X \overset{\text{déf.}}{=} \gamma_X(1) . \tag{3.64}$$

Ce théorème permet d'obtenir une transformation finie à partir d'une transformation infinitésimale.

Lorsque le groupe G admet une représentation matricielle, c.-à.-d. en termes de matrices $n \times n$ complexes pour un certain n, on peut donner une forme particulièrement explicite à cette procédure. En effet les sous-groupes à un paramètre sont tous de la forme d'une exponentielle matricielle :

$$\gamma(t) = e^{tM} , \quad M \in \mathrm{Mat}(n, n; \mathbb{C}) , \quad e^A \overset{\text{déf.}}{=} \sum_{k=0}^{\infty} \frac{1}{k!} A^k . \tag{3.65}$$

Le développement, pour $t = \epsilon$, $|\epsilon| \ll 1$,

$$e^{\epsilon M} = \mathbb{I}_n + \epsilon M + \mathcal{O}(\epsilon^2) \tag{3.66}$$

montre que M correspond effectivement au générateur d'une transformation infinitésimale. Pour le passage à une transformation finie de paramètre t on effectue un grand nombre de transformations infinitésimales successives de paramètre $\epsilon = t/n$ et on utilise

$$\lim_{n\to\infty}\left(1+\frac{tM}{n}\right)^N = e^{tM}. \tag{3.67}$$

Exemple 2 (générateur des rotations dans le plan.) Les éléments du groupe des rotations dans le plan correspondent aux matrices (3.59). Les éléments $g(\theta)$ ne dépendant que de θ il existe un unique sous-groupe à un paramètre, confondu avec le groupe lui même. Nous avons

$$g(\theta) = \mathbb{I}_2 + \theta\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix} + \mathcal{O}(\theta^2) \implies g(\theta) = \exp\left\{\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\theta\right\} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}. \tag{3.68}$$

En effet, nous vérifions que

$$\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}^{2k} = (-1)^k\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \quad \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}^{2k+1} = (-1)^k\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}. \tag{3.69}$$

3.3.3 Représentations linéaires d'un groupe ⋆

En mécanique quantique, l'espace des états possède la structure d'un espace de Hilbert $\mathcal{H}$ qui se réduit, lorsqu'il est de dimension finie, à un espace hermitien. L'espace de Hilbert d'un système physique admet naturellement une action du groupe de symétrie G du problème – nous préciserons plus bas ce que nous entendons par là. Il est naturel, pour préserver la structure de cet espace, de considérer que cette action est associée à des applications linéaires sur $\mathcal{H}$. L'étude de ce type d'action de groupe porte le nom de *théorie des représentations*. Il constitue l'aspect le plus important de la théorie des groupes pour la mécanique quantique.

Nous nous restreindrons dans ce cours à l'analyse des représentations de dimension finie, associées à l'action d'un groupe sur un espace hermitien $\mathcal{H}$. Se rappelant que tout espace hermitien de dimension n est isomorphe à $\mathbb{C}^n$, on introduit le groupe $GL(n;\mathbb{C})$ des matrices inversibles complexes de dimensions $n\times n$:

$$GL(n;\mathbb{C}) \overset{\text{déf.}}{=} \left\{M \in \mathrm{Mat}(n,n;\mathbb{C}),\ \det M \neq 0\right\}, \tag{3.70}$$

où $\mathrm{Mat}(n,n;\mathbb{C})$ désigne l'ensemble des matrices $n\times n$. L'action de ces matrices sur les vecteurs de $\mathcal{H}$ est naturellement une action linéraire.

Définition 7 (représentation linéaire d'un groupe). *Soit G un groupe et $\mathcal{H}$ un espace hermitien de dimension n. Une représentation linéaire complexe de G est un homomorphisme de groupe :*

$$D : \begin{cases} G & \to & GL(n;\mathbb{C}) \\ g & \mapsto & D(g)\end{cases} \tag{3.71}$$

Rappelons que la propriété d'homomorphisme, voir la définition 4, signifie qu'une représentation d'un groupe respecte la structure de groupe :

$$\forall g_1,\, g_2 \in G\ , \quad D(g_1 \cdot g_2) = D(g_1)D(g_2)\,. \tag{3.72}$$

En outre, l'espace d'arrivée étant le groupe linéaire $GL(n;\mathbb{C})$, l'action sur les vecteurs de $\mathcal{H}$ est linéaire : [7]

$$\forall g \in G\ , \quad \forall |\alpha\rangle\,, |\beta\rangle \in \mathcal{H}\ , \quad \forall \lambda, \mu \in \mathbb{C}\ , \quad D(g)\,(\lambda\,|\alpha\rangle + \mu\,|\beta\rangle) = \lambda D(g)\,|\alpha\rangle + \mu D(g)\,|\beta\rangle\ . \tag{3.73}$$

L'exemple le plus simple de représentation d'un groupe, qui existe toujours est une *représentation triviale* qui associe à tout $g \in G$ la matrice identité de taille $n \times n$.

Pour un groupe G donné, un problème important est la classification de ses différentes représentations. Dans le cadre de la mécanique quantique, différentes représentations correspondent à plusieurs manières possibles de se transformer sous un groupe de symétrie ; nous verrons par la suite que des particules de spin différents correspondent à différentes représentations du groupe des rotations à trois dimensions.

Il n'est naturellement pas nécessaire de distinguer plusieurs représentations qui se distinguent uniquement par un choix de base sur l'espace des états. Deux représentations D et D' de dimension n sont dites *équivalentes* s'il existe une matrice T inversible telle que

$$\forall g \in G\ , \quad D'(g) = T^{-1}D(g)T\,. \tag{3.74}$$

Remarquons la similitude avec la notion de matrices semblables en algèbre linéaire. La différence ici est que la *même* matrice T doit donner le changement de base pour *toutes* les matrices $D(g)$, quel que soit g.

En algèbre linéaire, l'action d'un opérateur linéaire se simplifie si celui-ci est diagonalisable ; il suffit alors de trouver une base de vecteurs propres. Dans le contexte de la théorie des représentations des groupes, le problème est plus complexe puisqu'il faudrait diagonaliser *simultanément* toutes les matrices $\{D(g)\,,\ g \in G\}$. Cela n'est – sauf exception – pas possible mais on peut envisager au moins de diagonaliser les matrices $D(g)$ par blocs, afin d'étudier séparément l'action du groupe sur des sous-espaces de $\mathcal{H}$.

Définition 8 (représentations reductibles et irréductibles). *Soit* D *une représentation d'un groupe* G *de dimension* n*. Cette représentation est réductible si toutes les matrices* $D(g)$ *sont de la forme :*

$$\forall g \in G\ , \quad D(g) = \begin{pmatrix} d_1(g) & b(g) \\ 0 & d_2(g) \end{pmatrix}\ , \tag{3.75}$$

où $d_1(g)$*,* $d_2(g)$ *et* $b(g)$ *sont respectivement des matrices* $k\times k$*,* $(n-k)\times(n-k)$ *et* $k\times(n-k)$ *pour un certain* $k \leqslant n$*, ou bien peuvent être toutes mises sous cette forme par un même*

7. Naturellement, on peut définir des représentations linéaires réelles qui sont des homomorphismes $G \to GL(n;\mathbb{R})$.

changement de base (cf. représentations équivalentes). Cette propriété implique qu'il existe un sous-espace de dimension k *stable par l'action du groupe :*

$$\forall g \in G\ , \quad D(g) \begin{pmatrix} v_1 \\ \vdots \\ v_n \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} d_1(g)\vec{v} \\ \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{3.76}$$

Si en outre $b(g) = 0$ *pour tout* $g \in G$ *la représentation est dite complètement réductible. Une représentation qui n'est pas réductible est appelée représentation irreductible, et ne possède donc pas de sous-espace invariant par l'action du groupe.*

On voit que le problème de la reductibilité des représentations des groupes est un problème de diagonalisation par blocs simultanée d'un ensemble de matrices, associées aux différents éléments du groupe. Comme pour la diagonalisation usuelle des matrices, la réductibilité peut être possible sur le corps des complexes mais pas sur le corps des réels.

Exemple 3 (représentation vectorielle du groupe des rotations 2d). Reprenons l'exemple 2, matrices de rotation dans le plan (3.59) correspondant à une représentation irreductible réelle de dimension 2 du groupe $SO(2)$. En revanche dans le corps des complexes nous pouvons écrire

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 \\ \pm i \end{pmatrix} = e^{\mp i\theta} \begin{pmatrix} 1 \\ \pm i \end{pmatrix}\ , \tag{3.77}$$

et on peut donc réduire complètement la représentation en deux représentations irreductibles complexes de dimension un :

$$\forall \theta\ ,\ g(\theta) = V \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} V^{-1}\ , \quad V = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}\ . \tag{3.78}$$

Exemple 4 (représentations de $\mathbb{Z}$). Le groupe $(\mathbb{Z}, +)$, c.-à.-d. le groupe des entiers muni de l'addition, est un groupe discret infini. Une représentation de $(\mathbb{Z}, +)$ de dimension d est un ensemble de matrices $d \times d$ inversibles $D(n) \in GL(d, \mathbb{C})$, associées à chaque $n \in \mathbb{Z}$, telles que

$$D(0) = \mathbb{I}_d\ , \quad D(\underbrace{1 + 1 + \cdots + 1}_{=n}) \stackrel{!}{=} D(1)^n\ . \tag{3.79}$$

Ainsi, une représentation donnée est entièrement caractérisée par la donnée d'une matrice $d \times d$ inversible $D(1)$. Si cette matrice est diagonalisable, alors la représentation est décomposable en somme directe de d représentations irreductibles de dimension un, [8] étant donné qu'il existe alors une matrice V telle que $D(1) = V^{-1}\mathrm{diag}\,(\lambda_1, \ldots, \lambda_d) \cdot V$ et donc, pour tout n,

$$D(n) = D(1)^n = V^{-1}\mathrm{diag}\,(\lambda_1, \ldots, \lambda_d) \cdot V \cdot V^{-1}\mathrm{diag}\,(\lambda_1, \ldots, \lambda_d) \cdot V \cdots = V^{-1}\mathrm{diag}\,(\lambda_1^n, \ldots, \lambda_d^n) \cdot V\ . \tag{3.80}$$

8. Rappelons que deux représentations D et D' de même dimension sont *équivalentes* s'il existe une matrice de passage V, associée à un changement de base de $\mathbb{R}^d$, telle que $\forall g \in G$, $D(g) = V^{-1} \cdot D'(g) \cdot V$.

Si la matrice $D(1)$ n'est pas diagonalisable, on peut toujours, par un changement de base, obtenir une matrice triangulaire supérieure – la forme canonique de Jordan, voir le cours d'algèbre linéaire de L2 – de telle sorte que la représentation soit effectivement réductible en d représentations de dimension un. En revanche la représentation n'est pas complètement réductible. Par exemple, une représentation de $(\mathbb{Z}, +)$ de dimension 2 est donnée par les matrices

$$D(n) = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n , \tag{3.81}$$

donc aucune, sauf évidemment $D(0) = \mathbb{I}_2$, n'est diagonalisable. Cette représentation est réductible mais ne peut se décomposer en somme directe de représentations (une telle décomposition correspondrait à la décomposition en sous-espaces propres de $D(1)$ si cette matrice était diagonalisable).

Représentations unitaires

Une classe de représentations particulièrement importante pour la mécanique quantique est donnée par les représentations qui préservent le produit hermitien sur $\mathcal{H}$; nous verrons que cette condition est nécessaire pour préserver les probabilités par transformation de symétrie.

Définition 9 (représentation unitaire d'un groupe). *Soit* G *un groupe et* $\mathcal{H}$ *un espace hermitien de dimension* n*. Une représentation linéaire* D *de* G *agissant sur* $\mathcal{H}$ *est dite unitaire si toutes les matrices* $D(g)$ *sont unitaires :*

$$\forall g \in G , \quad D(g)^\dagger D(g) = \mathbb{I}_n . \tag{3.82}$$

Si D est une représentation unitaire d'un groupe G on a alors :

$$\forall |\alpha\rangle , |\beta\rangle \in \mathcal{H} , \quad \forall g \in G , \quad \iota (D(g)|\alpha\rangle , D(g)|\beta\rangle) = \langle\alpha| D(g)^\dagger D(g) |\beta\rangle = \langle\alpha|\beta\rangle , \tag{3.83}$$

et l'action du groupe préserve le produit hermitien.

Exemple 5 (représentations unitaires de $\mathbb{Z}$). Nous avons constaté qu'une représentation de $\mathbb{Z}$ de dimension d est entièrement caractérisée par une matrice inversible $D(1) \in GL(d, \mathbb{C})$. Pour qu'une telle représentation soit unitaire, il faut et il suffit que la matrice $D(1)$ le soit aussi, c.-à.-d. que $D(1)^\dagger D(1) = \mathbb{I}_d$.

Toute matrice unitaire étant diagonalisable, on en déduit que la représentation est complètement réductible en d représentations de dimension un, car il existe une matrice unitaire V telle que $D(1) = V^{-1} \cdot \mathrm{diag}\,(\lambda_1, \ldots, \lambda_d) \cdot V$. La condition d'unitarité implique également que les valeurs propres sont toutes de module unité ; on peut donc écrire $\lambda_k = \exp i\theta_k$, avec $\theta_k \in [0, 2\pi[$ pour tout $k \in \{1, \ldots, d\}$, qui se résume de manière compacte comme :

$$D = \bigoplus_{k=1}^{d} D^{\theta_k} , \tag{3.84}$$

où D désigne la représentation unitaire de dimension d de $(\mathbb{Z}, +)$ considérée et D^{θ_k} la représentation complexe irreductible unitaire de dimension un de ce groupe associée au paramètre $\theta_k \in [0, 2\pi[$.

Les représentations unitaires jouent un rôle central dans la classification des représentations des groupes. Nous avons en particulier le très important théorème (dont il ne s'agit en réalité d'une version très partielle) :

Théorème 3 (théorème de Peter–Weyl). *Soit* G *un groupe compact, c.-à.-d. correspondant à un espace compact, et* D *une représentation unitaire de* G*. Alors* D *se décompose en somme directe de représentations irreductibles unitaires de dimension finie.*

Dans le cas des groupes finis, on démontre le théorème suivant :

Théorème 4 (théorème de Maschke). *Toute représentation d'un groupe fini se décompose en somme directe de représentations irreductibles.*

Contrairement au cas précédent la condition d'unitarité n'est pas nécessaire, car toute représentation d'un groupe fini est équivalente à une représentation unitaire.

Pour finir, le lemme suivant est capital pour l'étude des symétries en mécanique quantique :

Lemme 1 (lemme de Schur). *Soit* D *et* D' *deux représentations linéaires irreductibles complexes d'un groupe* G*. S'il existe un opérateur linéaire* V *tel que*

$$\forall \mathsf{g} \in \mathsf{G}\,,\ \mathsf{V}\mathsf{D}(\mathsf{g}) = \mathsf{D}'(\mathsf{g})\mathsf{V} \tag{3.85}$$

alors soit $\mathsf{V} = 0$*, soit* V *est inversible et les représentations* D *et* D' *sont alors équivalentes,*

Corollaire 1 *Soit* D *une représentation irreductible complexe d'un groupe* G *sur un espace hermitien* $\mathcal{H}$ *de dimension* n *et* V *un opérateur linéaire* $\mathcal{H} \to \mathcal{H}$*.*

$$\boxed{\forall \mathsf{g} \in \mathsf{G}\,,\ [\,\mathsf{V},\, \mathsf{D}(\mathsf{g})\,] = 0 \implies \exists \lambda \in \mathbb{C}^\star\,,\ \mathsf{V} = \lambda \mathbb{I}_n} \tag{3.86}$$

Ainsi, un opérateur commutant avec toutes les matrices d'une représentation irreductible est proportionnel à l'identité.

La demonstration de ce corollaire est aisée. Le lemme de Schur indique que V est inversible, donc admet au moins une valeur propre non-nulle $\lambda \in \mathbb{C}^\star$. On forme alors l'opérateur $\mathsf{V} - \lambda \mathbb{I}_n$ qui vérifie $[\,\mathsf{V} - \lambda \mathbb{I}_n,\, \mathsf{D}(\mathsf{g})\,] = 0$ pour tout $\mathsf{g} \in \mathsf{G}$. Ayant une valeur propre nulle, il n'est pas inversible donc $\mathsf{V} - \lambda \mathbb{I}_n = 0$ d'après le lemme de Schur. □

Corollaire 2 *Toute représentation irreductible complexe d'un groupe abélien (commutatif) est de dimension un.*

Nous laissons la preuve de cette propriété en exercice.

Exemple 6 (représentations irreductibles du groupe des rotations 2d). *Un exemple important est celui du groupe des rotations à deux dimensions,* $\{g(\theta), \theta \in [0, 2\pi[\}$*, voir l'éqn. (3.59). Nous avons vu que la variété de groupe correspondante est assimilée à un cercle, et les différentes représentations irreductibles unitaires correspondent au nombre de tours effectués dans le plan lorsque ce cercle est parcouru une fois. Les représentations irreductibles complexes sont donc classifiées par un entier,*[9] *appelé parfois nombre d'enroulement, et données par :*

$$\boxed{D^n(\theta) = e^{in\theta} \quad n \in \mathbb{Z}} \tag{3.87}$$

La décomposition en série de Fourier d'une fonction sur un cercle, $f(\theta) = \sum_n e^{in\theta} f_n$, peut ainsi être considérée comme une décomposition en représentations irreductibles.

3.4 Symétries en mécanique quantique

Avant d'aborder au cours n° 4 l'étude explicite des symétries usuelles (translations, rotations,...) nous présentons dans cette dernière section quelques résultats très généraux qui illustrent l'importance du concept de représentations de groupe en mécanique quantique.

3.4.1 Théorème de Wigner

L'action d'une symétrie sur un système physique donne, par définition, le même système physique ou du moins un système ayant exactement les mêmes propriétés observables. Une symétrie d'un système doit donc posséder la propriété essentielle de ne pas changer les prédictions physiques.

Dans le cadre du formalisme de la mécanique quantique exposé au cours n° 1, cela se traduit par la propriété suivante : *l'action d'une symétrie sur un système quantique doit préserver les probabilités associées aux mesures des observables du système.*

Postulons dans un premier temps que l'action d'une symétrie g (appartenant à un groupe de symétrie G) sur l'espace des états $\mathcal{H}$ corresponde à un opérateur linéaire $\mathcal{G}$, afin de préserver la structure d'espace de Hilbert ; l'action de la symétrie est alors identifiée à une représentation linéaire du groupe G contenant g, voir la sous-section 3.3.3 :

$$|\alpha\rangle \overset{G}{\mapsto} \mathcal{G}|\alpha\rangle \in \mathcal{H} \ , \quad \mathcal{G}(\mu|\alpha\rangle + \nu|\beta\rangle) = \mu\mathcal{G}|\alpha\rangle + \mathcal{G}\nu|\beta\rangle \ . \tag{3.88}$$

Les probabilités associées à la mesure des observables sont obtenues par le produit scalaire du vecteur d'état du système avec les vecteurs propres associés aux résultat possibles de la mesure. Pour que les propriétés associées à toutes les observables soit conservées, nous exigeons que

$$\boxed{\forall |\alpha\rangle, |\beta\rangle \in \mathcal{H} \ , \quad |\iota(\mathcal{G}|\alpha\rangle, \mathcal{G}|\beta\rangle)|^2 = |\iota(|\alpha\rangle, |\beta\rangle)|^2} \tag{3.89}$$

9. Si $n \notin \mathbb{Z}$, on obtient une discontinuité en $\theta = 0 \mod 2\pi$ incompatible avec la structure d'un groupe de Lie qui est une variété différentiable.

Une condition suffisante — mais certainement pas nécessaire – pour satisfaire l'équation (3.89) est d'imposer que le produit scalaire soit invariant :

$$\forall \mathsf{g} \in \mathsf{G} \ , \quad \forall \left|\alpha\right\rangle , \left|\beta\right\rangle \in \mathcal{H} \ , \iota \left(\mathcal{G} \left|\alpha\right\rangle , \mathcal{G} \left|\beta\right\rangle\right) = \iota \left(\mathcal{G}^\dagger \mathcal{G} \left|\alpha\right\rangle , \left|\beta\right\rangle\right) = \left\langle\alpha\right| \mathcal{G}^\dagger \mathcal{G} \left|\beta\right\rangle = \left\langle\alpha|\beta\right\rangle \ . \quad (3.90)$$

Cela signifie que l'opérateur $\mathcal{G}$ satisfait $\mathcal{G}^\dagger \mathcal{G} = \mathfrak{I}$ et est par définition un opérateur unitaire.

La solution la plus générale du problème est donnée par le théorème de Wigner, dont on pourra trouver une demonstration accessible dans [2] :

Théorème 5 (théorème de Wigner). *Soit* g *une symétrie d'un système quantique dont l'espace des états est donné par un espace de Hilbert* $\mathcal{H}$*. Cette symétrie est associée à un opérateur* $\mathcal{G} : \mathcal{H} \to \mathcal{H}$ *qui est soit unitaire, soit anti-unitaire.*

Le premier cas de figure a été discuté ci-dessus, voir l'équation (3.90). L'autre cas de figure correspond à une autre solution possible de la relation générale (3.89) de conservation de la probabilité : on exige que $\mathcal{G}$ satisfasse

$$\forall \left|\alpha\right\rangle , \left|\beta\right\rangle \in \mathcal{H} \ , \iota \left(\mathcal{G} \left|\alpha\right\rangle , \mathcal{G} \left|\beta\right\rangle\right) = \overline{\iota \left(\left|\alpha\right\rangle , \left|\beta\right\rangle\right)} = \overline{\left\langle\alpha|\beta\right\rangle} = \left\langle\beta|\alpha\right\rangle \ . \quad (3.91)$$

Un tel opérateur satisfaisant cette condition et étant également anti-linéaire, c.-à.-d.

$$\mathcal{G} \left(\lambda \left|\alpha\right\rangle + \mu \left|\beta\right\rangle\right) = \bar{\lambda} \mathcal{G} \left|\alpha\right\rangle + \bar{\mu} \mathcal{S} \left|\beta\right\rangle \ . \quad (3.92)$$

est dit *anti-unitaire*. En pratique, les toutes les symétries que nous rencontrerons dans ce cours seront associées à des opérateurs unitaires, à l'exception du renversement du temps qui est anti-unitaire (voir le cours n° 8).

Représentations unitaires et projectives $\star$

Examinons plus en détail la situation la plus courante, où l'action d'un groupe de symétrie G sur l'espace des états est associée à des operateurs unitaires.

En termes de théorie des groupes, cela signifie que la représentation du groupe de symétrie G sur l'espace des états est une représentation unitaire, voir la définition 9. D'après le théorème 3 dans le cas d'un groupe compact, ou du théorème 4 dans le cas d'un groupe discret, cette représentation est décomposable en représentations irreductibles.

Plus généralement, la relation (3.89) reste vraie si le produit scalaire est préservé à une phase près, c.-à.-d. si :

$$\forall \mathsf{g} \in \mathsf{G} \ , \quad \forall \left|\alpha\right\rangle , \left|\beta\right\rangle \in \mathcal{H} \ , \quad \iota \left(\mathcal{G} \left|\alpha\right\rangle , \mathcal{G} \left|\beta\right\rangle\right) = e^{\mathrm{i}\phi(\mathsf{g})} \left\langle\alpha|\beta\right\rangle \quad (3.93)$$

où $\phi(\mathsf{g}) \in \mathbb{R}$. Une formulation plus précise du théorème de Wigner énonce en effet qu'une symétrie est associée à un opérateur $\mathcal{G}$ agissant sur l'espace des rayons, c.-à.-d. l'espace projectif $\mathbb{P}[\mathcal{H}]$, compatible avec un opérateur $\mathcal{U}$ unitaire ou anti-unitaire, dans le sens que, pour tout $\left|\psi\right\rangle \in \mathcal{H}$, $\mathcal{U} \left|\psi\right\rangle$ appartienne à l'image du rayon de $\left|\psi\right\rangle$ par $\mathcal{G}$.

La structure d'espace projective de l'espace des états implique donc que des représentations des groupes de symétrie plus générales que les représentations linéaires, voir éqn. (3.71),

sont autorisées en mécanique quantique. On peut admettre des *représentations projectives*, qui satisfont à la relation

$$\forall g, h \in G\ , \quad D(g)D(h) = \epsilon(g,h)D(g\cdot h)\ , \quad |\epsilon(g,h)| = 1\,. \tag{3.94}$$

où le facteur de phase $\epsilon(g,h)$ est appelé *multiplicateur de Schur*.

Cette remarque n'est absolument pas anecdotique, étant donné que l'existence de représentations projectives en mécanique quantique, appliquée au groupes des rotations de l'espace, permet l'existence de spins demi-entiers !

Mentionnons pour finir que, moyennant quelques hypothèses sur la topologie du groupe G, il existe un groupe $\widehat{G}$, appelé *recouvrement universel* de G, tel que les représentations projectives de G correspondent à des représentations ordinaires de $\widehat{G}$. On peut ainsi, par exemple dans le cas des groupes des rotations, se ramener à l'étude des représentations « ordinaires » d'un groupe que nous construirons plus loin.

3.4.2 États propres du hamiltonien et représentations irreductibles

En mécanique classique hamiltonienne le générateur d'une symétrie continue (supposant qu'elle ne dépend pas explicitement du temps) est associé à une fonction $Q(q^i, p_i)$ correspondant à une quantité conservée – voir la sous-section 3.4.2 – c.-à.-d. dont le crochet de Poisson avec le hamiltonien est nul : $\{Q, H\} = 0$.

Le passage de la mécanique classique hamiltonienne à la mécanique quantique a pour point de départ, rappelons-le, le processus de quantification canonique, reposant sur le postulat suivant :

Postulat 1 *La quantification d'un système physique en mécanique hamiltonienne consiste à remplacer les fonctions* $f(q^i, p_i, t)$ *sur l'espace des phases par des opérateurs* $\mathcal{F}$ *agissant sur l'espace des états du système. Ces opérateurs obéissent à des relations de commutation*[10] *déduites de leurs crochets de Poisson via la correspondance*

$$\{\,.\ ,.\,\} \longleftrightarrow \frac{1}{i\hbar}[\,.\ ,.\,]\ , \quad [\,\mathcal{F}_1, \mathcal{F}_2\,] \overset{\text{déf.}}{=} \mathcal{F}_1\mathcal{F}_2 - \mathcal{F}_2\mathcal{F}_1\,. \tag{3.95}$$

On fait correspondre à la charge de Noether Q associée à une symétrie un opérateur $\mathcal{Q}$ qui commute avec l'opérateur hamiltonien du système :

$$\boxed{[\,\mathcal{Q}, \mathcal{H}\,] = 0} \tag{3.96}$$

Cet opérateur est compris comme le générateur de la symétrie associée.

*
* *

10. Les opérateurs de type *fermonique* obéissent pour leur part à des relations d'anti-commutation.

Nous pouvons maintenant faire une synthèse des différentes notions présentées dans ce cours, en particulier les représentations irreductibles et unitaires, les théorèmes de Peter–Weyl et Maschke, le lemme de Schur et le théorème de Wigner, pour obtenir un résultat capital concernant les symétries en mécanique quantique.

Soit G un groupe de symétrie, compact ou fini, d'un système quantique. D'après le théorème de Weyl, l'action du groupe sur l'espace des états $\mathcal{H}$ correspond à une représentation D unitaire (mettant de côté l'éventualité d'opérateurs anti-unitaires). Le théorème de Peter–Weyl, dans le cas d'un groupe continu compact, et le théorème de Maschke, dans le cas d'un groupe fini, nous indiquent que cette représentation est décomposable en somme directe de représentations unitaires irreductibles $\{\mathsf{D}_\ell, \ell = 1, 2, \ldots\}$. Cela signifie que l'espace des états se scinde en sous-espaces orthogonaux entre eux qui sont tous stables par l'action du groupe de symétrie :

$$\mathcal{H} = \bigoplus_\ell \mathcal{H}_\ell \quad \text{t.q.} \quad \forall \ell\,,\ \forall g \in \mathsf{G}\,,\ \forall\, |\alpha;\ell\rangle \in \mathcal{H}_\ell\,,\ \mathsf{D}_\ell(g)\,|\alpha;\ell\rangle \in \mathcal{H}_\ell\,. \tag{3.97}$$

Concentrons nous sur le sous-espace $\mathcal{H}_\ell$ de l'espace des états, qui se transforme sous une représentation irreductible et unitaire du groupe G. Étant donné que G est un groupe de symétrie du problème, nous avons nécessairement

$$\forall g \in \mathsf{G}\,,\ [\,\mathcal{H},\, \mathsf{D}_\ell(g)\,] = 0\,, \tag{3.98}$$

c.-à.-d. plus précisément que la restriction du hamiltonien au sous-espace $\mathcal{H}_\ell$ commute avec tous les opérateurs $\mathsf{D}_\ell(g)$ associés à la représentation D_ℓ. L'application du corollaire 1 du lemme de Schur nous indique alors que l'opérateur hamiltonien est proportionnel à l'identité dans ce sous-espace :

$$\boxed{\exists \mathsf{E}_\ell \in \mathbb{R} \quad \text{t.q.} \quad \forall\, |\alpha;\ell\rangle \in \mathcal{H}_\ell\,,\ \mathcal{H}\,|\alpha;\ell\rangle = \mathsf{E}_\ell\,|\alpha;\ell\rangle} \tag{3.99}$$

Ainsi *les vecteurs d'état appartenant à une représentation unitaire irreductible d'un groupe de symétrie sont vecteurs propres du hamiltonien et appartiennent au même sous-espace propre (ils possèdent tous la même énergie propre).*

Ce résultat fondamental expliquera à de nombreuses reprises l'existence de dégénérescences dans le spectre en énergie de systèmes physiques. Il illustre la puissance du concept de symétrie en mécanique quantique.

À retenir

— chaque symétrie continue classique correspond à une charge de Noether conservée ;

— ces charges génèrent de transformations canoniques infinitésimales dont le crochet de Poisson avec le hamiltonien est nul ;

— notions de groupe, groupe discret, groupe continu ;

— générateur infinitésimal d'un groupe continu et sous-groupe à un paramètre ;

— représentations d'un groupe, représentations irreductibles et unitaires ;

— théorème de Wigner : les symétries sont des opérateurs unitaires ou anti-unitaires ;

— si un opérateur de symmétrie commute avec le hamiltonien, tous les états d'une même représentation irreductible ont la même énergie.

Bibliographie

[1] L. Landau et E. Lifchitz, "Mécanique," éditions MIR, 1982

[2] V. Bargmann, "Note On Wigner's Theorem On Symmetry Operations," J. Math. Phys. **5** (1964) 862. doi :10.1063/1.1704188

Cours n°4

Symétries II : symétries spatiales

Dans ce cours nous examinerons plus en détail le rôle des symétries de l'espace (rotation, translation et inversion) en physique quantique. Les résultats généraux obtenus lors du cours n° 3 seront très utiles pour tirer les conséquences physiques de ces symétries.

4.1 Rotations dans l'espace

Soit $\vec{v}$ un vecteur de l'espace à trois dimensions et $\{\vec{e}_i,\, i = 1, \ldots, 3\}$ une base orthonormée de $\mathbb{R}^3$. On note $\vec{v} = \sum_i v^i \vec{e}_i$ la décomposition de $\vec{v}$ sur la base.

4.1.1 Points de vue actif et passif

Une rotation de l'espace autour de l'origine des coordonnées, ou tout autre transformation géométrique, peut se concevoir de deux manières équivalentes, appelées points de vue actif et passif.

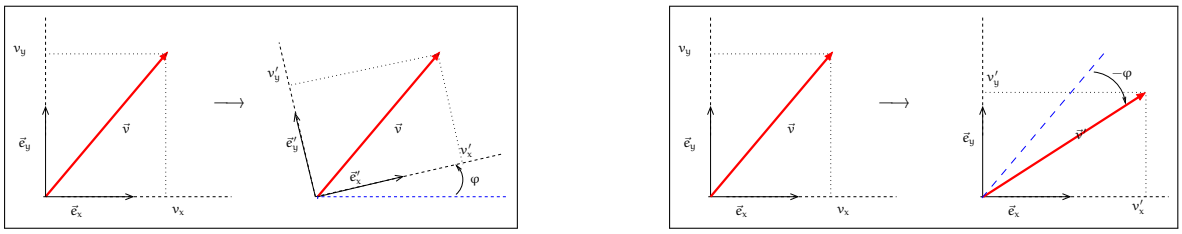

FIGURE 4.1 – Point de vue passif (gauche) et actif (droite)

Illustrons ces deux points de vue avec une rotation à deux dimensions, voir la figure 4.1 :

— *point de vue passif* : le vecteur $\vec{v}$ est inchangé mais on effectue une rotation de la base d'angle φ :

$$\vec{e}_x \mapsto \vec{e}'_x = \vec{e}_x \cos\varphi + \vec{e}_y \sin\varphi \;, \quad \vec{e}_y \mapsto \vec{e}'_y = -\vec{e}_x \sin\varphi + \vec{e}_y \cos\varphi \,. \tag{4.1}$$

Les composantes (v'_x, v'_y) dans la nouvelle base se déduisent alors de celles de l'ancienne base en exigeant l'invariance de $\vec{v}$, soit :

$$v_x\vec{e}_x + v_y\vec{e}_y = v'_x\vec{e}'_x + v'_y\vec{e}'_y \quad \Longrightarrow \quad \begin{pmatrix} v_x \\ v_y \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} v'_x \\ v'_y \end{pmatrix} \tag{4.2}$$

— *point de vue actif* : la base $\{\vec{e}_x, \vec{e}_y\}$ est inchangée mais le vecteur $\vec{v}$ est transformé en un vecteur $\vec{v}'$ par une rotation d'angle $-\varphi$:

$$\vec{v} = v_x\vec{e}_x + v_y\vec{e}_y \;\mapsto\; \vec{v}' = v'_x\vec{e}_x + v'_y\vec{e}_y \tag{4.3}$$

avec (voir le panneau de droite de la figure)

$$\begin{pmatrix} v'_x \\ v'_y \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} v_x \\ v_y \end{pmatrix} \tag{4.4}$$

Naturellement, les matrices apparaissant dans les équations (4.1) et (4.2) sont inverses l'une de l'autre. Le choix de l'un ou l'autre de ces points de vue ne doit évidemment rien changer à la physique étudiée. Dans l'exposé fait ici, le point de vue actif sera privilégié.

4.1.2 Groupe des rotations et générateurs

L'objectif de cette sous-section est de comprendre plus en détail la structure du groupe des rotations. Nous allons pour cela nous attarder plus en détail sur la *représentation vectorielle* de ce groupe, que nous noterons R, correspondant à l'action des rotations sur les vecteurs de $\mathbb{R}^3$.

Une rotation de l'espace est spécifiée premièrement par son axe de rotation, passant par l'origine, et deuxièmement par l'angle de rotation dans le plan perpendiculaire à cet axe. L'axe de rotation est caractérisé par un vecteur $\vec{n}$ de norme unité, qui est déterminé, en coordonnées sphériques, par la donnée de de deux angles $\theta \in [0, \pi]$ et $\phi \in [0, 2\pi[$:

$$\vec{n} = \begin{pmatrix} \cos\theta \\ \sin\theta\cos\phi \\ \sin\theta\sin\phi \end{pmatrix} \tag{4.5}$$

Reprenons l'analyse des générateurs de rotation entamée à la section 3.2 du cours n° 3. Une rotation infinitésimale d'angle α d'un vecteur est donnée, voir à partir de l'éqn. (3.23), par la matrice

$$R(\vec{n}, \alpha) = \mathbb{I} - \alpha\mathfrak{T} + \mathcal{O}(\alpha^2)\ , \quad \mathfrak{T}^i{}_j = \sum_k \epsilon^i{}_{jk} n^k\,. \tag{4.6}$$

On peut alors convertir la base canonique de $\mathbb{R}^3$ en base de l'espace des matrices de rotations

infinitésimales $\mathfrak{T}^{i}{}_{j}$:

$$\vec{n} = -\begin{pmatrix}1\\0\\0\end{pmatrix} \;\leftrightarrow\; j_x \overset{\text{déf.}}{=} \begin{pmatrix}0&0&0\\0&0&-1\\0&1&0\end{pmatrix} \tag{4.7a}$$

$$\vec{n} = -\begin{pmatrix}0\\1\\0\end{pmatrix} \;\leftrightarrow\; j_y \overset{\text{déf.}}{=} \begin{pmatrix}0&0&1\\0&0&0\\-1&0&0\end{pmatrix} \tag{4.7b}$$

$$\vec{n} = -\begin{pmatrix}0\\0\\1\end{pmatrix} \;\leftrightarrow\; j_z \overset{\text{déf.}}{=} \begin{pmatrix}0&-1&0\\1&0&0\\0&0&0\end{pmatrix} \tag{4.7c}$$

$$\tag{4.7d}$$

correspondant respectivement aux rotations infinitésimales autour des axes x, y et z. On peut alors écrire

$$R(\vec{n},\alpha) = \mathbb{I} + \alpha\,(n^x j_x + n^y j_y + n^z j_z) + \mathcal{O}(\alpha^2)\,, \tag{4.8}$$

soit, en introduisant le « vecteur de matrices » $\vec{j}$ dont les composantes sont les matrices j_x, j_y et j_z :

$$R(\vec{n},\alpha) = \mathbb{I} + \alpha\,\vec{n}\cdot\vec{j} + \mathcal{O}(\alpha^2)\,. \tag{4.9}$$

On remarque que, pour une direction $\vec{n}$ fixée, l'ensemble des éléments du groupe $R(\vec{n},\alpha)$ doit former un sous-groupe à un paramètre, voir la définition 6 du cours n° 3. Il est en effet aisé de réaliser que :

$$R(\vec{n},\alpha_1)R(\vec{n},\alpha_2) = R(\vec{n},\alpha_1+\alpha_2)\,, \tag{4.10}$$

s'agissant de rotations successives dans le même plan orthogonal à $\vec{n}$.

En utilisant les propriétés des sous-groupes à un paramètre, on peut alors donner la matrice d'une rotation arbitraire d'axe $\vec{n}$ et d'angle α en exponentiant le générateur infinitésimal :

$$\boxed{R(\theta,\phi,\alpha) \overset{\text{déf.}}{=} e^{\alpha\,\vec{n}\cdot\vec{j}}} \tag{4.11}$$

où la dépendance dans les angles θ et ϕ se lit dans la décomposition (4.5) du vecteur $\vec{n}$ en terme des angles sphériques. Remarquons que les générateurs des rotations (4.7) sont donnés par des matrices anti-hermitiennes, c.-à.-d. telles que $M^\dagger = -M$. Il s'agit de la convention habituelle en mathématiques, alors qu'en physique quantique on préférera des générateurs hermitiens.

On peut vérifier que les matrices $R(\theta,\phi,\alpha)$ données par l'éqn. (4.11) vérifient les propriétés :

$$\forall\theta,\ \phi,\ \alpha\ ,\quad R(\theta,\phi,\alpha)^{\mathsf{T}}R(\theta,\phi,\alpha) = \mathbb{I}\ ,\quad \det R(\theta,\phi,\alpha) = \mathbb{I}\,. \tag{4.12}$$

Ce sont donc bien des matrices de rotation, étant orthogonales et de déterminant unité. Le groupe des rotations à trois dimension est usuellement noté $SO(3)$, où O fait référence au caractère orthogonal des matrices et S, initiale de « spécial », précise que leur déterminant est égal à un.

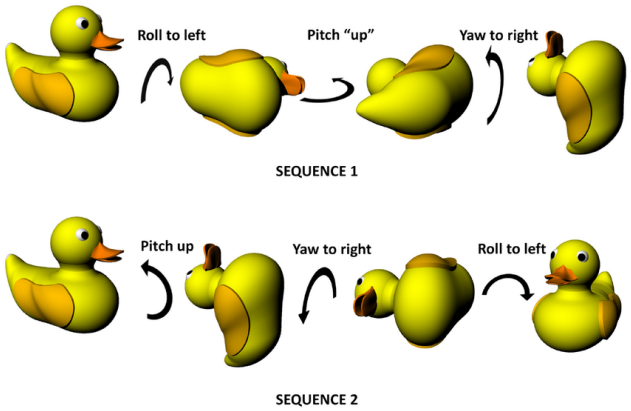

FIGURE 4.2 – Illustration de la non-commutativité des rotations *(source : Jeferey et al. 2015)*.

Le groupe des rotations en lui-même, comme nous avons insisté dans le cours n° 3, n'est pas défini par les matrices de rotation de vecteurs en elles-mêmes mais par les règles de multiplication qu'elles impliquent. Explicitement, on exprime

$$R(\theta_3, \phi_3, \alpha_3) = R(\theta_2, \phi_2, \alpha_2) R(\theta_1, \phi_1, \alpha_1) \tag{4.13}$$

à l'aide du produit matriciel ordinaire, et les fonctions $\theta_3(\theta_2, \phi_2, \alpha_2; \theta_1, \phi_1, \alpha_1)$, etc. donnent la loi de multiplication du groupe, correspondant à la composition de deux rotations arbitraires de l'espace successives ; il s'agit de l'équivalent de la « table de multiplication » d'un groupe fini, voir par exemple l'éqn. (3.58).

Relations de commutation

Les matrices j_x, j_y et j_z reflétent une propriété importante du groupe des rotations en dimension supérieure à 2 : les rotations d'axes différents ne commutent pas entre elles, voir la figure 4.2.

Par un calcul explicite nous trouvons les relations de commutation suivantes pour les générateurs infinitésimaux, à l'aide de la représentation matricielle donnée par les équations (4.7) :

$$[j_x, j_y] = j_z \tag{4.14a}$$
$$[j_y, j_z] = j_x \tag{4.14b}$$
$$[j_z, j_x] = j_y \tag{4.14c}$$

que nous pouvons résumer par (avec $(x, y, z) \leftrightarrow (1, 2, 3)$) :

$$\boxed{[j_a, j_b] = \sum_c \epsilon_{ab}{}^c j_c} \tag{4.15}$$

Ces relations de commutation entre les générateurs infinitésimaux de rotations contiennent l'information nécessaire pour « reconstruire » le groupe des rotations, mis à part les propriétés globales, plus précisément topologiques, sur le groupe. De telles considérations seront utiles pour comprendre l'existence de spins demi-entiers en mécanique quantique, comme nous le montrerons dans le cours n° 13.

De manière générale, les relations de commutation entre générateurs d'un groupe de Lie forment une *algèbre de Lie* ; l'algèbre de Lie du groupe des rotations se note $\mathfrak{so}_3$.

4.1.3 Rotations en mécanique quantique

Pour comprendre le rôle des rotations en mécanique quantique, commençons par étudier la représentation position. Un vecteur de l'espace $\vec{x}$ se transforme, par définition, dans la représentation vectorielle de $\mathsf{SO}(3)$ qui est unitaire – les matrices étant orthogonales et réelles – et irreductible. Notons les éléments de cette rotation $R(\theta, \phi, \alpha)$ comme précédemment.

La fonction d'onde d'une particule sans spin se transforme sous les rotations dans la représentation triviale. Cela signifie explicitement que la fonction d'onde transformée au point image est égale à la fonction d'onde d'origine évaluée en l'antécédent de ce point :

$$R : \begin{cases} \vec{x} & \mapsto & \vec{x}_r = R\vec{x} \\ \psi(\vec{x}) & \mapsto & \psi_r(\vec{x}_r) = \psi(\vec{x}) \end{cases} \tag{4.16}$$

D'où nous obtenons la « nouvelle » fonction d'onde en fonction des « anciennes » coordonnées (en omettant les paramètres) : $\psi_r(\vec{x}) = \psi(R^{-1}\vec{x})$. Cette relation signifie que pour connaître la valeur de la fonction en $\vec{x}$, il faut prendre sa valeur au point antécédent par la rotation.

Pour formaliser cette règle de transformation, associons aux rotations des opérateurs unitaires $\mathcal{R}(\theta, \psi, \phi)$ agissant sur l'espace des états tels que :

$$\langle\vec{x}|\,\mathcal{R}(\theta, \phi, \alpha) = \left\langle R(\theta, \phi, \alpha)^{-1}\vec{x}\right| \,. \tag{4.17}$$

Si on accepte d'utiliser les « kets de position » nous avons également, par unitarité de la représentation du groupe des rotations

$$\mathcal{R}(\theta, \phi, \alpha)\,|\vec{x}\rangle = |R(\theta, \phi, \alpha)\vec{x}\rangle \,. \tag{4.18}$$

La fonction d'onde associée à une particule sans spin se transforme alors sous l'action de l'opérateur de rotation selon :

$$\langle\vec{x}|\psi\rangle = \psi(\vec{x}) \;\mapsto\; \langle\vec{x}|\,\mathcal{R}\,|\psi\rangle = \left\langle R^{-1}\vec{x}\middle|\psi\right\rangle = \psi\left(R^{-1}\vec{x}\right) \,, \tag{4.19}$$

comme nous avions déduit de l'éqn. (4.16).

Recherchons maintenant l'expression de l'opérateur associé à une rotation infinitésimale. Reprenons le développement de l'éqn. (4.6). Nous avons :

$$(R\vec{x})^i = \sum_j \left(\delta^i{}_j - \alpha \sum_k \epsilon^i{}_{jk} n^k + \mathcal{O}(\alpha^2)\right) x^j \,. \tag{4.20}$$

On reconnaît dans le second terme un produit vectoriel et on peut écrire

$$R\vec{x} = \vec{x} + \alpha\,\vec{n}\wedge\vec{x} + \mathcal{O}(\alpha^2)\,, \tag{4.21}$$

et pour la transformation inverse, à cet ordre

$$R^{-1}\vec{x} = \vec{x} - \alpha\,\vec{n}\wedge\vec{x} + \mathcal{O}(\alpha^2)\,. \tag{4.22}$$

Nous pouvons alors développer l'image de la fonction d'onde comme

$$\begin{aligned}\psi(R^{-1}\vec{x}) &= \psi\left(\vec{x} - \alpha\,\vec{n}\wedge\vec{x} + \mathcal{O}(\alpha^2)\right)\\ &= \psi(\vec{x}) - \alpha\,(\vec{n}\wedge\vec{x})\cdot\vec{\nabla}\psi(\vec{x}) + \mathcal{O}(\alpha^2)\\ &= \psi(\vec{x}) - \frac{i\alpha}{\hbar}\,(\vec{n}\wedge\vec{x})\cdot\left(-i\hbar\vec{\nabla}\right)\psi(\vec{x}) + \mathcal{O}(\alpha^2)\,,\end{aligned} \tag{4.23}$$

où nous avons fait apparaître à la dernière ligne l'opérateur impulsion, ou plus précisément la quantité $\langle\vec{x}|\,\vec{\mathcal{P}}\,|\psi\rangle$. Nous avons donc obtenu l'identité suivante :

$$\langle\vec{x}|\,\mathcal{R}\,|\psi\rangle = \langle\vec{x}|\left(\mathcal{I} - \frac{i\alpha}{\hbar}\left(\vec{n}\wedge\vec{\mathcal{X}}\right)\cdot\vec{\mathcal{P}}\right)|\psi\rangle + \mathcal{O}(\alpha^2)\,. \tag{4.24}$$

En utilisant l'invariance du produit mixte par permutation circulaire, nous avons $\vec{\mathcal{P}}\cdot(\vec{n}\wedge\vec{\mathcal{X}}) = \vec{n}\cdot(\vec{\mathcal{X}}\wedge\vec{\mathcal{P}})$. En introduisant l'opérateur *moment cinétique* :

$$\boxed{\vec{\mathcal{L}} \overset{\text{déf.}}{=} \vec{\mathcal{X}}\wedge\vec{\mathcal{P}}} \tag{4.25}$$

On a obtenu finalement le développement de l'opérateur de rotation comme :

$$\boxed{\mathcal{R} = \mathcal{I} - \frac{i\alpha}{\hbar}\,\vec{n}\cdot\vec{\mathcal{L}} + \mathcal{O}(\alpha^2)} \tag{4.26}$$

où nous rappelons que $\vec{n}$ est le vecteur de norme unité spécifiant l'axe de rotation. Ce résultat fondamental, rappelant naturellement l'expression classique (3.32), indique que le moment cinétique $\vec{\mathcal{L}}$ est le générateur des rotations pour des particules sans spin. Le cas des particules avec spin sera considéré en détail au cours n° 13.

Étant donné que les opérateurs position $\vec{\mathcal{X}}$ et impulsion $\vec{\mathcal{P}}$ sont hermitiens, l'opérateur moment cinétique $\vec{\mathcal{L}}$ l'est également et

$$\mathcal{R}^\dagger = \mathcal{I} + \frac{i\alpha}{\hbar}\,\vec{n}\cdot\vec{\mathcal{L}} + \mathcal{O}(\alpha^2) = \mathcal{R}^{-1} + \mathcal{O}(\alpha^2) \tag{4.27}$$

donc l'opérateur de rotation infinitésimale est unitaire.

En exponentiant le générateur infinitésimal de cette représentation du groupe des rotations, nous obtenons l'opérateur unitaire associé à une rotation quelconque :

$$\boxed{\mathcal{R}(\theta,\phi,\alpha) = e^{-\frac{i\alpha}{\hbar}\,\vec{n}\cdot\vec{\mathcal{L}}}} \tag{4.28}$$

Remarquons que, étant donné que l'opérateur moment cinétique $\vec{\mathcal{L}}$ est hermitien, les opérateurs $\mathcal{R}(\theta, \phi, \alpha)$ sont unitaires. L'action des rotations sur l'espace de Hilbert $\mathcal{L}^2(\mathbb{R}^3)$ associé à une particule sans spin, qui peut s'exprimer comme $\psi(\vec{x}) \mapsto \exp(-\alpha\vec{n} \cdot (\vec{x} \wedge \nabla))\psi(\vec{x})$, correspond à une représentation unitaire de $\mathsf{SO}(3)$ de dimension infinie. Suivant le théorème de Peter–Weyl, cette représentation peut se décomposer en somme directe de représentations irréductibles unitaires de dimension finie ; nous y reviendrons lors de l'étude des harmoniques sphériques.

Dérivons maintenant les relations de commutation canoniques entre les composantes $\mathcal{L}_x$, $\mathcal{L}_y$ et $\mathcal{L}_z$ du moment cinétique, à l'aide des relations de commutation canoniques, éqn. (2.35) et la propriété (1.36) du commutateur :

$$\begin{aligned}
[\,\mathcal{L}_x,\, \mathcal{L}_y\,] &= [\,y\mathcal{P}_z - z\mathcal{P}_y,\, z\mathcal{P}_x - x\mathcal{P}_z\,] = i\hbar\,(x\mathcal{P}_y - y\mathcal{P}_x) = i\hbar\mathcal{L}_z \\
[\,\mathcal{L}_z,\, \mathcal{L}_x\,] &= [\,x\mathcal{P}_y - y\mathcal{P}_x,\, y\mathcal{P}_z - z\mathcal{P}_y\,] = i\hbar\,(z\mathcal{P}_x - x\mathcal{P}_z) = i\hbar\mathcal{L}_y \\
[\,\mathcal{L}_y,\, \mathcal{L}_z\,] &= [\,z\mathcal{P}_x - x\mathcal{P}_z,\, x\mathcal{P}_y - y\mathcal{P}_x\,] = i\hbar\,(y\mathcal{P}_x - x\mathcal{P}_y) = i\hbar\mathcal{L}_x
\end{aligned} \tag{4.29}$$

qui peuvent se regrouper comme (avec $(x, y, z) \leftrightarrow (1, 2, 3)$) :

$$\boxed{[\,\mathcal{L}_a,\, \mathcal{L}_b\,] = i\hbar \sum_c \epsilon_{ab}{}^c\, \mathcal{L}_c} \tag{4.30}$$

On remarque que ces relations de commutation sont similaires aux relations de commutation (4.15) entre les générateurs « géométriques » des rotations, au facteur $i\hbar$ près. Le facteur $\hbar$ s'explique par des raisons dimensionnelles, alors que le facteur i reflète le choix suivant : les générateurs des rotations (4.14) sont définis, comme il est d'usage pour les mathématiciens, comme des opérateurs anti-hermitiens, alors que les composantes de l'opérateur moment cinétique $\vec{\mathcal{L}}$ sont, par construction, des opérateurs hermitiens.

Opérateurs vectoriels

Les composantes du moment cinétique obéissent au relations de commutation (4.30) qui encodent les propriétés du groupe du rotation dont ils sont elles sont les générateurs infinitésimaux.

Ces relations de commutation peuvent s'interpréter d'une manière un peu différente. Remarquons que les opérateurs position et impulsion satisfont à des relations de commutation avec le moment cinétique d'un type analogue :

$$[\,\mathcal{L}_i,\, \mathcal{X}_j\,] = \left[\sum_{kl} \epsilon_i{}^{kl} \mathcal{X}_k \mathcal{P}_l,\, \mathcal{X}_j\right] = \sum_{kl} \epsilon_i{}^{kl} \mathcal{X}_k\, [\,\mathcal{P}_l,\, \mathcal{X}_j\,] = -i\hbar \sum_{kl} \epsilon_i{}^{kl} \mathcal{X}_k \delta_{l,j} = i\hbar \sum_k \epsilon_{ij}{}^k\, \mathcal{X}_k \tag{4.31a}$$

$$[\,\mathcal{L}_i,\, \mathcal{P}_j\,] = \left[\sum_{kl} \epsilon_i{}^{kl} \mathcal{X}_k \mathcal{P}_l,\, \mathcal{P}_j\right] = \sum_{kl} \epsilon_i{}^{kl}\, [\,\mathcal{X}_k,\, \mathcal{P}_j\,]\, \mathcal{P}_l = i\hbar \sum_{kl} \epsilon_i{}^{kl} \delta_{k,j} \mathcal{P}_l = i\hbar \sum_l \epsilon_{ij}{}^l\, \mathcal{P}_l \tag{4.31b}$$

Les trois relations de commutation (4.30) et (4.31) signifient que, vis à vis des rotations infinitésimales de l'espace-temps, la position, l'impulsion et le moment cinétique se transforment comme des vecteurs, dans un sens que nous préciserons plus bas.

On définit de manière complètement générale un opérateur vectoriel en mécanique quantique par :

Définition 1 (opérateur vectoriel). *Soit* $\{\mathcal{O}_i,\, i = 1,2,3\}$ *un triplet d'opérateurs. Ils forment un opérateur vectoriel s'ils satisfont aux relations de commutation*

$$\boxed{[\,\mathcal{L}_i,\, \mathcal{O}_j\,] = i\hbar \sum_k \epsilon_{ij}{}^k\, \mathcal{O}_k} \tag{4.32}$$

Considérons la valeur moyenne d'un opérateur vectoriel, par exemple l'opérateur position, pour une particule sans spin caractérisée par un état $|\psi\rangle$. Dans le point de vue actif l'action d'une rotation sur le système va changer l'état du système (en un état physiquement équivalent, s'agissant d'une symétrie) ; le ket d'état $|\psi\rangle$ va se transfomer suivant l'éqn. (4.16). Nous avons :

$$\langle\psi|\,\vec{\mathcal{X}}\,|\psi\rangle \;\mapsto\; \langle\psi|\,\mathcal{R}^\dagger\vec{\mathcal{X}}\mathcal{R}\,|\psi\rangle = \langle\psi|\,\mathcal{R}^{-1}\vec{\mathcal{X}}\mathcal{R}\,|\psi\rangle\ , \tag{4.33}$$

où nous avons utilisé l'unitarité des opérateurs $\mathcal{R}(\theta,\phi,\alpha)$. Explicitement nous avons en insérant une relation de fermeture (2.27) :

$$\begin{aligned}
\langle\psi|\,\mathcal{R}^{-1}\vec{\mathcal{X}}\mathcal{R}\,|\psi\rangle &= \int \mathrm{d}^3x\,\langle\psi|\vec{x}\rangle\,\langle\vec{x}|\,\mathcal{R}^{-1}\vec{\mathcal{X}}\mathcal{R}\,|\psi\rangle = \int \mathrm{d}^3x\,\overline{\psi(\vec{x})}\,\langle R\vec{x}|\,\vec{\mathcal{X}}R\,|\psi\rangle \\
&= \int \mathrm{d}^3x\,\overline{\psi(\vec{x})}R\vec{x}\,\big\langle R^{-1}(R\vec{x})\big|\psi\big\rangle = R\int \mathrm{d}^3x\ |\psi(\vec{x})|^2\,\vec{x} \\
&= R\langle\vec{\mathcal{X}}\rangle_\psi\,.
\end{aligned} \tag{4.34}$$

Comme attendu, la valeur moyenne de l'opérateur position se transforme comme un vecteur de l'espace ; ceci se généralise à tous les « éléments de matrice » de tout opérateur vectoriel.

Un autre moyen, complètement équivalent, d'arriver à la même conclusion est de postuler que le ket d'état $|\psi\rangle$ ne se transforme pas par la rotation mais qu'en revanche l'opérateur position se transforme selon

$$\vec{\mathcal{X}} \;\mapsto\; \mathcal{R}^{-1}\vec{\mathcal{X}}\mathcal{R}\,. \tag{4.35}$$

Ce point de vue correspond au *point de vue passif* suivant la discussion au début de ce cours, si on considère que les « vecteurs propres » de position associés à l'opérateur $\vec{\mathcal{X}}$ comme des vecteurs de base de l'espace des états. Dans ce point de vue passif cette loi de transformation doit s'appliquer à tous les autres opérateurs vectoriels comme l'impulsion ou le moment cinétique.

Si on considère une rotation infinitésimale, voir l'éqn. (4.26), nous obtenons au premier ordre :

$$\vec{\mathcal{X}} \;\mapsto \vec{\mathcal{X}}_R \overset{\text{déf.}}{=} \left(\mathcal{I} + \frac{i\alpha}{\hbar}\,\vec{n}\cdot\vec{\mathcal{L}}\right)\vec{\mathcal{X}}\left(\mathcal{I} - \frac{i\alpha}{\hbar}\,\vec{n}\cdot\vec{\mathcal{L}}\right) + \mathcal{O}(\alpha^2)\,, \tag{4.36}$$

impliquant que la variation de $\vec{\mathcal{X}}$ au premier ordre soit donnée par :

$$\delta\vec{\mathcal{X}} \overset{\text{déf.}}{=} \vec{\mathcal{X}}_R - \vec{\mathcal{X}} = \frac{i\alpha}{\hbar}\left[\vec{n}\cdot\vec{\mathcal{L}},\, \vec{\mathcal{X}}\right] . \tag{4.37}$$

Ainsi, dans le point de vue passif, les relations de commutation (4.32) s'interprètent comme la transformation au premier ordre des opérateurs vectoriels sous une rotation. Mathématiquement, ces lois de transformations correspondent à la *représentation adjointe* de $\mathfrak{so}_3$, l'algèbre de Lie du groupe des rotations.

On peut également définir les opérateurs scalaires comme les opérateurs commutant avec les trois composantes du moment cinétique, soit

Définition 2 (opérateur scalaire). *Un opérateur $\mathcal{O}$ un appelé opérateur scalaire s'il satisfait aux relations de commutation :*

$$\forall i = 1,2,3\ , \quad [\mathcal{L}_i,\, \mathcal{O}] = 0\,, \tag{4.38}$$

Par exemple l'opérateur $\mathcal{X}^2 \overset{\text{déf.}}{=} \vec{\mathcal{X}}\cdot\vec{\mathcal{X}}$ est un opérateur scalaire (exercice : le montrer explicitement).

Lorsqu'un système quantique est invariant par rotation, son hamiltonien $\mathcal{H}$ est un opérateur scalaire. Ce n'est pas vrai par exemple si on place un système dans un champ électrique $\vec{E}$ ou un champ magnétique $\vec{B}$ extérieurs, ce qui brise l'isotropie de l'espace en donnant une direction privilégiée ; il ne reste plus que le sous-groupe de symétrie correspondant aux rotations dans le plan perpendiculaire à cette direction.

4.2 Groupes de translations

Une translation de l'espace à trois dimensions correspond à une transformation de la forme :

$$\vec{x} \mapsto \tau_{\vec{a}}(\vec{r}) \overset{\text{déf.}}{=} \vec{x} + \vec{a}\ , \quad \vec{a} \in \mathbb{R}^3\,. \tag{4.39}$$

Il s'agit naturellement d'un groupe abélien, étant donné que

$$\tau_{\vec{a}}\tau_{\vec{b}} = \tau_{\vec{a}+\vec{b}}\,, \tag{4.40}$$

qui est de plus isomorphe à $(\mathbb{R}^3, +)$.

L'étude des translations en mécanique quantique suit les mêmes étapes que celles des rotations. Une fonction d'onde se transforme trivialement par translation, c.-à.-d. :

$$\tau_{\vec{a}} : \begin{cases} \vec{x} & \mapsto & \vec{x}_t = \vec{x} + \vec{a} \\ \psi(\vec{x}) & \mapsto & \psi_T(\vec{x}_T) = \psi(\vec{x}) \end{cases} \tag{4.41}$$

d'où on tire que $\psi_T(\vec{x}) = \psi(\vec{x} - \vec{a})$.

Associons aux translations des opérateurs unitaires $\mathcal{T}_{\vec{a}}$ agissant sur l'espace des états tels que :

$$\langle\vec{x}|\,\mathcal{T}_{\vec{a}} = \langle\vec{x} - \vec{a}|\ . \tag{4.42}$$

Si on accepte d'utiliser les « kets de position » nous avons également, par unitarité

$$\mathcal{T}_{\vec{a}}\,|\vec{x}\rangle = |\vec{x}+\vec{a}\rangle\ . \tag{4.43}$$

Une fonction d'onde se transforme alors sous l'action de l'opérateur de rotation selon :

$$\langle\vec{x}|\psi\rangle = \psi(\vec{x}) \ \mapsto\ \langle\vec{x}|\,\mathcal{T}_{\vec{a}}\,|\psi\rangle = \langle\vec{x}-\vec{a}|\psi\rangle = \psi\left(\vec{x}-\vec{a}\right)\,, \tag{4.44}$$

comme nous avions déduit de l'éqn. (4.41).

Recherchons maintenant l'expression de l'opérateur associé à une translation infinitésimale. Nous pouvons développer l'image de la fonction d'onde comme

$$\begin{aligned}\psi(\vec{x}-\vec{a}) &= -\vec{a}\cdot\vec{\nabla}\psi\left(\vec{x}\right) + \mathcal{O}(a^2)\\ &= -\frac{i\vec{a}}{\hbar}\left(-i\hbar\vec{\nabla}\right)\psi\left(\vec{x}\right) + \mathcal{O}(a^2)\end{aligned} \tag{4.45}$$

On a obtenu finalement le développement de l'opérateur de translation comme :

$$\boxed{\mathcal{T}_{\vec{a}} = \mathcal{I} - \frac{i\vec{a}\cdot\vec{\mathcal{P}}}{\hbar} + \mathcal{O}(a^2)} \tag{4.46}$$

Ce résultat rappelle naturellement le résultat obtenu à la section 3.2 concernant la mécanique classique : le générateur des translations spatiales est associé à l'opérateur impulsion. Étant donné que l'opérateur impulsion $\vec{\mathcal{P}}$ est hermitien,

$$\mathcal{T}_{\vec{a}}^{\dagger} = \mathcal{I} + \frac{i\vec{a}\cdot\vec{\mathcal{P}}}{\hbar} + \mathcal{O}(a^2) = \mathcal{T}_{-\vec{a}} + \mathcal{O}(a^2) = \mathcal{T}_{\vec{a}}{}^{-1} + \mathcal{O}(a^2)\,, \tag{4.47}$$

et l'opérateur de translation infinitésimale est bien unitaire.

L'identification entre le générateur des translations et l'opérateur impulsion implique que les « états propres » d'impulsion $\vec{p}$ sont également des états propres de l'opérateur de translation infinitésimale ; cela reflète une propriété évidente des ondes planes, $\exp\frac{i\vec{p}\cdot(\vec{x}+\vec{a})}{\hbar} = \exp\frac{i\vec{p}\cdot\vec{a}}{\hbar}\,\exp\frac{i\vec{p}\cdot\vec{x}}{\hbar}$.

On peut déduire de cette étude la forme d'une translation finie, c.-à.-d. par un vecteur non-infinitésimal. En exponentiant le générateur des translations nous trouvons l'opérateur unitaire

$$\boxed{\mathcal{T}_{\vec{a}} = e^{-\frac{i\vec{a}\cdot\vec{\mathcal{P}}}{\hbar}}} \tag{4.48}$$

dont les « éléments de matrice » dans la représentation d'impulsion sont

$$\langle\vec{p}\,'|\,\mathcal{T}_{\vec{a}}\,|\vec{p}\rangle = e^{-\frac{i\vec{a}\cdot\vec{p}}{\hbar}}\,\delta(\vec{p}\,'-\vec{p})\,. \tag{4.49}$$

Dans la perspective de la théorie des groupes, une onde plane correspond à une représentation irreductible du groupe des translations, qui est de dimension complexe un comme on s'y attend d'après le corollaire 2 du cours n° 3.

Dans la représentation de position on peut également écrire formellement une translation finie par exponentiation :

$$\langle \vec{x}| \mathcal{T}_{\vec{a}} |\psi\rangle = e^{-\vec{a}\cdot\vec{\nabla}}\psi(x) = \psi(\vec{x}) - \vec{a}\cdot\vec{\nabla}\psi(\vec{x}) + \frac{1}{2}\left(-\vec{a}\cdot\vec{\nabla}\right)^2\psi(\vec{x}) + \cdots \tag{4.50}$$

où on reconnaît le développement en série entière de $\psi(\vec{x}-\vec{a})$ (supposant ce dernier convergent).

Suivant le raisonnement développé à la section 4.1, on peut considérer les translations suivant un point de vue passif dans lequel les observables se transforment alors que les kets d'état sont fixes. On peut s'attendre à ce que les opérateurs position $\vec{\mathcal{X}}$ et impulsion $\vec{\mathcal{P}}$ se comportent différemment, la vitesse d'une particule ne changeant pas par translation. Nous avons au premier ordre

$$\begin{aligned}\vec{\mathcal{X}} \;\mapsto\; \mathcal{T}_{-\vec{a}}\,\vec{\mathcal{X}}\,\mathcal{T}_{\vec{a}} &= \left(\mathcal{I} + \frac{i\vec{a}\cdot\vec{\mathcal{P}}}{\hbar}\right)\left(\vec{\mathcal{X}}\mathcal{I} - \frac{i\vec{a}\cdot\vec{\mathcal{P}}}{\hbar}\right)\\ &= \vec{\mathcal{X}} + \frac{i}{\hbar}\left[\,\vec{a}\cdot\vec{\mathcal{P}},\,\vec{\mathcal{X}}\,\right] = \vec{\mathcal{X}} + \vec{a}\end{aligned} \tag{4.51}$$

ainsi que

$$\vec{\mathcal{P}} \;\mapsto\; \mathcal{T}_{-\vec{a}}\,\vec{\mathcal{P}}\,\mathcal{T}_{\vec{a}} = \vec{\mathcal{P}} + \frac{i}{\hbar}\left[\,\vec{a}\cdot\vec{\mathcal{P}},\,\vec{\mathcal{P}}\,\right] = \vec{\mathcal{P}}\,. \tag{4.52}$$

Ces deux lois de transformations reflètent l'intuition que l'on peut avoir de la transformation des grandeurs classiques associées et se généralisent trivialement à des transformations finies, par exemple en décomposant une translation finie en translations infinitésimales succcessives ou en développant l'exponentielle dans (4.48) en série entière.

4.3 Parité

Les isométries de l'espace vectoriel $\mathbb{R}^3$ sont les transformations linéaires qui préservent les longueurs. Munissant $\mathbb{R}^3$ du produit scalaire usuel, l'action d'une isométrie sur les composantes d'un vecteur en coordonnées cartésiennes est donnée par une matrice réelle 3×3 orthogonale, c.-à.-d. vérifiant $R^T R = \mathbb{I}_3$. L'ensemble de ces matrices forme le groupe $O(3)$, groupe orthogonal de degré trois.

La condition $R^T R = \mathbb{I}_3$ implique que, pour tout $R \in O(3)$, $\det R = \pm 1$. Les éléments de $O(3)$ pour lesquels $\det R = 1$ forment le sous-groupe $SO(3) \subset O(3)$, groupe des rotations étudié à la section (4.1).

Les transformations $S \in O(3)$ pour lesquelles $\det S = -1$ sont des *rotation impropres*, qui peuvent se décrire comme une combinaison d'une rotation ordinaire R et de la symétrie discrète de parité ou inversion spatiale :

$$P : \vec{x} \mapsto\; -\vec{x}, \tag{4.53}$$

qui vérifie évidemment $P^2 = e$, e étant l'élément neutre. L'opération de parité inverse l'orientation de l'espace.

Pour tout $S \in O(3)$ tel que $\det S = -1$, l'élément du groupe $P \cdot S$ a pour déterminant $\det P \times \det S = 1$, et correspond donc à une certaine rotation $R \in O(3)$; on peut ainsi, en multipliant $P \cdot S = R$ à gauche par P, paramétrer toute rotation impropre S comme $S = P \cdot R$ où P est la parité (4.53) et R une rotation ordinaire. Une conséquence évidente de ce raisonnement est que le produit de deux rotations impropres est une rotation ordinaire ; elles ne forment donc pas un sous-groupe de $O(3)$.

La parité en elle-même génère un sous-groupe fini de $O(3)$ à deux élements, $\{e, P\}$ qui est isomorphe à $\mathbb{Z}_2$. Étant donné que, pour tout $g \in O(3)$, $P \cdot g = g \cdot P$, les éléments de ce groupe commutent avec tous les éléments de $O(3)$: on dit que $\{e, P\}$ est le *centre* du groupe $O(3)$.

La théorie des représentations du groupe $\mathbb{Z}_2$ est extrêmement simple. Il n'existe en tout et pour tout que deux représentations irreductibles inéquivalentes :

— la représentation triviale D^0, telle que $D^0(e) = D^0(P) = 1$;

— la représentation signature D^1, telle que $D^1(e) = 1$ et $D^1(P) = -1$.

Une quantité se transformant dans la représentation triviale du groupe des rotations $SO(3)$ peut avoir en conséquence deux comportements possibles par rapport à la partité :

— Un *scalaire* se transforme dans la représentation triviale ;

— Un *pseudo-scalaire* se transforme dans la représentation signature.

De même, une quantité se transformant dans la représentation vectorielle du groupe des rotations $SO(3)$ peut avoir en conséquence deux comportements possibles par rapport à la partité :

— Un *vecteur* se transforme dans la représentation signature ;

— Un *pseudo-vecteur* se transforme dans la représentation triviale.

Parmi les quantités usuelles, la position $\vec{x}$, l' impulsion $\vec{p}$ ou le champ électrique $\vec{E}$ sont des vecteurs alors que le champ magnétique $\vec{B}$ et le moment cinétique $\vec{L}$ sont des pseudo-vecteurs. On en déduit, par exemple, que la quantité $\vec{E} \cdot \vec{B}$ est un pseudo-scalaire ou que la quantité $\vec{B} \cdot \vec{L}$ est un scalaire.

Passons maintenant à la description quantique de la parité. On introduit un opérateur $\mathfrak{P}$ tel que :

$$\forall \vec{x} \in \mathbb{R}^3 \ , \quad \langle \vec{x} | \, \mathfrak{P} = \langle -\vec{x} | \ . \tag{4.54}$$

Il satisfait naturellement la propriété $\mathfrak{P}^2 = \mathfrak{I}$. On en déduit l'action de la parité sur les fonctions d'onde,

$$\psi(\vec{x}) \ \mapsto \ \langle \vec{x} | \, \mathfrak{P} \, | \psi \rangle = \langle -\vec{x} | \psi \rangle = \psi(-\vec{x}) \, . \tag{4.55}$$

Pour comprendre comment obtenir des fonctions d'onde se transformant dans des représentations irreductibles de la parité, on rappelle que toute fonction peut se décomposer en fonctions paire et impaire :

$$f(x) = f_+(x) + f_-(x) \ , \quad f_\pm(x) = \frac{f(x) \pm f(-x)}{2} \ , \quad f_\pm(-x) = \pm f(x) \, . \tag{4.56}$$

Par analogie on peut projeter tout vecteur d'état en un vecteur se transformant dans la représentation triviale et un vecteur se transformant dans la représentation signature :

$$|\psi\rangle = |\psi\rangle_+ + |\psi\rangle_- \ , \quad |\psi\rangle_\pm = \frac{1 \pm \mathfrak{P}}{2} |\psi\rangle \ , \quad \mathfrak{P} |\psi\rangle_\pm = \pm |\psi\rangle_\pm \ . \tag{4.57}$$

Examinons maintenant le comportement des observables position et impulsion par l'action de la parité. Nous avons premièrement

$$\begin{aligned} \langle \vec{x}| \mathfrak{P} \vec{\mathfrak{X}} &= \langle -\vec{x}| \vec{\mathfrak{X}} = -\vec{x} \langle -\vec{x}| \\ \langle \vec{x}| \vec{\mathfrak{X}} \mathfrak{P} &= \vec{x} \langle \vec{x}| \mathfrak{P} = \vec{x} \langle -\vec{x}| \end{aligned} \tag{4.58}$$

d'où nous déduisons

$$\forall \vec{x}, \langle \vec{x}| \mathfrak{P} \vec{\mathfrak{X}} = - \langle \vec{x}| \vec{\mathfrak{X}} \mathfrak{P} \, . \tag{4.59}$$

D'où on déduit, l'opérateur $\mathfrak{P}$ étant unitaire

$$\boxed{\mathfrak{P}^\dagger \vec{\mathfrak{X}} \mathfrak{P} = -\vec{\mathfrak{X}}} \tag{4.60}$$

Son action sur les kets de position est donc naturellement

$$\mathfrak{P} |\vec{x}\rangle = |-\vec{x}\rangle \tag{4.61}$$

et $\mathfrak{P}$ est non seulement unitaire mais également hermitien.

Comme nous l'avions fait précédemment on peut examiner comment la valeur moyenne de la position se transforme par parité. Nous avons

$$\begin{aligned} \langle \vec{\mathfrak{X}} \rangle_\psi = \langle \psi| \vec{\mathfrak{X}} |\psi\rangle \ \mapsto \langle \psi| \mathfrak{P}^\dagger \vec{\mathfrak{X}} \mathfrak{P} |\psi\rangle &= - \langle \psi| \vec{\mathfrak{X}} |\psi\rangle \\ &= -\langle \vec{\mathfrak{X}} \rangle_\psi \, . \end{aligned} \tag{4.62}$$

Passons maintenant à l'opérateur impulsion. Nous avons

$$\mathfrak{X}_i \mathcal{P}_j - \mathfrak{X}_j \mathcal{P}_i = i\hbar \delta_{ij} \tag{4.63a}$$

$$\mathfrak{X}_i \mathfrak{P} + \mathfrak{P} \mathfrak{X}_i = 0 \tag{4.63b}$$

Si on multiplie le commutateur canonique (4.63a) à gauche par $\mathfrak{P}^\dagger$ et à droite par $\mathfrak{P}$ nous avons, en utilisant l'unitarité de $\mathfrak{P}$ et (4.63b) :

$$\left(\mathfrak{P}^\dagger \mathfrak{X}_i \mathfrak{P}\right) \left(\mathfrak{P}^\dagger \mathcal{P}_j \mathfrak{P}^\dagger\right) - \left(\mathfrak{P}^\dagger \mathcal{P}_j \mathfrak{P}^\dagger\right) \left(\mathfrak{P}^\dagger \mathfrak{X}_i \mathfrak{P}\right) = i\hbar \delta_{ij} \tag{4.64}$$

soit

$$-\mathfrak{X}_i \left(\mathfrak{P}^\dagger \mathcal{P}_j \mathfrak{P}^\dagger\right) + \left(\mathfrak{P}^\dagger \mathcal{P}_j \mathfrak{P}^\dagger\right) \mathfrak{X}_i = i\hbar \delta_{ij} \ \Leftrightarrow \ \left[\mathfrak{X}_i, \left(\mathfrak{P}^\dagger \mathcal{P}_j \mathfrak{P}^\dagger\right) \right] = -i\hbar \delta_{ij} \, . \tag{4.65}$$

Cela permet d'identifier

$$\mathfrak{P}^\dagger \vec{\mathcal{P}} \mathfrak{P} = -\vec{\mathcal{P}} \, . \tag{4.66}$$

Si on se rappelle que l'opérateur impulsion est le générateur des translations, voir éqn. (4.46), on peut faire l'observation suivante :

$$\mathfrak{P}^\dagger \left(\mathfrak{I} - \frac{i\vec{a}\cdot\vec{\mathcal{P}}}{\hbar} \right) \mathfrak{P} = \mathfrak{I} + \frac{i\vec{a}\cdot\vec{\mathcal{P}}}{\hbar}\,, \tag{4.67}$$

donc la parité transforme, comme on pourrait s'y attendre, un générateur de translation $\mathcal{T}_{\vec{a}}$ en le générateur de la translation opposée, $\mathcal{T}_{-\vec{a}}$.

Examinons finalement la transformation du moment cinétique sous l'action de la parité. Nous avons

$$\mathfrak{P}^\dagger \mathcal{L}_i \mathfrak{P} = \sum_{jk} \epsilon_{ijk} \mathfrak{P}^\dagger \mathcal{X}_j \mathcal{P}_k \mathfrak{P} = \sum_{jk} \epsilon_{ijk} \left(\mathfrak{P}^\dagger \mathcal{X}_j \mathfrak{P} \right) \left(\mathfrak{P}^\dagger \mathcal{P}_k \mathfrak{P} \right) = \mathcal{L}_i\,, \tag{4.68}$$

ce qui montre que l'opérateur moment cinétique, comme son analogue classique, est un pseudo-vecteur.

Pour terminer cette étude nous pouvons énoncer le théorème suivant (voir par exemple [1], page 273) :

Théorème 1 *Soit $\mathcal{O}$ un opérateur hermitien admettant un sous-espace propre de dimension un pour la valeur propre λ engendré par un vecteur propre $|\lambda\rangle$. Si $[\,\mathcal{O},\,\mathfrak{P}\,] = 0$ alors $|\lambda\rangle$ est aussi un vecteur propre de $\mathfrak{P}$.*

La preuve en est immédiate. On décompose $|\lambda\rangle = \frac{1+\mathfrak{P}}{2}|\lambda\rangle + \frac{1-\mathfrak{P}}{2}|\lambda\rangle$. Les deux vecteurs $|\lambda\rangle_\pm = \frac{1\pm\mathfrak{P}}{2}|\lambda\rangle$ sont orthogonaux car :

$${}_+\langle\lambda|\lambda\rangle_- = \frac{1}{4}\langle\lambda|\left(1+\mathfrak{P}^\dagger\right)(1-\mathfrak{P})|\lambda\rangle = \frac{1}{4}\langle\lambda|\left(1-\mathfrak{P}^\dagger\mathfrak{P}+\mathfrak{P}^\dagger-\mathfrak{P}\right)|\lambda\rangle = 0\,, \tag{4.69}$$

où nous avons utilisé l'unitarité et l'hermiticité de $\mathfrak{P}$. Comme le sous-espace propre de $\mathcal{O}$ est de dimension un par hypothèse, il vient que soit $|\psi\rangle_+ = 0$ soit $|\psi\rangle_- = 0$. □

Prenons l'exemple du hamiltonien H de l'oscillateur harmonique, à une dimension dont le spectre n'est pas dégénéré. On considère l'opérateur de parité à une dimension vérifiant $\mathfrak{P}|x\rangle = |-x\rangle$, et il est aisé de vérifier que

$$[\,\mathfrak{P},\,H\,] = 0\,, \tag{4.70}$$

car $\mathcal{X}^2$ et $\mathcal{P}^2$ sont des opérateurs pairs. La fonction d'onde de l'état fondamental $|0\rangle$, voir l'éqn. (2.140), étant une gaussienne, fonction paire, nous avons

$$\mathfrak{P}|0\rangle = |0\rangle\ . \tag{4.71}$$

Étant donné que

$$a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\mathcal{X} - \frac{i}{\sqrt{2m\hbar\omega}}\mathcal{P} \tag{4.72}$$

nous avons, d'après (4.60,4.66)

$$\mathfrak{P}^\dagger a^\dagger \mathfrak{P} = -a^\dagger \tag{4.73}$$

d'où nous déduisons que

$$\mathfrak{P}\,|n\rangle = \frac{1}{\sqrt{n!}}\left(\mathfrak{P}a^\dagger\mathfrak{P}^\dagger\right)^n \mathfrak{P}\,|0\rangle = (-1)^n\,|n\rangle\ . \tag{4.74}$$

Les états propres de l'oscillateur harmonique sont ainsi des états propres de $\mathfrak{P}$, alternativement pairs et impairs par rapport à l'inversion.

4.4 Le groupe euclidien

Les symétries de l'espace que nous avons étudiées jusqu'à présent forment d'une part le groupe orthogonal $O(3)$ (rotations propres et impropres) et d'autre part le groupe des translations, isomorphe à $(\mathbb{R}^3, +)$. On peut en réalité englober ces deux types de transformations dans un seul groupe, appelé *groupe euclidien*, groupes des isométries de l'espace euclidien (qui est un espace affine), et noté $E(3)$ ou $ISO(3)$. Tout élément de ce groupe se décompose de manière unique en une rotation (propre ou impropre) et une translation :

$$\forall g \in E(3)\ ,\quad \exists!(r,\vec{a}) \in O(3)\times\mathbb{R}^3\ ,\quad \text{t.q.}\quad \forall \vec{x}\in\mathbb{R}^3\ ,\quad g(\vec{x}) = r(\vec{x}+\vec{a})\,. \tag{4.75}$$

On vérifie en outre que, pour tout $g \in E(3)$, si $\tau_{\vec{a}}$ est une translation de vecteur $\vec{a}$, alors $g^{-1}\circ\tau_{\vec{a}}\circ g$ est une translation également. En effet, tout $g \in E(3)$ étant soit un élément de $O(3)$ (rotation propre ou impropre) soit un élément de $\mathbb{R}^3$ (c.-à.-d. une translation), nous avons soit

$$\forall r \in O(3)\ ,\quad \forall \vec{x}\in\mathbb{R}^3\ ,\quad \left(r^{-1}\circ\tau_{\vec{a}}\circ r\right)(\vec{x}) = r^{-1}\left(r\vec{x}+\vec{a}\right) = \vec{x} + r^{-1}\vec{a} = \tau_{r^{-1}\vec{a}}(\vec{x})\,, \tag{4.76}$$

soit :

$$\forall \vec{b} \in \mathbb{R}^3\ ,\quad \forall \vec{x}\in\mathbb{R}^3\ ,\quad \left(\tau_{\vec{b}}^{-1}\circ\tau_{\vec{a}}\circ\tau_{\vec{b}}\right)(\vec{x}) = \tau_{-\vec{b}}\left(\vec{x}+\vec{a}+\vec{b}\right) = \vec{x}+\vec{a} = \tau_{\vec{a}}(\vec{x})\,. \tag{4.77}$$

Cette propriété indique que le groupe des translations $\mathbb{R}^3$ est, par définition, un *sous-groupe normal* du groupe euclidien $E(3)$. Combinée avec la propriété précédente, cela signifie, par définition, que le groupe $E(3)$ possède la structure d'un *produit semi-direct* :

$$E(3) = O(3) \ltimes \mathbb{R}^3\,. \tag{4.78}$$

Cette structure de groupe sera surtout utile pour l'étude des systèmes périodiques au chapitre 5.

À retenir

— Les points de vue passif et actif ;
— les générateurs des rotations en mécanique quantique sont les composantes du moment cinétique ;
— les relations de commutation entre les composantes du moment cinétique et leur signification ;
— la notion d'opérateur vectoriel ;
— les composantes de l'impulsion sont les générateurs des translations dans l'espace ;
— l'opérateur parité et son action sur les observables.

Bibliographie

[1] J. J. Sakurai, "Modern Quantum mechanics," Cambridge University Press, 1982

Cours n°5

Systèmes périodiques

Dans le chapitre 4 nous avons étudié en détail les symétries de l'espace et leurs conséquences physiques. Plus précisément, cette analyse s'appliquait à une particule dans le vide, ou dans un milieu pouvant être considéré, aux échelles d'énergies pertinentes pour le système, homogène (identique en tout point) et isotrope (identique dans toutes les directions). Dans ce cas le groupe de symétrie du système est celui de l'espace euclidien, $\mathsf{E}(3) = \mathsf{O}(3) \ltimes \mathbb{R}^3$.

Un grand nombre de solides se trouvent sous une forme cristalline, c.-à.-d. dans laquelle les atomes ou groupes d'atomes qui le constituent sont arrangés périodiquement dans l'espace. Pour un matériau *monocristallin*, cette structure périodique persiste sur tout le volume du solide, alors qu'un matériau *polycristallin* est constitué d'un assemblage de petits cristaux appelés *cristallites*.

Aux échelles de longueur auxquelles ces structures périodiques sont respectées – c.-à.-d. aux échelles inférieures à la taille du solide pour les matériaux monocristallins et inférieures à la taille des cristallites pour les matériaux polycristallins – un sous-groupe discret du groupe de $\mathsf{E}(3)$, groupe de symétrie de l'espace euclidien, sera préservé par le système. Ces symétries ont dés conséquences physiques directement observables, en particulier optiques. Les expériences de diffraction des rayons X ont permis de mettre en évidence la structure cristalline [1], travaux récompensés par le prix Nobel décerné à Max von Laue en 1914.

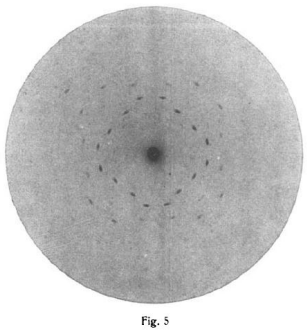

FIGURE 5.1 – Diffraction de rayons X par un échantillon de Sulfure de Zinc, publication historique de Friedrich et al. [1].

Plus récemment le développement des microscopes à effet tunnel a permis d'observer directement la structure cristalline de la matière [2], avancée pour laquelle Gerd Binnig et Heinrich Rohrer ont reçu le prix Nobel en 1986.

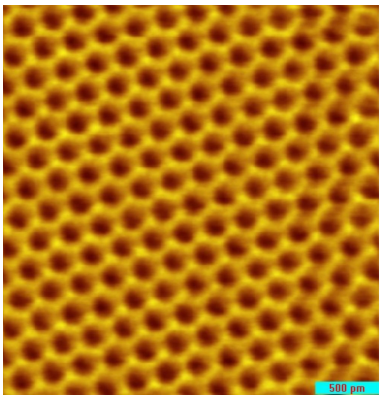

FIGURE 5.2 – Image de graphène obtenue par microscope à effet tunnel (crédit A. Luican-Mayer)

5.1 Réseau unidimensionnel

Avant d'aborder la description des structures périodiques à deux et trois dimensions, nous étudierons tout d'abord brièvement le cas d'une chaîne d'atomes à une dimension, voir la figure 5.3. Ce réseau à une dimension possède une symétrie discrète de translation donnée

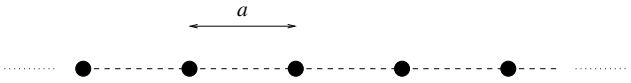

FIGURE 5.3 – Chaîne d'atomes à une dimension

par la transformation :

$$\tau_n \,:\, x \;\mapsto\; x + n\,a\,, \quad n \in \mathbb{Z}\,, \tag{5.1}$$

où a est le pas fixé du réseau. Il s'agit naturellement d'un groupe abélien qui est isomorphe au groupe $(\mathbb{Z}, +)$ des entiers naturels, sous-groupe discret du groupe des translations à une dimension (qui est lui isomorphe à $(\mathbb{R}, +)$), voir la sous-section 3.3.2 du cours 3. Notons que ce système peut correspondre soit à une chaîne infinie, soit à une chaîne de longueur finie $L = Na$, $N \in \mathbb{N}^*$, avec des conditions aux limites périodiques.

Le groupe $(\mathbb{Z}, +)$ étant abélien, ses représentations complexes unitaires et irreductibles sont toutes de dimension un, voir la section 3.3.3 du chapitre 3. Rappelons qu'elles sont paramétrées par un nombre réel $\theta \in [0, 2\pi[$, tel que

$$\forall n \in \mathbb{Z}\,, \quad D^\theta(n) = e^{in\theta}\,, \tag{5.2}$$

et il est évident que $D^\theta(n)D^\theta(n') = e^{i(n+n')\theta} = D^\theta(n+n')$. Dans la suite, plutôt que le paramètre θ sans dimension, nous introduirons le nombre d'onde k tel que

$$\theta = ka\,. \tag{5.3}$$

L'angle θ paramétrant les représentations de $(\mathbb{Z},+)$ étant défini modulo 2π, le nombre d'onde k est défini modulo $2\pi/a$; on peut le choisir conventionnellement dans l'intervalle :

$$k \in \left]-\frac{\pi}{a},\,\frac{\pi}{a}\right]. \tag{5.4}$$

Du point de vue de la théorie des groupes, en raison de la périodicité du paramètre θ, le nombre d'onde k appartient à un réseau à une dimension associé à la symétrie :

$$\tilde\tau_\ell \,:\, k \;\mapsto\; k + \ell\,\tfrac{2\pi}{a}\,, \quad \ell \in \mathbb{Z}\,, \tag{5.5}$$

qui forme naturellement un groupe abélien isomorphe également à $(\mathbb{Z},+)$, appelé *réseau dual*.

Même s'il n'existe pas de notion de générateur infinitésimal pour un groupe symétrie discret, l'analyse effectuée dans le cours n° 4 reste en partie applicable, car nous considérons ici un sous-groupe discret d'un groupe continu : les éléments de $(\mathbb{Z},+)$ peuvent être considérés comme des éléments particuliers de $(\mathbb{R},+)$. L'opérateur unitaire réalisant une translation de n pas du réseau se déduit ainsi de l'éqn. (4.48) :

$$\mathcal{T}_n = e^{-\frac{ina\mathcal{P}}{\hbar}}\,. \tag{5.6}$$

Recherchons l'expression des vecteurs propres de l'opérateur de translation se transformant dans la représentation D^θ. On doit avoir

$$\mathcal{T}_n\,|\psi_\theta\rangle = e^{in\theta}\,|\psi_\theta\rangle\,, \tag{5.7}$$

soit, en représentation position comme $\langle x|\,\mathcal{T}_n\,|\psi_\theta\rangle = \psi_\theta(x-na)$, en posant $k = \theta/a$ comme précédemment,

$$\boxed{\forall n \in \mathbb{Z}\,, \quad \psi_k(x-na) = e^{inka}\psi_k(x)} \tag{5.8}$$

Les solutions de cette équation fonctionnelle sont données en termes d'une fonction périodique u_k arbitraire :

$$\boxed{\psi_k(x) = e^{-ikx}u_k(x)\,, \quad u_k(x+a) = u_k(x)} \tag{5.9}$$

En effet, si on définit $u_k(x) = e^{ikx}\psi_k(x)$, la condition (5.8) satisfaite par ψ_k impose que u_k soit périodique de période a.

Ce résultat, sous une forme plus générale que nous étudierons dans la suite de ce chapitre, porte le nom de *théorème de Bloch* et les fonction d'ondes ψ_k appelées *fonctions de Bloch*. Ces fonctions d'onde peuvent être interprétées comme des ondes planes (premier facteur), appelées *ondes de Bloch*, modulées à l'échelle interatomique par la fonction $u_k(\vec{x})$, voir la fig. 5.4.

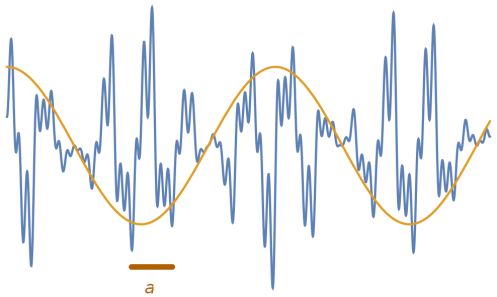

FIGURE 5.4 – Exemple de fonction de Bloch à une dimension. *Seule la partie réelle de la fonction d'onde est représentée.*

Remarque 1 (chaîne d'atomes de longueur finie). *Pour une chaîne d'atomes de longueur finie, constituée de* N *atomes, il est naturel– comme nous l'avions déjà évoqué – de choisir des conditions aux limites périodiques, qui respectent la symétrie de translation sur le réseau. Pour une chaîne de* N *atomes nous devons imposer aux solutions données par l'éqn. (5.46) la condition suivante :*

$$\forall x \in \mathbb{R}\ , \quad \psi_k(x + Na) = \psi_k(x) \quad \Longrightarrow \quad e^{ikNa} = 1 \ \Leftrightarrow \ k \in \tfrac{2\pi}{Na}\,\mathbb{Z}\,. \tag{5.10}$$

5.2 Aspects mathématiques des réseaux

La structure mathématique sous-jacente à la description des solides cristallins, du point de vue des symétries, est la notion de réseau, c.-à.-d. un ensemble de points disposés périodiquement dans le plan ou dans l'espace. Dans le contexte de la cristallographie, ce réseau s'appelle *réseau de Bravais* et les points *nœuds* du réseau.

On peut définir un réseau comme sous-groupe du groupes des translations à d, dimensions, le groupe additif $(\mathbb{R}^d, +)$; rappelons qu'un élément de ce groupe est paramétré par un vecteur $\vec{a} \in \mathbb{R}^d$.

Définition 1 (réseau). *Un réseau* Λ *est un sous-groupe du groupe additif* $(\mathbb{R}^d, +)$*, qui constitue un groupe abélien libre discret, c.-à.-d. un groupe abélien engendré par un nombre fini d'éléments, et qui peut être muni d'une base.*

Une base $\mathcal{B} = \{\vec{a}_1, \ldots, \vec{a}_d\}$ étant choisie, où chaque $\vec{a}_i$ est un vecteur de $\mathbb{R}^d$, tout élément $x \in \Lambda$ peut se décomposer de manière unique sur cette base, avec des coefficients entiers :

$$\forall x \in \Lambda\ , \quad \exists!(n_1, \ldots, n_d) \in \mathbb{Z}^d \text{ t.q.} \quad x = \sum_{i=1}^{d} n_i \vec{a}_i\,. \tag{5.11}$$

Nous avons implicitement supposé que le réseau était de rang maximal, c.-à.-d. que le nombre de vecteurs de base est égal à la dimension de $\mathbb{R}^d$. Nous pouvons associer à une base $\mathcal{B}$ une matrice réelle B de dimension $d \times d$, telle que la p-ième colonne contient les p composantes de $\vec{a}_p \in \mathbb{R}^n$. On peut alors décrire le réseau comme

$$\Lambda = B \cdot \mathbb{Z}^d = \left\{B \cdot z\,,\ z \in \mathbb{Z}^d\right\}. \tag{5.12}$$

5.2.1 Bases d'un réseau

Comme pour les espaces vectoriels, le choix de base pour un réseau n'est pas unique.[1] Avant de l'illustrer par des exemples, démontrons le lemme suivant donnant la relation entre les différentes bases :

Lemme 1 (changement de base pour un réseau). *Deux bases $\mathcal{B}_1$ et $\mathcal{B}_2$, associées à des matrices B_1 et B_2, engendrent le même réseau $\Lambda \in \mathbb{R}^n$ si et seulement si il existe une matrice U à coefficients entiers et de déterminant $\det U = \pm 1$, dite matrice unimodulaire, telle que les matrices B_1 et B_2 satisfont :*

$$B_1 = B_2 \cdot U\,. \tag{5.13}$$

Supposons que le réseau Λ soit engendré à la fois par la base $\mathcal{B}_1$ et la base $\mathcal{B}_2$. Chaque colonne de B_1 est, par définition, un élément de $\Lambda = B_2 \cdot \mathbb{Z}^d$, c.-à.-d. une combinaison linéaire des colonnes de B_2 avec des coefficients entiers, et inversement. Il existe donc des matrices U et V à coefficients entiers telles que $B_1 = B_2 \cdot U$ d'une part et $B_2 = B_1 \cdot V$ d'autre part. On en déduit que $B_2 = B_2 \cdot U \cdot V$ puis, comme B_2 est inversible (le réseau étant de rang maximal), que $U \cdot V = \mathbb{I}_d$ donc $V = U^{-1}$ et $\det U = \pm 1$ (pour garantir que la matrice V se compose d'entiers). Réciproquement, supposons qu'il existe une matrice U unimodulaire telle que $B_1 = B_2 \cdot U$. Nous avons alors

$$\Lambda = \left\{B_1 \cdot z,\, z \in \mathbb{Z}^d\right\} = \left\{B_2 \cdot (U \cdot z),\, z \in \mathbb{Z}^d\right\}. \tag{5.14}$$

Montrons finalement que $U\mathbb{Z}^d = \mathbb{Z}^d$. Pour tout $z \in \mathbb{Z}^d$, $U \cdot z \in \mathbb{Z}^d$, et inversement si $y = U \cdot z$ avec $z \in \mathbb{Z}^d$, $z = U^{-1} \cdot y \in \mathbb{Z}^d$ (car $\det U = \pm 1$ implique que la matrice U n'a que des coefficients entiers). □

Remarque 2 *L'ensemble des matrices unimodulaires $d \times d$ forme un groupe. En particulier, comme nous l'avons remarqué dans la démonstration précédente, la condition $\det U = \pm 1$ implique que l'inverse d'une matrice unimodulaire est également unimodulaire. Ce groupe, appelé groupe général linéaire de degré d sur les entiers, est noté $\mathrm{GL}(d, \mathbb{Z})$.*

Exemple 1 (réseaux bidimensionnels).
Un réseau bidimensionnel Λ est un ensemble de points du plan euclidien disposés périodiquement selon deux vecteurs de base $\vec{a}_1$ et $\vec{a}_2$ non colinéaires, c.-à.-d. $\mathbb{Z}\vec{a}_1 + \mathbb{Z}\vec{a}_2 \subset \mathbb{R}^2$, voir la figure 5.5.

Une autre base du réseau Λ est donnée sur la fig. 5.5 par $\{\vec{b}_1, \vec{b}_2\}$; on vérifie en effet que $(b_1 b_2) = (a_1 a_2) \cdot U$, avec $U = \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)$ qui est une matrice unimodulaire. En revanche $\{\vec{c}_1, \vec{c}_2\}$ ne forme pas une base du reseau – bien que ce soit une base de $\mathbb{R}^2$. Nous avons $(c_1 c_2) = (a_1 a_2) \cdot V$ avec $V = \left(\begin{smallmatrix} 2 & 1 \\ 0 & -1 \end{smallmatrix}\right)$ qui n'est pas une matrice unimodulaire (car $\det V = -2$). On voit facilement sur la figure que $\{\vec{c}_1, \vec{c}_2\}$ engendre un réseau Λ' qui ne contient qu'une partie des points du réseau Λ ; il s'agit d'un *sous-réseau* de ce dernier, c.-à.-d. d'une partie de Λ qui forme elle-même un réseau.

1. Un réseau n'est pas un espace vectoriel car $\mathbb{Z}$ n'est pas un corps (comme $\mathbb{R}$ ou $\mathbb{C}$) mais un anneau. La structure mathématique pertinente ici est celle de module, dont les propriétés sont différentes de celles d'un espace vectoriel. Par exemple, l'existence d'une base n'est pas garantie.

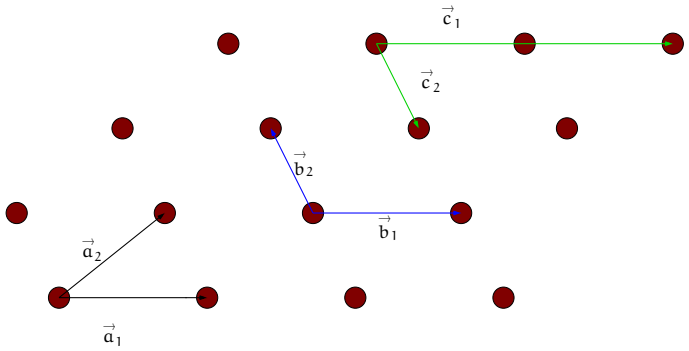

FIGURE 5.5 – Réseau bidimensionnel

Définition 2 (sous-réseau). *Soit $\Lambda \subset \mathbb{R}^d$ un réseau et $\Lambda' \subset \Lambda$ une partie de Λ. Λ' est un sous réseau de Λ s'il constitue lui-même un réseau, c.-à.-d. que la somme d'éléments de Λ' et l'opposé de tout élement de Λ' appartiennent également à Λ'.*

Par exemple, $2\mathbb{Z}$ est un sous réseau de $\mathbb{Z}$ alors que $2\mathbb{Z}+1$ ne l'est pas : la somme de deux nombres impairs étant un nombre pair, elle n'appartient pas à $2\mathbb{Z}+1$.

5.2.2 Réseau dual

À tout réseau peut être associé un *réseau dual*, ou *réseau réciproque*, qui joue un rôle important dans la formulation du théorème de Bloch et, au-delà, concernant les propriétés optiques des cristaux (diffraction des rayons X).

Définition 3 (réseau dual).
Soit $\Lambda \subset \mathbb{R}^d$ un réseau. Le réseau dual, noté $\Lambda^\vee$, est l'ensemble des applications linéaires de Λ à valeurs dans $2\pi\mathbb{Z}$:[2]

$$\Lambda^\vee = \{\phi : \Lambda \subset \mathbb{R}^d \to \mathbb{R} \quad \text{t.q.} \quad \forall \vec{v} \in \Lambda,\ \phi(\vec{v}) \in 2\pi\mathbb{Z}\} \tag{5.15}$$

En munissant $\mathbb{R}^d$ du produit scalaire euclidien usuel, $\langle \vec{u}, \vec{v} \rangle = \vec{u} \cdot \vec{v}$ il est possible de construire un isomorphisme entre $\Lambda^\vee$ et une partie de $\mathbb{R}^d$, formant un réseau que nous noterons également $\Lambda^\vee$ (voir la sous-section 1.1.2 du cours 1) :

$$\boxed{\Lambda^\vee = \{\vec{w} \in \mathbb{R}^d \quad \text{t.q.} \quad \forall \vec{v} \in \Lambda,\ \langle \vec{w}, \vec{v} \rangle \in 2\pi\mathbb{Z}\}} \tag{5.16}$$

$\Lambda^\vee$ peut donc être également considéré comme un groupe, sous-groupe discret de $(\mathbb{R}^d, +)$.

Propriété 1 (base duale).
Soit $\Lambda \subset \mathbb{R}^d$ un réseau, muni d'une base $\mathcal{B} = \{\vec{a}_1, \ldots, \vec{a}_d\}$ associée à la matrice B. Une base du réseau dual $\Lambda^\vee$ est donnée par la base duale de $\mathcal{B}$, notée $\mathcal{B}^\vee$ et définie par :

$$\mathcal{B}^\vee = \left\{\vec{a}_1^\vee, \ldots, \vec{a}_d^\vee \in \mathbb{R}^d \quad \text{t.q.} \quad \langle \vec{a}_i^\vee, \vec{a}_j \rangle = 2\pi\delta_{ij}\right\}, \tag{5.17}$$

ou de manière équivalente associée à la matrice $D = 2\pi(B^{-1})^T$.

2. Dans la littérature mathématique, le facteur de 2π est omis.

Si la matrice B est associée à la base $\mathcal{B}$, nous avons, avec $\{\vec{e}_j, j = 1, \dots, d\}$ la base orthonormée canonique de $\mathbb{R}^n$,

$$\vec{a}_\ell = \sum_{i=1}^{d} \vec{e}_i \, B^i{}_\ell \,. \tag{5.18}$$

De même, si la matrice D est associée à la base $\mathcal{B}^\vee$, nous avons

$$\vec{a}_k^\vee = \sum_{j=1}^{d} \vec{e}_j \, D^j{}_k \,. \tag{5.19}$$

Nous avons donc, en choisissant $D = 2\pi(B^{-1})^T$,

$$\langle \vec{a}_k^\vee, \, \vec{a}_\ell \rangle = \sum_{i,j} \vec{e}_j \cdot \vec{e}_i \, D^j{}_k B^i{}_\ell = 2\pi \sum_j (B^{-1})^k{}_i B^i{}_\ell = 2\pi \delta_{k,\ell} \,, \tag{5.20}$$

En conséquence pour tout $\vec{v} \in \Lambda$ et pour tout $k \in \{1, \dots, d\}$, $\langle \vec{a}_k^\vee, \vec{v} \rangle \in 2\pi\mathbb{Z}$, donc

$$\mathbb{Z}\,\vec{a}_1^\vee + \cdots + \mathbb{Z}\,\vec{a}_d^\vee \subseteq \Lambda^\vee \,, \tag{5.21}$$

c.-à.-d. $2\pi(B^T)^{-1}\mathbb{Z}^d \subseteq \Lambda^\vee$. Réciproquement, pour tout $\vec{w} \in \Lambda^\vee$, nous avons

$$\forall k \in \{1, \dots, d\} \,, \quad \langle \vec{w}, \, \vec{a}_k \rangle \in 2\pi\mathbb{Z} \implies B^T \vec{w} \in 2\pi\mathbb{Z}^d \implies \vec{w} \in 2\pi\,(B^T)^{-1}\mathbb{Z}^d \,, \tag{5.22}$$

donc $\Lambda^\vee \subseteq 2\pi(B^T)^{-1}\mathbb{Z}^d$. □

Propriété 2 (propriétés du réseau dual).
Le réseau dual $\Lambda^\vee$ d'un réseau Λ satisfait les propriétés suivantes :

1. $\left(\Lambda^\vee\right)^\vee = \Lambda$;
2. *pour tout* $\alpha \in \mathbb{R}^\star$, $(\alpha\Lambda)^\vee = \frac{1}{\alpha}\Lambda^\vee$;
3. *Si* $\Lambda \subseteq M$ *alors* $M^\vee \subseteq \Lambda^\vee$.

Pour la première propriété, il suffit de choisir un base B de Λ et de constater que

$$\left(2\pi\big(2\pi(B^{-1})^T\big)^{-1}\right)^T = \left(B^T\right)^T = B \,. \tag{5.23}$$

Pour la seconde, on constate que, si B est une base de Λ, αB est une base de $\alpha\Lambda$, donc la base duale est donnée par $2\pi\big(\alpha B)^{-1}\big)^T = \frac{2\pi}{\alpha}(B^{-1})^T$. Pour la troisième, si $y \in M^\vee$, alors, pour tout $x \in M$, $\langle y, x \rangle \in 2\pi\mathbb{Z}$. Étant donné que $\Lambda \subseteq M$, ceci est vrai en particulier si $x \in \Lambda$, donc $M^\vee$ est contenu dans $\Lambda^\vee$. □

Les propriétés 2. et 3. sont liées. Si Λ est un réseau, $M = 2\Lambda \subset \Lambda$ est un sous-réseau de Λ tel que l'espacement entre les points est doublé. Son réseau dual est donné par $M^\vee = \frac{1}{2}\Lambda^\vee$, dont les points sont deux fois plus rapprochés que dans $\Lambda^\vee$, avec évidemment $\Lambda^\vee \subset M^\vee$.

Exemple 2 (réseaux duaux à deux et trois dimensions).
Soit un réseau à deux dimensions, engendré par une base $\{\vec{a}_1, \vec{a}_2\}$. La base duale $\{\vec{a}_1^\vee, \vec{a}_2^\vee\}$ satisfait les relations :

$$\vec{a}_1^\vee \cdot \vec{a}_1 = \vec{a}_2^\vee \cdot \vec{a}_2 = 2\pi\,, \quad \vec{a}_1^\vee \cdot \vec{a}_2 = \vec{a}_2^\vee \cdot \vec{a}_1 = 0\,, \tag{5.24}$$

que l'on peut résoudre explicitement comme :

$$\vec{a}_1^\vee = \frac{2\pi}{\det^2(\vec{a}_1, \vec{a}_2)}\Big(\vec{a}_2^2\vec{a}_1 - \vec{a}_1\cdot\vec{a}_2\vec{a}_2\Big) \quad , \qquad \vec{a}_2^\vee = \frac{2\pi}{\det^2(\vec{a}_1, \vec{a}_2)}\Big(\vec{a}_1^2\vec{a}_2 - \vec{a}_1\cdot\vec{a}_2\vec{a}_1\Big)\,, \tag{5.25}$$

avec $\det^2(\vec{a}_1, \vec{a}_2) = \vec{a}_1^2\vec{a}_2^2 - (\vec{a}_1\cdot\vec{a}_2)^2$ (à vérifier !).

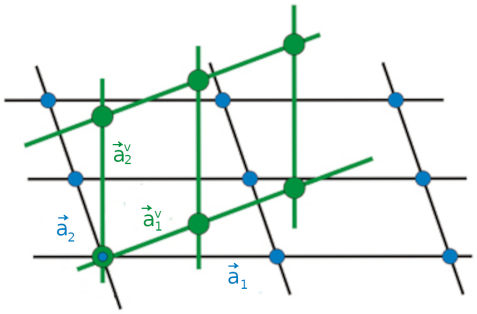

FIGURE 5.6 – Réseau bidimensionnel et son réseau dual

Pour un réseau à trois dimensions, la base duale d'une base $\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}$ s'exprime simplement comme :

$$\boxed{\vec{a}_1^\vee = \frac{2\pi\,\vec{a}_2\wedge\vec{a}_3}{\det(\vec{a}_1,\vec{a}_2,\vec{a}_3)}\,, \quad \vec{a}_2^\vee = \frac{2\pi\,\vec{a}_3\wedge\vec{a}_1}{\det(\vec{a}_1,\vec{a}_2,\vec{a}_3)}\,, \quad \vec{a}_3^\vee = \frac{2\pi\,\vec{a}_1\wedge\vec{a}_2}{\det(\vec{a}_1,\vec{a}_2,\vec{a}_3)}} \tag{5.26}$$

On vérifie en effet que $\vec{a}_1\cdot\vec{a}_1^\vee = 2\pi\,\vec{a}_1\cdot(\vec{a}_2\wedge\vec{a}_3)/\det(\vec{a}_1,\vec{a}_2,\vec{a}_3) = 2\pi$, puis que $\vec{a}_2\cdot\vec{a}_1^\vee = 2\pi\vec{a}_2\cdot(\vec{a}_2\wedge\vec{a}_3)/\det(\vec{a}_1,\vec{a}_2,\vec{a}_3) = 0$, etc.

5.3 Structures cristallines

Un cristal infini se construit par répétition d'une entité élémentaire, atome ou ou groupe d'atomes, appelée *motif*, à laquelle sont appliquées des opérations de symétrie formant un sous-groupe discret du groupe d'isométrie de l'espace ; leur étude aboutit à la classification des différentes structures cristallines possibles.

5.3.1 Groupe ponctuel et groupe d'espace

La structure cristalline d'un solide préserve, comme nous l'avons étudié ci-dessus, un sous-groupe discret du groupe des translation spatiales $\Lambda \subset (\mathbb{R}^d, +)$, associé à la notion mathématique de *réseau*.

Le groupe de symétrie de l'espace euclidien à trois dimensions $E_3 = O(3) \ltimes \mathbb{R}^3$ contient, outre des translations, des rotations propres ou impropres – voir le chapitre 4 – et différents types de cristaux peuvent préserver différents sous-groupes discrets du groupe $O(3)$. Dans le contexte de la cristallographie ce sous-groupe de $O(3)$ qui, par définition, laisse l'origine des coordonnées invariante, est appelé *groupe ponctuel*; nous le noterons P_Λ. Le produit semi-direct du groupe ponctuel avec le réseau Λ, sous-groupe discret du groupe des translation spatiales, forme un sous-groupe discret du groupe euclidien $E(3)$ appelé *groupe d'espace*; nous le noterons $E_\Lambda = P_\Lambda \ltimes \Lambda$.

Théorème 1 (restriction cristallographique). *Le groupe ponctuel P_Λ d'un réseau Λ à deux ou trois dimensions ne peut contenir des rotations que d'ordre 1, 2, 3, 4 ou 6.*

Soit un réseau Λ engendré par une base $\{\vec{e}_a,\, a = 1, \ldots, d\}$, avec $d = 2$ ou $d = 3$. On considère l'action d'une rotation $r \in SO(d)$ autour d'un nœud du réseau donné, laissant ce réseau invariant. Par définition, l'image de tout nœud du réseau doit également faire partie du réseau. En exprimant l'image d'un vecteur $\vec{e}_a$ sous forme matricielle :

$$r(\vec{e}_a) = \sum_b R^a{}_b \vec{e}_b \,, \tag{5.27}$$

et en remarquant que

$$r(\vec{e}_a) \in \Lambda \implies \exists\, n_{ab} \in \mathbb{Z} \quad \text{t.q.} \qquad r(\vec{e}_a) = \sum n_{ab} \vec{e}_b \,, \tag{5.28}$$

voir l'eqn. (5.11), nous en déduisons que la matrice de rotation $R^a{}_b$ ne doit contenir que des coefficients entiers. Cette matrice étant de déterminant unité, il s'agit d'ailleurs d'une matrice unimodulaire associée à un changement de base du réseau, voir le lemme 1.

Dans la base orthonormée associée aux coordonnées cartésiennes, la matrice d'une rotation à deux dimensions et celle d'une rotation à trois dimensions d'axe Oz, toutes deux d'angle θ, s'expriment respectivement comme

$$R_2(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \,, \quad R_3(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \,, \tag{5.29}$$

dont les traces sont respectivement données par

$$\operatorname{Tr} R_2(\theta) = 2\cos\theta \,, \quad \operatorname{Tr} R_3(\theta) = 1 + 2\cos\theta \,. \tag{5.30}$$

La trace étant invariante par changement de base, nous devons obtenir la même expression dans une base engendrant le réseau. Nous obtenons ainsi dans les deux cas, bidimensionnel et tridimensionnel, la condition

$$2\cos\theta \in \mathbb{Z} \quad \implies \quad \cos\theta \in \{-1, -1/2, 0, 1/2, 1\} \tag{5.31}$$

Les valeurs possibles de l'angle de rotation θ correspondantes sont, dans l'intervalle $[0, \pi]$, données par $\theta \in \{2\pi/2, 2\pi/3, 2\pi/4, 2\pi/6, 0\}$, complétant la preuve. □

Exemple 3 (groupes ponctuels bidimensionnels). Classifier les groupes ponctuels possibles dans le plan revient dans un premier temps à classifier les sous-groupes discrets du groupe orthogonal $\mathsf{O}(2)$ qui ne contiennent pas de rotations d'ordres distincts de 2, 3, 4 ou 6.

Une première famille de sous-groupes discrets de $\mathsf{O}(2)$ est donnée par les *groupes cycliques* d'ordre n, notés C_n. Le groupe C_n – qui est en réalité un sous-groupe de $\mathsf{SO}(2)$ – est engendré par une rotation r d'ordre n c.-à.-d. que n est le plus petit entier positif tel que $r^n = e$. On peut donc associer à r une rotation d'angle $2\pi/n$. Ce groupe contient les éléments $\{e, r, r^2, \ldots, r^{n-1}\}$. Les groupes cycliques C_1, C_2, C_3, C_4 et C_6 sont donc des groupes ponctuels a priori possibles pour un réseau bidimensionnel, le premier étant le groupe trivial.

La deuxième famille de sous-groupes discrets de $\mathsf{O}(2)$ est constituée des *groupes dihédraux*. Le groupe dihédral D_n, d'ordre $2n$, peut être vu comme le groupe de symétrie d'un polygône régulier à n côtés (triangle équilatéral, carré, etc.). Le groupe D_n est engendré par deux éléments, une rotation r d'ordre n (avec, comme précédemment, $r^n = e$), et une symétrie axiale a telle que $a^2 = e$ et $r \cdot a \cdot r \cdot a = e$.[3] Le groupe D_n contient donc les $2n$ éléments $\{e, r, \ldots r^{n-1}, a, r \cdot a, r^2 \cdot a, \ldots, r^{n-1} \cdot a\}$. Les groupes dihédraux D_1, D_2, D_3, D_4 et D_6 sont donc des groupes ponctuels a priori possibles pour un réseau bidimensionnel.

En réalité, tous ces sous-groupes discrets de $\mathsf{O}(2)$ ne sont pas réalisés comme groupes ponctuels de réseaux bidimensionnels. Premièrement à deux dimensions l'inversion spatiale correspond à une rotation d'angle π. Nous avons ainsi, pour tout élément $\vec{x}$ d'un réseau,

$$\vec{x} = n_1\vec{a}_1 + n_2\vec{a}_2 \in \Lambda \quad \xrightarrow{r(\pi)} \quad -\vec{x} = (-n_1)\vec{a}_1 + (-n_2)\vec{a}_2 \in \Lambda\,, \tag{5.32}$$

et le groupe ponctuel de tout réseau contient au moins une rotation d'ordre 2, excluant C_1, C_3, D_1 et D_3. Deuxièmement, si le groupe ponctuel contient une rotation d'angle $2\pi/4$, le réseau admet nécessairement une base orthogonale telle que les deux vecteurs de base $\vec{a}_1$ et $\vec{a}_2$ ont même longueur. Le réseau admet donc un axe de symétrie contenant $\vec{a}_1$, agissant sur le réseau suivant

$$\vec{x} = n_1\vec{a}_1 + n_2\vec{a}_2 \in \Lambda \quad \xrightarrow{a} \quad n_1\vec{a}_1 - n_2\vec{a}_2 \in \Lambda\,, \tag{5.33}$$

ainsi que les trois autres axes obtenus par rotation de celui-ci ; le groupe C_4 est ainsi exclu. Par un raisonnement analogue, on observe que le groupe C_6 est aussi exclu, car un réseau possédant cette symétrie possède nécessairement 6 axes de symétries également. *Au final, un réseau bidimensionnel admet l'un des quatre groupes ponctuels suivants :* C_2, D_2, D_4 *ou* D_6.

Une étude analogue pour des réseaux tridimensionnels donne sept groupes ponctuels possibles, voir par exemple [3].

Propriété 3 (groupe ponctuel du réseau dual). *Soit Λ un réseau et $\Lambda^\vee$ son réseau dual. Si P_Λ est le groupe ponctuel associé au réseau Λ, P_Λ est également le groupe ponctuel associé au réseau dual.*

3. On peut comprendre cette dernière propriété en constant que $r \cdot a$ est une symétrie axiale obtenue à partir de a par une rotation de l'axe de symétrie d'un angle $2\pi/n$.

En effet, le groupe ponctuel est, par définition, un sous-groupe discret du groupe orthogonal $O(d)$, qui est l'ensemble des transformations linéaires laissant invariant le produit scalaire entre deux vecteurs de $\mathbb{R}^d$. La condition définissant les vecteurs du réseau dual, voir l'éqn. (5.16), est donc invariante sous une action simultanée des éléments de P_Λ sur les vecteurs de Λ et sur les vecteurs de $\Lambda^\vee$. □

5.3.2 Domaine fondamental et mailles

Un cristal infini est décrit par un choix de *maille* et de groupe de translations qui lui sont appliquées. Suivant le choix de maille, il s'agit de tout le réseau de Bravais Λ ou de l'un de ses sous-réseaux.

Définition 4 (domaine fondamental).
Soit $\Lambda \subset \mathbb{R}^d$ un réseau de rang maximal et $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_d\}$ une base de Λ. Un domaine fondamental du réseau est le parallélépipède défini par :[4]

$$\mathcal{F}_\mathcal{B}(\Lambda) \overset{\text{déf.}}{=} \{t_1\vec{b}_1 + t_2\vec{b}_2 + \cdots + t_d\vec{b}_d \ , \quad 0 \leqslant t_i \leqslant 1\}. \tag{5.34}$$

Chaque choix de base définit un choix de domaine fondamental. Par translation d'un domaine fondamental selon le réseau, on couvre entièrement $\mathbb{R}^d$, sans recouvrement. Dit autrement, nous obtenons une *partition* de $\mathbb{R}^d$:

$$\mathbb{R}^d = \left\{\mathcal{F}_\mathcal{B} + \lambda \ , \quad \lambda \in \Lambda\right\}. \tag{5.35}$$

Dans le contexte de la cristallographie, le domaine fondamental est une *maille élémentaire* du réseau de Bravais, c.-à.-d. une maille ne contenant qu'un seul nœud.

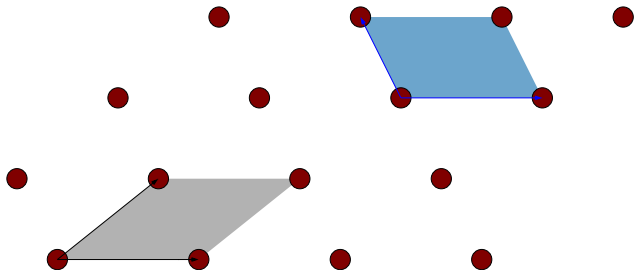

FIGURE 5.7 – Volume fondamental d'un réseau calculé dans deux bases différentes.

Propriété 4 (volume du domaine fondamental).
Soit $\Lambda \subset \mathbb{R}^d$ un réseau de rang maximal et $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_d\}$ une base de Λ. Le volume d'un domaine fondamental ou déterminant du réseau est donné par :

$$\mathcal{V}(\Lambda) \overset{\text{déf.}}{=} \left|\det\left(\vec{b}_1, \ldots, \vec{b}_d\right)\right|. \tag{5.36}$$

4. En dimension quelconque, il faudrait plutôt parler de parallélotope.

Comme la notation le suggère, il ne dépend pas de la base choisie. Soit B la matrice associée à la base $\mathcal{B}$ et $\hat{\mathsf{B}}$ la matrice associée à une autre base $\hat{\mathcal{B}}$ du même réseau. D'après le lemme 1 nous savons qu'il existe une matrice U unimodulaire telle que $\mathsf{B} = \hat{\mathsf{B}} \cdot \mathsf{U}$. Nous avons alors $\det\left(\vec{b}_1, \ldots, \vec{b}_d\right) = \det \mathsf{B} = \det(\hat{\mathsf{B}} \cdot \mathsf{U}) = \pm \det \hat{\mathsf{B}}$ en utilisant $\det \mathsf{U} = \pm 1$. □

Un inconvénient de la description d'un réseau de Bravais à partir d'une maille élémentaire est que le groupe de symétrie de celle-ci – constitué des opérations de symétrie (rotations, réflexions, etc.) laissant la maille invariante – ne correspond pas nécessairement au groupe ponctuel du réseau, dans le cas où l'image d'un nœud de la maille élémentaire par un élément du groupe ponctuel appartient à une maille voisine. Dans la classification des cristaux, il est fait dans ces cas usage d'une *maille conventionnelle*, non élémentaire, dont le groupe de symétrie coïncide avec la symétrie ponctuelle du réseau, voir les figures 5.8 (deuxième colonne) et 5.9 ci-dessous.

5.3.3 Systèmes cristallins et classification des réseaux de Bravais

Pour classifier les réseaux de Bravais on les regroupe en *systèmes cristallins*, dont les éléments possèdent donc le même groupe ponctuel ou, dit autrement, la même maille conventionnelle.

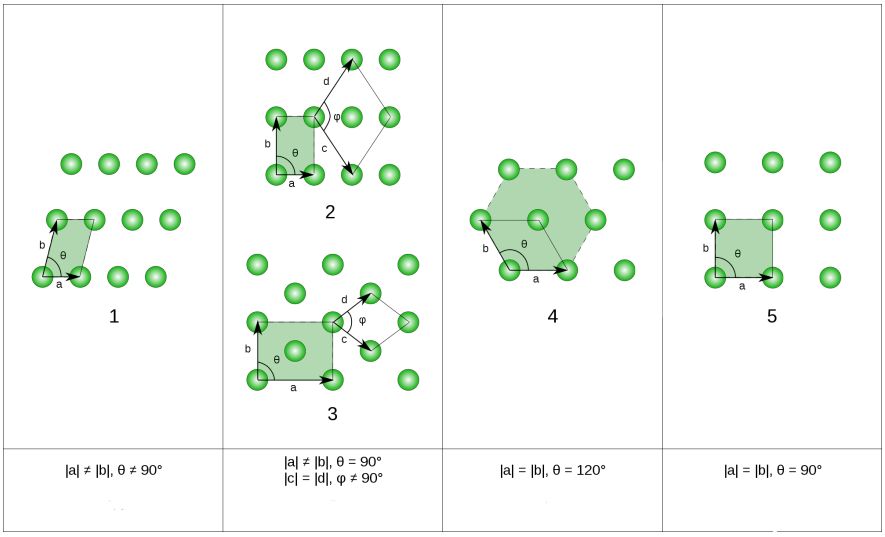

FIGURE 5.8 – Réseaux de Bravais bidimensionnels. *Source : Wikicommons.*

À deux dimensions, nous avons observé qu'il existe quatre groupes ponctuels possibles, donc quatre systèmes cristallins, voir la fig. 5.8, regroupant cinq types de réseaux :

1. le système oblique, correspondant au plus petit groupe ponctuel possible, C_2. Il n'existe qu'un seul réseau de ce type, le réseau oblique ;
2. le système rectangulaire, correspondant au groupe ponctuel D_2. Il existe deux réseaux de Bravais correspondant, le réseau rectangulaire et le réseau rectangulaire centré ;

3. le système hexagonal, correspondant au groupe ponctuel D_6. Il n'existe qu'un seul réseau de ce type, le réseau hexagonal ;
4. le système carré, correspondant au groupe ponctuel D_4. Il n'existe qu'un seul réseau de ce type, le réseau carré ;

À trois dimensions, on distingue sept systèmes cristallins, associés aux groupes ponctuels possibles, et quatorze réseaux de Bravais. Par exemple le système cubique est associé au groupe ponctuel O_h, groupe de symétrie du cube. Ce système regroupe les réseaux cubique, cubique centré et cubique faces centrées, dont les mailles conventionnelles sont illustrées par la fig. 5.9.

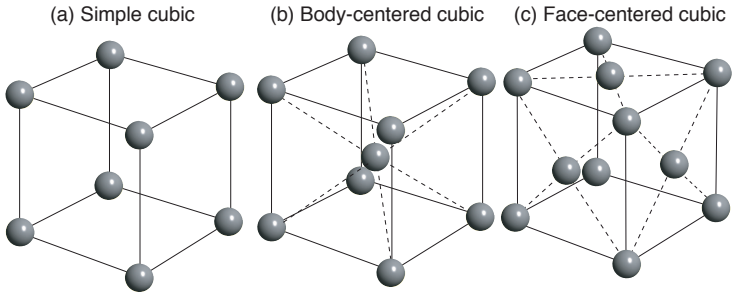

FIGURE 5.9 – Réseaux de Bravais cubiques. *Source : Hofmann [3]*

5.3.4 Maille de Wigner–Seitz et première zone de Brillouin

Pour tout réseau il existe un type de maille élémentaire qui présente l'avantage d'avoir comme groupe de symétrie le groupe ponctuel du réseau mais n'a pas nécessairement la géométrie d'un parallélotope.

Définition 5 (maille de Wigner–Seitz).
Soit $\Lambda \subset (\mathbb{R}^\mathrm{d}, +)$ un réseau de Bravais. La maille de Wigner–Seitz associée, attachée à un nœud donné du réseau, est le lieu géométrique des points de $\mathbb{R}^\mathrm{d}$ plus proches de ce nœud que des autres nœuds du réseau.

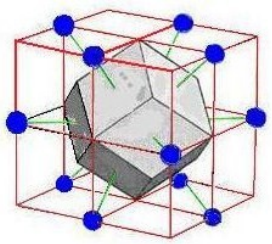

FIGURE 5.10 – Maille de Wigner–Seitz d'un réseau cubique centré. *Source : Wikicommons*

En deux dimensions (resp. trois dimensions) la maille de Wigner–Seitz est donc un polygône (resp. un polyèdre) dont les côtés (resp. les faces) correspondent aux médiatrices (resp. aux plans médiateurs) des segments joignant le nœud choisi avec les nœuds voisins du réseau, voir un exemple sur la fig. 5.10. Par construction, elle est une maille élémentaire car elle engendre une partition de l'espace par translation.

Si on considère un nœud donné du réseau, un élément du groupe ponctuel du cristal en ce point (comme une rotation autour de ce nœud) va transformer un nœud voisin en un nœud voisin, et agira donc comme une certaine permutation entre les plans médiateurs définissant la maille de Wigner–Seitz ; cette dernière sera donc invariante par la transformation considérée ; ainsi *le groupe ponctuel du réseau laisse la maille de Wigner–Seitz invariante.*

Dans la pratique, la notion de maille de Wigner–Seitz d'un réseau est plutôt utilisée pour le réseau dual $\Lambda^\vee$ que pour le réseau Λ d'un cristal lui-même. On définit ainsi

Définition 6 (première zone de Brillouin).
Soit Λ un réseau de Bravais et $\Lambda^\vee$ son réseau dual. La première zone de Brillouin est la maille de Wigner–Seitz du réseau dual $\Lambda^\vee$.

Comme nous l'avions montré, le réseau dual $\Lambda^\vee$ possède le même groupe ponctuel que le réseau Λ, et la première zone de Brillouin est ainsi invariante sous l'action du groupe ponctuel P_Λ du réseau de Bravais considéré. Son utilité sera manifeste à la section suivante.

5.4 Théorème de Bloch

Dans la section 5.1 nous avions formulé le théorème de Bloch à une dimension en cherchant la forme générale des vecteurs de l'espace de Hilbert se transformant dans une certaine représentation irreductible du groupe des translations discrètes le long de la chaine d'atomes.

Nous considérons ici un système invariant par translation discrète selon un réseau Λ. Concrètement, ce peut être un électron dans un solide soumis au potentiel périodique $\mathcal{V}(\vec{x})$ créé par les noyaux. Étant donné que

$$\forall \vec{\lambda} \in \Lambda\,, \quad \mathcal{T}_{\vec{\lambda}}^\dagger\, \mathcal{V}(\vec{x})\, \mathcal{T}_{\vec{\lambda}} = \mathcal{V}(\vec{x}+\vec{\lambda}) = \mathcal{V}(\vec{x})\,, \tag{5.37}$$

le hamiltonien $\mathcal{H}$ du système satisfait :

$$\forall \vec{\lambda} \in \Lambda\,, \quad \big[\, \mathcal{T}_{\vec{\lambda}},\, \mathcal{H}\,\big] = 0\,. \tag{5.38}$$

Suivant les considérations générales sur les symétries en mécanique quantique (voir en particulier la propriété exposée dans la sous-section 3.4.2 du chapitre 3) l'action du groupe sur l'espace des états $\mathcal{H}$ correspond à une représentation D unitaire de Λ, d'après le théorème de Weyl. Une telle représentation est complètement réductible, comme nous l'avons vu.

Cette décomposition permet de construire une base orthonormée de vecteurs propres communs au hamiltonien $\mathcal{H}$ et aux opérateurs de translation $\mathcal{T}_{\vec{\lambda}}$. Les fonctions d'onde associées satisfont :

$$\forall \vec{x} \in \mathbb{R}^{\mathrm{d}}\,, \quad \forall \vec{\lambda} \in \Lambda\,, \quad \psi(\vec{x}-\vec{\lambda}) = \mathsf{D}(\vec{\lambda})\psi(\vec{x})\,, \tag{5.39}$$

où D est une représentation unitaire irreductible de Λ, vu comme un groupe. Les représentations unitaires irreductibles complexes sont de dimension un ; on peut en effet obtenir simplement ces représentations à partir de celles de $(\mathbb{Z}^d, +)$, en rappelant que $\Lambda = B\mathbb{Z}^d$, c.-à.-d. qu'il existe un isomorphisme de groupe entre tout réseau (de rang maximal) et le groupe $(\mathbb{Z}^d, +)$. Les représentations unitaires irreductibles de $(\mathbb{Z}^d, +)$ sont une généralisation directe de celles de $(\mathbb{Z}, +)$, paramétrées par un vecteur $\vec{\theta}$ de composantes $\theta_i \in [0, 2\pi[$, et données par

$$\forall \vec{n} \in \mathbb{Z}^d \ , \quad D^{\vec{\theta}}(\vec{n}) = e^{i(n_1\theta_1 + \cdots + n_d\theta_d)} = e^{i\vec{n}\cdot\vec{\theta}} \, . \tag{5.40}$$

La périodicité du membre de droite de l'éqn. (5.40) dans les variables θ_i signifie que $\vec{\theta}$ est défini à un vecteur du réseau $2\pi\mathbb{Z}^d$ près.

Pour un réseau arbitraire de rang maximal $\Lambda = B\mathbb{Z}^d$, c.-à.-d. associé à une matrice B, nous pouvons écrire :

$$\vec{n}\cdot\vec{\theta} = \langle \vec{n}, \vec{\theta} \rangle = \langle B^{-1}\cdot B\cdot\vec{n}, \vec{\theta} \rangle = \langle \underbrace{B\cdot\vec{n}}_{\in\Lambda} , \underbrace{(B^{-1})^T\vec{\theta}}_{\in\mathbb{R}^d/\Lambda^\vee} \rangle \, . \tag{5.41}$$

On remarque en effet que la translation $\vec{\theta} \mapsto \vec{\theta} + 2\pi(n_1\vec{e}_1 + \cdots + n_d\vec{e}_d)$, reflétant la périodicité des « angles » θ_i, engendre la translation du second facteur par un vecteur du réseau dual, associé rappelons-le à la matrice $2\pi(B^{-1})^T$.

Ainsi les représentations irreductibles de Λ sont paramétrées par un vecteur $\vec{k} \in \mathbb{R}^d/\Lambda^\vee$, de telle sorte que :

$$\forall \vec{\lambda} \in \Lambda \ , \quad D^{\vec{k}}(\vec{\lambda}) = e^{i\vec{k}\cdot\vec{\lambda}} \, . \tag{5.42}$$

Les opérateurs de translation sur le réseau s'obtiennent, comme dans le cas unidimensionnel, en considérant $(\Lambda, +)$ comme un sous-groupe de $(\mathbb{R}^d, +)$. Pour tout $\vec{\lambda} \in \Lambda$, nous avons :

$$\mathcal{T}_{\vec{\lambda}} = e^{-\frac{i\vec{\lambda}\cdot\vec{p}}{\hbar}} \, . \tag{5.43}$$

Les vecteurs propres de l'opérateur de translation se transformant dans la représentation $D^{\vec{k}}$ doivent satisfaire :

$$\forall \vec{\lambda} \in \Lambda \ , \quad \mathcal{T}_{\vec{\lambda}} \left|\psi_{\vec{k}}\right\rangle = e^{i\vec{k}\cdot\vec{\lambda}} \left|\psi_{\vec{k}}\right\rangle \, , \tag{5.44}$$

soit, en représentation position,

$$\forall \vec{x} \in \mathbb{R}^d \ , \quad \forall \vec{\lambda} \in \Lambda \ , \quad \psi_{\vec{k}}(\vec{x} - \vec{\lambda}) = e^{i\vec{k}\cdot\vec{\lambda}}\psi_{\vec{k}}(\vec{x}) \, . \tag{5.45}$$

Les solutions de cette équation fonctionnelle sont données en termes d'une fonction $u_{\vec{k}}$ périodique sur le réseau Λ :

$$\boxed{\psi_{\vec{k}}(\vec{x}) = e^{-i\vec{k}\cdot\vec{x}} u_{\vec{k}}(\vec{x}) \ , \quad \text{t.q.} \quad \forall \vec{\lambda} \in \Lambda \ , \quad u_{\vec{k}}(\vec{x} + \vec{\lambda}) = u_{\vec{k}}(\vec{x})} \tag{5.46}$$

Dans cette formulation générale du théorème de Bloch sur un réseau Λ, les fonctions d'onde ou *ondes de Bloch* sont donc caractérisées par un vecteur d'onde $\vec{k} \in \mathbb{R}^d/\Lambda^\vee$, c.-à.-d. défini à un vecteur du réseau réciproque $\Lambda^\vee$ près. Il est ainsi possible, pour l'analyse des états propres du hamiltonien du système, de se restreindre à une maille du réseau réciproque. On pourra toujours choisir comme maille la première zone de Brillouin qui présente l'avantage, comme nous l'avons remarqué, de respecter les symétries ponctuelles du système.

Remarque 3 (cristal de taille finie). *Comme dans le cas unidimensionel, voir la remarque 1, on choisira pour un cristal de taille finie des conditions aux limites périodiques pour préserver la symétrie de translation associée au réseau. Si* $\{\vec{a}_1, \ldots, \vec{a}_3\}$ *est une base du réseau il sera également préférable, pour préserver également le groupe ponctuel, de choisir, pour un cristal constitué de* N^3 *atomes, les conditions suivantes :*

$$\forall \vec{x} \in \mathbb{R}^3 \ , \quad \psi_{\vec{k}}(\vec{x} + N\vec{a}_1) = \psi_{\vec{k}}(\vec{x} + N\vec{a}_2) = \psi_{\vec{k}}(\vec{x} + N\vec{a}_3) \, , \tag{5.47}$$

de telles sorte que

$$e^{iN\vec{k}\cdot\vec{a}_\ell} = 1 \ , \ell \in \{1, 2, 3\} \implies \vec{k} \in \tfrac{1}{N} \Lambda^\vee \, . \tag{5.48}$$

Le réseau $\frac{1}{N}\Lambda^\vee$ *possède comme le réseau dual* $\Lambda^\vee$ *le même groupe ponctuel que le réseau* Λ*.*

Bibliographie

[1] W. Friedrich, P. Knipping, et M. Laue, "Interferenzerscheinungen bei Röntgenstrahlen," Annalen der Physik **346**, 10 (1913).

[2] G. Binning, H. Roher, Ch. Gerber et E. Weibel, "Surface Studies by Scanning Tunneling Microscopy," Phys. Rev. Lett. **49**, 1 (1982).

[3] P. Hofmann, "Solid State Physics : An Introduction, 2nd Edition," John Wiley and Sons, 2015

Cours n°6

Produit tensoriel et particules identiques

L'objectif de ce cours est de présenter, de manière approfondie, les outils permettant de traiter de systèmes quantiques composés de plusieurs sous-systèmes qui sont, soit indépendants, soit faiblement couplés. Un cas particulièrement important est celui d'un système composé de plusieurs particules identiques, pour lequel des effets quantiques extrêmement spectaculaires vont se manifester, conéquence, là encore, de l'apparition d'une symétrie dans le problème.

6.1 Produit tensoriel

La notion de produit tensoriel permet de décrire l'espace des états d'un système qui se scinde en deux sous systèmes, ou en deux jeux de degrés de liberté.

On peut considérer l'exemple de deux particules séparées spatialement, l'une dont l'évolution est régie par un hamiltonien $\mathcal{H}_1$, et l'autre particule un hamiltonien $\mathcal{H}_2$; chaque particule possédera son espace des états :

— l'espace de Hilbert $\mathcal{H}_1$, dont une base est $\{|\ell_1\rangle\,,\ \ell_1 \in \mathcal{I}_1\}$, décrit les états possibles de la première particule ;

— l'espace de Hilbert $\mathcal{H}_2$, dont une base est $\{|\ell_2\rangle\,,\ \ell_2 \in \mathcal{I}_2\}$, décrit les états possibles de la seconde particule ;

Si les deux particules interagissent faiblement, on peut supposer que les vecteurs de base de $\mathcal{H}_1$ d'une part et les vecteurs de base de $\mathcal{H}_2$ d'autre part continuent de décrire tous les états possibles de chacune des particules, mais néanmoins l'étude de leur interaction nécessite une description combinée de l'espace des états des deux particules sous la forme d'un unique espace des états plus grand qui peut se concevoir comme une forme de « produit » de $\mathcal{H}_1$ avec $\mathcal{H}_2$. Cet espace que nous noterons $\mathcal{H}$, doit avoir des propriétés compatibles avec les axiomes de la mécanique quantique, et doit donc avoir lui-même la structure d'un espace de Hilbert.

6.1.1 Construction du produit tensoriel

La structure pertinente pour construire un espace de Hilbert à partir du produit de deux espaces de Hilbert et celle du *produit tensoriel*. On considère deux espaces de Hilbert $\mathcal{H}_1$ et $\mathcal{H}_2$, qui peuvent être associés, dans un contexte physique, aux états d'une particule et d'une autre. On souhaite construire un espace des états, noté $\mathcal{H}_1 \otimes \mathcal{H}_2$, permettant de décrire les états du système constitué de ces deux particules. Pour satisfaire aux axiomes de la mécanique quantique, cet espace doit évidemment être lui-même un espace de Hilbert.

Une contrainte supplémentaire provient de l'action des opérateurs, comme les observables, sur l'espace $\mathcal{H}_1 \otimes \mathcal{H}_2$ ainsi engendré. Comme pour tout espace des états, ces opérateurs doivent être linéaires, alors que dans la description d'origine du système, les opérateurs sont d'une part linéaires sur $\mathcal{H}_1$ et d'autre part linéaires sur $\mathcal{H}_2$, et donc *bilinéaires* sur le produit cartésien $\mathcal{H}_1 \times \mathcal{H}_2$.

Propriété 1 (propriété universelle du produit tensoriel). *Soit* $\mathfrak{f}$ *une application bilinéaire sur* $\mathcal{H}_1 \times \mathcal{H}_2$ *et* $\mathcal{H}_1 \otimes \mathcal{H}_2$ *le produit tensoriel de ces deux espaces de Hilbert. Il existe une unique application linéaire* $\tilde{\mathfrak{f}}$ *sur* $\mathcal{H}_1 \otimes \mathcal{H}_2$ *telle que, pour tout* $(|\nu_1\rangle, |\nu_2\rangle) \in \mathcal{H}_1 \times \mathcal{H}_2$, $\tilde{\mathfrak{f}}\big(|\nu_1\rangle \otimes |\nu_2\rangle\big) = \mathfrak{f}(|\nu_1\rangle, |\nu_2\rangle)$.

Cette propriété définit essentiellement le produit tensoriel de manière unique. Construisons tout d'abord l'espace vectoriel associé au produit cartésien $\mathcal{H}_1 \times \mathcal{H}_2$. Si $\{|\ell_1\rangle, \ell_1 \in \mathcal{I}_1\}$ et $\{|\ell_2\rangle, \ell_2 \in \mathcal{I}_2\}$ sont des bases respectivement de $\mathcal{H}_1$ et $\mathcal{H}_2$, il s'agit de l'ensemble des combinaisons linéaires des éléments de la forme

$$\sum_{\ell_1,\ell_2} c_{\ell_1,\ell_2}\big(|\ell_1\rangle, |\ell_2\rangle\big) \ , \quad c_{\ell_1,\ell_2} \in \mathbb{C} \, .$$

Définition 1 (produit tensoriel). *Soit* $\mathcal{H}_1$ *et* $\mathcal{H}_2$ *deux espaces de Hilbert. On construit le produit tensoriel* $\mathcal{H}_1 \otimes \mathcal{H}_2$ *en identifiant les couples ordonnés* $(|\alpha_1\rangle, |\alpha_2\rangle)$ *du produit Cartésien* $\mathcal{H}_1 \times \mathcal{H}_2$ *selon la relation d'équivalence :*

- $(\lambda|\alpha_1\rangle, |\alpha_2\rangle) \sim \lambda(|\alpha_1\rangle, |\alpha_2\rangle)$ *et* $(|\alpha_1\rangle, \lambda|\alpha_2\rangle) \sim \lambda(|\alpha_1\rangle, |\alpha_2\rangle)$ *;*
- $(|\alpha_1\rangle + |\beta_1\rangle, |\alpha_2\rangle) \sim (|\alpha_1\rangle, |\alpha_2\rangle) + (|\beta_1\rangle, |\alpha_2\rangle)$ *et* $(|\alpha_1\rangle, |\alpha_2\rangle + |\beta_2\rangle) \sim (|\alpha_1\rangle, |\alpha_2\rangle) + (|\alpha_1\rangle, |\beta_2\rangle)$.

Les classes d'équivalence ainsi obtenues sont notées $|\alpha_1\rangle \otimes |\alpha_2\rangle$.

Exprimé autrement, on considère le sous-espace de Hilbert $\mathcal{N}$ de $\mathcal{H}_1 \times \mathcal{H}_2$, engendré par tous les vecteurs de la forme $(\lambda|\alpha_1\rangle, |\alpha_2\rangle) - \lambda(|\alpha_1\rangle, |\alpha_2\rangle)$, $(|\alpha_1\rangle + |\beta_1\rangle, |\alpha_2\rangle) - (|\alpha_1\rangle, |\alpha_2\rangle) + (|\beta_1\rangle, |\alpha_2\rangle)$, $(|\alpha_1\rangle + |\beta_1\rangle, |\alpha_2\rangle) - (|\alpha_1\rangle, |\alpha_2\rangle) - (|\beta_1\rangle, |\alpha_2\rangle)$ et $(|\alpha_1\rangle, |\alpha_2\rangle + |\beta_2\rangle) - (|\alpha_1\rangle, |\alpha_2\rangle) - (|\alpha_1\rangle, |\beta_2\rangle)$.

On identifie alors tous les vecteurs de $\mathcal{H}_1 \times \mathcal{H}_2$ qui diffèrent par un vecteur de $\mathcal{N}$:

$$(|\nu_1\rangle, |\nu_2\rangle) \sim (|w_1\rangle, |w_2\rangle) \ \Leftrightarrow \ \exists |\nu\rangle \in \mathcal{N} \text{ t.q.} \quad (|\nu_1\rangle, |\nu_2\rangle) - (|w_1\rangle, |w_2\rangle) = |\nu\rangle \ . \tag{6.1}$$

On peut ainsi comprendre le produit tensoriel $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ comme le quotient

$$\mathcal{H} = \big(\mathcal{H}_1 \times \mathcal{H}_2\big)/\mathcal{N} \, . \tag{6.2}$$

La propriété universelle énoncée ci-dessus est bien vérifiée par le produit tensoriel ainsi défini. Soit $\mathfrak{f} : \mathcal{H}_1 \times \mathcal{H}_2 \to \mathcal{H}_3$ une application bilinéaire. On définit alors une application linéaire $\tilde{\mathfrak{f}}$ sur l'ensemble des combinaisons linéaires des éléments de $\mathcal{H}_1 \times \mathcal{H}_2$ par :

$$\tilde{\mathfrak{f}}\big(\lambda(|v\rangle_1, |v_2\rangle) + \mu(|w\rangle_1, |w_2\rangle) + \cdots\big) = \lambda \mathfrak{f}(|v\rangle_1, |v_2\rangle) + \mu \mathfrak{f}(|w\rangle_1, |w_2\rangle) + \cdots \tag{6.3}$$

On remarque, par bilinéarité de $\mathfrak{f}$, que

$$\tilde{\mathfrak{f}}\big((\lambda |v_1\rangle, |v_2\rangle) - \lambda(|v_1\rangle, |v_2\rangle) = \mathfrak{f}(\lambda |v_1\rangle, |v_2\rangle) - \lambda \mathfrak{f}(|v_1\rangle, |v_2\rangle) = 0_{\mathcal{H}_3}, \tag{6.4}$$

et il en est de même pour tous les vecteurs appartenant au sous-espace $\mathcal{N}$ défini plus haut. Ainsi, deux vecteurs de $\mathcal{H}_1 \times \mathcal{H}_2$ qui diffèrent par un vecteur de $\mathcal{N}$ ont la même image par $\tilde{\mathfrak{f}}$, qui est donc bien une application linéaire sur le produit tensoriel $\mathcal{H}_1 \otimes \mathcal{H}_2$. □

De manière plus élémentaire, la construction du produit tensoriel implique que le produit tensoriel de deux vecteurs est linéaire par rapport à chacun de ses constituants, c.-à.-d. bilinéaire :

$$(\lambda |\alpha_1\rangle + \mu |\beta_1\rangle) \otimes |\alpha_2\rangle = \lambda |\alpha_1\rangle \otimes |\alpha_2\rangle + \mu |\beta_1\rangle \otimes |\alpha_2\rangle \tag{6.5a}$$

$$|\alpha_1\rangle \otimes (\lambda |\alpha_2\rangle + \mu |\beta_2\rangle) \otimes = \lambda |\alpha_1\rangle \otimes |\alpha_2\rangle + \mu |\alpha_1\rangle \otimes |\beta_2\rangle \tag{6.5b}$$

il est néanmoins lui-même un espace vectoriel, c.-à.-d. que

$$\forall |\psi\rangle, |\psi'\rangle \in \mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 , \quad \forall \lambda, \mu \in \mathbb{C} , \quad \lambda |\psi\rangle + \mu |\psi'\rangle \in \mathcal{H} . \tag{6.6}$$

Il est important de noter que l'ordre des facteurs dans le produit tensoriel est déterminant : si $|\alpha_1\rangle \otimes |\alpha_2\rangle$ est un vecteur de $\mathcal{H}$, $|\alpha_2\rangle \otimes |\alpha_1\rangle$ n'a pas nécessairement de sens car les espaces de Hilbert $\mathcal{H}_1$ et $\mathcal{H}_2$ sont a priori distincts ; s'ils ne le sont pas $|\alpha_1\rangle \otimes |\alpha_2\rangle$ a un sens mais représente un autre vecteur que $|\alpha_2\rangle \otimes |\alpha_1\rangle$ si $\alpha_1 \neq \alpha_2$.

Le produit tensoriel de deux espaces de Hilbert $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ étant un espace de Hilbert, il admet un produit hermitien $\iota_{\mathcal{H}}$. Il se déduit naturellement de ceux sur $\mathcal{H}_1$ et $\mathcal{H}_2$ (respectivement $\iota_{\mathcal{H}_1}$ et $\iota_{\mathcal{H}_2}$) par :

$$\boxed{\iota_{\mathcal{H}} (|\alpha_1\rangle \otimes |\alpha_2\rangle , |\beta_1\rangle \otimes |\beta_2\rangle) \overset{\text{déf.}}{=} \iota_{\mathcal{H}_1} (|\alpha_1\rangle , |\beta_1\rangle) \, \iota_{\mathcal{H}_2} (|\alpha_2\rangle , |\beta_2\rangle)} \tag{6.7}$$

À noter que si $|\alpha_1\rangle$ et $|\alpha_2\rangle$ sont normalisés, leur produit tensoriel l'est également.

Il ne faut surtout pas confondre les notions de produit tensoriel d'espaces de Hilbert et de somme directe de sous-espace de Hilbert. Physiquement, dans le premier cas, on considère les états accessibles à deux types de degrés de liberté (par exemple, deux particules) et dans le second cas deux familles d'états accessibles aux mêmes degrés de liberté, par exemple des sous-espace propres du hamiltonien associés à des énergies différentes.

6.1.2 Espace dual d'un espace produit tensoriel

Nous avons vu dans le cours n° 1 comment associer à tout vecteur d'un espace de Hilbert $\mathcal{H}$ une forme linéaire, élément de son espace dual $\mathcal{H}^\star$. De la même manière que nous avons défini l'espace de Hilbert $\mathcal{H}$ produit tensoriel de deux espaces de Hilbert $\mathcal{H}_1$ et $\mathcal{H}_2$, nous pouvons définir son espace dual $\mathcal{H}^\star$ comme produit tensoriel des espaces duaux $\mathcal{H}_1^\star$ et $\mathcal{H}_2^\star$. Pour un vecteur de $\mathcal{H}$ pouvant s'écrire sous forme produit $|\alpha_1\rangle \otimes |\alpha_2\rangle$, nous avons la correspondance :

$$|\alpha_1\rangle \otimes |\alpha_2\rangle \in \mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \;\longleftrightarrow\; \langle\alpha_2| \otimes \langle\alpha_1| \in \mathcal{H}^\star = \mathcal{H}_2^\star \otimes \mathcal{H}_1^\star \,. \tag{6.8}$$

On note ici que, passant du vecteur $|\alpha_1\rangle \otimes |\alpha_2\rangle$ à son dual $\langle\alpha_2| \otimes \langle\alpha_1|$, l'ordre des facteurs a été inversé. Cette convention est choisie pour écrire le produit hermitien sur l'espace produit tensoriel (6.7) en notation « brakets » de manière naturelle :

$$\big(\,\langle\alpha_2| \otimes \langle\alpha_1|\,\big)\big(\,|\beta_1\rangle \otimes |\beta\rangle_2\,\big) = \langle\alpha_1|\beta_1\rangle\ \langle\alpha_2|\beta_2\rangle \ , \tag{6.9}$$

où on commence par contracter le bra et le ket situés « à l'intérieur » avant de contracter le bar et le ket situés à « l'extérieur ».

6.1.3 Base d'un espace produit tensoriel

Pour construire une base orthonormée d'un espace de Hilbert obtenu par produit tensoriel, nous pouvons évoquer la propriété suivante :

Lemme 1 (base orthonormée d'un espace produit tensoriel). *Soit $\mathcal{H}_1$ (resp. $\mathcal{H}_2$) un espace de Hilbert muni d'une base orthonormée $\{|\ell_1\rangle\,,\ \ell_1 \in \mathcal{I}_1\}$ (resp. $\{|\ell_2\rangle\,,\ \ell_2 \in \mathcal{I}_2\}$). Une base orthonormée de l'espace de Hilbert produit tensoriel $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ est donnée par*

$$\boxed{\big\{\,|\ell_1\rangle \otimes |\ell_2\rangle\,,\ \ell_1 \in \mathcal{I}_1,\ ,\ell_2 \in \mathcal{I}_2\big\}} \tag{6.10}$$

Si $\mathcal{H}_1$ (resp. $\mathcal{H}_2$) est de dimension finie d_1 (resp. d_2) alors $\mathcal{H}$ est de dimension $\mathrm{d}_1\,\mathrm{d}_2$.

L'orthonormalité de la famille $\{|\ell_1\rangle \otimes |\ell_2\rangle\}$ est évidente en utilisant la définition (6.7) du produit hermitien sur un espace produit tensoriel ; il reste à montrer que cette famille est dense dans $\mathcal{H}$, ce qui demande un peu plus de travail.

Tout vecteur de $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ se décompose alors dans une base orthonormée $\{|\ell_1\rangle \otimes |\ell_2\rangle\}$ de $\mathcal{H}$ comme :

$$\boxed{|\alpha\rangle = \sum_{\ell_1,\ell_2} \alpha_{\ell_1\,\ell_2}\,|\ell_1\rangle \otimes |\ell_2\rangle} \tag{6.11}$$

Malgré la notation $\alpha_{\ell_1\,\ell_2}$ à deux indices, il faut comprendre ces quantités comme les composantes d'un vecteur dans une base. [1]

1. Prenons le produit tensoriel de deux espaces hermitiens de dimension 2. On peut considérer, par exemple, α_{11} comme la première composante de $|\alpha\rangle$ dans la base de $\mathcal{H}$, $\alpha_{1,2}$ comme la deuxième, $\alpha_{2,1}$ comme la troisième et $\alpha_{2,2}$ comme la quatrième ; ce sont bien les composantes de la décomposition d'un vecteur sur une base d'un espace hermitien à 4 dimensions.

6.1.4 Opérateurs sur espace produit tensoriel

Soit $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ le produit tensoriel de deux espaces de Hilbert. Un opérateur agissant sur le premier espace de Hilbert, $\mathcal{O}_1 : \mathcal{H}_1 \to \mathcal{H}_1$ se prolonge naturellement en opérateur sur $\mathcal{H}$ de la manière suivante :

$$\mathcal{O}_1 \otimes \mathcal{I}_2 \ : \quad \left\{ \begin{array}{ccc} \mathcal{H}_1 \otimes \mathcal{H}_2 & \to & \mathcal{H}_1 \otimes \mathcal{H}_2 \\ |\alpha_1\rangle \otimes |\alpha_2\rangle & \mapsto & \big(\mathcal{O}_1 |\alpha_1\rangle\big) \otimes |\alpha_2\rangle \end{array} \right. \tag{6.12}$$

On peut naturellement faire de même pour un operateur agissant sur le deuxième espace de Hilbert, $\mathcal{O}_2 : \mathcal{H}_2 \to \mathcal{H}_2$. Dans ces deux cas de figure il s'agit d'opérateurs agissant uniquement sur un des deux facteurs.

Plus généralement on peut obtenir, à partir d'un opérateur $\mathcal{O}_1$ sur $\mathcal{H}_1$ et d'un opérateur $\mathcal{O}_2$ sur $\mathcal{H}_2$, un opérateur agissant sur le produit tensoriel des espaces de Hilbert :

$$\mathcal{O}_1 \otimes \mathcal{O}_2 \ : \quad \left\{ \begin{array}{ccc} \mathcal{H}_1 \otimes \mathcal{H}_2 & \to & \mathcal{H}_1 \otimes \mathcal{H}_2 \\ |\alpha_1\rangle \otimes |\alpha_2\rangle & \mapsto & \big(\mathcal{O}_1 |\alpha_1\rangle\big) \otimes \big(\mathcal{O}_2 |\alpha_2\rangle\big) \end{array} \right. \tag{6.13}$$

La propriété universelle du produit tensoriel, voir la prop. 1, implique l'unicité de l'opérateur $\mathcal{O}_1 \otimes \mathcal{O}_2$ ainsi obtenu, qui est un opérateur linéaire sur $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Crucialement, si les deux opérateurs $\mathcal{O}_1$ et $\mathcal{O}_2$ sont auto-adjoints (respectivement sur $\mathcal{H}_1$ et sur $\mathcal{H}_2$), on peut démontrer le théorème suivant :

Théorème 1 (produit tensoriel d'opérateurs auto-adjoints). *Soit $\mathcal{H}_1$ et $\mathcal{H}_2$ deux espaces de Hilbert et $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ leur produit tensoriel. Si $\mathcal{O}_1$ est un opérateur auto-adjoint sur $\mathcal{H}_1$ et si $\mathcal{O}_2$ est un opérateur auto-adjoint sur $\mathcal{H}_2$ alors $\mathcal{O}_1 \otimes \mathcal{O}_2$ est un opérateur auto-adjoint sur $\mathcal{H}$ et son spectre est donné par*

$$\boxed{\sigma(\mathcal{O}) = \sigma(\mathcal{O}_1)\,\sigma(\mathcal{O}_2)} \tag{6.14}$$

Ce théorème permet, en mécanique quantique, d'obtenir des observables sur $\mathcal{H}$ à partir des observables sur $\mathcal{H}_1$ et $\mathcal{H}_2$.

Exemple 1 (système de deux particules libres). Prenons par exemple un système composé de deux particules libres de masses m_1 et m_2 dans une boîte cubique de volume L^3 avec des conditions aux limites périodiques. Nous appellerons $\mathcal{H}_1$ et $\mathcal{H}_2$ leurs hamiltoniens respectifs. L'espace de Hilbert du système est le produit tensoriel des espaces de Hilbert de chacune des particules, et une base de l'espace des états du système est donnée par le produit tensoriel des états propres d'impulsion :

$$\left\{ |\vec{p}_1\rangle \otimes |\vec{p}_2\rangle \ , \quad \vec{p}_i \in \left(\frac{2\pi\hbar}{L}\mathbb{Z}\right)^{\times 3} \right\} . \tag{6.15}$$

Si l'état du système est donné par un de ces vecteurs de base, l'énergie de l'une ou l'autre ce des particules est donnée par

$$(\mathcal{H}_1 \otimes \mathcal{I}_2)\,|\vec{p}_1\rangle \otimes |\vec{p}_2\rangle = (\mathcal{H}_1\,|\vec{p}_1\rangle) \otimes |\vec{p}_2\rangle = \frac{p_1^2}{2m_1}\,|\vec{p}_1\rangle \otimes |\vec{p}_2\rangle \tag{6.16a}$$

$$(\mathcal{I}_2 \otimes \mathcal{H}_2)\,|\vec{p}_1\rangle \otimes |\vec{p}_2\rangle = |\vec{p}_1\rangle \otimes (\mathcal{H}_2\,|\vec{p}_2\rangle) = \frac{p_2^2}{2m_2}\,|\vec{p}_1\rangle \otimes |\vec{p}_2\rangle \ , \tag{6.16b}$$

et l'énergie totale est valeur propre du hamiltonien total

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{I}_2 + \mathcal{I}_2 \otimes \mathcal{H}_2 \tag{6.17}$$

De même l'impulsion totale de la particule est :

$$\vec{\mathcal{P}} = \vec{\mathcal{P}}_1 \otimes \mathcal{I}_2 + \mathcal{I}_1 \otimes \vec{\mathcal{P}}_2 \,. \tag{6.18}$$

Le spectre de cet opérateur d'impulsion totale est, suivant le théorème 1 :

$$\begin{aligned} \sigma(\vec{\mathcal{P}}) &= \sigma(\vec{\mathcal{P}}_1 \otimes \mathcal{I}_2 + \mathcal{I}_1 \otimes \vec{\mathcal{P}}_2) = \sigma(\vec{\mathcal{P}}_1 \otimes \mathcal{I}_2) + \sigma(\mathcal{I}_1 \otimes \vec{\mathcal{P}}_2) \\ &= \underbrace{\sigma(\vec{\mathcal{P}}_1)}_{\frac{2\pi\hbar}{L}\mathbb{Z}^{\times 3}} \times \underbrace{\sigma(\mathcal{I}_2)}_{1} + \underbrace{\sigma(\mathcal{I}_2)}_{1} \times \underbrace{\sigma(\vec{\mathcal{P}}_2)}_{\frac{2\pi\hbar}{L}\mathbb{Z}^{\times 3}} = \frac{2\pi\hbar}{L}\mathbb{Z}^{\times 3} \,. \end{aligned} \tag{6.19}$$

On peut également considérer des opérateurs ayant une action non-triviale sur $\mathcal{H}_1$ et sur $\mathcal{H}_2$, comme l'opérateur $\mathcal{O} = \mathcal{P}_1^x \otimes \mathcal{P}_2^x$. Nous avons

$$\mathcal{O}\,|\vec{p}_1\rangle \otimes |\vec{p}_2\rangle = \left(\mathcal{P}_1^x\,|\vec{p}_1\rangle\right) \otimes \left(\mathcal{P}_2^x\,|\vec{p}_2\rangle\right) = p_1^x p_2^x\,|\vec{p}_1\rangle \otimes |\vec{p}_2\rangle \ , \tag{6.20}$$

dont le spectre est

$$\sigma(\mathcal{O}) = \tfrac{2\pi\hbar}{L}\mathbb{Z}\,\tfrac{2\pi\hbar}{L}\mathbb{Z} = \left(\tfrac{2\pi\hbar}{L}\right)^2 \mathbb{Z} \,. \tag{6.21}$$

Mesure partielle et projection

Une classe importante d'exemples d'opérateurs sur un espace produit tensoriel correspond aux projecteurs agissant sur un seul des deux espaces du produit.

Pour en comprendre l'utilité, on considère un système quantique dont l'espace des états possède la structure d'un produit tensoriel $\mathcal{H}_1 \otimes \mathcal{H}_2$. On prépare le système dans un état arbitraire $|\alpha\rangle \in \mathcal{H}$.

On effectue alors une mesure physique associée à une observable $\mathcal{O}_1$ de l'espace de Hilbert $\mathcal{H}_1$. En continuant de raisonner sur l'exemple 1, on peut considérer une mesure de l'impulsion de la première des deux particules uniquement. Si la mesure donne une certaine impulsion $\vec{p}_1 \in (\frac{2\pi\hbar}{L}\mathbb{Z})^{\times 3}$, alors l'état $|\alpha\rangle$, suivant le postulat 5 de la mesure du cours n° 1 (réduction du paquet d'ondes), est modifié juste après la mesure suivant :

$$|\alpha\rangle \xrightarrow{\text{impulsion } \vec{p}_1 \text{ pour la particule 1}} \frac{1}{\|\,(\Pi_{\vec{p}_1} \otimes \mathcal{I}_2)\,|\alpha\rangle\,\|}\,(\Pi_{\vec{p}_1} \otimes \mathcal{I}_2)\,|\alpha\rangle \ , \tag{6.22}$$

où

$$\Pi_{\vec{p}_1} \overset{\text{déf.}}{=} |\vec{p}_1\rangle\,\langle\vec{p}_1| \tag{6.23}$$

est l'opérateur de $\mathcal{H}_1$ projetant sur le vecteur propre d'impulsion $\vec{p}_1$. Étant donné que l'opérateur $\Pi_{\vec{p}_1} \otimes \mathcal{I}_2$ agit trivialement sur $\mathcal{H}_2$, nous appelerons cette opération une *mesure partielle* sur le système. La probabilité associée à une telle mesure est donnée par la norme carrée de la projection partielle associée :

$$\boxed{p(\vec{p}_1) = \|\,(\Pi_{\vec{p}_1} \otimes \mathcal{I}_2)\,|\alpha\rangle\,\|^2} \tag{6.24}$$

La manière dont une mesure partielle change l'état du système est l'objet du paragraphe suivant.

6.1.5 Intrication

L'espace produit tensoriel $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ est engendré par les vecteurs d'une base de vecteurs de la forme $|\ell_1\rangle \otimes |\ell_2\rangle$, voir l'éqn. (6.10), composé chacun du produit tensoriel d'un vecteur de base de $\mathcal{H}_1$ avec un vecteur de base de $\mathcal{H}_2$.

Soit $|\alpha\rangle$ un vecteur quelconque de $\mathcal{H}$; il s'exprime alors comme combinaison linéaire de ces vecteurs de base, voir l'éqn. (6.11). Deux cas de figure sont alors possibles :

— il existe $|\alpha_1\rangle \in \mathcal{H}_1$ et $|\alpha_2\rangle \in \mathcal{H}_2$ tel que $|\alpha\rangle = |\alpha_1\rangle \otimes |\alpha_2\rangle$; on a alors

$$|\alpha\rangle = \sum_{\ell_1,\ell_2} \alpha_{\ell_1\,\ell_2} |\ell_1\rangle \otimes |\ell_2\rangle = \left(\sum_{\ell_1} \alpha_{\ell_1} |\alpha_1\rangle\right) \otimes \left(\sum_{\ell_2} \alpha_{\ell_2} |\alpha_2\rangle\right) , \tag{6.25}$$

et un tel état est appelé *état factorisable* ;

— il n'existe *aucun* couple $(|\alpha_1\rangle , |\alpha_2\rangle) \in \mathcal{H}_1 \times \mathcal{H}_2$ tel que $|\alpha\rangle = |\alpha_1\rangle \otimes |\alpha_2\rangle$; un tel état est appelé *état intriqué*.

Pour comprendre la signification de cette terminologie, prenons un système très simple composé de deux systèmes à deux niveaux, que nous assimilerons à deux spins 1/2 dans un champ magnétique avec comme états propres d'énergie les états « haut » et « bas » (respectivement $|\uparrow\rangle$ et $|\downarrow\rangle$). Les hamiltoniens des deux systèmes sont donnés par :

$$\mathcal{H}_1 = \frac{1}{2}\hbar\omega_1 \left(|\uparrow_1\rangle \langle\uparrow_1| - |\downarrow_1\rangle \langle\downarrow_1|\right) , \tag{6.26a}$$

$$\mathcal{H}_2 = \frac{1}{2}\hbar\omega_2 \left(|\uparrow_2\rangle \langle\uparrow_2| - |\downarrow_2\rangle \langle\downarrow_2|\right) . \tag{6.26b}$$

Supposons que ce soit préparé dans un état factorisable générique :

$$|\theta_1;\theta_2\rangle = (\cos\theta_1 |\uparrow_1\rangle + \sin\theta_1 |\downarrow_1\rangle) \otimes (\cos\theta_2 |\uparrow_2\rangle + \sin\theta_2 |\downarrow_2\rangle) . \tag{6.27}$$

La probabilité de mesurer la première particule dans l'état de spin $|\uparrow_1\rangle$ – exemple de mesure partielle – est alors donnée par le module carré du vecteur obtenu par l'action du projecteur $|\uparrow_1\rangle \langle\uparrow_1| \otimes \mathfrak{I}_2$:

$$\begin{aligned} \mathrm{p}(\uparrow_1) &= \| \, |\uparrow_1\rangle \langle\uparrow_1| \, (\cos\theta_1 |\uparrow_1\rangle + \sin\theta_1 |\downarrow_1\rangle) \otimes (\cos\theta_2 |\uparrow_2\rangle + \sin\theta_2 |\downarrow_2\rangle) \, \|^2 \\ &= \cos^2\theta_1 , \end{aligned} \tag{6.28}$$

et l'état du système après une mesure positive de l'état $\uparrow_1$ est

$$|\theta_1;\theta_2\rangle \;\longrightarrow\; |\uparrow_1\rangle \otimes (\cos\theta_2 |\uparrow_2\rangle + \sin\theta_2 |\downarrow_2\rangle) . \tag{6.29}$$

Si on effectue *ensuite* une mesure sur la deuxième particule, la probabilité de l'observer dans l'état $\uparrow_2$ est alors

$$\mathrm{p}(\uparrow_2) = \| \, |\uparrow_1\rangle \otimes |\uparrow_2\rangle \langle\uparrow_2| \, (\cos\theta_2 |\uparrow_2\rangle + \sin\theta_2 |\downarrow_2\rangle) \, \|^2 = \cos^2\theta_2 . \tag{6.30}$$

On trouverait exactement la même probabilité si on effectuait *d'abord* sur la deuxième particule, avec un raisonnement similaire à celui conduisant à l'éqn. (6.28). Le point crucial est que la mesure (6.29) faite d'abord sur la première particule n'affecte en rien l'état de la deuxième particule, et donc la probabilité de la trouver dans l'un ou l'autre de ses deux états possibles.

On peut donc considérer un état factorisé comme un produit de deux états de deux systèmes complètement indépendants l'un de l'autre, du point de vue des probabilités de les mesurer dans un état ou dans un autre.

Si le système se trouve, au contraire, dans un état *intriqué*, la mesure effectuée sur le sous-système associé à l'espace de Hilbert $\mathcal{H}_1$ va affecter l'état du sous-système associé à $\mathcal{H}_2$. Prenons le cas extrême d'un état *complètement intriqué*. On considère le système préparé dans l'état

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow_1\rangle \otimes |\downarrow_2\rangle + |\downarrow_1\rangle \otimes |\uparrow_2\rangle \right) . \tag{6.31}$$

On effectue une mesure sur le premier spin. Nous avons une probabilité $\mathfrak{p}(\uparrow_1) = 1/2$ de trouver le premier spin dans l'état $\uparrow_1$ et, si cette mesure est positive, le vecteur d'état est alors projeté selon

$$|\psi\rangle \longrightarrow \frac{1}{\left\| \left(|\uparrow_1\rangle \langle\uparrow_1| \otimes \mathfrak{I}_2 \right) |\psi\rangle \right\|} \left(|\uparrow_1\rangle \langle\uparrow_1| \otimes \mathfrak{I}_2 \right) |\psi\rangle = |\uparrow_1\rangle \otimes |\downarrow_2\rangle . \tag{6.32}$$

Nous voyons que le deuxième spin est alors dans l'état $\downarrow_2$ de manière certaine !

Cet effet quantique est très spectaculaire car il n'est absolument pas réservé aux systèmes microscopiques. Si les deux particules s'éloignent l'une de l'autre après être préparées ensemble sans être perturbées par leur environnement extérieur, et que la mesure a lieu alors qu'elles sont séparées de plusieurs mètres – ou même plusieurs kilomètres ! – la « corrélation » entre leurs états subsistera malgré la distance.

Notons également, pour éviter toute confusion, que l'intrication de deux particules ne présuppose pas d'interaction entre celles-ci, comme on le voit dans l'exemple précédent : on peut construire un état intriqué pour deux particules libres.

Entre les cas extrêmes des états complètement factorisés – comme l'état (6.27) – et les états complètement intriqués – comme l'état (6.31) – il existe un continuum d'états partiellement intriqués.

Exemple 2 (expérience de Stern–Gerlach). Étudions pour illustrer le concept d'intrication un exemple un peu différent d'espace des états produit tensoriel, en reprenant l'expérience de Stern–Gerlach décrite dans l'appendice 1.B du cours n° 1.

Les atomes d'argent sont assimilables à des particules portant un spin 1/2, pouvant donc se trouver dans deux états de spin différents $|\uparrow\rangle$ et $|\downarrow\rangle$. On considère un atome sortant du four et se trouvant dans l'état factorisé :

$$|\psi_0\rangle \otimes \frac{1}{\sqrt{2}} \left(|\uparrow\rangle + |\downarrow\rangle \right) , \tag{6.33}$$

où le ket $|\psi_0\rangle \in \mathcal{L}^2(\mathbb{R}^3)$ est associé à la fonction d'onde $\psi_0(\vec{x}) = \langle \vec{x} | \psi_0 \rangle$, paquet d'ondes de largeur σ et centré, selon l'axe Oz portant le champ magnétique, en $z = 0$.

Dans cet exemple les deux espaces de Hilbert intervenant dans le produit tensoriel décrivant l'espace des états de l'atome d'argent correspondent à deux types de degrés de liberté d'une même particule :

$$\mathcal{H} = \mathcal{L}^2(\mathbb{R}^3) \otimes \mathbb{C}^2 \,, \tag{6.34}$$

où le premier terme correspond aux degrés de liberté « orbitaux », associés à la densité de probabilité de présence, et le second aux degrés de liberté de spin.

Lorsque l'atome traverse la cavité contenant le champ magnétique, les atomes sont déviés selon l'axe Oz selon leur spin, disons vers les z croissants pour le spin $\uparrow$ et vers les z décroissants pour le spin $\downarrow$. À la sortie de la cavité, l'atome est schématiquement dans un état de la forme

$$\frac{1}{\sqrt{2}} \left(|\psi_d\rangle \otimes |\uparrow\rangle + |\psi_{-d}\rangle \, |\downarrow\rangle \right) \,, \tag{6.35}$$

où la fonction d'onde $\psi_d(\vec{x}) = \langle \vec{x} | \psi_d \rangle$ (resp. $\psi_{-d}(\vec{x}) = \langle \vec{x} | \psi_{-d} \rangle$) est un paquet d'onde centré selon l'axe Oz en $z = d$ (resp. $z = -d$), d étant la magnitude de la déviation du faisceau.

On voit que l'atome, initialement dans l'état factorisé (6.33), devient sous l'action du champ magnétique l'état *intriqué* (6.35). En effet la mesure du point d'impact de l'atome sur l'écran nous renseigne sur le spin de l'atome car la densité de probabilité de présence et l'état de spin sont maintenant corrélés. Dans la limite ou $d \gg \sigma$, les paquets d'ondes correspondant aux deux états de spin $\uparrow$ et $\downarrow$ ont un recouvrement très faible et l'état de l'atome est complètement intriqué.

6.2 Particules identiques

Le formalisme de produit tensoriel que nous avons développé à la section 6.1 permet d'étudier l'espace des états d'un système composé de plusieurs constituants.

Nous avons étudié dans l'exemple 1 le cas de deux particules libres dans une boîte. Si ces deux particules sont identiques (c.-à.-d. deux électrons, deux atomes d'argent, etc...) nous nous heurtons, en mécanique quantique, a une difficulté immédiate. Supposons que l'on puisse préparer le système dans un état

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \,, \tag{6.36}$$

$\psi_1(\vec{x}_1)$ et $\psi_2(\vec{x}_2)$ étant les fonctions d'onde, a priori différentes, donnant respectivement la densité de probabilité de présence pour la première et pour la deuxième particule.

On effectue alors une mesure sur le système et on détecte une particule dans un petit volume D de la cavité, correspondant à l'emplacement du détecteur (voir le postulat 2 du cours n° 2). Quel est l'état du système juste après la mesure ? Il faut distinguer deux cas de figure :

- si l'expérience a détecté la particule 1, c'est son paquet d'ondes qui est réduit et l'état du système est donc décrit par les fonctions d'onde $\delta_D(\vec{x}_1)\psi_1(\vec{x}_1)$ et $\psi_2(\vec{x}_2)$.
- si l'expérience a détecté la particule 2, c'est son paquet d'ondes qui est réduit et l'état du système est donc décrit par les fonctions d'onde $\psi_1(\vec{x}_1)$ et $\delta_D(\vec{x}_2)\psi_2(\vec{x}_2)$.

Ce raisonnement pose un problème immédiat ; les particules étant en tout point identiques, il est *impossible* pour l'observateur de déterminer si la particule qu'il a détectée est la particule 1 (premier cas) ou la particule 2 (second cas). [2] Le problème ne se poserait pas en mécanique classique car on pourrait, au moins sur le principe, savoir à chaque instant laquelle est la particule 1 et laquelle est la particule 2 en suivant leurs trajectoires respectives.

Ce problème ne se pose que pour des particules pouvant effectivement échanger leurs positions. Si on considère des atomes d'un cristal, le recouvrement de leurs fonctions d'onde est tellement faible qu'en pratique on peut les considérer comme des particules discernables pouvant être distinguées les unes des autres.

6.2.1 Groupe symétrique

Pour comprendre comment résoudre ce paradoxe, il est utile de se pencher plus avant sur la nouvelle symétrie du problème apparaissant lorsqu'on considère un système constitué de N particules identiques.

Un système composé de N particules libres identiques de masse m est décrit par un hamiltonien de la forme :

$$\mathcal{H} = \sum_{k=1}^{n} \frac{1}{2m} \vec{\mathcal{P}}_k^{\,2} \overset{\text{déf.}}{=} \sum_{k=1}^{n} \mathcal{H}_k \,. \tag{6.37}$$

Une base de l'espace des états physiques semble être donnée par le produit tensoriel des vecteurs propres d'impulsion de chaque particule :

$$|\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_N\rangle \overset{?}{=} |\vec{p}_1\rangle \otimes |\vec{p}_2\rangle \otimes \cdots \otimes |\vec{p}_N\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_N \,. \tag{6.38}$$

Cependant, aucune expérience ne permettrait de distinguer le vecteur (6.38) de, disons, un autre vecteur où les deux premières impulsions auraient été permutées :

$$|\vec{p}_2, \vec{p}_1, \ldots, \vec{p}_N\rangle \overset{?}{=} |\vec{p}_2\rangle \otimes |\vec{p}_1\rangle \otimes \cdots \otimes |\vec{p}_N\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_N \,, \tag{6.39}$$

ou n'importe quelle combinaison linéaire de (6.38) et (6.39). Le système comporte ainsi une nouvelle dégénerescence dans le spectre de ses états physiques, appelée *dégénéresence d"échange*, qui semblerait impossible de lever.

L'origine de cette difficulté est l'existence d'une nouvelle symétrie du problème. Introduisons un opérateur de *transposition* (ou opérateur d'échange) $\mathcal{T}_{ij}$, échangeant la i-ème particule avec la j-ème :

$$\mathcal{T}_{ij}\ |\vec{p}_1\rangle \otimes \cdots \otimes \textcircled{$|\vec{p}_i\rangle$} \otimes \cdots \otimes \textcircled{$|\vec{p}_j\rangle$} \otimes \cdots \otimes |\vec{p}_N\rangle = |\vec{p}_1\rangle \otimes \cdots \otimes \textcircled{$|\vec{p}_j\rangle$} \otimes \cdots \otimes \textcircled{$|\vec{p}_i\rangle$} \otimes \cdots \otimes |\vec{p}_N\rangle \tag{6.40}$$

Cet opérateur satisfait évidemment $\mathcal{T}_{ij}^2 = \mathfrak{I}$ et est unitaire, s'agissant d'une symétrie ; comme $\mathcal{T}_{ij}^\dagger = \mathcal{T}_{ij}^{-1} = \mathcal{T}_{ij}$, il est donc également hermitien ; ses valeurs propres sont ± 1.

2. Dans le cas de deux particules différentes, il faut imaginer, dans cette expérience de pensée, que la détection de la particule s'accompagne de la mesure de ses caractéristiques physiques (masse, charge,...) permettant de distinguer les différentes espèces de particules entre elles.

La forme particulière, factorisée, du hamiltonien (6.37) nous indique que ses éléments de matrice satisfont aux relations :

$$\begin{aligned}\forall i,j \quad \langle \vec{p}_1,\dots,\vec{p}_i,\dots,\vec{p}_j,\dots,\vec{p}_N|\,\mathcal{T}_{ij}^\dagger \mathcal{H}\mathcal{T}_{ij}\,|\vec{p}'_1,\dots,\vec{p}'_i,\dots,\vec{p}'_j,\dots,\vec{p}'_N\rangle \\ = \langle \vec{p}_1,\dots,\vec{p}_j,\dots,\vec{p}_i,\dots,\vec{p}_N|\,\mathcal{H}\,|\vec{p}'_1,\dots,\vec{p}'_j,\dots,\vec{p}'_i,\dots,\vec{p}'_N\rangle \\ = \delta_{\vec{p}_1,\vec{p}'_1}\cdots\delta_{\vec{p}_N,\vec{p}'_N}\sum_i \frac{\vec{p}_i^{\,2}}{2m} \\ = \langle \vec{p}_1,\dots,\vec{p}_i,\dots,\vec{p}_j,\dots,\vec{p}_N|\,\mathcal{H}\,|\vec{p}'_1,\dots,\vec{p}'_i,\dots,\vec{p}'_j,\dots,\vec{p}'_N\rangle \end{aligned} \tag{6.41}$$

Nous avons ainsi

$$\boxed{\forall i,j \quad \mathcal{T}_{ij}^\dagger\,\mathcal{H}\,\mathcal{T}_{ij} = \mathcal{H}} \tag{6.42}$$

Cette relation signifie que le système possède une symétrie discrète associée aux permutations de deux particules. L'opérateur $\mathcal{T}_{ij}$ étant unitaire, en multipliant à gauche cette égalité par $\mathcal{T}_{ij}$ nous obtenons

$$\mathcal{H}\,\mathcal{T}_{ij} = \mathcal{T}_{ij}\,\mathcal{H} \implies \boxed{[\,\mathcal{T}_{ij},\,\mathcal{H}\,] = 0} \tag{6.43}$$

Remarquons qu'il est possible d'avoir un hamiltonien satisfaisant à la relation (6.42), et donc invariant par échange de deux particules, sans pour autant qu'il soit factorisable ; prenons par exemple un hamiltonien de deux particules identiques avec interaction dépendant de leur séparation :

$$\mathcal{H} = \frac{\vec{\mathcal{P}}_1^{\,2}}{2m} + \frac{\vec{\mathcal{P}}_2^{\,2}}{2m} + V\left(\left|\vec{\mathcal{X}}_1 - \vec{\mathcal{X}}_2\right|\right), \tag{6.44}$$

qui est clairement invariant sous l'échange des positions et impulsions des deux particules même si celles-ci sont en interaction ; en ne prenant pas en compte le spin, un système de deux électrons interagissant par répulsion coulombienne entre dans cette catégorie et le hamiltonien correspondant est invariant par permutation des deux électrons.

L'ensemble de toutes transpositions d'un ensemble de N objets, $\mathcal{T}_{ij}$ pour $i,j \in \{1,\dots,N\}$, ainsi que toutes leurs compositions (c.-à.-d. transpositions successives) forment un groupe appelé groupe symétrique :

Définition 2 (groupe symétrique). *Le groupe symétrique S_N est le groupe de toutes les permutations possibles de N objets. Sa dimension est $\sharp S_N = N!$. Le groupe S_2 est abélien et isomorphe à $\mathbb{Z}_2$, alors que les groupes S_N pour $N > 2$ sont non-abéliens.*

Une propriété importante du groupe symétrique est la décomposition d'une permutation en transpositions. On montre le résultat suivant :

Propriété 2 (décomposition d'une permutation). *Toute permutation $\sigma \in S_N$ se décompose en produit de transpositions, $\sigma = T_{ij} \circ T_{i'j'} \circ \cdots$. Cette décomposition n'est pas unique mais la parité du nombre de transpositions est fixe : une permutation est dite paire ou impaire.*

On introduit alors la signature d'une permutation, définie par

$$\epsilon_\sigma = \begin{cases} +1 & \text{si} \quad \sigma \text{ paire} \\ -1 & \text{si} \quad \sigma \text{ impaire} \end{cases} \tag{6.45}$$

qui satisfait aux propriétés

$$\epsilon_{\sigma^{-1}} = \epsilon_\sigma \ , \quad \epsilon_{\sigma\circ\tau} = \epsilon_\sigma \, \epsilon_\tau \tag{6.46}$$

Suivant la logique du cours n° 3, on s'attend à ce que les états d'un système quantique présentant une symétrie de permutation, comme celui décrit par le hamiltonien (6.37), puissent se décomposer en représentations irreductibles du groupe symétrique S_N (toutes équivalentes à des représentations unitaires, le groupe étant fini).

La théorie des représentations irreductibles des groupes symétriques est bien connue (voir par exemple [1]) ; une propriété importante est que toutes les représentations irreductibles sont réelles. Mentionnons trois représentations irreductibles importantes et génériques :

- La *représentation triviale* D^0 de dimension 1, telle que $\forall \sigma \in S_N$, $D^0(\sigma) = 1$;
- La *représentation signature* D^ϵ de dimension 1, telle que $\forall \sigma \in S_N$, $D^1(\sigma) = \epsilon_\sigma$;
- La *représentation standard* D^s de dimension réelle $N-1$ qui se construit comme suit ($N > 2$). Une représentation de dimension N est obtenue en associant à une permutation σ donnée, la matrice $N \times N$ qui envoie la base orthonormée $\{\vec{e}_i,\ i = 1, \dots N\}$ sur la base $\{\vec{e}_{\sigma(i)},\ i = 1, \dots N\}$. Cette représentation est réductible car elle contient la représentation triviale, l'action de toute permutation sur le vecteur $\vec{s} = \sum_{i=1}^{N} \vec{e}_i$ étant triviale. L'action sur le sous-espace supplémentaire à $\vec{s}$ définit une représentation irreductible D^s de dimension $N-1$.

La combinaison de la représentation signature avec la représentation standard fournit une autre représentation irreductible générique de dimension $N-1$. Le détail des autres représentations irreductibles dépend de N ; néanmoins, pour tout N, les seules représentations de dimension 1 inéquivalentes sont la représentation triviale et la représentation signature.

6.2.2 Postulat d'(anti)-symétrisation : bosons et fermions

Un système quantique de N particules identiques possède une symétrie de permutation entre les particules, associée à l'action du groupe symétrique S_N.

Il est naturel de postuler que chaque type de particule se transforme dans une certaine représentation irreductible et unitaire du groupe symétrique ; remarquons que, par particule, nous ne désignons pas uniquement des particules élémentaires comme l'électron mais aussi des particules « composites » comme un atome qui, pour des énergies suffisamment basses pour ne pas exciter de degrés de liberté internes, peut être considéré comme une particule.

Pour lever complètement la dégénerscence d'échange, les représentations irreductibles dans lesquelles se transforment les particules devraient être de dimension un. On s'attend également à ce que les propriétés d'un système de N particules de ne dépendent pas fortement de la valeur précise de N, alors que la structure détaillée des représentations irreductibles de S_N change avec N, à l'exception des quatre représentations mentionnées plus haut.

Ce raisonnement incite à privilégier les représentations triviales et signature et on est amené a formuler le postulat suivant :

Postulat 1 (postulat de symétrisation et antisymétrisation). *Pour un système quantique de* N *particules identiques il existe deux comportements possibles vis à vis des permutations entre les particules :*

- *pour les particules suivant la statistique de Bose–Einstein, appelées bosons, les états du système se transforment dans la représentation triviale du groupe symétrique* S_N *; autrement dit les vecteurs d'états sont invariants par toute transposition de deux particules et, par extension, par toute permutation ;*
- *pour les particules suivant la statistique de Fermi–Dirac, appelées fermions, les états du système se transforment dans la représentation signature du groupe symétrique* S_N *; autrement dit les vecteurs d'états sont impairs par toute transposition de deux particules et, par extension, totalement antisymétriques par toute permutation.*

Concrètement, si $|\psi\rangle$ est un vecteur d'état pour un système de N particules identiques, nous avons

$$\boxed{\forall \sigma \in \mathsf{S}_\mathsf{N} \ , \quad \mathcal{Q}(\sigma)\,|\psi\rangle = \begin{cases} |\psi\rangle & \text{pour des bosons} \\ \epsilon_\sigma\,|\psi\rangle & \text{pour des fermions} \end{cases}} \tag{6.47}$$

où $\mathcal{Q}(\sigma)$ est l'opérateur associé à la permutation σ.

Cette différence de comportement des bosons et fermions vis à vis des permutations a des conséquences dramatiques, et est responsable d'effets macroscopiques importants.

Propriété 3 (principe d'exclusion de Pauli). *Dans un système quantique de* N *fermions identiques, il ne peut y avoir au plus qu'un fermion dans chacun des états physiques possibles pour chaque particule.*

En effet, si deux fermions se trouvaient dans le même état, le vecteur d'état serait symétrique par permutation de ces deux particules.

Ce principe d'exclusion se manifeste en physique du solide, en particulier dans l'étude des électrons libres dans un cristal. L'état fondamental du système, qui décrit l'état du système à température nulle, s'obtient en remplissant tous les niveaux d'énergie à partir du niveau le plus bas, avec un seul électron par niveau.

À une échelle bien plus grande, les *naines blanches*, astres très denses correspondant à des résidus d'étoiles de masse voisine de celle du Soleil ayant épuisées leur ressources en hydrogène, résistent à l'effondrement gravitationnel grâce à la *pression de dégénérescence* des électrons due au principe d'exclusion de Pauli (cf. cours de mécanique statistique).

6.2.3 Construction des états (anti)-symétriques

Soit un système de N bosons ou N fermions identiques et $|\psi;\mathsf{N}\rangle$ un vecteur de l'espace de Hilbert associé $\mathcal{H}$ sur lequel nous n'avons pas encore appliqué le postulat 1. Un vecteur

d'état physique est obtenu en symétrisant ou antisymétrisant ce vecteur suivant qu'il s'agisse de bosons ou de fermions.

Soit $\mathcal{Q}(\sigma)$ l'opérateur unitaire associé à une permutation $\sigma \in S_N$. On introduit les opérateurs de symétrisation $\mathfrak{S}_N$ et d'antisymétrisation $\mathfrak{A}_N$:

$$\mathfrak{S}_N \overset{\text{déf.}}{=} \frac{1}{N!} \sum_{\sigma \in S_N} \mathcal{Q}(\sigma) \tag{6.48a}$$

$$\mathfrak{A}_N \overset{\text{déf.}}{=} \frac{1}{N!} \sum_{\sigma \in S_N} \epsilon_\sigma \, \mathcal{Q}(\sigma) \tag{6.48b}$$

En d'autres termes, $\mathfrak{S}_N$ est le projecteur sur la représentation triviale de S_N et $\mathfrak{A}_N$ le projecteur sur la représentation signature. On a naturellement :

$$\forall \tau \in S_N, \quad \mathcal{Q}(\tau)\mathfrak{S}_N = \frac{1}{N!} \sum_{\sigma \in S_N} \mathcal{Q}(\tau \circ \sigma) = \frac{1}{N!} \sum_{\sigma' \in S_N} \mathcal{Q}(\sigma') = \mathfrak{S}_N \tag{6.49a}$$

$$\forall \tau \in S_N, \quad \mathcal{Q}(\tau)\mathfrak{A}_N = \frac{1}{N!} \sum_{\sigma \in S_N} \epsilon_\sigma \mathcal{Q}(\tau \circ \sigma) = \frac{1}{N!} \sum_{\sigma' \in S_N} \epsilon_{\tau^{-1} \circ \sigma'} \, \mathcal{Q}(\sigma') = \epsilon_\tau \, \mathfrak{A}_N \, , \tag{6.49b}$$

ou on a posé dans la somme $\sigma' = \tau \circ \sigma$ et utilisé le fait que sommer sur tous les $\sigma \in S_N$ est équivalent à sommer sur tous les $\tau \circ \sigma$ à τ fixé. On voit ainsi que, pour tout opérateur de transposition $\mathcal{T}_{ij}$, le projecteur $\mathfrak{S}_N$ (resp. $\mathfrak{A}_N$) projette sur le sous-espace propre de $\mathcal{T}_{ij}$ de valeur propre $+1$ (resp. -1).

Le facteur de normalisation $1/N!$ garantit que $\mathfrak{S}_N^2 = \mathfrak{S}_N$ et que $\mathfrak{A}_N^2 = \mathfrak{A}_N$, $N!$ étant la dimension de S_N. Nous avons en effet :

$$\mathfrak{S}_N^2 = \frac{1}{N!} \sum_{\sigma \in S_N} \mathcal{Q}(\sigma)\mathfrak{S}_N = \frac{1}{N!} \sum_{\sigma \in S_N} \mathfrak{S}_N = \mathfrak{S}_N \tag{6.50a}$$

$$\mathfrak{A}_N^2 = \frac{1}{N!} \sum_{\sigma \in S_N} \epsilon_\sigma \mathcal{Q}(\sigma)\mathfrak{S}_N = \frac{1}{N!} \sum_{\sigma \in S_N} \epsilon_\sigma^2 \mathfrak{A}_N = \frac{1}{N!} \sum_{\sigma \in S_N} \mathfrak{A}_N = \mathfrak{A}_N \tag{6.50b}$$

En outre les opérateurs $\mathfrak{S}_N$ et $\mathfrak{A}_N$ sont hermitiens car, par unitarité, $\mathcal{Q}(\sigma)^\dagger = \mathcal{Q}(\sigma^{-1})$ et la somme sur tous les $\sigma \in S_N$ coïncide avec la somme sur tous leurs inverses σ^{-1}.

On construit alors les vecteurs d'état physiques pour les bosons et les fermions de la manière suivante :

Propriété 4 *Soit un système de N bosons (resp. N fermions) identiques dont les états, avant (anti-)symétrisation, appartiennent à un espace de Hilbert $\mathcal{H}$. Les états physiques sont obtenus en symétrisant (resp. antisymétrisant) les vecteurs de $\mathcal{H}$, c.-à.-d. en associant à tout $|\psi\rangle \in \mathcal{H}$ les vecteurs*

$$|\psi\rangle_S \overset{\text{déf.}}{=} \frac{1}{\| \, \mathfrak{S}_N \, |\psi\rangle \, \|} \mathfrak{S}_N \, |\psi\rangle \qquad \textit{(bosons)} \tag{6.51a}$$

$$|\psi\rangle_A \overset{\text{déf.}}{=} \frac{1}{\| \, \mathfrak{A}_N \, |\psi\rangle \, \|} \mathfrak{A}_N \, |\psi\rangle \qquad \textit{(fermions)} \tag{6.51b}$$

Les vecteurs symétrisés et antisymétrisés obtenus sont vecteurs propres des opérateurs de transposition, respectivement pour la valeur propre +1 et la valeur propre -1 :

$$\forall i,j \in \{1,\dots,N\}\,, \qquad \begin{array}{l} \mathcal{T}_{ij}\,|\psi\rangle_S = +\,|\psi\rangle_S \\ \mathcal{T}_{ij}\,|\psi\rangle_A = -\,|\psi\rangle_A \end{array} \tag{6.52}$$

Si le hamiltonien se scinde en somme de hamiltoniens individuels pour les particules, c.-à.-d.

$$\mathcal{H} = \sum_{k=1}^{N} \mathcal{H}_k \tag{6.53}$$

où les hamiltoniens $\mathcal{H}_k$ sont tous identiques, on peut donner une forme plus explicite à cette procédure. On note $\{|\ell_k\rangle\,, \ell \in \mathfrak{I}\}$ une base orthonormée de $\mathcal{H}_k$.

Fermions. On obtient :

$$\begin{aligned} |\psi\rangle = |\ell_1\rangle \otimes |\ell_2\rangle \cdots |\ell_N\rangle \;\to |\psi\rangle_A &\overset{\text{déf.}}{=} \sqrt{N!}\,\mathfrak{A}_N\,|\ell_1\rangle \otimes \cdots \otimes |\ell_N\rangle \\ &= \boxed{\frac{1}{\sqrt{N!}} \sum_{\sigma\in S_N} \epsilon_\sigma \left|\ell_{\sigma(1)}\right\rangle \otimes \cdots \otimes \left|\ell_{\sigma(N)}\right\rangle} \end{aligned} \tag{6.54}$$

où les ℓ_k sont nécessairement tous distincts par le principe d'exclusion de Pauli. Ces états sont bien normalisés car nous avons premièrement :

$$\begin{aligned} &\frac{1}{N!} \sum_{\sigma\in S_N} \epsilon_\sigma \sum_{\tau\in S_N} \epsilon_\tau \big(\left\langle \ell_{\sigma(N)}\right| \otimes \cdots \otimes \left\langle \ell_{\sigma(1)}\right| \big) \big(\left|\ell_{\tau(1)}\right\rangle \otimes \cdots \otimes \left|\ell_{\tau(N)}\right\rangle \big) \\ &\qquad = \frac{1}{N!} \sum_{\sigma\in S_N} \sum_{\tau\in S_N} \epsilon_{\sigma\circ\tau} \prod_{k=1}^{N} \left\langle \ell_{\sigma(k)}\middle|\ell_{\tau(k)}\right\rangle \end{aligned} \tag{6.55}$$

Les ℓ_k étant tous distincts par antisymétrie, les produits scalaires $\left\langle \ell_{\sigma(k)}\middle|\ell_{\tau(k)}\right\rangle$ sont nuls sauf si $\sigma(k) = \tau(k)$ et on obtient

$$ {}_A\langle\psi|\,\psi\rangle_A = \frac{1}{N!} \sum_{\sigma\in S_N} \sum_{\tau\in S_N} \epsilon_{\sigma^{-1}\circ\tau} \prod_{k=1}^{N} \delta_{k,\sigma^{-1}\circ\tau(k)}\,, \tag{6.56}$$

et la somme sur τ donne un seul terme non-nul, celui pour lequel on $\sigma^{-1}\circ\tau$ est la permutation identité, d'où on tire ${}_A\langle\psi|\,\psi\rangle_A = 1$. Dans le cas d'un système de deux particules on a :

$$|\ell_1,\ell_2\rangle_A = \frac{1}{\sqrt{2}} \big(|\ell_1\rangle\,|\ell_2\rangle - |\ell_2\rangle\,|\ell_1\rangle \big) \qquad \ell_1 \neq \ell_2\,. \tag{6.57}$$

Prenons un système de deux fermions libres ; l'état de chaque particule $|\psi_k\rangle$ est associé à une fonction d'onde $\langle\vec{x}_k|\psi_k\rangle = \psi_k(\vec{x}_k)$. La fonction d'onde associée à l'état antisymétrisé (6.57) est donnée par

$$\begin{aligned}\psi(\vec{x}_1,\vec{x}_2) &= \big(\langle\vec{x}_2|\otimes\langle\vec{x}_1|\big)\,|\psi_1,\psi_2\rangle_A = \frac{1}{\sqrt{2}}\big(\langle\vec{x}_2|\otimes\langle\vec{x}_1|\big)\big(|\psi_1\rangle\otimes|\psi_2\rangle - |\psi_2\rangle\otimes|\psi_1\rangle\big)\\ &= \frac{1}{\sqrt{2}}\big(\psi_1(\vec{x}_1)\psi_2(\vec{x}_2) - \psi_1(\vec{x}_2)\psi_2(\vec{x}_1)\big).\end{aligned} \tag{6.58}$$

Cette construction se généralise facilement à un système de N fermions (on néglige pour l'instant le rôle du spin).

Définition 3 (déterminant de Slater) *Soit un système de* N *fermions identiques de hamiltonien* $\mathcal{H} = \sum_{k=1}^{N}\mathcal{H}_k$. *Les états d'une des particules, notés* $|\psi_k\rangle$, *sont associés à des fonctions d'onde* $\psi_k(\vec{x}_k) = \langle\vec{x}_k|\psi_k\rangle$. *Les fonctions d'ondes physiques du système sont données par le déterminant de Slater :*

$$\psi(\vec{x}_1,\ldots,\vec{x}_N) = \frac{1}{\sqrt{N!}}\begin{vmatrix}\psi_1(\vec{x}_1) & \psi_2(\vec{x}_1) & \cdots & \psi_N(\vec{x}_1)\\ \psi_1(\vec{x}_2) & \psi_2(\vec{x}_2) & \cdots & \psi_N(\vec{x}_2)\\ \vdots & \vdots & \vdots & \vdots\\ \psi_1(\vec{x}_N) & \psi_2(\vec{x}_N) & \cdots & \psi_N(\vec{x}_N)\end{vmatrix} \tag{6.59}$$

obéissant, grâce aux propriétés d'antisymétrie totale du déterminant, à la condition d'antisymétrie pour des fermions identiques.

Bosons. Pour des bosons plusieurs particules peuvent être dans le même état individuel $|\ell\rangle$. On regroupe alors les états à une particule occupés correspondant au même état individuel en sous-ensembles de taile n_j, avec $\sum_{j=1}^{r} n_j = N$. On obtient alors :

$$\begin{aligned}|\psi\rangle = |\ell_1\rangle\otimes|\ell_2\rangle\cdots|\ell_N\rangle \;\to |\psi\rangle_S &\overset{\text{déf.}}{=} \sqrt{\frac{N!}{n_1!\cdots n_r!}}\,\mathfrak{S}_N\,|\ell_1\rangle\otimes\cdots\otimes|\ell_N\rangle\\ &= \boxed{\sqrt{\frac{1}{N!n_1!\cdots n_r!}}\sum_{\sigma\in S_N}\big|\ell_{\sigma(1)}\big\rangle\otimes\cdots\otimes\big|\ell_{\sigma(N)}\big\rangle}\end{aligned} \tag{6.60}$$

La différence avec les fermions, voir l'éqn. (6.54), est que plusieurs termes contribuent à la somme sur les permutations de la même manière si plusieurs particules sont dans le même état. Prenons l'exemple d'un système de trois particules avec deux particules dans le même état. Nous avons

$$\mathfrak{S}_3\,|1\rangle\otimes|1\rangle\otimes|2\rangle = \frac{1}{6}\Big(2\,|1\rangle\otimes|1\rangle\otimes|2\rangle + 2\,|1\rangle\otimes|2\rangle\otimes|1\rangle + 2\,|2\rangle\otimes|1\rangle\otimes|1\rangle\Big) \tag{6.61}$$

Dont la norme carrée est égale à $\|\mathfrak{S}_3\,|1\rangle\otimes|1\rangle\otimes|2\rangle\,\|^2 = \left(\frac{2}{6}\right)^2\times 3 = 1/3 = 2!/3!$ alors que si les trois particules sont dans des états distincts, nous avons après symétrisation la somme sur six vecteurs distincts (un par permutation) et $\|\mathfrak{S}_3\,|1\rangle\otimes|2\rangle\otimes|3\rangle\,\|^2 = \frac{1}{6^2}\times 6 = 1/6 = 1/3!$.

6.3 Théorème spin-statistique

Les considérations développées dans la section 6.2 nous indiquent que les particules se rangent en deux catégories, les bosons et les fermions, suivant qu'elles suivent la statistique de Bose–Einstein ou de Fermi–Dirac.

Une autre caractéristique importante des particules massives en mécanique quantique est leur *spin*, dont nous développerons en détail la théorie à partir du cours n° 13 (le spin d'une particule indique essentiellement dans quelle représentation irreductible du groupe des rotations elle se transforme). [3]

Le spin d'une particule d'une part, et sa nature bosonique et fermonique d'autre part, semblent deux caractéristiques complètement indépendantes, et il n'est en effet pas possible de trouver un lien théorique entre les deux dans le cadre de la mécanique quantique non-relativiste. En revanche en théorie quantique des champs, théorie quantique relativiste (dans laquelle le nombre de particules n'est pas fixé) on démontre le théorème suivant :

Théorème 2 (théorème spin-statistique). *Toute particule de spin entier est un boson, et toute particule de spin demi-entier est un fermion.*

Par exemple l'électron, de spin 1/2, est un fermion et le boson de Higgs, de spin 0, est un boson comme son nom l'indique. Pour démontrer ce théorème, on part de l'action classique relativiste pour des particules de spin entier et on montre que la théorie quantique correspondante est cohérente – en particulier, que tous les états physique sont de norme carrée positive – uniquement pour une statistique de Bose–Einstein ; de même si on part de l'action de particules de spin demi-entier la cohérence de la théorie quantique impose la statistique de Fermi–Dirac.

Notons encore une fois que ce théorème s'applique aussi aux particules composites. Deux isotopes du lithium, le ^{6}Li et le ^{7}Li, se comportent à basse température respectivement comme des particules de spin demi-entier et entier. Le premier suit donc la statistique de Fermi–Dirac et le second la statistique de Bose–Einstein. Il est possible, par refroidissement laser d'un mélange des deux isotopes, d'obtenir une « mixture » de gaz de Fermi fortement dégénéré et de condensat de Bose–Einstein [2].

6.4 Espaces de Fock

Dans un système de N particules identiques – qu'il s'agisse de bosons ou de fermions – l'information pertinente pour décrire le système correspond au nombre de particules dans chacun des états propres du hamiltonien à une particule. Le formalisme adapté a cette description, qui est surtout utilisé lorsque le nombre de particules dans le système n'est pas fixé, utilise la notion d'*espace de Fock* que nous présentons brièvement dans cette section.

Exemple 3 (système de deux bosons libres). Considérons pour illustrer cette idée un système de deux particules massives bosoniques à une dimension, sur un intervalle de longueur L, avec

3. Pour des particules de masse nulle, dans le cadre d'une théorie quantique relativiste, il faudrait introduire la notion d'hélicité au lieu de la notion de spin.

des conditions aux limites périodiques. Le hamiltonien de ce système est naturellement $\mathcal{H} = \mathcal{P}_1^2/2m + \mathcal{P}_2^2/2m$, où le spectre de chaque opérateur impulsion $\mathcal{P}_i$ est donné par $\sigma(\mathcal{P}_i) = \frac{2\pi\hbar}{L}\mathbb{Z}$.

Un état propre de $\mathcal{P}_i$ d'impulsion $2\pi\hbar k/L$ est simplement noté $|k\rangle$, et, dans l'espace de Hilbert produit tensoriel $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, les états propres d'impulsion sont notés $|k\rangle \otimes |\ell\rangle$ avec $k, \ell \in \mathbb{Z}$. Un état physique quelconque du système est alors donné par :

$$|\psi\rangle = \sum_{k,\ell\in\mathbb{Z}} a_{k\ell}\, \frac{|k\rangle \otimes |\ell\rangle + |\ell\rangle \otimes |k\rangle}{\sqrt{2}} \;, \quad a_{k\ell} = a_{\ell k} \;, \quad \sum_{k,\ell\in\mathbb{Z}} |a_{k\ell}|^2 = 1 \,. \tag{6.62}$$

Une manière complèment équivalente de décrire ce système est de spécifier, pour chaque valeur propre de l'opérateur d'impulsion à une particule, ou, autrement dit, pour tout entier k tel que l'impulsion soit égale à $2\pi\hbar k/L$, le nombre de particules dans un tel état. Nous noterons n_k le *nombre d'occupation* associé. Un état du système caractérisé par n_1 particules dans l'état d'impulsion avec $k = 1$, n_2 particules dans l'état d'impulsion avec $k = 2$, etc... sera représenté par le vecteur normalisé :

$$|\mathbf{n}\rangle \overset{\text{déf.}}{=} |\cdots, n_{-1}, n_0, n_1, n_2, \cdots\rangle \;. \tag{6.63}$$

Dans le cas présent, le nombre de particule est fixé à deux, ce qui limite fortement le nombre de possibilités. D'une part, chaque nombre d'occupation satisfait $n_k \in \{0, 1, 2\}$, et d'autre part $\sum_{k\in\mathbb{Z}} n_k = 2$. On peut ainsi décrire un état physique quelconque du système comme

$$|\psi\rangle = \sum_{\substack{n_k\in\{0,1,2\} \\ \sum n_k = 2}} c_{\mathbf{n}}\, |\mathbf{n}\rangle \;, \qquad \sum_{\substack{n_k\in\{0,1,2\} \\ \sum n_k = 2}} |c_{\mathbf{n}}|^2 = 1 \,. \tag{6.64}$$

Cette manière de décrire les états physiques est complètement équivalente à la précédente. En effet l'état symétrisé $(|k\rangle \otimes |\ell\rangle + |\ell\rangle \otimes |k\rangle)/\sqrt{2}$ dans la première description correspond dans la seconde description à un état $|\mathbf{n}\rangle$ avec $n_k = n_\ell = 1$ et $n_i = 0$ pour $i \neq k, \ell$.

La construction illustrée ci-dessus s'étend sans difficulté à un système composé de trois ou plus particules, et également aux particules suivant la statistique de Fermi–Dirac. Cependant, les contraintes imposées aux nombres d'occupation n_k, à savoir que le nombre de particules est fixé à une valeur N donnée, rend son utilisation malcommode. Ce formalisme est plus adapté aux systèmes pour lesquels le nombre de particules n'est pas fixé, comme les systèmes relativistes. Cependant, dans la limite où la valeur moyenne du nombre de particules est très grande, la distribution statistique du nombre de particules autour de cette valeur moyenne est suffisamment « piquée » pour que ce formalisme puisse être utilisé sans inconvénient pour des systèmes fermés non relativistes, à nombre de particules fixé. Des considérations similaires, en physique statistique, peuvent être invoquées pour justifier l'utilisation de l'ensemble grand-canonique par commodité, même pour des systèmes fermés.

Définition 4 (espace de Fock). *Soit $\mathcal{H}$ un espace de Hilbert, associé aux états physiques d'une particule. Si ce type de particule suit la statistique de Bose–Einstein, l'espace de Fock associé est défini formellement comme :*

$$\mathcal{F}(\mathcal{H}) \overset{\text{déf.}}{=} \bigoplus_{n=0}^{\infty} \mathfrak{S}_n\big(\mathcal{H}^{\otimes n}\big) = \mathbb{C} + \mathcal{H} \oplus \mathfrak{S}_2(\mathcal{H} \otimes \mathcal{H}) \oplus \cdots \tag{6.65}$$

ou $\mathfrak{S}_n$ est l'opérateur de symétrisation, voir l'éqn. (6.48a). Si ce type de particule suit la statistique de Fermi–Dirac, l'espace de Fock associé est défini formellement comme :

$$\mathcal{F}(\mathcal{H}) \overset{\text{déf.}}{=} \bigoplus_{n=0}^{\infty} \mathfrak{A}_n\big(\mathcal{H}^{\otimes n}\big) = \mathbb{C} \oplus \mathcal{H} \oplus \mathfrak{A}_2(\mathcal{H} \otimes \mathcal{H}) \oplus \cdots \tag{6.66}$$

ou $\mathfrak{A}_n$ est l'opérateur d'antisymétrisation, voir l'éqn. (6.48b).

En d'autres termes, l'espace de Fock associé à un hamiltonien $\mathcal{H}$ à une particule est la somme directe des espaces de Hilbert associé à des systèmes de zéro, une, deux, etc. particules identiques de ce type. Soit $\{|i\rangle\,,\, i \in \mathcal{I}\}$ une base de $\mathcal{H}$. À cette base il faut ajouter un état unique, l'*état fondamental* noté $|0\rangle$, qui représente un état du système ne contenant aucune particule. L'espace $\mathcal{F}(\mathcal{H})$ contient alors l'ensemble des vecteurs de la forme

$$|\psi\rangle = c_0\,|0\rangle + \sum_i c_i\,|i\rangle + \sum_{i,j} c_{ij}\,|i\rangle \otimes |j\rangle + \sum_{i,j,k} c_{ijk}\,|i\rangle \otimes |j\rangle \otimes |k\rangle + \cdots\,, \tag{6.67}$$

où les propriétés des coefficients complexes $c_{i_1 i_2 \cdots i_n}$ reflètent la statistique suivie par les particules, étant symétriques ou antisymétriques par permutation :

$$c_{\sigma(i_1)\sigma(i_2)\cdots\sigma(i_n)} = \epsilon_\sigma c_{i_1 i_2 \cdots i_n}\,, \quad \sigma \in S_N \tag{6.68}$$

L'espace de Fock $\mathcal{F}(\mathcal{H})$, que ce soit pour des bosons ou des fermions, hérite d'un produit hermitien découlant de celui sur $\mathcal{H}$. Les sous-espaces associés à des nombres de particules différents étant en somme directe, nous posons :

$$\big\langle \psi \big| \tilde{\psi} \big\rangle_{\mathcal{F}(\mathcal{H})} \overset{\text{déf.}}{=} \sum_{n=0}^{\infty} {}_n\big\langle \psi \big| \tilde{\psi} \big\rangle_n = \overline{c_0}\,\tilde{c}_0 + \sum_{i,j} \overline{c_{ij}}\,\tilde{c}_{ij} + \sum_{i,j,k} \overline{c_{ijk}}\,\tilde{c}_{ijk} + \cdots \tag{6.69}$$

c.-à.-d. que le produit hermitien de deux vecteurs de l'espace de Fock est la somme des produits hermitiens sur chaque sous-espace à n particules. De ce produit il découle naturellement une norme :

$$\big|\big|\psi\big|\big|_{\mathcal{F}(\mathcal{H})} \overset{\text{déf.}}{=} \sqrt{\sum_{n=0}^{\infty} {}_n\langle \psi | \psi \rangle_n}\,. \tag{6.70}$$

Les vecteurs de la forme (6.67) contenant une infinité de termes, la convergence de telles séries se pose ; nous imposerons que leur norme est finie. [4]

4. Mathématiquement parlant, l'espace vectoriel complexe (6.65) ou (6.66), muni du produit hermitien (6.69) ne constitue pas a priori un espace de Hilbert en lui-même, mais peut être complété de manière unique en un espace complet, qui est un espace de Hilbert. L'espace de Fock $\mathcal{F}(\mathcal{H})$ est alors défini rigoureusement comme étant donné par cette complétion.

6.4.1 Opérateurs de création et d'annihilation

Dans le formalisme des espaces de Fock, le nombre de particules n'étant pas fixé, il peut exister des processus physiques créant ou détruisant des particules ; cela est particulièrement évident dans les systèmes relativistes, pour lesquels la collision de particules peut en créer de nouvelles, si l'énergie de la collision dans le réferentiel du centre de masse est suffisamment élevée. Nous pouvons définir en toute généralité, en dehors d'un contexte physique particulier, des opérateurs sur $\mathcal{F}(\mathcal{H})$ associés à de tels processus.

Définition 5 (opérateurs de création et d'annihilation bosoniques.) *Soit $\mathcal{H}$ un espace de Hilbert, associé aux états physiques d'un boson, et $\{|i\rangle\,,\, i \in \mathcal{I}\}$ une base de $\mathcal{H}$. On considère l'espace de Fock $\mathcal{F}(\mathcal{H})$ associé à un système de tels bosons identiques, et $\mathfrak{S}_n\big(\mathcal{H}^{\otimes n}\big)$ le sous-espace associé aux états contenant n particules. L'opérateur de création associé à un état $|i\rangle$, noté $a_i^\dagger$, est l'application :*

$$a_i^\dagger \,:\, \left\{ \begin{array}{lcl} \mathfrak{S}_n\big(\mathcal{H}^{\otimes n}\big) & \to & \mathfrak{S}_{n+1}\big(\mathcal{H}^{\otimes (n+1)}\big) \\ \mathfrak{S}_n\big(\,|j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle\,\big) & \mapsto & \sqrt{n+1}\,\mathfrak{S}_{n+1}\big(\,|i\rangle \otimes |j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle\,\big) \end{array} \right. \tag{6.71}$$

L'opérateur d'annihilation associé à un état $|i\rangle$, noté a_i, est l'application :

$$a_i \,:\, \left\{ \begin{array}{lcl} \mathfrak{S}_n\big(\mathcal{H}^{\otimes n}\big) & \to & \mathfrak{S}_{n-1}\big(\mathcal{H}^{\otimes (n-1)}\big) \\ \mathfrak{S}_n\big(\,|j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle\,\big) & \mapsto & \frac{1}{\sqrt{n}} \sum\limits_k \langle i|j_k\rangle\, \mathfrak{S}_{n-1}\big(\,|j_1\rangle \otimes |j_2\rangle \otimes \cdots \cancel{|j_k\rangle} \cdots \otimes |j_n\rangle\,\big) \end{array} \right. \tag{6.72}$$

En d'autres termes, l'opérateur de création $a_i^\dagger$ ajoute au système une particule dans l'état $|i\rangle$, alors que l'opérateur d'annihilation a_i retire du système une particule dans cet état. Si l'état d'origine ne comprenait pas de particules dans un tel état, alors le résultat sera donné par le vecteur nul. En agissant sur le vide, état sans particules, nous avons évidemment, pour tout i, $a_i\,|0\rangle = 0$.

Définition 6 (opérateurs de création et d'annihilation fermioniques.) *Soit $\mathcal{H}$ un espace de Hilbert, associé aux états physiques d'un fermion, et $\{|i\rangle\,,\, i \in \mathcal{I}\}$ une base de $\mathcal{H}$. On considère l'espace de Fock $\mathcal{F}(\mathcal{H})$ associé à un système de tels fermions identiques, et $\mathfrak{A}_n\big(\mathcal{H}^{\otimes n}\big)$ le sous-espace associé aux états contenant n particules. L'opérateur de création associé à un état $|i\rangle$, noté $a_i^\dagger$, est l'application :*

$$a_i^\dagger \,:\, \left\{ \begin{array}{lcl} \mathfrak{A}_n\big(\mathcal{H}^{\otimes n}\big) & \to & \mathfrak{A}_{n+1}\big(\mathcal{H}^{\otimes (n+1)}\big) \\ \mathfrak{A}_n\big(\,|j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle\,\big) & \mapsto & \sqrt{n+1}\,\mathfrak{A}_{n+1}\big(\,|i\rangle \otimes |j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle\,\big) \end{array} \right. \tag{6.73}$$

L'opérateur d'annihilation associé à un état $|i\rangle$, noté a_i, est l'application :

$$a_i \,:\, \left\{ \begin{array}{lcl} \mathfrak{A}_n\big(\mathcal{H}^{\otimes n}\big) & \to & \mathfrak{A}_{n-1}\big(\mathcal{H}^{\otimes (n-1)}\big) \\ \mathfrak{A}_n\big(\,|j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle\,\big) & \mapsto & \frac{1}{\sqrt{n}} \sum\limits_k (-1)^k\, \langle i|j_k\rangle\, \mathfrak{A}_{n-1}\big(\,|j_1\rangle \otimes |j_2\rangle \otimes \cdots \cancel{|j_k\rangle} \cdots \otimes |j_n\rangle\,\big) \end{array} \right. \tag{6.74}$$

Suivant la statistique de Fermi–Dirac, si un état $|\psi\rangle_n$ à n particules contient déjà une particule dans l'état $|i\rangle$, alors cet état sera annihilé par l'opérateur de création associé à ce même état, car deux fermions ne peuvent se trouver dans le même état quantique : $a_i^\dagger |\psi\rangle_n = 0$. Cela découle immédiatement de l'opérateur d'antisymétrisation apparaissant dans le membre de droite de l'éqn. (6.73).

Que ce soit pour des bosons ou des fermions, l'action des opérateurs de création et d'annihilation, définie sur les sous-espaces à n particules, s'étend immédiatement à une action sur l'espace de Fock, c.-à.-d. à des applications de $\mathcal{F}(\mathcal{H})$ dans lui-même. Nous laissons en exercice la preuve de la propriété suivante, suggérée par les notations utilisées :

Propriété 5 *Les opérateurs de création et d'annihilation associés au même état* $|i\rangle$, $a_i^\dagger$ *et* a_i, *sont adjoints l'un de l'autre.*

Prenons le cas bosonique. Nous avons d'une part :

$$\iota\Big(\mathfrak{S}_n |j_1\rangle \otimes \cdots \otimes |j_n\rangle\,,\, a_i^\dagger \mathfrak{S}_{n-1} |k_1\rangle \otimes \cdots \otimes |k_{n-1}\rangle\Big) = \sqrt{n} \sum_i \iota\Big(\mathfrak{S}_n |j_1\rangle \otimes \cdots \otimes |j_n\rangle\,,\, \mathfrak{S}_n |i\rangle \otimes |k_1\rangle \otimes \cdots \otimes |k_{n-1}\rangle\Big) \quad (6.75)$$

et d'autre part :

$$\iota\Big(a_i \mathfrak{S}_n |j_1\rangle \otimes \cdots \otimes |j_n\rangle\,,\, \mathfrak{S}_{n-1} |k_1\rangle \otimes \cdots \otimes |k_{n-1}\rangle\Big) \frac{1}{\sqrt{n}} \sum_k \langle i|j_k\rangle\, \iota\Big(\mathfrak{S}_{n-1}\big(|j_1\rangle \otimes |j_2\rangle \otimes \cdots \cancel{|j_k\rangle} \cdots \otimes |j_n\rangle\big),\, \mathfrak{S}_{n-1} |k_1\rangle \otimes \cdots \otimes |k_{n-1}\rangle\Big) \quad (6.76)$$

Pour passer de la première expression à la seconde on peut, dans l'éqn. (6.75), pour chacun des termes apparaissant dans la symétrisation de $|i\rangle \otimes |k_1\rangle \otimes \cdots \otimes |k_{n-1}\rangle$, contracter $|i\rangle$ avec l'un quelconque des $\langle j_k|$, donnant une somme similaire à l'eqn. (6.76). Le passage du préfacteur $\sqrt{n}$ à $1/\sqrt{n}$ s'explique par la différence de normalisation entre $\mathfrak{S}_n$ et $\mathfrak{S}_{n-1}$, voir l'éqn. (6.48a). Naturellement le raisonnement est identique pour les fermions.

Les opérateurs de création et d'annihilation bosoniques rappellent évidemment ceux de l'oscillateur harmonique, et satisfont des propriétés similaires. Nous avons en particulier :

Propriété 6 *Soit* $\mathcal{F}(\mathcal{H})$ *associé à un système de bosons identiques. Les opérateurs de création* $a_j^\dagger$ *et d'annihilation* a_i, *associés respectivement aux états* $|j\rangle\,,\, |i\rangle \in \mathcal{H}$, *satisfont les relations de commutation suivantes :*

$$\Big[\, a_i,\, a_j^\dagger \,\Big] \;=\; \langle i|j\rangle\, \mathcal{I}\,, \quad \Big[\, a_i,\, a_j \,\Big] = \Big[\, a_i^\dagger,\, a_j^\dagger \,\Big] = 0\,. \quad (6.77)$$

En effet, nous avous premièrement

$$\begin{aligned} a_i a_j^\dagger \mathfrak{S}_n\big(|j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle\big) &= \sqrt{n+1}\, a_i \mathfrak{S}_{n+1}\big(|j\rangle \otimes |j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle\big) \\ &= \langle i|j\rangle\, \mathfrak{S}_n\big(|j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle\big) + \sum_k \langle i|j_k\rangle\, \mathfrak{S}_n\big(|j\rangle \otimes |j_1\rangle \otimes |j_2\rangle \otimes \cdots \cancel{|j_k\rangle} \cdots \otimes |j_n\rangle\big) \end{aligned} \quad (6.78)$$

et deuxièmement

$$a_j^\dagger a_i \mathfrak{S}_n\big(|j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle\big) = \tfrac{1}{\sqrt{n}} a_j^\dagger \sum_k \langle i|j_k\rangle\, \mathfrak{S}_{n-1}\big(|j_1\rangle \otimes |j_2\rangle \otimes \cdots \cancel{|j_k\rangle} \cdots \otimes |j_n\rangle\big)$$
$$= \sum_k \langle i|j_k\rangle\, \mathfrak{S}_n\big(|j\rangle \otimes |j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes \cancel{|j_k\rangle} \otimes \cdots \otimes |j_n\rangle\big) \tag{6.79}$$

prouvant la première propriété demandée ; les deux autres sont laissées en exercice □

Si les particules concernées suivent la statistique de Fermi–Dirac, nous obtenons des relations d'*anti-commutation* à la place des relation de commutation pour les bosons.

Propriété 7 *Soit $\mathcal{F}(\mathcal{H})$ associé à un système de fermions identiques. Les opérateurs de création $a_j^\dagger$ et d'annihilation a_i, associés respectivement aux états $|j\rangle, |i\rangle \in \mathcal{H}$, satisfont les relations d'anti-commutation suivantes :*

$$\Big\{ a_i,\, a_j^\dagger \Big\} = \langle i|j\rangle\, \mathcal{I}\,, \quad \Big\{ a_i,\, a_j \Big\} = \Big\{ a_i^\dagger,\, a_j^\dagger \Big\} = 0\,. \tag{6.80}$$

où

$$\{\, a,\, b \,\} \overset{\text{déf.}}{=} ab + ba\,. \tag{6.81}$$

En effet, nous avous premièrement

$$a_i a_j^\dagger \mathfrak{A}_n\big(|j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle\big) = \sqrt{n+1}\, a_i \mathfrak{A}_{n+1}\big(|j\rangle \otimes |j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle\big)$$
$$= \langle i|j\rangle\, \mathfrak{A}_n\big(|j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle\big) + \sum_k (-1)^{k+1} \langle i|j_k\rangle\, \mathfrak{A}_{n-1}\big(|j\rangle \otimes |j_1\rangle \otimes |j_2\rangle \otimes \cdots \cancel{|j_k\rangle} \cdots \otimes |j_n\rangle\big) \tag{6.82}$$

le vecteur « retiré » du produit tensoriel par l'opérateur d'annihilation dans le second terme étant en $(k+1)$-ième position, et deuxièmement

$$a_j^\dagger a_i \mathfrak{S}_n\big(|j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes |j_n\rangle\big) = \tfrac{1}{\sqrt{n}} a_j^\dagger \sum_k (-1)^k \langle i|j_k\rangle\, \mathfrak{S}_{n-1}\big(|j_1\rangle \otimes |j_2\rangle \otimes \cdots \cancel{|j_k\rangle} \cdots \otimes |j_n\rangle\big)$$
$$= \sum_k (-1)^k \langle i|j_k\rangle\, \mathfrak{S}_n\big(|j\rangle \otimes |j_1\rangle \otimes |j_2\rangle \otimes \cdots \otimes \cancel{|j_k\rangle} \otimes \cdots \otimes |j_n\rangle\big) \tag{6.83}$$

prouvant la première propriété demandée ; les deux autres sont laissées en exercice. □

Propriété 8 *Les opérateurs de création et d'annihilation fermioniques sont nilpotents :*

$$\big(a_i\big)^2 = \big(a_i^\dagger\big)^2 = 0\,. \tag{6.84}$$

Cette propriété découle trivialement des relations d'anti-commutation ci-dessus. Notons en particulier que la nilpotence de l'opérateur de création $a_i^\dagger$ reflète fidèlement le principe d'exclusion de Pauli : il n'est pas possible de créer deux particules fermioniques dans le même état quantique.

6.4.2 Nombres d'occupation

Comme nous l'avions évoqué pour un système de deux particules, voir l'exemple 3, il est plus commode de caractériser les états simplement par le nombre de particules, bosons ou fermions, qui se trouve dans un état $|i\rangle \in \mathcal{H}$ donné. La collection de ces *nombres d'occupation* n_i caractérise entièrement l'état du système, et un vecteur de base normalisé de $\mathcal{F}(\mathcal{H})$, associé à la base $\{|i\rangle\,,\, i \in \mathcal{I}\}$ de $\mathcal{H}$, sera noté

$$|\mathbf{n}\rangle \overset{\text{déf.}}{=} |n_{i_1}, n_{i_2}, n_{i_3}, \ldots\rangle \tag{6.85}$$

c.-à.-d. etiqueté par la collection des nombres d'occupation associés à tous les vecteurs de base de $\mathcal{H}$. Ce formalisme est plus simple car il n'est pas nécessaire de faire apparaître explicitement l'opérateur de symétrisation (resp. anti-symétrisation) pour les bosons (resp. les fermions).

Propriété 9 *Si l'espace de Fock $\mathcal{F}(\mathcal{H})$ est associé à un système de bosons identiques, les nombres d'occupation n_i caractérisant un vecteur de base sont des entiers positifs quelconques, reflétant la statistique de Bose–Einstein. Pour des fermions identiques, nous avons, pour tout $i \in \mathcal{I}$, $n_i \in \{0, 1\}$, reflétant la statistique de Fermi–Dirac.*

Les opérateurs de création et d'annihilation ont une action très simple dans ce formalisme. En tenant compte de la normalisation états symétrisés et antisymétrisés, nous avons pour des **bosons** :

$$a_{i_k}^\dagger\, |n_{i_1}, n_{i_2}, \ldots, n_{i_k}, \ldots\rangle = \sqrt{n_{i_k} + 1}\, |n_{i_1}, n_{i_2}, \ldots, n_{i_k} + 1, \ldots\rangle\ , \tag{6.86a}$$

$$a_{i_k}\, |n_{i_1}, n_{i_2}, \ldots, n_{i_k}, \ldots\rangle = \sqrt{n_{i_k}}\, |n_{i_1}, n_{i_2}, \ldots, n_{i_k} - 1, \ldots\rangle\ , \tag{6.86b}$$

si $n_{i_k} \pm 1$ est dans le domaine autorisé c.-à.-d. $n_{i_k} \pm 1 \in \mathbb{N}$ et le vecteur nul sinon. Le facteur de normalisation s'explique de la manière suivante. D'après l'eq. (6.60), le vecteur d'état symétrisé normalisé à $N = n_{i_1} + n_{i_2} + \cdots$ particules est :

$$|n_{i_1}, n_{i_2}, \ldots, n_{i_k}, \ldots\rangle = \sqrt{\tfrac{N!}{n_{i_1}! n_{i_2}! \ldots}}\, \mathfrak{S}_N\, (|i_1\rangle \otimes \cdots)\ . \tag{6.87}$$

Par exemple, l'action de $a_{i_1}^\dagger$ sur cet état normalisé donne, d'après l'eqn. (6.71),

$$a_{i_1}^\dagger\, |n_{i_1}, n_{i_2}, \ldots, n_{i_k}, \ldots\rangle = \sqrt{\tfrac{N!}{n_{i_1}! n_{i_2}! \ldots}}\, a_{i_1}^\dagger\, \mathfrak{S}_N\, (|i_1\rangle \otimes \cdots) = \sqrt{\tfrac{(N+1)!}{n_{i_1}! n_{i_2}! \ldots}}\, \mathfrak{S}_{N+1}\, (|i_1\rangle \otimes |i_1\rangle \cdots) \tag{6.88}$$

$$= \sqrt{n_{i_1} + 1}\, \sqrt{\tfrac{(N+1)!}{(n_{i_1}+1)! n_{i_2}! \ldots}}\, \mathfrak{S}_{N+1}\, (|i_1\rangle \otimes |i_1\rangle \cdots) \tag{6.89}$$

$$= \sqrt{n_{i_1} + 1}\, |n_{i_1} + 1, n_{i_2}, \ldots, n_{i_k}, \ldots\rangle\ . \tag{6.90}$$

Pour des **fermions** on a de même, en utilisant l'eqn. (6.54) :

$$a_{i_k}^\dagger\, |n_{i_1}, n_{i_2}, \ldots, n_{i_k}, \ldots\rangle = (-1)^{\sum_{\ell<k} n_{i_\ell}} \sqrt{n_{i_k} + 1}\, |n_{i_1}, n_{i_2}, \ldots, n_{i_k} + 1, \ldots\rangle\ , \tag{6.91a}$$

$$a_{i_k}\, |n_{i_1}, n_{i_2}, \ldots, n_{i_k}, \ldots\rangle = (-1)^{\sum_{\ell<k} n_{i_\ell}} \sqrt{n_{i_k} + 1}\, \sqrt{n_{i_k}}\, |n_{i_1}, n_{i_2}, \ldots, n_{i_k} - 1, \ldots\rangle\ , \tag{6.91b}$$

si $n_{i_k} \pm 1$ est dans le domaine autorisé c.-à.-d. $n_{i_k} \pm 1 \in \{0, 1\}$ et le vecteur nul sinon. Le signe peut se comprendre de la manière suivante. L'état $|n_{i_1}, n_{i_2}, \ldots, n_{i_k}, \ldots\rangle$ s'obtient par action des opérateurs de création sur l'état fondamental. On peut par convention les ranger dans l'ordre des étiquettes du vecteur :

$$|n_{i_1}, n_{i_2}, \ldots, n_{i_k}, \ldots\rangle = \left(a_{i_1}^\dagger\right)^{n_{i_1}} \left(a_{i_2}^\dagger\right)^{n_{i_2}} \cdots \left(a_{i_k}^\dagger\right)^{n_{i_k}} \cdots |0\rangle \ , \quad n_{i_k} \in \{0, 1\} . \tag{6.92}$$

Si on fait agir l'opérateur d'annihilation a_{i_k} sur cet état, nous avons, en utilisant les relations d'anticommutation (6.80),

$$\begin{aligned} a_{i_k} \left(a_{i_1}^\dagger\right)^{n_{i_1}} \left(a_{i_2}^\dagger\right)^{n_{i_2}} \cdots \left(a_{i_k}^\dagger\right)^{n_{i_k}} \cdots |0\rangle &= (-1)^{n_{i_1}} \left(a_{i_1}^\dagger\right)^{n_{i_1}} a_{i_k} \left(a_{i_2}^\dagger\right)^{n_{i_2}} \cdots \left(a_{i_k}^\dagger\right)^{n_{i_k}} \cdots |0\rangle \\ &= (-1)^{n_{i_1}+n_{i_2}+\cdots+n_{i_{k-1}}} \left(a_{i_1}^\dagger\right)^{n_{i_1}} \left(a_{i_2}^\dagger\right)^{n_{i_2}} \cdots a_{i_k} \left(a_{i_k}^\dagger\right)^{n_{i_k}} \cdots |0\rangle \end{aligned} \tag{6.93}$$

Si $n_{i_k} = 0$, on obtient le vecteur nul car $a_{i_k} |0\rangle = 0$. Si $n_{i_k} = 1$, on utilise

$$\left(\cdots a_{i_k} a_{i_k}^\dagger \cdots\right) |0\rangle = -\cancel{\left(\cdots a_{i_k}^\dagger a_{i_k} \cdots |0\rangle\right)} + \left(\cdots \underbrace{\left\{ a_{i_k},\, a_{i_k}^\dagger \right\}}_{=1} \cdots\right) |0\rangle \ . \tag{6.94}$$

Comme pour l'oscillateur harmonique, il est possible de créer un état arbitraire à n particules par action itérée des opérateurs de création sur l'état fondamental $|0\rangle$ ne contenant pas de particules. Le nombre moyen de particules qui se trouvent dans un état $|i\rangle$ est donné par la valeur moyenne de l'observable

$$N_i \overset{\text{déf.}}{=} a_i^\dagger a_i \, , \tag{6.95}$$

et le nombre moyen de particules par la valeur moyenne de l'observable

$$N \overset{\text{déf.}}{=} \sum_{i \in \mathcal{I}} N_i \, . \tag{6.96}$$

L'application du formalisme exposé dans cette section au problème de la quantification du champ électromagnétique sera développée dans le chapitre 20.

Exemple 4 (phonons). Dans la limite de basse énergie, les degrés de liberté d'un cristal isolant sont associés aux modes de vibration de celui-ci. Ces modes de vibrations peuvent être considérés comme des *quasi-particules*, c.-à.-d. des excitations élémentaires du système, valables dans un certain domaine d'énergie, ne correspondant pas à des particules élémentaires, appelées *phonons*. Le cristal peut être modélisé comme un réseau d'oscillateur harmoniques, les excitations du système étant de faible amplitude dans la limite de basse énergie considérée.

Un modèle très simple est donné par une chaîne unidimensionnelle d'oscillateurs couplés, composée de N masses ponctuelles de masse m, séparées à l'équlibre par une distance a, et reliées par des ressorts de même raideur R. Rappelons que les modes propres du système,

associés à des déplacements autour de la position d'équilibre de la forme $u_n = e^{ikna-\omega_k t}$ pour $n = 1, \ldots, N$, satisfont la relation de dispersion

$$\omega_k = 2\sqrt{\frac{R}{m}} \sin\left(\tfrac{|k|a}{2}\right) \ , \quad k \in [-\pi/a, \pi/a] \, , \tag{6.97}$$

avec en outre, pour des conditions aux limites périodiques, $k \in \frac{2\pi}{Na}\mathbb{Z}$. Après quantification canonique, le système est décrit comme un ensemble d'oscillateurs harmoniques indépendants, de hamiltonien

$$\mathcal{H} = \sum_k \mathcal{H}_k = \sum_k \hbar\omega_k (a_k^\dagger a_k + 1/2) \, . \tag{6.98}$$

Une base de l'espace des états, qui est également une base d'états propres du hamiltonien, est donnée par l'ensemble des vecteurs :

$$|\mathbf{n}\rangle \overset{\text{déf.}}{=} |\cdots, n_k, \cdots\rangle \ , \quad n_k \in \mathbb{N} \ , \quad a_k^\dagger a_k \, |\mathbf{n}\rangle = n_k \, |\mathbf{n}\rangle \ , \tag{6.99}$$

caractérisés par le nombre d'excitations associé à chacun des oscillateurs harmoniques correspondant aux différents modes propres du système.

Du point de vue adopté ici, on considère les phonons comme des (quasi-)particules identiques, suivant la statistique de Bose–Einstein, et l'espace des états associé au hamiltonien (6.98) correspond à l'espace de Fock pour ces bosons. Chaque terme dans la somme, ou dit autrement chaque pulsation propre ω_k, correspond à l'un des états physiques dans lequel peut se trouver un phonon.

À retenir

— La construction d'un espace produit tensoriel ;
— opérateurs sur un espace produit tensoriel ;
— la notion de mesure partielle ;
— la notion d'intrication quantique ;
— le problème posé par les particules identiques ;
— la structure du groupe symétrique et son action sur l'espace des états ;
— le postulat de symétrisation et antisymétrisation pour les bosons et fermions respectivement ;
— la construction de vecteurs d'état symétrisés et antisymétrisés.
— la notion d'espace de Fock, les opérateurs de création et d'annihilation.

Bibliographie

[1] W. Fulton et J. Harris, "Representation Theory : A First Course," Springer, 1991

[2] M. O. Mewes, G. Ferrari, F. Schreck, A. Sinatra et C. Salomon, "Simultaneous Magneto-Optical Trapping of Two Lithium Isotopes," Physical Review A. 61, 1999

Cours n°7

Électrons dans les solides

Un grand nombre des propriétés physiques des solides, en particulier leurs propriétés de conduction électrique, sont liées au comportement des électrons. Ces derniers peuvent être modélisés, lorsque les interactions inter-électrons ne jouent pas un rôle important, comme un gaz de fermions identiques, soit libres en première approximation, soit soumis au potentiel effectif périodique engendré par les ions du cristal, qui pourra être pris en compte de manière perturbative.

7.1 Gaz de fermions libres, surface de Fermi

En première approximation, il est possible de considérer que les électrons sont non seulement sans interactions entre eux, mais ne sont soumis qu'à un potentiel moyen *constant* dans le volume du solide, insensibles de ce fait à la structure périodique du cristal ; ce modèle est en particulier pertinent pour les électrons de valence des métaux alcalins voir la section 2.4 du chapitre 2.

En absorbant le potentiel constant dans le cristal par une redéfinition de l'énergie, le système constitué de N_e électrons correspond donc à N_e fermions libres indiscernables, associé à l'espace des états anti-symétrisé $\mathfrak{A}_{N_e}\big(\mathcal{H}_e^{\otimes N_e}\big)$, où $\mathcal{H}_e$ est l'espace de Hilbert d'un électron libre, associé au hamiltonien individuel

$$\mathcal{H}_e = \frac{1}{2m_e}\vec{\mathcal{P}}^{\,2}\,, \tag{7.1}$$

voir la section 6.2 du chapitre 6. Nous considérerons ce gaz d'électrons dans une boîte cubique de côté L, et imposerons des conditions aux limites périodiques dans toutes les directions. Les vecteurs d'ondes des électrons prennent ainsi les valeurs discrètes

$$\vec{k} = \frac{2\pi}{L}\big(\ell_x\vec{e}_x + \ell_y\vec{e}_y + \ell_z\vec{e}_z\big)\ , \quad \ell_{x,y,z} \in \mathbb{Z}\,, \tag{7.2}$$

et le spectre discret du hamiltonien individuel est donné par

$$\sigma\big(\mathcal{H}_e\big) = \left\{ \frac{\hbar^2\vec{k}^2}{2m_e}\,,\ \vec{k} \in \frac{2\pi}{L}\big(\mathbb{Z}\vec{e}_x + \mathbb{Z}\vec{e}_y + \mathbb{Z}\vec{e}_z\big) \right\}. \tag{7.3}$$

En réalité les électrons sont des particules de spin 1/2, et les degrés de liberté de spin sont décrits par un espace hermitien de dimension 2, donc isomorphe à $\mathbb{C}^2$, voir le chapitre 13. L'espace des états du gaz d'électrons est donc plus correctement donné par

$$\boxed{\mathcal{H}_{N_e} \overset{\text{déf.}}{=} \mathfrak{A}_{N_e}\Big(\big(\mathcal{H}_e \otimes \mathbb{C}^2\big)^{\otimes N_e}\Big)} \tag{7.4}$$

7.1.1 Densité d'états

On considère, dans un espace à d dimensions, une particule libre sans spin confinée dans un (hyper-)cube de côté L, avec des conditions aux limites périodiques. Les vecteurs d'onde $\vec{k}$ sont donc à valeurs dans $\frac{2\pi}{L}\big(\mathbb{Z} \oplus \mathbb{Z} \oplus \cdots \oplus \mathbb{Z}\big)$.

Propriété 1 (volume de l'hypersphère unité). *Le volume d'une hypersphère de rayon unité en dimension d est donné par :*

$$\boxed{\Omega_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}} \tag{7.5}$$

Considérons une intégrale gaussienne à d dimensions. Nous avons d'une part, en coordonnées cartésiennes,

$$I = \int_{-\infty}^{\infty} dx^1 \cdots \int_{-\infty}^{\infty} dx^d e^{-(x^1)^2-(x^2)^2-\cdots-(x^d)^2} = \Big(\int_{-\infty}^{\infty} dx e^{-x^2}\Big)^d = \pi^{d/2}\,, \tag{7.6}$$

et d'autre part en coordonnées hyper-sphériques

$$I = \int_0^{\infty} r^{d-1} dr \int d\Omega_d e^{-r^2} = \Omega_d \int_0^{\infty} r^{d-1} dr\, e^{-r^2}\,, \tag{7.7}$$

où Ω_d est le volume de l'hyper-sphère. L'intégrale restante se calcule par un changement de variables,

$$\int_0^{\infty} r^{d-1} dr\, e^{-r^2} \overset{t=r^2}{=\!=} \frac{1}{2}\int_0^{\infty} t^{d/2-1} e^{-t} dt = \frac{1}{2}\Gamma(d/2)\,. \tag{7.8}$$

Nous avons ainsi par identification

$$\Omega_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}\,, \tag{7.9}$$

qui est le résultat attendu. Rappelons que $\Gamma(z+1) = \Gamma(z)$ et qu'ainsi, pour n entier, $\Gamma(n+1) = n\Gamma(n) = n!\Gamma(1) = n!$. On montre aussi facilement que $\Gamma(1/2) = \sqrt{\pi}$.

Les modes étant espacés de $2\pi/L$ dans chaque direction de l'espace, le nombre de modes d'impulsion $\vec{k}$ de norme inférieure à K, dans la limite où $KL \gg 1$, c.-à.-d. pour laquelle

les modes forment un « quasi-continuum », est donné avec une bonne approximation par l'intégrale suivante :

$$\Phi(K) = \left(\frac{L}{2\pi}\right)^d \int_{\|\vec{k}\|<K} d^d k = \left(\frac{L}{2\pi}\right)^d \int_0^K \kappa^{d-1} d\kappa \int d\Omega_d = \left(\frac{L}{2\pi}\right)^d \frac{2\pi^{d/2}}{\Gamma(d/2)} \frac{1}{d} K^d \,. \tag{7.10}$$

La densité d'états correspondante est donc donnée par :

$$\rho(K) dK = d\Phi(K) = \left(\frac{L}{2\pi}\right)^d \frac{2\pi^{d/2}}{\Gamma(d/2)} K^{d-1} dK \,. \tag{7.11}$$

Nous pouvons en déduire également le nombre de modes d'énergie inférieure à $E_K = \frac{\hbar^2 K^2}{2m}$. Nous avons

$$\Phi(K) = \int_0^K \rho(\kappa) d\kappa = \int_0^{E_K} \rho(E) dE \tag{7.12}$$

D'où nous obtenons, avec $V = L^d$ le volume de l'échantillon,

$$\rho(E) = \frac{V}{(4\pi)^{d/2}\Gamma(d/2)} \left(\frac{2m}{\hbar^2}\right)^{d/2} E^{d/2-1} \,. \tag{7.13}$$

Dans le cas bidimensionnel, cela donne

$$\rho_{2d}(E) = \frac{V}{2\pi} \frac{m}{\hbar^2} \,, \tag{7.14}$$

qui est constante, et à trois dimensions,

$$\rho_{3d}(E) = \frac{V}{4\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} \sqrt{E} \,, \tag{7.15}$$

qui croît avec l'énergie.

7.1.2 Surface de Fermi, température de Fermi

Considérons un modèle de gaz parfait pour les électrons d'un solide. À température nulle, le système va se trouver dans l'état fondamental, c.-à.-d. la borne inférieure du spectre du hamiltonien du système dans l'espace des états $\mathcal{H}_N$ donné par l'éqn. (7.4), antisymétrisé en raison de la statistique de Fermi–Dirac suivie par les électrons. Les vecteurs de $\mathcal{H}_N$ étant antisymétriques par échange de des vecteurs d'états de deux électrons,le électrons doivent tous se trouver dans des états associés à des vecteurs distincts de $\mathcal{H}_e \otimes \mathbb{C}^2$. L'état fondamental est donc obtenu en remplissant les niveaux d'énergie de $\mathcal{H}_e$ à partir de celui de plus basse énergie, avec deux électrons par niveau.

En utilisant la densité d'états en énergie (7.13), et en tenant compte de la dégénérescence de spin, nous obtenons ainsi la condition :

$$N_e = 2 \int_0^{E_F} \rho(E) dE \,, \tag{7.16}$$

où la borne supérieure correspond à l'énergie du niveau le plus élevé contenant un électron dans cet état fondamental ; cette énergie E_F est appelée *énergie de Fermi* du système. Explicitement, nous trouvons, pour un gaz d'électrons de densité $n_e = N_e/V$, la valeur suivante de l'énergie de Fermi,

$$n_e = \frac{2}{(4\pi)^{d/2}\Gamma(d/2)}\left(\frac{2m_e}{\hbar^2}\right)^{d/2}\int_0^{E_F} E^{d/2-1}dE = \frac{2}{(4\pi)^{d/2}\Gamma(1+d/2)}\left(\frac{2m_e}{\hbar^2}\right)^{d/2} E_F^{d/2}$$

$$\Longrightarrow \quad E_F = \frac{\hbar^2}{2m_e}\underbrace{\left(2^{d-1}\pi^{d/2}\Gamma(1+d/2)n_e\right)^{2/d}}_{k_F^2} \tag{7.17}$$

exprimée en fonction du nombre d'onde k_F correspondant. Dans le cas tridimensionnel nous en déduisons que :

$$\boxed{E_F = \frac{\hbar^2 k_F^2}{2m_e}\,, \quad k_F = \left(3\pi^2 n_e\right)^{1/3}} \tag{7.18}$$

La sphère de rayon k_F dans l'espace des vecteurs d'onde est appelée *surface de Fermi* du système. Plus généralement, en physique du solide, il s'agit de la surface, dans l'espace des impulsions, décrite par les niveaux occupés de plus haute énergie lorsque la température est nulle. Étant donné que tous les niveaux d'énergie inférieure sont complètement occupés, les excitations du système à température non-nulle se situent au voisinage de cette surface, dont l'étude est cruciale dans la compréhension de la physique du système.

À l'échelle d'énergie (7.18) on peut également faire correspondre une échelle de température, appelée température de Fermi :

$$T_F = \frac{\hbar^2 k_F^2}{2m_e k_B}\,, \quad k_F = \left(3\pi^2 n_e\right)^{1/3}. \tag{7.19}$$

Le gaz d'électrons devient faiblement dégénéré (c.-à.-d. au comportement proche d'un gaz parfait classique) pour des températures $T \gg T_F$.

Dans l'exemple du sodium métallique, la densité électornique étant de l'ordre de $n_e \simeq 10^{28} m^{-3}$, on trouve une énergie de Fermi de l'ordre de 5 eV, donnant une température de Fermi de l'ordre de $T_F = 12000$ K. Ainsi, à température ambiante, le gaz d'électrons est très fortement dégénéré et l'état du système très proche de l'état fondamental.

L'énergie moyenne du gaz d'électrons dans l'état fondamental est donc, pour un gaz tridimensionnel,

$$\langle E\rangle = 2\int_0^{k_F}\rho(k)dk\,\frac{\hbar^2k^2}{2m} = \frac{V}{\pi^2}\frac{\hbar^2}{2m_e}\int_0^{k_F}k^4dk = \frac{V}{\pi^2}\frac{\hbar^2}{2m_e}\frac{1}{5}k_F^5 = \frac{3}{5}N_eE_F\,. \tag{7.20}$$

À cette énergie peut être associée une pression, appelée *pression de dégénérescence*, d'origine purement quantique :

$$P = -\frac{\partial}{\partial V}\langle E\rangle = \frac{2}{3V}\langle E\rangle = \frac{2}{5}n_eE_F\,. \tag{7.21}$$

Dans les astres appelés *naines blanches*, stade ultime d'évolution d'étoiles de masses voisines de celle du soleil, la contraction gravitationnelle est compensée par la pression de dégénérescence du gaz d'électrons fortement dégénéré. Bien qu'ayant une température très élevée (de l'ordre de 10000 K), la densité extrême qui y règne donne une température de Fermi bien plus grande que élevée que celle-ci, voir l'éqn. (7.19).

7.2 Potentiel périodique et structure de bandes

On considérera par la suite, tout en continuant à négliger les interactions entre électrons, le rôle de la structure cristalline sur la physique du système constitué par les électrons dans le solide. Les interactions entre les ions et les électrons vont alors se manifester, dans cette description approchée, par un *potentiel effectif* engendré par les ions dont la périodicité reflète celle du réseau cristallin, voir le chapitre 5. Le hamiltonien individuel d'un électron est alors de la forme :

$$\mathcal{H}_e = \frac{1}{2m_e}\vec{\mathcal{P}}^{\,2} + \mathcal{V}\big(\vec{\mathcal{X}}\big)\ , \quad \text{t.q.} \quad \forall \vec{\lambda} \in \Lambda\ , \quad \mathcal{T}_{\vec{\lambda}}^{\dagger}\mathcal{V}(\vec{\mathcal{X}})\mathcal{T}_{\vec{\lambda}} = \mathcal{V}\big(\vec{\mathcal{X}} + \vec{\lambda}\big) = \mathcal{V}\big(\vec{\mathcal{X}}\big)\,. \tag{7.22}$$

D'après le théorème de Bloch, il est alors possible de trouver des vecteurs propres de $\mathcal{H}_e$ se transformant dans une représentation irréductible du réseau Λ, associés à des fonctions d'onde de la forme :

$$\psi_{\vec{k}}(\vec{x}) = e^{i\vec{k}\cdot\vec{x}}u_{\vec{k}}(\vec{x})\ , \quad \text{t.q.} \quad \forall \vec{\lambda} \in \Lambda\ , \quad u_{\vec{k}}(\vec{x} + \vec{\lambda}) = u_{\vec{k}}(\vec{x})\,, \tag{7.23}$$

où $\vec{k}$ est défini à un vecteur du réseau réciproque $\Lambda^\vee$ près ; on peut choisir un représentant appartenant à $\text{PZB}(\Lambda^\vee)$, première zone de Brillouin du réseau réciproque.

Dans le contexte actuel (système de taille finie) le vecteur d'onde $\vec{k}$ ne prend que les valeurs discrètes compatibles avec les conditions aux limites périodiques sur les faces du cristal. Choisissant comme au chapitre 5 des conditions aux limites compatibles avec la symétrie du réseau, nous avons

$$\vec{k} \in \tfrac{1}{N}\Lambda^\vee \cap \text{PZB}\big(\Lambda^\vee\big)\,. \tag{7.24}$$

Les différents états propres obtenus pour un vecteur $\vec{k}$ donné choisi dans cet ensemble discret sont repérés par un entier n, ou plusieurs entiers, appelés *indices de bande*. Pour un indice de bande donné, dans la limite où le nombre de mailles du système est très grand (cristal macroscopique), l'ensemble des énergies $\vec{E}_{\vec{k}}$ obtenues pour tous les vecteurs d'onde $\vec{k} \in \text{PZB}(\Lambda)$, qui forment un quasi-continuum, est une *bande d'énergie* du système.

7.2.1 Structure de bandes

La fonction à valeurs complexes $u_{\vec{k}}$ apparaissant dans la décomposition (7.23) étant périodique sur le réseau Λ, il est possible de la décomposer en série de Fourier sur le réseau réciproque correspondant :

$$u_{\vec{k}}(\vec{x}) = \sum_{\vec{q}\in\Lambda^\vee} e^{i\vec{q}\cdot\vec{x}}\,\widehat{u}_{\vec{k}}(\vec{q})\ , \quad \widehat{u}_{\vec{k}}(\vec{q}) = \tfrac{1}{\mathcal{V}(\Lambda)}\int_{\mathcal{F}_{\mathcal{B}}(\Lambda)} d^3x\, e^{-i\vec{q}\cdot\vec{x}}\,u_{\vec{k}}(\vec{x})\,, \tag{7.25}$$

l'intégrale s'effectuant sur un domaine fondamental du réseau, voir l'éqn. (5.34). Il est naturellement possible de développer le potentiel périodique créé par les ions de la même manière :

$$V(\vec{x}) = \sum_{\vec{q}\in\Lambda^\vee} e^{i\vec{q}\cdot\vec{x}}\,\widehat{V}(\vec{q})\,, \quad \widehat{V}(\vec{q}) = \tfrac{1}{\mathcal{V}(\Lambda)}\int_{\mathcal{F}_{\mathcal{B}}(\Lambda)} d^3x\, e^{-i\vec{q}\cdot\vec{x}}\,V(\vec{x})\,, \tag{7.26}$$

avec ici, le potentiel étant réel, la relation $\widehat{V}(-\vec{q}) = \widehat{V}(\vec{q})^*$. Rappelons que pour un produit de fonctions, nous avons

$$f(\vec{x})g(\vec{x}) = \sum_{\vec{p},\vec{q}\in\Lambda^\vee} e^{i(\vec{q}+\vec{p})\cdot\vec{x}}\,\widehat{f}(\vec{p})\,\widehat{g}(\vec{q}) = \sum_{\vec{P}\in\Lambda^\vee} e^{i\vec{P}\cdot\vec{x}}\left(\sum_{\vec{q}\in\Lambda^\vee}\widehat{f}(\vec{P}-\vec{q})\,\widehat{g}(\vec{q})\right), \tag{7.27}$$

c.-à.-d. un produit de convolution discret, et ainsi

$$\widehat{f\,g}(\vec{p}) = \sum_{\vec{q}\in\Lambda^\vee}\widehat{f}(\vec{p}-\vec{q})\,\widehat{g}(\vec{q})\,. \tag{7.28}$$

La recherche des états propres du hamiltonien (7.22) peut alors se ramener, par cette décomposition en série de Fourier, aux équations suivantes (une pour chaque $\vec{q}\in\Lambda^\vee$) :

$$\forall\vec{q}\in\Lambda^\vee\,, \quad \boxed{\frac{\hbar^2\big(\vec{k}+\vec{q}\big)^2}{2m_e}\widehat{u}_{\vec{k}}(\vec{q}) + \sum_{\vec{p}\in\Lambda^\vee}\widehat{V}(\vec{p})\widehat{u}_{\vec{k}}(\vec{q}-\vec{p}) = E_{\vec{k}}\widehat{u}_{\vec{k}}(\vec{q})} \tag{7.29}$$

On remarque que ces équations ne dépendent du vecteur $\vec{k}$ choisi dans la décomposition de Bloch (7.23) uniquement via le premier terme en $(\vec{k}+\vec{q})$.

En conséquence, une transformation $\vec{k}\mapsto\vec{k}+\vec{K}$ avec $\vec{K}\in\Lambda^\vee$ va laisser le système d'équations (7.29) invariant, ne faisant que réorganiser ces équations ; on remarque en effet que $\widehat{u}_{\vec{k}+\vec{K}}(\vec{q}-\vec{K})$ satisfait la même équation que $\widehat{u}_{\vec{k}}(\vec{q})$. Si on indexe les solutions par leur indice de bande n, nous avons donc :

$$\widehat{u}^n_{\vec{k}+\vec{K}}(\vec{q}-\vec{K}) = \widehat{u}^{n'}_{\vec{k}}(\vec{q})\,, \tag{7.30}$$

où n' est l'indice de bande d'une autre solution de la même équation.

Dans l'espace des positions, nous avons alors :

$$\begin{aligned} u^n_{\vec{k}+\vec{K}}(\vec{x}) &= \sum_{\vec{q}\in\Lambda^\vee} e^{i\vec{q}\cdot\vec{x}}\widehat{u}^n_{\vec{k}+\vec{K}}(\vec{q}) = \sum_{\vec{q'}\in\Lambda^\vee} e^{i(\vec{q'}-\vec{K})\cdot\vec{x}}\widehat{u}^n_{\vec{k}+\vec{K}}(\vec{q'}-\vec{K}) = e^{-i\vec{K}\cdot\vec{x}}\sum_{\vec{q'}\in\Lambda^\vee} e^{i\vec{q'}\cdot\vec{x}}\widehat{u}^n_{\vec{k}+\vec{K}}(\vec{q'}-\vec{K}) \\ &= e^{-i\vec{K}\cdot\vec{x}}u^{n'}_{\vec{k}}(\vec{x})\,. \end{aligned} \tag{7.31}$$

Nous en déduisons premièrement que

$$\forall\vec{K}\in\Lambda^\vee\,, \quad \big\{\psi_{n,\vec{k}+\vec{K}}(\vec{x})\big\} = \big\{\psi_{n,\vec{k}}(\vec{x})\big\}\,, \tag{7.32}$$

c.-à.-d. que les fonctions de Bloch sont périodiques sur le réseau réciproque $\Lambda^\vee$, dans le sens où l'ensemble des solutions du problème obtenues ne dépend du choix du vecteur $\vec{k}$ dans la décomposition de Bloch (7.23) qu'à un vecteur du réseau réciproque près. On obtient la périodicité correspondante pour les énergies :

$$\forall \vec{K} \in \Lambda^\vee \ , \quad \left\{ E_{n,\vec{k}} \right\} = \left\{ E_{n,\vec{k}+\vec{K}} \right\} \tag{7.33}$$

signifiant là encore que les énergies propres obtenues pour différentes valeur de n en choisissant des fonctions de Bloch de la forme (7.23) pour un certain $\vec{k}$ seront identiques à celles obtenues en ayant choisi $\vec{k}+\vec{K}$ avec $\vec{K} \in \Lambda^\vee$.

En raison de cette périodicité, que nous avions déjà anticipée du point de vue de la théorie des groupes, il est possible de se restreindre, pour l'étude du spectre, à un vecteur d'onde $\vec{k}$ dans la première zone de Brillouin – notée par la suite PZB (Λ) – qui correspond, rappelons-le, à une maille de Wigner–Seitz du réseau réciproque $\Lambda^\vee$.

7.2.2 Réseau vide

Le formalisme développé dans la sous-section précédente s'applique à tout potentiel dont la périodicité est associée à un réseau Λ, même formellement à un potentiel identiquement nul dans le cristal. L'étude d'un tel système peut sembler futile, néanmoins elle permet d'une part d'appréhender la structure de bandes du spectre, et d'autre part forme le point de départ du traitement perturbatif du potentiel ionique, que vous étudierez en travaux dirigés. Dans ce cadre, il est naturel de choisir à l'ordre zéro d'organiser les états du système d'une manière compatible avec la périodicité de la perturbation.

On choisit donc un réseau $\Lambda \subset (\mathbb{R}^3, +)$, et on considère un système de N_e électrons libres, sans interactions ni entre eux ni avec le réseau cristallin, avec des conditions aux limites périodiques dans les trois directions,

$$\forall \vec{x} \in \mathbb{R}^3 \ , \quad \psi_{\vec{k}}(\vec{x} + N\vec{a}_1) = \psi_{\vec{k}}(\vec{x} + N\vec{a}_2) = \psi_{\vec{k}}(\vec{x} + N\vec{a}_3) \, , \tag{7.34}$$

où $\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}$ constitue une base de Λ, voir le chapitre 5, imposant la condition de quantification $\vec{k} \in \frac{1}{N}\Lambda^\vee$.

On considère alors, pour un certain $\vec{k} \in \text{PZB}(\Lambda) \cap \frac{1}{N}\Lambda^\vee$, les fonctions de Bloch $\psi_{\vec{k}}(\vec{x})$ associées :

$$\psi_{\vec{k}}(\vec{x}) = e^{i\vec{k}\cdot\vec{x}} u_{\vec{k}}(\vec{x}) \, . \tag{7.35}$$

Comme nous l'avons vu précédemment, les coefficients de Fourier de la fonction périodique $u_{\vec{k}}$ satisfont alors, voir l'éqn. (7.29) avec $\widehat{V} = 0$,

$$\frac{\hbar^2 \left(\vec{k} + \vec{q}\right)^2}{2m_e} \widehat{u}_{\vec{k}}(\vec{q}) = E_{\vec{k}} \widehat{u}_{\vec{k}}(\vec{q}) \, . \tag{7.36}$$

Cette équation est évidemment trivialement satisfaite, donnant, pour chaque $\vec{q} \in \Lambda^\vee$, une solution d'énergie $\hbar^2(\vec{k}+\vec{q})^2/2m_e$. En décomposant $\vec{q}$, sur la base duale de la base choisie pour le réseau Λ,

$$\vec{q} = n_1 \vec{a}_1^\vee + n_2 \vec{a}_2^\vee + n_3 \vec{a}_3^\vee \, , \tag{7.37}$$

On obtient ainsi une famille de solutions

$$\psi_{\vec{k};n_1,n_2,n_3}(\vec{x}) = \tfrac{1}{\sqrt{V}} e^{i\vec{k}\cdot\vec{x}} e^{i(n_1\vec{a}_1^\vee + n_2\vec{a}_2^\vee + n_3\vec{a}_3^\vee)\cdot\vec{x}} \ , \quad E = \tfrac{\hbar^2}{2m}\big(\vec{k} + n_1\vec{a}_1^\vee + n_2\vec{a}_2^\vee + n_3\vec{a}_3^\vee\big)^2$$
$$\vec{k} \in \mathrm{PZB}(\Lambda) \cap \tfrac{1}{N}\Lambda^\vee \ , \quad n_{1,2,3} \in \mathbb{Z} \, . \quad (7.38)$$

Il ne s'agit bien sûr que d'une manière particulièrement alambiquée de réexprimer les états propres d'énergie d'une particule libre dans une boîte !

Si, plutôt que $\vec{k} \in \mathrm{PZB}(\Lambda)$, on considère $\vec{k} \mapsto \vec{k} + \vec{K}$ avec $\vec{K} \in \Lambda^\vee$, ce changement peut être absorbé par un choix différent d'entiers (n_1, n_2, n_3), en d'autre termes d'indices de bande. Explicitement,

$$\psi_{\vec{k}+\vec{K};n_1,n_2,n_3}(\vec{x}) = \psi_{\vec{k};n_1+\ell_1,n_2+\ell_2,n_3+\ell_3}(\vec{x}) \ , \quad \vec{K} = \ell_1\vec{a}_1^\vee + \ell_2\vec{a}_2^\vee + \ell_3\vec{a}_3^\vee \ , \quad \ell_i \in \mathbb{Z} \, . \quad (7.39)$$

Réseau vide unidimensionnel

Il est utile pour visualiser les résultats de cette analyse de considérer un réseau vide à une dimension, c.-à.-d. une particule confinée à un intervalle de longueur $L = Na$. Naturellement, les solutions du système sont données par

$$\psi(x) = \tfrac{1}{\sqrt{L}} \, e^{ikx} \ , \quad E = \frac{\hbar^2 k^2}{2m} \ , \quad k \in \frac{2\pi}{a}\mathbb{Z} \, . \quad (7.40)$$

On peut réexprimer ces résultats, suivant les considérations précédentes, sous une forme compatible avec l'existence d'un « réseau vide » de pas a,

$$\psi_{k,n}(x) = \tfrac{1}{\sqrt{L}} \, e^{ikx} e^{\frac{2i\pi}{a}nx} \ , \quad E = \tfrac{\hbar^2}{2m}\big(k + \tfrac{2\pi n}{a}\big)^2 \ , \quad k \in \left]-\tfrac{\pi}{a}, \tfrac{\pi}{a}\right] \cap \tfrac{2\pi}{Na}\mathbb{Z} \ , \quad n \in \mathbb{Z} \, . \quad (7.41)$$

En comparant avec l'éqn. (7.40), cela revient à « replier » la relation de dispersion $E = f(k)$ lorsque k sort de la première zone de Brillouin, en appliquant une translation de $2\pi n/a$ dans le réseau réciproque, pour un certain entier n, afin de s'y ramener, voir la figure 7.1

On obtient ainsi un « empilement » de bandes d'énergies jointives, paramétrées par un entier n (strictement parlant, un entier positif $|n|$). À l'intérieur de chaque bande, le cristal étant de taille finie, on trouve $2N$ états distincts, paramétrés par une valeur de k parmi N et un état de spin.

Cette étude du réseau vide doit être comprise encore une fois comme un préliminaire à une étude de l'effet du potentiel périodique créé par les ions du cristal par une approche perturbative. Le réseau Λ étant associé à la périodicité du réseau cristallin, les énergies $E_{\vec{k}}$ sont principalement perturbées au voisinage du bord de la première zone de Brillouin – les fréquences spatiales dominantes dans la décomposition de Fourier (7.26) étant associées aux plus petites valeurs de $\|\vec{k}\|$.

Qualitativement, cette perturbation du spectre due au potentiel d'interaction entre l'électron et le réseau va engendrer une séparation entre les bandes, créant de ce fait des lacunes en énergie ou *bandes interdites* dans le spectre (*energy gaps* en anglais), voir la figure 7.2.

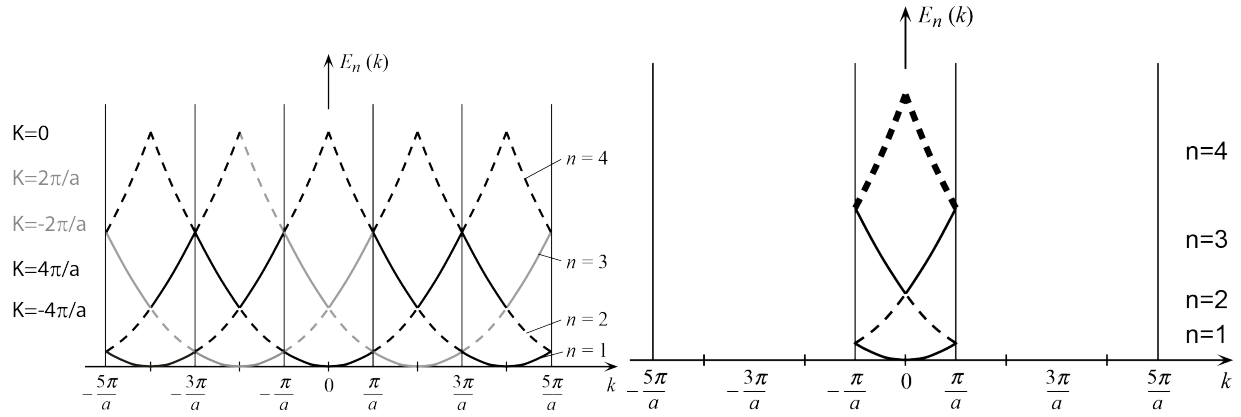

FIGURE 7.1 – *À gauche : repésentation des courbes de dispersion* $\mathsf{E} = \mathsf{f}(\mathsf{k})$ *pour différentes valeurs de* $\mathsf{K} \in \Lambda^\vee$ *dans le cas d'un potentiel nul ; les bandes d'énergie sont représentées en gras. À droite : représentation des bandes d'énergie dans la première de Brillouin.*

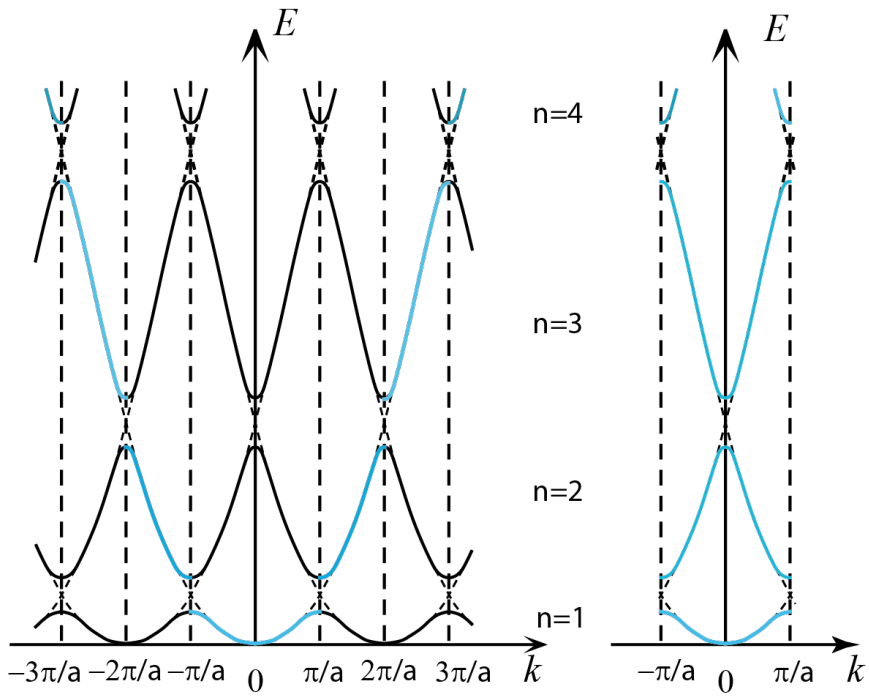

FIGURE 7.2 – *Modifications de la structure de bande dues à un potentiel périodique non-nul.*

7.2.3 Conducteurs et isolants

L'analyse précédente concernait un unique électron soumis à un potentiel périodique créé par les ions du cristal. Nous considérons dorénavant l'ensemble des N_e électrons présents dans le cristal, qui contient N^3 mailles élémentaires. En tenant compte des degrés de liberté de spin, chaque bande d'énergie contient $2\mathsf{N}^3$ niveaux.

Au zéro absolu, le remplissage des niveaux d'énergie par les N_e électrons est tel que l'énergie de la configuration obtenue est minimale, tout en respectant le principe d'exclusion de Pauli (pas plus d'un électron par état accessible). Plusieurs situations sont possibles, suivant en particulier le nombre d'électrons par maille élementaire, voir la figure 7.3. On peut alors qualitativement distinguer les cas suivants, concernant les propriétés de conduction électrique du solide considéré.

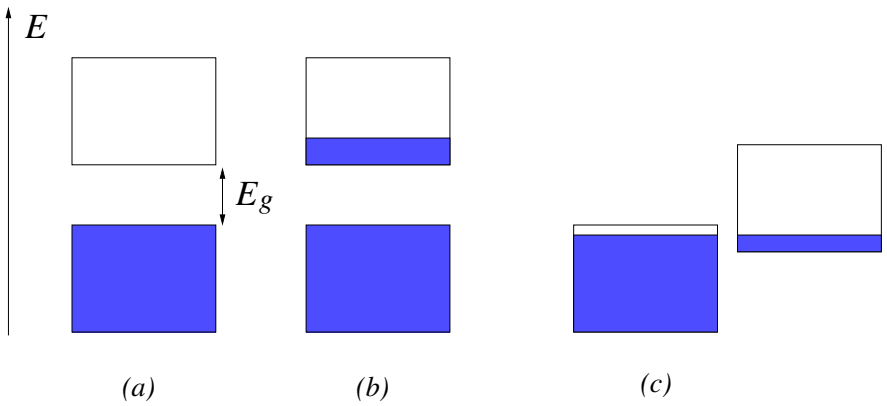

FIGURE 7.3 – *Bandes d'énergie pour un (a) isolant (b) métal (c) semi-métal.*

Dans un **métal** la bande de plus haute énergie contenant des électrons est seulement partiellement ; cette bande est alors appelée *bande de conduction*. Dans cette bande les électrons proches dont l'énergie est proche de l'énergie de Fermi peuvent alors être excités pour assurer la conduction d'électricité dans le solide. Toutes les bandes d'énergie inférieure sont entièrement remplies, la bande située juste en-dessous de la bande de conduction étant appelée *bande de valence*.

Dans un **semi-métal**, comme le bismuth, les deux bandes d'énergie les plus élevées (bande de valence et de conduction) se recouvrent faiblement et la bande supérieure peut être partiellement remplie sans que la précédente le soit complètement : un tel solide est également un conducteur électrique.

Dans le cas d'un **isolant**, la bande de plus haute énergie remplie (bande de valence) l'est entièrement et la bande de conduction juste au-dessus. Il n'existe alors pas d'électrons « disponibles » pour assurer la conduction d'électricité.

Un **semi-conducteur** est un isolant à température nulle pour lequel l'écart en énergie E_g entre la bande de valence et la bande de conduction (largeur de la bande interdite) est a minima de l'ordre de $k_B T$, où T est la température d'utilisation du matériau. Le nombre moyen d'électrons présents dans la bande de conduction est alors non-nul, voir le cours de physique statistique.

À retenir

— Notion de gaz de Fermi

— Structure en bandes

— Isolant, conducteur, semi-conducteur

Bibliographie

[1] H. Alloul, “Physique des électrons dans les solides,” Ecole Polytechnique, 2007

[2] Harald Ibach et Hans Lüth, “Solid-State Physics An Introduction to Principles of Materials Science,” Springer, 2009

[3] C. Kittel, “Introduction to Solid State Physics,” John Wiley & Sons, 2004

Cours n°8

Évolution temporelle

Dans les chapitres précédents nous avons étudié en détail la structure de l'espace des états d'un système quantique et le rôle des symétries. À tout instant l'état du système est déterminé par un vecteur d'état appartenant à un espace des états. Le rôle de ce chapitre est de comprendre comment déterminer le vecteur d'état associé au système à tout instant ultérieur. Notre approche sera relativement proche de celle de l'ouvrage de Sakurai [1].

Remarquons pour commencer la dissymétrie entre la translation dans l'espace étudiée au cours n° 4 et la translation dans le temps. Si on peut définir un opérateur position, qui se transforme par translation comme la quantité classique associée, il n'existe pas d'opérateur « temps » en mécanique quantique : le temps joue le rôle de paramètre externe, dont peut dépendre le hamiltonien.

8.1 Opérateur d'évolution

Soit un système quantique dont l'état est donné à un instant t_0 par un vecteur d'état $|\alpha; t_0)\rangle$ appartenant à un espace de Hilbert $\mathcal{H}_0$.

Si le hamiltonien $\mathcal{H}$ du système ne dépend pas explicitement du temps, cet espace des états est évidemment le même à tout instant et on peut choisir une base de vecteurs propres de $\mathcal{H}$ – ou plus généralement une base vecteurs propres communs au hamiltonien et à d'autres observables commutant avec $\mathcal{H}$ – comme base orthonormée de l'espace des états.

Si le hamiltonien dépend explicitement du temps, on supposera que l'évolution dans le temps est suffisamment « douce » pour que l'espace des états reste le même, ou plus précisément que l'espace des états pour $t_1 > t_0$ soit isomorphe à l'espace des états $\mathcal{H}_0$; on peut alors en particulier utiliser la même base orthonormée que pour l'espace de Hilbert $\mathcal{H}_0$ à tout instant.

Pour préserver le principe de superposition en mécanique quantique (c.-à.-d. que l'évolution d'une combinaison linéaire de vecteurs donne la combinaison linéaire du résultat de l'évolution de chaque vecteur) on suppose que l'évolution du système d'un état $|\alpha; t_0\rangle$ à $t = t_0$ vers un état $|\alpha; t_1\rangle$ à $t = t_1$ s'effectue par un opérateur linéaire, appelé opérateur d'évolution :

$$\boxed{|\alpha; t_1\rangle = \mathcal{U}(t_1, t_0)\,|\alpha; t_0\rangle} \tag{8.1}$$

Un état initial normé, c.-à.-d. tel que $\langle\alpha;t_0|\alpha;t_0\rangle = 1$ doit rester normé au cours de son évolution, par conservation de la probabilité. Si on considère en effet une observable quelconque $\mathcal{O}$ associée à une base orthonormée de vecteurs propres $|\lambda_i\rangle$, la distribution de probabilité associée à ses valeurs propres à $t = t_0$ est : est :

$$p(\lambda_i;t_0) = |\langle\lambda_i|\alpha;t_0\rangle|^2 \ , \quad \sum_{\lambda_i} p(\lambda_i;t_0) = \sum_i |\langle\lambda_i|\alpha;t_0\rangle|^2 = 1. \tag{8.2}$$

À tout instant ultérieur t_1, on doit vérifier également

$$\sum_{\lambda_i} p(\lambda_i;t_1) = \sum_i |\langle\lambda_i|\alpha;t_1\rangle|^2 = \sum_i |\langle\lambda_i|\,\mathcal{U}(t_1,t_0)\,|\alpha;t_0\rangle|^2 = 1. \tag{8.3}$$

Une condition suffisante pour que cette condition soit vérifiée est que l'opérateur d'évolution soit un opérateur unitaire :

$$\mathcal{U}(t_1,t_0)^\dagger \mathcal{U}(t_1,t_0) = \mathcal{I}\,. \tag{8.4}$$

Cette analyse rappelle le raisonnement ayant conduit au théorème de Wigner, voir la sous-section 3.4.1 du cours n° 3, mais, différence importante, nous n'avons pas supposé pour l'instant que le système admettait une symétrie de translation dans le temps. Notons également que nous avons uniquement imposé la contrainte « faible » $\sum p(\lambda_i) = 1$, voir l'éqn. (8.3), et non pas de contrainte « forte » associée à la conservation des probabilités individuelles $p(\lambda_i)$ car ces dernières peuvent évoluer au cours du temps suivant l'observable considérée.

La deuxième condition que nous imposerons à l'opérateur d'évolution est intuitive. On considère la propagation d'un instant t_0 à un instant t_1, et un instant intermédiaire $t_0 < t < t_1$. L'évolution du système de t_0 à t puis de t à t_1 doit être équivalente à l'évolution directe de t_0 à t_1. Explicitement nous avons donc :

$$\forall t \in]t_0,t_1[\ , \quad \mathcal{U}(t_1,t_0) = \mathcal{U}(t_1,t)\,\mathcal{U}(t,t_0)\,. \tag{8.5}$$

Considérons une évolution sur un intervalle de temps infinitésimal δt. Lorsque $\delta t \to 0$ nous devons, par continuité de l'évolution temporelle, retrouver l'opérateur identité. Posons :

$$\mathcal{U}(t_0+\delta t,t_0) = \mathcal{I} - \frac{i\delta t}{\hbar}\mathcal{G}(t_0)\,, \tag{8.6}$$

où $\mathcal{G}(t_0)$ est un opérateur à déterminer, qui a la dimension d'une énergie, correspondant au générateur de l'évolution temporelle et dépendant du paramètre t_0.

La condition d'unitarité de l'opérateur d'évolution $\mathcal{U}(t_1,t_0)$, voir éqn. (8.4), se traduit au niveau infinitésimal par :

$$\left(\mathcal{I} + \frac{i\delta t}{\hbar}\mathcal{G}^\dagger\right)\left(\mathcal{I} - \frac{i\delta t}{\hbar}\mathcal{G}\right) = \mathcal{I} + \mathcal{O}[(\delta t)^2] \implies \mathcal{G}^\dagger = \mathcal{G}\,, \tag{8.7}$$

donc le générateur $\mathcal{G}$ doit être un opérateur hermitien. La condition (8.5) est trivialement satisfaite à cet ordre ; posons $t_0 = t - \delta t$ et $t_1 = t + \delta t$. Nous avons :

$$\begin{aligned} \mathcal{U}(t_1, t_0) = \mathcal{U}(t + \delta t, t - \delta t) &= \mathcal{I} - \frac{2i\delta t}{\hbar}\mathcal{G}(t - \delta t) + \mathcal{O}[(\delta t)^2] \\ &= \mathcal{I} - \frac{2i\delta t}{\hbar}\mathcal{G}(t) + \mathcal{O}[(\delta t)^2] \end{aligned} \tag{8.8a}$$

$$\mathcal{U}(t_1, t) = \mathcal{U}(t + \delta t,) = \mathcal{I} - \frac{i\delta t}{\hbar}\mathcal{G}(t) + \mathcal{O}[(\delta t)^2] \tag{8.8b}$$

$$\mathcal{U}(t, t_0) = \mathcal{U}(t, t - \delta t) = \mathcal{I} - \frac{i\delta t}{\hbar}\mathcal{G}(t - \delta t) + \mathcal{O}[(\delta t)^2] = \mathcal{I} - \frac{i\delta t}{\hbar}\mathcal{G}(t) + \mathcal{O}[(\delta t)^2] \tag{8.8c}$$

et

$$\mathcal{U}(t_1, t)\,\mathcal{U}(t, t_0) = \left(\mathcal{I} - \frac{i\delta t}{\hbar}\mathcal{G}(t)\right)\left(\mathcal{I} - \frac{i\delta t}{\hbar}\mathcal{G}(t - \delta t)\right) + \mathcal{O}[(\delta t)^2] = \mathcal{I} - \frac{2i\delta t}{\hbar}\mathcal{G}(t) + \mathcal{O}[(\delta t)^2]\,. \tag{8.9}$$

8.1.1 Hamiltonien indépendant du temps

Considérons un système classique pour lequel le lagrangien ne dépend pas explicitement du temps. Nous avons montré dans le cours n° 3 l'existence d'une quantité conservée, ou charge de Noether, associée à la *symétrie de translation dans le temps*

$$t \mapsto t + \tau\,, \tag{8.10}$$

et correspondant au hamiltonien du système. Du point de vue de la mécanique hamiltonienne, on peut alors considérer le hamiltonien comme générateur de l'évolution dans le temps, voir l'éqn. (3.53).

Il est alors naturel d'associer, dans le cas d'un système admettant une symétrie de translation dans le temps, l'opérateur hamiltonien avec le générateur hermitien de l'évolution temporelle, voir l'éqn. (8.6) :

$$\mathcal{U}(t_0 + \delta t, t_0) = \mathcal{I} - \frac{i\delta t}{\hbar}\,\mathcal{H}\,. \tag{8.11}$$

Dans le cas présent l'opérateur d'évolution infinitésimale est évidemmment indépendant de l'instant t_0 correspondant à l'instant de début de l'évolution, par symétrie de translation dans le temps. Nous pourrons donc choisir sans restriction $t_0 = 0$ et nous poserons $t_1 = t$; l'opérateur d'évolution sera noté plus simplement $\mathcal{U}(t)$.

Le hamiltonien peut être assimilé au générateur du groupe des translations dans le temps, isomorphe à $(\mathbb{R}, +)$, et on peut donner la forme de l'opérateur d'évolution pour un intervalle de temps arbitraire en utilisant le théorème suivant :

Théorème 1 (théorème de Stone). *Soit* $U(t)$ *un groupe unitaire continu à un paramètre. Il admet un générateur* A *auto-adjoint tel que*

$$\forall t \in \mathbb{R}\ , \quad U(t) = e^{iAt}\,. \tag{8.12}$$

L'opérateur d'évolution est donc obtenu par exponentiation du hamiltonien :

$$\boxed{\mathcal{U}(t) = e^{-i\frac{\mathcal{H}}{\hbar}t}} \tag{8.13}$$

On peut le montrer également, comme nous l'avions déjà fait, en posant $\delta t = t/N$ dans (8.11) et en appliquant plusieurs évolutions infinitésimales successives :

$$\lim_{N\to\infty}\left(\mathcal{I} - \frac{it}{N\hbar}\,\mathcal{H}\right)^N = e^{-i\frac{\mathcal{H}t}{\hbar}}\,. \tag{8.14}$$

La relation de composition (8.5) est évidemment satisfaite par (8.13), car il s'agit simplement de la loi de composition sur le groupe abélien :

$$\mathcal{U}(t)\mathcal{U}(t') = e^{-i\frac{i\mathcal{H}}{\hbar}t}e^{-i\frac{i\mathcal{H}}{\hbar}t'} = e^{-i\frac{i\mathcal{H}}{\hbar}(t+t')} = \mathcal{U}(t+t)\,. \tag{8.15}$$

Finalement l'opérateur d'évolution (8.13) satisfait à l'équation différentielle du premier ordre suivante :

$$\boxed{i\hbar\,\frac{d}{dt}\mathcal{U}(t) = \mathcal{H}\,\mathcal{U}(t)} \tag{8.16}$$

forme opérationnelle de l'équation de Schrödinger pour un hamiltonien indépendant du temps. [1]

États stationnaires

Choisissons, toujours pour un hamiltonien indépendant du temps, une base de l'espace des états $\{|\ell\rangle\ ,\ \ell \in \mathcal{I}\}$ composée de vecteurs propres du hamiltonien :

$$\forall \ell \in \mathcal{I}\,,\ \mathcal{H}\,|\ell\rangle = E_\ell\,|\ell\rangle\ . \tag{8.17}$$

On suppose l'état du système donné à l'instant $t = 0$ par le vecteur d'état $|\psi_0\rangle$. On décompose ce vecteur sur la base (8.17) :

$$|\psi_0\rangle = \sum_{\ell\in\mathcal{I}} \alpha_\ell\,|\ell\rangle\ \ ,\quad \alpha_\ell = \langle\ell|\psi_0\rangle\ . \tag{8.18}$$

L'état à un instant t est alors donné en appliquant l'opérateur d'évolution (8.13) à ce vecteur :

$$|\psi(t)\rangle = e^{-\frac{i\mathcal{H}}{\hbar}t}\,|\psi_0\rangle = \sum_{\ell\in\mathcal{I}} \alpha_\ell\, e^{-\frac{i\mathcal{H}}{\hbar}t}\ |\ell\rangle\ . \tag{8.19}$$

1. Du point de vue de la théorie des groupes, cette équation est à rapprocher de la définition d'un sous-groupe à un paramètre, voir l'éqn. (3.63) du cours n° 3. L'opérateur hamiltonien correspond au générateur d'une transformation infinitésimale au voisinage de l'élément $\mathcal{U}(t)$ du groupe abélien des translations temporelles.

Les vecteurs $|\ell\rangle$ sont vecteurs propres de $\mathcal{H}$ et donc de toute fonction de $\mathcal{H}$ (on peut s'en convaincre ici par développement en série entière de la fonction exponentielle). On trouve donc :

$$\boxed{|\psi(t)\rangle = \sum_{\ell\in\mathcal{I}} \alpha_\ell \, e^{-i\frac{E_\ell}{\hbar}t} \, |\ell\rangle} \tag{8.20}$$

On voit qu'un état propre du hamiltonien évolue dans le temps comme $|\ell\rangle \mapsto e^{-i\frac{E_\ell}{\hbar}t}\,|\ell\rangle$, c.-à.-d. qu'il reste identique à une phase près sans signification physique.

Pour cette raison, les états propres de $\mathcal{H}$ sont appelés *états stationnaires* du système. Pour connaître l'évolution de n'importe quel état, il suffit de décomposer le vecteur d'état correspondant sur une base des vecteurs propres du hamiltonien, puis de faire évoluer chaque état stationnaire $|\ell\rangle$ par le facteur de phase $\exp -\frac{it}{\hbar}E_\ell$.

Exemple 1 (système à deux niveaux et oscillations de Rabi). On considère un système à deux niveaux $\uparrow$, $\downarrow$, de hamiltonien

$$\mathcal{H} = \frac{\hbar\omega}{2}\big(\,|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|\,\big)\,, \tag{8.21}$$

pouvant représenter un spin 1/2 dans un champ magnétique. On prépare le système à l'instant $t = 0$ dans l'état :

$$|\psi_0\rangle = \frac{|\uparrow\rangle + |\downarrow\rangle}{\sqrt{2}}\,. \tag{8.22}$$

L'état du système à un instant $t > 0$ est donc donné par :

$$|\psi(t)\rangle = e^{-\frac{i\mathcal{H}}{\hbar}t}\,|\psi_0\rangle = \frac{1}{\sqrt{2}}\big(e^{-i\frac{\omega t}{2}}\,|\uparrow\rangle + e^{+i\frac{\omega t}{2}t}\,|\downarrow\rangle\,\big)\,. \tag{8.23}$$

La probabilité de trouver le système dans l'état $\uparrow$ est donné à tout instant par :

$$p(\uparrow;t) = |\langle\uparrow|\,\psi(t)|^2 = \left|\frac{1}{\sqrt{2}}\big(e^{-i\frac{\omega t}{2}}\underbrace{\langle\uparrow|\uparrow\rangle}_{=0} + e^{+i\frac{\omega t}{2}}\underbrace{\langle\uparrow|\downarrow\rangle}_{=1}\big)\right|^2 = \frac{1}{2} \tag{8.24}$$

La probabilité obtenue est indépendante du temps.

Comparons ce résultat avec la probabilité que le système reste dans l'état initial $|\psi_0\rangle$, donné par l'éqn. (8.22) :

$$\begin{aligned} p(\psi_0;t) = |\langle\psi_0|\,\psi(t)|^2 &= \left|\frac{1}{2}\big(\,\langle\uparrow| + \langle\downarrow|\,\big)\big(e^{-i\frac{\omega t}{2}}\,|\uparrow\rangle + e^{+i\frac{\omega t}{2}}\,|\downarrow\rangle\,\big)\right|^2 = \left|\frac{1}{2}\left(e^{-i\frac{\omega t}{2}} + e^{i\frac{\omega t}{2}}\right)\right|^2 \\ &= \cos^2\left(\frac{\omega t}{2}\right)\,. \end{aligned} \tag{8.25}$$

Nous voyons que cette probabilité oscille entre 0 et 1, avec une pulsation ω ; il s'agit du phénomène connu sous le nom d'*oscillations de Rabi*. La différence cruciale avec le cas précédent est que $|\psi_0\rangle$ n'est pas un état propre du hamiltonien et ne se conserve donc pas (à une phase près) au cours du temps.

Constantes du mouvement

L'exemple précédent peut être généralisé de la manière suivante. On considère un système quantique dont le hamiltonien ne dépend pas explicitement du temps. Une observable $\mathcal{O}$ est appelée constante du mouvement, par analogie avec la mécanique classique (voir la sous-section 3.4.2 du cours n° 3) si cet opérateur satisfait :

$$[\,\mathcal{O},\,\mathcal{H}\,] = 0\,. \tag{8.26}$$

Les générateurs des groupes de symétries rentrent naturellement dans cette catégorie.

On alors peut construire (voir le théorème (9) du cours n° 1) une base de vecteurs propres de $\mathcal{O}$ qui sont également vecteurs propres de $\mathcal{H}$:

$$\forall \ell \in \mathcal{I}\ , \quad \mathcal{O}\,|\ell\rangle = \lambda_\ell\,|\ell\rangle\ , \quad \mathcal{H}\,|\ell\rangle = E_\ell\,|\ell\rangle\ . \tag{8.27}$$

On supposera le spectre ponctuel de $\mathcal{O}$ non dégénéré, c.-à.-d. qu'il n'existe qu'un vecteur propre pour une valeur propre donnée λ_ℓ. Sinon le sous-espace propre associé à λ_ℓ pourrait contenir des états propres du hamiltonien d'énergies différentes et le raisonnement suivant ne tiendrait pas.

On considère que le système est dans un état initial $|\psi_0\rangle$, que l'on décompose sur la base commune de vecteurs propres $\{|\ell\rangle\,,\ \ell \in \mathcal{I}\}$:

$$|\psi_0\rangle = \sum_{\ell\in\mathcal{I}} \alpha_\ell\,|\ell\rangle\ , \quad \alpha_\ell = \langle\ell|\psi_0\rangle\ . \tag{8.28}$$

L'état du système à l'instant t, $|\psi(t)\rangle$, est donc donné par l'éqn. (8.20).

La probabilité qu'une mesure à l'instant t associée à l'observable $\mathcal{O}$ donne la valeur λ_ℓ est alors donnée par :

$$p(\lambda_\ell; t) = |\langle\ell|\psi(t)\rangle|^2 = \left|\sum_{\ell'\in\mathcal{I}} \alpha_{\ell'}\, e^{-i\frac{E_{\ell'}}{\hbar}t}\ \langle\ell|\ell'\rangle\right|^2 = |\alpha_\ell|^2\ , \tag{8.29}$$

car chaque valeur propre λ_ℓ de $\mathcal{O}$ est associé à exactement un vecteur propre du hamiltonien (par hypothèse de non-dégénérescence du spectre), et que l'évolution temporelle de ce dernier ne fait que lui ajouter un facteur de phase $\exp -\frac{it}{\hbar}E_\ell$.

Nous arrivons ainsi à un résultat important : *pour un système quantique dont le hamiltonien ne dépend pas explicitement du temps, la distribution de probabilité associée une constante de mouvement dont le spectre n'est pas dégenéré n'évolue pas avec le temps.*

Naturellement ce principe s'applique au hamiltonien lui même, sans devoir faire d'hypothèse concernant la dégénerescence du spectre en énergie.

Propriété 1 *Soit un système quantique dont le hamiltonien ne dépend pas explicitement du temps. La distribution de probabilité associée aux énergies accessibles au système est indépendante du temps.*

8.1.2 Propagateur d'une particule libre

Nous pouvons définir une quantité appelée *propagateur*, noté $U(x_0^i, x_T^i; T)$, de la manière suivante : [2]

Définition 1 (propagateur). *Le propagateur* $U(x_0^i, x_T^i; T)$ *représente l'amplitude de probabilité du système d'évoluer du point de coordonnées* x_0^i *à l'instant* $t = 0$ *au point de coordonnées* x_T^i *à* $t = T$.

En notation de Dirac le propagateur est donc donné par :

$$U(\vec{x}_0, \vec{x}_T; T) = \langle \vec{x}_T | \, \mathcal{U}(T) \, | \vec{x}_0 \rangle \ . \tag{8.30}$$

Pour une particule libre, le propagateur est donné en termes de l'opérateur d'évolution (8.13) construit avec le hamiltonien :

$$\mathcal{H} = \frac{1}{2m} \vec{\mathcal{P}}^2 . \tag{8.31}$$

Afin de le calculer explicitement nous insérons deux relations de fermeture sur les impulsions :

$$U(\vec{x}_0, \vec{x}_T; T) = \int d^3p \int d^3p' \ \langle \vec{x}_T | \vec{p} \rangle \, \langle \vec{p} | \, \mathcal{U}(T) \, | \vec{p}' \rangle \, \langle \vec{p}' | \vec{x}_0 \rangle \ . \tag{8.32}$$

Les « éléments de matrice » pour l'opérateur d'évolution se calculent simplement dans la représentation impulsion :

$$\langle \vec{p} | \, \mathcal{U}(T) \, | \vec{p}' \rangle = \langle \vec{p} | \, e^{-\frac{iT}{2m\hbar} \vec{\mathcal{P}}^2} \, | \vec{p}' \rangle = \delta(\vec{p} - \vec{p}') e^{-\frac{iT}{2m\hbar} \vec{p}^2} , \tag{8.33}$$

d'où on obtient, en utilisant l'expression (2.54) pour le passage de la représentation position à la représentation impulsion :

$$U(\vec{x}_0, \vec{x}_T; T) = \frac{1}{(2\pi\hbar)^3} \int d^3p \, e^{-\frac{iT}{2m\hbar} \vec{p}^2} e^{\frac{i}{\hbar} \vec{p} \cdot (\vec{x}_T - \vec{x}_0)} . \tag{8.34}$$

Cette transformée de Fourier se calcule par une intégrale triple de Fresnel donnée par (avec ici $\lambda > 0$) :

$$\int dx \, e^{i \frac{\lambda x^2}{2} + i\mu x} = \sqrt{\frac{2\pi}{\lambda}} e^{-i \frac{\mu^2}{2\lambda} + i \frac{\pi}{4}} . \tag{8.35}$$

On obtient ainsi :

$$U(\vec{x}_0, \vec{x}_T; T) = \frac{1}{(2\pi\hbar)^3} \left(\frac{2\pi m \hbar}{iT} \right)^{3/2} e^{-\frac{i}{2} \left(\frac{\vec{x}_T - \vec{x}_0}{\hbar} \right)^2 \left(-\frac{m\hbar}{T} \right)} \tag{8.36}$$

soit finalement :

$$\boxed{U(\vec{x}_0, \vec{x}_T; T) = \left(\frac{m\hbar}{2i\pi\hbar T} \right)^{3/2} e^{\frac{im}{2\hbar T} (\vec{x}_T - \vec{x}_0)^2}} \tag{8.37}$$

2. Notons que, strictement parlant, les conditions initiales où la particule est complèment localisée ne sont pas physiques. Une manière plus rigoureuse d'interpréter cette quantité est donnée un peu plus loin, voir l'éqn. (8.41).

Remarquons une propriété importante du propagateur. Lorsque l'intervalle de temps tend vers zéro on montre que :

$$\lim_{T\to 0^+} \left(\frac{m}{2i\pi\hbar T}\right)^{3/2} e^{\frac{im(\vec{x}_T-\vec{x}_0)^2}{2T\hbar}} = \delta(\vec{x}_T-\vec{x}_0) \tag{8.38}$$

dont l'interprétation physique est évidente : une particule se trouvant en $\vec{x}_0$ à $t = 0$ sera toujours au même point au bout d'un intervalle de temps nul.

On peut interpréter ce résultat d'une manière différente. Le propagateur représente la fonction d'onde de la particule au temps t si, initialement, elle est localisée « exactement » en $\vec{x} = \vec{x}_0$, soit [3]

$$\psi(\vec{x},t) = \left(\frac{m}{2i\pi\hbar t}\right)^{3/2} e^{\frac{im(\vec{x}-\vec{x}_0)^2}{2t\hbar}} . \tag{8.39}$$

Étant donné que toute fonction d'onde $\psi_0(\vec{x})$ peut se représenter comme

$$\psi_0(\vec{x}) = \int d^3y\, \delta(\vec{x}-\vec{y})\psi_0(\vec{y}) , \tag{8.40}$$

on obtient l'évolution temporelle d'une fonction d'onde générale en termes du propagateur :

$$\psi(\vec{x},t) = \int d^3y\, \psi_0(\vec{y}) U(\vec{y},\vec{x},;t) , \tag{8.41}$$

qui satisfait bien la propriété

$$\lim_{t\to 0^+} \psi(\vec{x},t) = \psi_0(\vec{x}) . \tag{8.42}$$

La fonction d'onde (8.41) satisfait évidemment à l'équation de Schrödinger.

8.1.3 Hamiltonien dépendant du temps

Dans la sous-section précédente nous avons étudié les systèmes quantiques invariants par translation dans le temps. Le hamiltonien étant, en mécanique classique, le générateur associé à cette symétrie, nous avons pu identifier l'identifier avec l'opérateur d'évolution infinitésimal apparaissant dans l'éqn. (8.6). De manière complètement équivalente, nous aurions pu imposer directement l'équation différentielle (8.16).

Dans un système quantique général, le hamiltonien peut dépendre explicitement du temps, si on considère par exemple des spins dans un champ magnétique extérieur dépendant du temps. Nous aurons typiquement à considérer des systèmes dont le hamiltonien est de la forme :

$$\mathcal{H}(t) = \mathcal{H}_0 + \mathcal{V}(t) , \tag{8.43}$$

où $\mathcal{H}_0$ est un hamiltonien indépendant du temps, par exemple le hamiltonien d'une particule libre, et $\mathcal{V}(t)$ un potentiel d'interaction dépendant explicitement du temps.

3. Comme nous l'avons vu en détail au cours n° 2, une fonction d'onde parfaitement localisée n'est pas physique. Rigoureusement, il faut considérer uniquement l'éqn. (8.41).

Appelons Δt l'échelle caractéristique de variation de temporelle de $\mathcal{V}$. En d'autre termes, pour des échelles de temps $\delta t \ll \Delta t$, on pourra considérer que le hamiltonien complet $\mathcal{H}$ est approximativement indépendant du temps. On peut alors supposer que l'analyse précédente reste valable « localement » et postuler que l'opérateur d'évolution infinitésimal, au voisinage de $t = t_0$, est donné par :

$$\mathcal{U}(t_0 + \delta t, t_0) = \mathcal{I} - \frac{i\delta t}{\hbar}\,\mathcal{H}(t_0)\,, \tag{8.44}$$

c.-à.-d. que le hamiltonien évalué en $t = t_0$ donne le générateur des translations infinitésimales dans le temps autour de ce point.

Le fait que l'évolution infinitésimale au voisinage de l'instant initial $t = t_0$ soit engendrée par l'opérateur $\mathcal{H}$ évalué au même instant t_0 doit se vérifier pour tout autre instant $t > t_0$, la propagation se faisant de proche en proche d'après l'équation (8.5).

Ce raisonnement nous amène au postulat fondamental d'évolution temporelle en mécanique quantique :

Postulat 1 (évolution temporelle). *Soit un système quantique associé à un hamiltonien $\mathcal{H}(t)$ pouvant dépendre explicitement du temps. L'opérateur d'évolution $\mathcal{U}(t, t_0)$ du système entre un instant t_0 et un instant $t > t_0$ est un opérateur unitaire obéissant à l'équation differentielle :*

$$\boxed{i\hbar\frac{d}{dt}\mathcal{U}(t, t_0) = \mathcal{H}(t)\mathcal{U}(t, t_0)\,.} \tag{8.45}$$

appelée forme opérationnelle de l'équation de Schrödinger.

En effet, la condition (8.5) puis le développement (8.44) évalué en t permettent d'écrire :

$$\mathcal{U}(t + \delta t, t_0) = \mathcal{U}(t + \delta t, t)\mathcal{U}(t, t_0) = \left(\mathcal{I} - \frac{i\delta t}{\hbar}\,\mathcal{H}(t)\right)\mathcal{U}(t, t_0)\,, \tag{8.46}$$

d'où

$$\lim_{\delta t \to 0^+} \frac{\mathcal{U}(t + \delta t, t_0) - \mathcal{U}(t, t_0)}{\delta t} = \frac{1}{i\hbar}\mathcal{H}(t)\mathcal{U}(t, t_0)\,. \tag{8.47}$$

L'opérateur d'évolution étant unitaire, il transforme une base orthonormée de l'espace des états en une autre base orthonormée (voir le cours 1, éqn. (1.64)). L'évolution du système peut ainsi être vue, d'après également la propriété (8.5) comme une succession continue de changements de bases orthonormées dépendant d'un paramètre « extérieur » t ; nous y reviendrons dans la section 8.2.

L'intégration explicite de l'équation différentielle (8.45) est relativement simple si les hamiltoniens évalués à différents instants commutent entre eux :

$$\forall t, t' \in [t_0, t_1]\,, \quad [\,\mathcal{H}(t),\, \mathcal{H}(t')\,] = 0\,. \tag{8.48}$$

On peut alors directement obtenir comme il est aisé de vérifier

$$\boxed{\mathcal{U}(t_1, t_0) = \exp\left(-\frac{i}{\hbar}\int_{t_0}^{t_1} dt\,\mathcal{H}(t)\right)} \tag{8.49}$$

Nous avons en effet

$$\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t_1} dt\,\mathcal{H}(t)\right) = \mathcal{I} - \frac{i}{\hbar}\int_{t_0}^{t_1} dt\,\mathcal{H}(t) + \frac{1}{2!}\left(-\frac{i}{\hbar}\right)^2\left(\int_{t_0}^{t_1} dt\,\mathcal{H}(t)\right)^2 + \cdots, \tag{8.50}$$

et la dérivation terme à terme de cette expression par rapport à t_1 donne bien l'équation (8.45).

Lorsque les hamiltoniens évalués à différents instants ne commutent pas entre eux ce raisonnement ne s'applique pas (essentiellement pour la même raison que, si $[A, B] \neq 0$, $\exp A \exp B \neq \exp A + B$). On peut alors montrer que la solution est donnée par une *série de Dyson* :

$$\mathcal{U}(t, t_0) = \mathcal{I} + \sum_{k=1}^{\infty}\left(-\frac{i}{\hbar}\right)^k \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \int_{t_0}^{t_{k-1}} dt_k\,\mathcal{H}(t_1)\cdots\mathcal{H}(t_k)\,. \tag{8.51}$$

Cette série formelle de Dyson un peu complexe à manipuler en pratique, puisqu'il faut resommer une infinité de termes avec de plus en plus d'intégrales à calculer, ne sera utilisée par la suite que dans le contexte de la théorie perturbations dépendantes du temps, dans le cours n° 18.

8.2 Représentations de Schrödinger et Heisenberg

Dans cette section nous ferons le lien entre le formalisme utilisé jusqu'à présent pour décrire l'évolution temporelle, l'opérateur d'évolution, avec les formalismes plus standard de Schrödinger et Heisenberg.

8.2.1 Théorème d'Eherenfest

Premièrement, l'opérateur d'évolution $\mathcal{U}(t_1, t_0)$ permet d'obtenir la variation au cours du temps des valeurs moyennes d'observables, qui sont les quantités physiques de la mécanique quantique indépendamment du formalisme choisi.

On considère un système préparé dans l'état $|\psi_0\rangle \in \mathcal{H}$ en $t = t_0$ et une observable $\mathcal{O}$ associée à ce système. La valeur moyenne de $\mathcal{O}$ à l'instant $t > t_0$ s'obtient alors comme :

$$\boxed{\langle\mathcal{O}\rangle_{\psi_0}(t) = \langle\psi_0|\,\mathcal{U}^\dagger(t, t_0)\mathcal{O}\mathcal{U}(t, t_0)\,|\psi_0\rangle} \tag{8.52}$$

L'équation différentielle (8.45) satisfaite par l'opérateur d'évolution $\mathcal{U}$ donne alors une équation différentielle satisfaite par toute valeur moyenne d'observable.

Théorème 2 (théorème d'Ehrenfest). *Soit un système quantique associé à un hamiltonien $\mathcal{H}$ et $\mathcal{O}(t)$ une observable du système dépendant éventuellement du temps. La valeur moyenne de $\mathcal{O}$, pour un système dans l'état initial $|\psi_0\rangle$, satisfait à l'équation différentielle :*

$$\boxed{\frac{d}{dt}\langle\mathcal{O}\rangle_{\psi_0} = \frac{1}{i\hbar}\langle[\mathcal{O}, \mathcal{H}]\rangle_{\psi_0} + \left\langle\frac{\partial\mathcal{O}}{\partial t}\right\rangle_{\psi_0}} \tag{8.53}$$

La démonstration de ce théorème est évidente. La dérivée par rapport au temps de l'éqn. (8.52) donne, en utilisant la règle de Leibniz, l'équation d'évolution (8.45) et l'hermiticité de $\mathcal{H}$:

$$\begin{aligned}\frac{d\langle\mathcal{O}\rangle_{\psi_0}}{dt} &= \langle\psi_0|\left(\frac{d}{dt}\mathcal{U}^\dagger(t,t_0)\right)\mathcal{O}\mathcal{U}(t,t_0)\,|\psi_0\rangle + \langle\psi_0|\,\mathcal{U}^\dagger(t,t_0)\mathcal{O}\left(\frac{d}{dt}\mathcal{U}(t,t_0)\right)|\psi_0\rangle + \left\langle\frac{\partial\mathcal{O}}{\partial t}\right\rangle_{\psi_0}\\ &= \langle\psi_0|\left(\frac{i}{\hbar}\mathcal{U}^\dagger(t,t_0)\mathcal{H}^\dagger\right)\mathcal{O}\mathcal{U}(t,t_0)\,|\psi_0\rangle + \langle\psi_0|\,\mathcal{U}^\dagger(t,t_0)\mathcal{O}\left(-\frac{i}{\hbar}\mathcal{H}\mathcal{U}(t,t_0)\right)|\psi_0\rangle + \left\langle\frac{\partial\mathcal{O}}{\partial t}\right\rangle_{\psi_0}\\ &= \frac{i}{\hbar}\,\langle\psi_0|\,\mathcal{U}^\dagger(t,t_0)\big(\mathcal{H}\mathcal{O}-\mathcal{O}\mathcal{H}\big)\mathcal{U}(t,t_0)\,|\psi_0\rangle + \left\langle\frac{\partial\mathcal{O}}{\partial t}\right\rangle_{\psi_0}\\ &= \frac{i}{\hbar}\,\langle\psi_0|\,\mathcal{U}^\dagger(t,t_0)\,[\,\mathcal{H},\,\mathcal{O}\,]\,\mathcal{U}(t,t_0)\,|\psi_0\rangle + \left\langle\frac{\partial\mathcal{O}}{\partial t}\right\rangle_{\psi_0}. \end{aligned} \tag{8.54}$$

d'où on tire le résultat en reconnaissant dans le premier terme la valeur moyenne du commutateur de $\mathcal{O}$ et $\mathcal{H}$. □

Exemple 2 (particule dans un potentiel). Considérons un système constitué d'une particule dans un potentiel, dont le hamiltonien est de la forme standard :

$$\mathcal{H} = \frac{1}{2m}\vec{\mathcal{P}}^2 + V(\vec{\mathcal{X}}) \tag{8.55}$$

En application du théorème d'Ehrenfest on peut donner l'équation d'évolution de la position et de l'impulsion moyennes, deux observables ne dépendant pas explicitement du temps. Nous avons premièrement :

$$\frac{d}{dt}\langle\vec{\mathcal{X}}\rangle_\psi = \frac{1}{i\hbar}\langle[\,\mathcal{X},\,\mathcal{H}\,]\rangle_\psi = \frac{1}{i\hbar}\frac{1}{2m}\langle\Big[\,\vec{\mathcal{X}},\,\vec{\mathcal{P}}^2\,\Big]\rangle_\psi = \frac{1}{i\hbar}\frac{1}{2m}\langle 2i\hbar\vec{\mathcal{P}}\rangle_\psi = \frac{1}{m}\langle\vec{\mathcal{P}}\rangle_\psi \tag{8.56a}$$

$$\frac{d}{dt}\langle\vec{\mathcal{P}}\rangle_\psi = \frac{1}{i\hbar}\langle\Big[\,\vec{\mathcal{P}},\,\mathcal{H}\,\Big]\rangle_\psi = \frac{1}{i\hbar}\langle\Big[\,\vec{\mathcal{P}},\,V(\vec{\mathcal{X}})\,\Big]\rangle_\psi = \frac{1}{i\hbar}\langle -i\hbar\vec{\nabla}V(\vec{\mathcal{X}})\rangle_\psi = \langle -\vec{\nabla}V(\vec{\mathcal{X}})\rangle_\psi\,, \tag{8.56b}$$

donnant les *relations d'Ehrenfest*, analogue quantique des équations du mouvement la mécanique classique pour une force dérivant d'un potentiel :

$$\boxed{m\frac{d}{dt}\langle\vec{\mathcal{X}}\rangle_\psi = \langle\vec{\mathcal{P}}\rangle_\psi\,,\quad \frac{d}{dt}\langle\vec{\mathcal{P}}\rangle_\psi = \langle -\vec{\nabla}V(\vec{\mathcal{X}})\rangle_\psi} \tag{8.57}$$

Exemple 3 (oscillateur harmonique). On considère un oscillateur harmonique à une dimension, voir l'appendice 2.B du cours 2. On rappelle que le hamiltonien du système peut s'écrire :

$$\mathcal{H} = \hbar\omega(a^\dagger a + 1/2)\,,\quad \big[\,a,\,a^\dagger\,\big] = 1\,, \tag{8.58}$$

en termes des opérateurs d'échelle a et $a^\dagger$.

D'après la définition des opérateurs d'échelle, voir éqn. (2.126), l'opérateur position s'exprime comme :

$$X = \sqrt{\frac{\hbar}{2m\omega}}\left(a + a^\dagger\right) . \tag{8.59}$$

On considère le système dans un certain état $|\psi\rangle$ à l'instant $t = 0$.[4] L'évolution de la valeur moyenne de X, observable de dépendant pas explicitement du temps, est donnée par

$$\begin{aligned}\frac{d}{dt}\langle X\rangle_\psi = \frac{1}{i\hbar}\langle[X, \mathcal{H}]\rangle_\psi = \frac{1}{i\hbar}\sqrt{\frac{\hbar}{2m\omega}}\left(\langle[a, \mathcal{H}]\rangle_\psi + \langle\left[a^\dagger, \mathcal{H}\right]\rangle_\psi\right)\\ = \frac{1}{i\hbar}\sqrt{\frac{\hbar}{2m\omega}}\hbar\omega\langle(a - a^\dagger)\rangle_\psi = \frac{1}{m}\langle\mathcal{P}\rangle_\psi ,\end{aligned} \tag{8.60}$$

L'opérateur impulsion est donné par :

$$\mathcal{P} = -i\sqrt{\frac{\hbar m\omega}{2}}(a - a^\dagger) \tag{8.61}$$

et nous trouvons de même

$$\begin{aligned}\frac{d}{dt}\langle\mathcal{P}\rangle_\psi = \frac{1}{i\hbar}\langle[\mathcal{P}, \mathcal{H}]\rangle_\psi = -\frac{1}{\hbar}\sqrt{\frac{\hbar m\omega}{2}}\left(\langle[a, \mathcal{H}]\rangle_\psi - \langle\left[a^\dagger, \mathcal{H}\right]\rangle_\psi\right)\\ = -\sqrt{\frac{m\omega}{2\hbar}}\hbar\omega\langle(a + a^\dagger)\rangle_\psi = \frac{1}{m}\langle\mathcal{P}\rangle_\psi = -m\omega^2\langle X\rangle_\psi .\end{aligned} \tag{8.62}$$

En dérivant (8.60) par rapport au temps nous trouvons donc :

$$\frac{d^2}{dt^2}\langle X\rangle_\psi = -m\omega^2\frac{d}{dt}\langle\mathcal{P}\rangle_\psi = -\omega^2\langle X\rangle_\psi \tag{8.63}$$

Dont la solution générale est bien la solution classique de l'oscillateur harmonique :

$$\langle X\rangle_\psi(t) = X_0\cos(\omega t + \varphi) , \tag{8.64}$$

où

$$\langle X\rangle_\psi(0) = X_0\cos\varphi , \quad \langle\mathcal{P}\rangle_\psi(0) = -m\omega X_0\sin\varphi . \tag{8.65}$$

8.2.2 Représentation de Schrödinger

Le théorème de Ehrenfest nous donne, voir éqn. (8.53), la manière dont évoluent dans le temps les valeurs moyennes d'observables, qui sont les grandeurs accessibles à l'expérience. L'expression (8.52) donnant la valeur moyenne d'une observable en fonction du temps peut s'interpréter de differents points de vue équivalents ; nous commencerons par le plus célèbre d'entre eux.

4. Si le système est préparé dans un état propre $|n\rangle$ du hamiltonien (8.58) alors la valeur moyenne de X est nulle car $\langle n|\, a\,|n\rangle = \sqrt{n}\,\langle n|n-1\rangle = 0$ et $\langle n|\, a^\dagger\,|n\rangle = \sqrt{n+1}\,\langle n|n+1\rangle = 0$.

Dans la première représentation, appelée *représentation de Schrödinger* et la plus couramment employée, on interprète la valeur moyenne d'une observable, éqn. (8.52), de la manière suivante :

$$\langle \mathcal{O} \rangle_{\psi_0}(t) = \underbrace{\Big(\langle \psi_0 | \, \mathcal{U}^\dagger(t, t_0) \Big)}_{\langle \psi(t) |} \; \mathcal{O} \; \underbrace{\Big(\mathcal{U}(t, t_0) \, |\psi_0\rangle \Big)}_{|\psi(t)\rangle} , \tag{8.66}$$

c.-à.-d. que le vecteur d'état du système à l'instant t, $|\psi(t)\rangle$, est donné en fonction de l'état initial $|\psi_0\rangle$ comme

$$|\psi(t)\rangle = \mathcal{U}(t, t_0) \, |\psi_0\rangle \ . \tag{8.67}$$

Il s'agit du point de vue déjà employé au début de ce cours. Dans la représentation de Schrödinger les observables qui ne dépendent pas explicitement du temps (comme la position et l'impulsion) sont indépendantes du temps.

Cela signifie en particulier que, si on choisit comme base de l'espace des états une base de vecteurs propres $\{|\ell\rangle \, , \, \ell \in \mathcal{I}\}$ d'une observable sans dépendance temporelle (ou d'un ensemble d'observables commutantes), cette base reste fixe au cours du temps mais les coefficients du développement d'un vecteur d'état dans cette base dépendent du temps :

$$|\psi(t)\rangle = \sum_{\ell \in \mathcal{I}} \alpha_\ell(t) \, |\ell\rangle \ . \tag{8.68}$$

L'équation différentielle satisfaite par les vecteurs d'état est immédiatement déduite de l'équation différentielle (8.45) donnant la dérivée temporelle de l'opérateur d'évolution $\mathcal{U}(t, t_0)$. Nous avons :

$$\frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = \frac{\mathrm{d}}{\mathrm{d}t} \mathcal{U}(t, t_0) \, |\psi_0\rangle = \left(\frac{1}{i\hbar} \mathcal{H}(t) \mathcal{U}(t, t_0) \right) |\psi_0\rangle \tag{8.69}$$

soit

$$\boxed{i\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = \mathcal{H}(t) \, |\psi(t)\rangle} \tag{8.70}$$

la célébrissime équation de Schrödinger pour les vecteurs d'état.

De manière équivalente, on obtient l'équation d'évolution pour les coefficients $\alpha_\ell(t)$ dans la décomposition (8.68) sur une base de l'espace des états en projetant l'équation de Schrödinger sur un vecteur d'état $|\ell\rangle$:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \langle \ell | \psi(t) \rangle = \langle \ell | \, \mathcal{H}(t) \, |\psi(t)\rangle \quad \Longrightarrow \quad \boxed{i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \alpha_\ell(t) = \sum_{\ell' \in \mathcal{I}} \alpha_{\ell'} \, \langle \ell | \, \mathcal{H}(t) \, |\ell'\rangle} \tag{8.71}$$

Il s'agit d'un système d'équations différentielles couplées sur lequel nous reviendrons dans le cours sous une forme légèrement modifiée dans le cours n° 18.

Considérons une particule dans un potentiel, dont le hamiltonien est donné par l'équation (8.55). L'espace des états du système est alors l'espace de Hilbert des fonctions de carré sommable $\mathcal{L}^2(\mathbb{R}^3)$, et l'application à l'équation (8.70) de la distribution de Dirac $\langle \vec{x}|$ donne l'équation d'évolution d'une fonction d'onde :

$$\boxed{i\hbar \frac{\partial}{\partial t} \psi(\vec{x}, t) = \left(-\frac{\hbar^2 \Delta}{2m} + V(\vec{x}) \right) \psi(\vec{x}, t)} \tag{8.72}$$

Cette *équation d'onde* est souvent prise comme point de départ de la mécanique quantique.

8.2.3 Représentation de Heisenberg

Nous avons remarqué plus haut que l'opérateur d'évolution, étant un opérateur unitaire, peut être vu comme un changement de bases orthonormées sur l'espace des états. De même que nous avions développé dans le cours n° 4 l'analogue quantique des points de vue actif et passif pour les symétries spatiales, nous pouvons considérer la représentation de Schrödinger comme l'analogue d'un point de vue actif car le vecteur d'état change avec le temps mais la base orthonormée, associée à une ou plusieurs observables, reste fixe au cours du temps.

La *représentation de Heisenberg* est l'analogue du point de vue passif, c.-à.-d. que le vecteur d'état reste fixe mais les observables évoluent. Nous reprenons donc l'expression (8.52) de la valeur moyenne d'une observable et lui donnons une interprétation différente :

$$\langle \mathcal{O} \rangle_{\psi_0}(t) = \langle \psi_0 | \underbrace{\Big(\mathcal{U}^\dagger(t, t_0)\, \mathcal{O}\, \mathcal{U}(t, t_0) \Big)}_{\mathcal{O}_{\text{H}}(t)} | \psi_0 \rangle \tag{8.73}$$

c.-à.-d. que les observables à l'instant t s'obtiennent à partir des observables à l'instant initial t_0 par une transformation unitaire donnée par l'opérateur d'évolution :

$$\boxed{\mathcal{O}_{\text{H}}(t) \overset{\text{déf.}}{=} \mathcal{U}^\dagger(t, t_0)\, \mathcal{O}\, \mathcal{U}(t, t_0)} \tag{8.74}$$

On obtient alors l'équation d'évolution des observables à partir de l'éqn. (8.45), appelée *équation de Heisenberg* :

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{O}_{\text{H}}(t) &= \left(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{U}(t, t_0)\right)^\dagger \mathcal{O}\mathcal{U}(t, t_0) + \mathcal{U}^\dagger(t, t_0)\mathcal{O}\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{U}(t, t_0)\right) + \mathcal{U}^\dagger(t, t_0)\frac{\partial \mathcal{O}}{\partial t}\,\mathcal{U}(t, t_0) \\
&= \left(\frac{1}{i\hbar}\mathcal{H}\mathcal{U}(t, t_0)\right)^\dagger \mathcal{O}\,\mathcal{U}(t, t_0) + \mathcal{U}^\dagger(t, t_0)\,\mathcal{O}\left(\frac{1}{i\hbar}\mathcal{H}\mathcal{U}(t, t_0)\right) + \frac{\partial \mathcal{O}}{\partial t}\Big|_{\text{H}}(t) \\
&= -\frac{1}{i\hbar}\Big(\mathcal{U}^\dagger(t, t_0)\mathcal{H}\mathcal{O}\mathcal{U}(t, t_0) - \mathcal{U}^\dagger(t, t_0)\mathcal{O}\mathcal{H}\mathcal{U}(t, t_0)\Big) + \frac{\partial \mathcal{O}}{\partial t}\Big|_{\text{H}}(t)
\end{aligned}$$

soit

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{O}_{\text{H}}(t) = \frac{1}{i\hbar}\,[\,\mathcal{O},\,\mathcal{H}\,]\,\Big|_{\text{H}}(t) + \frac{\partial \mathcal{O}}{\partial t}\Big|_{\text{H}}(t)} \tag{8.75}$$

Ce résultat est très semblable au théorème d'Eherenfest, voir l'éqn. (8.53), sauf qu'ici il s'agit d'une relation entre opérateurs et non pas entre valeurs moyennes dans un certain état.

En pratique nous serons souvent confrontés à une version simplifiée de l'équation de Heisenberg pour les raisons suivantes :

— si l'observable $\mathcal{O}$ ne dépend pas explicitement du temps, le dernier terme dans (8.75) est évidemment nul ;

— si le hamiltonien ne dépend pas expliciment du temps, l'opérateur d'évolution prend la forme (8.13) et $\mathcal{H}_{\text{H}} = \mathcal{U}^\dagger(t)\,\mathcal{H}\,\mathcal{U}(t) = \mathcal{H}$: l'opérateur hamiltonien prend la même forme dans les représentations de Schrödinger et Heisenberg.

Si ces deux hypothèse sont satisfaites, nos avons simplement :

$$\boxed{\frac{d}{dt}\mathcal{O}_{\text{H}}(t) = \frac{1}{i\hbar}\,[\,\mathcal{O}_{\text{H}}(t),\,\mathcal{H}\,]} \tag{8.76}$$

Dans le cas d'une particule dans un potentiel, de hamiltonien (8.55) nous trouvons ainsi par un calcul analogue à (8.56a,8.56b) :

$$m\frac{d}{dt}\vec{\mathcal{X}}_{\text{H}} = \vec{\mathcal{P}}_{\text{H}}\ , \quad \frac{d}{dt}\vec{\mathcal{P}}_{\text{H}} = -\vec{\nabla}V(\vec{\mathcal{X}})_{\text{H}} \tag{8.77}$$

8.3 Renversement du temps

Les équations de la mécanique, au moins pour un système de particules sans champ extérieur, sont invariantes par la symétrie discrète de renversement du temps $t \mapsto -t$, qui revient à inverser le mouvement des particules.

En mécanique quantique, nous définirons un opérateur χ réalisant le renversement du temps sur l'espace des états. :

$$\chi\ :\quad \left\{\begin{array}{ccc} \mathcal{H} & \to & \mathcal{H} \\ |\alpha\rangle & \mapsto & \chi\,|\alpha\rangle \end{array}\right. \tag{8.78}$$

L'opérateur χ realisant une symétrie il doit être, suivant le théorème de Wigner du cours n° 3, soit unitaire soit anti-unitaire.

Notons que nous ne pouvons utiliser la même approche que pour la parité, voir éqn. (4.54), étant donné qu'il n'existe pas d'opérateur « temps » dont nous pourrions considérer des vecteurs propres.

Pour obtenir les propriétés de cet opérateur, considérons un système quantique dont le hamiltonien ne dépend pas explicitement du temps, qui est donc associé à un système classique invariant par renversement du temps.

L'évolution temporelle infinitésimale d'un état est donnée par (voir éqn. (8.13)) :

$$|\alpha;t\rangle\ \mapsto\ |\alpha;t+\delta t\rangle = \left(\mathcal{I} - \frac{i\delta t}{\hbar}\mathcal{H}\right)|\alpha;t\rangle\ . \tag{8.79}$$

Nous attendons de l'opérateur de renversement du temps la propriété suivante :

$$\chi\,|\alpha;t+\delta t\rangle = |\alpha,-t-\delta t\rangle\ , \tag{8.80}$$

ce qui implique que

$$\chi\left(\mathcal{I} - \frac{i\delta t}{\hbar}\mathcal{H}\right)|\alpha;t\rangle = \left(\mathcal{I} + \frac{i\delta t}{\hbar}\mathcal{H}\right)|\alpha;-t\rangle = \left(\mathcal{I} + \frac{i\delta t}{\hbar}\mathcal{H}\right)\chi\,|\alpha;t\rangle\ . \tag{8.81}$$

Cette propriété devant être vérifiée pour tout $|\alpha;t\rangle \in \mathcal{H}$, nous arrivons à l'égalité sur l'espace des états :

$$\chi i\mathcal{H} = -i\mathcal{H}\chi \tag{8.82}$$

Si on supposait que χ est un opérateur unitaire, on obtiendrait alors que χ *anticommuterait* avec le hamiltonien, c.-à.-d.

$$\chi\mathcal{H} + \mathcal{H}\chi = 0\,. \tag{8.83}$$

Cette hypothése poserait des difficultés importantes. En effet considérant un état propre $|E\rangle$ de $\mathcal{H}$ d'énergie E nous aurions :

$$\chi\mathcal{H}\,|E\rangle = E\chi\,|E\rangle = -\mathcal{H}\chi\,|E\rangle \implies \mathcal{H}\,(\chi\,|E\rangle) = -E\,(\chi\,|E\rangle) \tag{8.84}$$

impliquant qu'à tout état stationnaire $|E\rangle$ d'énergie E on pourrait faire correspondre un état stationnaire $\chi\,|E\rangle$ d'énergie $-E$. Ce raisonnement conduirait à un spectre en énergie non borné inférieurement, ce qui est physiquement inacceptable. Par exemple nous savons que le spectre du hamiltonien d'une particule libre (voir la section 2.3.2 du cours n° 2) est $\sigma(\mathcal{H}_0) = \mathbb{R}^+$.

Nous sommes donc amenés à considérer l'autre possibilité offerte par le théorème 5, c.-à.-d. associer la symétrie de renversement du temps à un opérateur anti-unitaire :

Propriété 2 (opérateur de renversement du temps) *Soit un système quantique régi par un hamiltonien $\mathcal{H}$ ne dépendant pas explicitement du temps. Ce système possède une symétrie de renversement du temps $t \mapsto -t$, dont l'action sur l'espace des états $\mathcal{H}$ est donnée par un opérateur <u>anti-unitaire</u> χ tel que :*

$$[\mathcal{H}, \chi] = 0\,. \tag{8.85}$$

En effet, un opérateur anti-unitaire étant en particulier anti-linéraire – voir l'éqn. (3.92) – nous avons $\chi i\mathcal{H} = -i\chi\mathcal{H}$ et l'équation (8.82) implique bien la propriété (8.85).

Examinons le comportement des observables sous l'action du renversement du temps (rappelant que l'usage de la notation de Dirac est dangereuse avec les opérateurs anti-linéraires). Nous avons, pour tout opérateur

$$\begin{aligned}\langle\alpha|\,\mathcal{O}\,|\beta\rangle = \iota(|\alpha\rangle\,, \mathcal{O}\,|\beta\rangle) \overset{t\mapsto -t}{\longmapsto} \iota(\chi\,|\alpha\rangle\,, \mathcal{O}\chi\,|\beta\rangle) = \iota(\chi\,|\alpha\rangle\,, \chi\chi^{-1}\mathcal{O}\chi\,|\beta\rangle)\\ = \overline{\iota(|\alpha\rangle\,, \chi^{-1}\mathcal{O}\chi\,|\beta\rangle)} = \iota(\left(\chi^{-1}\mathcal{O}\chi\right)|\beta\rangle\,, |\alpha\rangle) = \iota(|\beta\rangle\,, \left(\chi^{-1}\mathcal{O}\chi\right)^\dagger|\alpha\rangle)\end{aligned} \tag{8.86}$$

Ce calcul nous indique que l'action du renversement du temps sur les opérateurs (utilisée dans le point de vue passif) est donnée par :

$$\mathcal{O} \overset{t\mapsto -t}{\longmapsto} \left(\chi^{-1}\mathcal{O}\chi\right)^\dagger \tag{8.87}$$

Nous pouvons définir deux types d'opérateurs suivant leur comportement sous le renversement du temps :

— les opérateurs pairs vérifient $\chi^{-1}\mathcal{O}_{\text{P}}\chi = \mathcal{O}$;

— les opérateurs impairs vérifient $\chi^{-1}\mathcal{O}_{\text{I}}\chi = -\mathcal{O}$.

Nous avons ainsi le comportement

$$\mathcal{O}_{\substack{\mathrm{P}\\\mathrm{I}}} \overset{t\mapsto -t}{\longmapsto} \pm\mathcal{O}^\dagger\,, \tag{8.88}$$

qui se simplifie, pour des opérateurs hermitiens, en

$$\mathcal{O}_{\substack{\mathrm{P}\\\mathrm{I}}} \overset{t\mapsto -t}{\longmapsto} \pm\mathcal{O}\,. \tag{8.89}$$

Exemple 4 (opérateurs impulsion et position). Il est raisonnable de postuler que les opérateurs impulsion et position sont respectivement impairs et paris, le renversement du temps en mécanique classique inversant le mouvement des corps. Nous posons

$$\chi^{-1}\vec{\mathcal{P}}\chi = -\vec{\mathcal{P}}\,, \tag{8.90a}$$
$$\chi^{-1}\vec{\mathcal{X}}\chi = +\vec{\mathcal{X}}\,. \tag{8.90b}$$

Nous pouvons vérifier que les relations de commutation canoniques sont bien préservées. En effet :

$$\begin{aligned}[\mathcal{X}_i,\,\mathcal{P}_j] \;&\mapsto\; \left(\chi^{-1}\,[\mathcal{X}_i,\,\mathcal{P}_j]\,\chi\right)^\dagger = \left(\chi^{-1}\mathcal{X}_i\mathcal{P}_j\chi - \chi^{-1}\mathcal{P}_j\mathcal{X}_i\chi\right)^\dagger \\ &= \left(\chi^{-1}\mathcal{X}_i\chi\chi^{-1}\mathcal{P}_j\chi - \chi^{-1}\mathcal{P}_j\chi\chi^{-1}\mathcal{X}_i\chi\right)^\dagger = (-\mathcal{X}_i\mathcal{P}_j + \mathcal{P}_j\mathcal{X}_i)^\dagger = (-i\hbar\mathcal{I})^\dagger = i\hbar\mathcal{I}\end{aligned} \tag{8.91}$$

Nous postulerons en conséquence que :

$$\chi\,|\vec{p}\rangle = |-\vec{p}\rangle\ , \tag{8.92a}$$
$$\chi\,|\vec{x}\rangle = |\vec{x}\rangle\ . \tag{8.92b}$$

Une phase est autorisée dans ces transformations, que nous choisirons triviale.

Examinons pour finir l'action de l'opération de renversement du temps sur les fonctions d'onde. Nous avons

$$\psi(\vec{x};t) = \iota(|\vec{x}\rangle\,,|\psi;t\rangle) \mapsto \iota(|\vec{x}\rangle\,,\chi\,|\psi;t\rangle) = \overline{\iota(\chi^{-1}\,|\vec{x}\rangle\,,\chi^{-1}\chi\,|\psi;t\rangle)} = \overline{\iota(|\vec{x}\rangle\,,|\psi;t\rangle)} = \overline{\psi(\vec{x};t)} \tag{8.93}$$

La fonction d'onde est donc envoyée sur son complexe conjugué par renversement du temps.

Considérons par exemple une onde plane, associée au hamiltonien (8.31) d'une particule libre. Nous avons suivant cette règle :

$$\frac{1}{(2\pi\hbar)^{3/2}}e^{-i\frac{\vec{p}\cdot\vec{x}}{\hbar}}e^{-i\frac{p^2}{2m\hbar}t} \;\longmapsto\; \frac{1}{(2\pi\hbar)^{3/2}}e^{+i\frac{\vec{p}\cdot\vec{x}}{\hbar}}e^{+i\frac{p^2}{2m\hbar}t} \tag{8.94}$$

conformément au comportement attendu sous le renversement du temps, c.-à.-d. que $t \mapsto -t$ et $\vec{p} \mapsto -\vec{p}$ alors que $\vec{x}$ est invariant.

Finissons cette étude par un théorème portant sur les fonctions d'onde associées à un système invariant par renversement du temps.

Théorème 3 *Soit un système quantique associé à un hamiltonien $\mathcal{H}$ indépendant du temps dont le spectre est non-dégénéré et qui est invariant par renversement du temps. Les fonctions d'ondes associées aux états propres du système sont réelles, à une phase constante près.*

Ce théorème se démontre simplement de la manière suivante. Soit $|\ell\rangle$ tel que $\mathcal{H}\,|\ell\rangle = E_\ell\,|\ell\rangle$ et $\langle\ell|\ell\rangle = 1$. Nous avons par (8.85)

$$\mathcal{H}\chi\,|\ell\rangle = \chi\mathcal{H}\,|\ell\rangle = E_\ell\chi\,|\ell\rangle\ , \tag{8.95}$$

le spectre de l'opérateur auto-adjoint $\mathcal{H}$ étant réel. Cela indique que $\chi\,|\ell\rangle$ est un vecteur propre de $\mathcal{H}$ d'énergie E_ℓ. L'hypothèse de non dégénerescence du spectre indique alors que

$$\exists\lambda \in \mathbb{C}^\star\ , \quad \chi\,|\ell\rangle = \lambda\,|\ell\rangle \tag{8.96}$$

Le vecteur $|\ell\rangle$ étant normalisé, nous avons d'une part

$$\iota(\chi\,|\ell\rangle\,,\chi\,|\ell\rangle) = \overline{\iota(|\ell\rangle\,,|\ell\rangle)} = 1 \tag{8.97}$$

et d'autre part

$$\iota(\chi\,|\ell\rangle\,,\chi\,|\ell\rangle) = \iota(\lambda\,|\ell\rangle\,,\lambda\,|\ell\rangle) = |\lambda|^2 \tag{8.98}$$

donc

$$\exists\varphi \in [0,2\pi[\ , \quad \lambda = e^{i\varphi}\,. \tag{8.99}$$

Nous utilisons ensuite, d'après (8.93)

$$\overline{\psi_\ell(\vec{x})} = \iota(|\vec{x}\rangle\,,\chi\,|\psi\rangle) = \iota(|\vec{x}\rangle\,,e^{i\varphi}\,|\psi\rangle) = e^{i\varphi}\psi_\ell(\vec{x}) \tag{8.100}$$

qui est le résultat annoncé. En effet, cette équation se résoud comme :

$$\psi_\ell(\vec{x}) = e^{-i\varphi/2}f_\ell(\vec{x}) \tag{8.101}$$

avec f_ℓ une fonction à valeurs dans $\mathbb{R}$. □

Théorème CPT et violation de CP

En théorie quantique des champs , théorie quantique relativiste des particules élémentaires, on peut envisager trois symétries discrètes fondamentales :

— la parité ou inversion spatiale $P : \vec{x} \mapsto -\vec{x}$, étudiée dans le cours n° 4, section 4.3 ;

— le renversement du temps $T : t \mapsto -t$, objet de la présente section ;

— la conjugaison de charge, qui associe à toute particule son anti-particule (qui possède en particulier la même masse et une charge électrique opposée).

Une question importante est de savoir si le modèle standard des interactions fondamentales respecte ces trois symétries discrètes. On peut démontrer de manière très générale le *théorème CPT* : un théorie quantique des champs, moyennant quelques hypothèses raisonnables comme la localité, est invariante par action simultanée de la conjugaison de charge, de la parité et du renversement du temps.

Ce théorème ne signifie absolument pas que chacune de ces opérations, prise individuellement, est une symétrie du modèle standard !

L'électromagnétisme et la gravitation respectent a priori la symétrie de parité, s'agissant de forces à grandes distances pouvant s'étudier dans le cadre de la physique classique et les lois de Newton étant invariantes sous la transformation $\vec{x} \mapsto -\vec{x}$. Toute les expériences de physique des particules indiquent que les interactions fortes ont aussi une symétrie de parité, propriété manifeste de la théorie utilisée dans le modèle standard pour les décrire (chromodynamique quantique).

Expérimentalement, il a été observé en 1957, à la grande surprise des théoriciens des particules, que les interactions faibles ne respectaient pas la symétrie de parité [2]. Ils ont étudié la désintégration béta du cobalt 60 :

$$^{60}\mathrm{Co} \;\rightarrow\; ^{60}\mathrm{Ni} + e^- + \bar{\nu}_e + 2\gamma\,, \tag{8.102}$$

dans laquelle un neutron d'un noyau de cobalt radioactif décroît en un proton émettant un électron et un anti-neutrino. Le noyau de nickel obtenu, n'étant pas dans son état fondamental, émet à son tour deux photons gamma. La distribution spatiale de l'émission gamma a alors montré clairement que la symétrie P n'était pas respectée.

Il a ensuite été suggéré que, si la parité n'était par respectée par les interactions faibles, la combinaison CP, c.-à.-d. l'inversion spatiale combinée avec la conjugaison de charge échangeant particules et anti-particules, pouvait être une symétrie du modèle standard. D'autres expériences menées en 1964 [3] ont montré que la symétrie CP également était violée par les interactions faibles.

Le fait que la symétrie CP ne soit pas respectée par les interactions fondamentales a des conséquences considérables en cosmologie, puisqu'elle explique la prédominance de matière sur l'antimatière dans l'Univers.

Finalement, mentionnons un problème ouvert de la physique des particules. Il n'a été observé aucune violation de la symétrie CP par les interactions fortes, avec des contraintes expérimentales très strictes. Or, dans la théorie quantique des champs associée, la chromodynamique quantique, on peut ajouter un terme à l'action qui brise explicitement cette symétrie. La confrontation avec l'expérience nous montre que, soit ce terme n'apparaît pas, soit le coefficient devant ce terme est extraordinairement petit.

Comme nous l'avions mentionné au début du cours n° 3, dans une théorie quantique toutes les interactions susceptibles d'apparaître existent a priori, sauf si une symétrie l'interdit. L'absence de violation de la symétrie CP par les interactions fortes suggère ainsi l'existence d'une nouvelle symétrie, liée à l'existence de nouvelles particules, les *axions*, jamais observés à ce jour.

À retenir

— la notion d'opérateur d'évolution unitaire et son équation différentielle ;

— le théorème d'Eherenfest ;

— les points de vue de Schrödinger et de Heisenberg ;

— l'opérateur anti-unitaire de renversement du temps.

Bibliographie

[1] J. J. Sakurai, "Modern Quantum mechanics," Cambridge University Press, 1982

[2] C. S. Wu, E. Ambler,R. W. Hayward, D. D. Hoppes et R. P. Hudsoon, "Experimental Test of Parity Conservation in Beta Decay," Physical Review, 1957

[3] J. H. Christenson, J. W. Cronin, V. L. et R. Turlay, "Evidence for the 2π Decay of the K0 2 Meson System," Physical Review Letters. 1964

Cours n°9

Opérateur densité et entropies

L'objectif de ce chapitre est d'introduire le formalisme de l'opérateur densité, faisant le lien entre mécanique quantique et physique statistique, et de l'appliquer à l'étude de l'entropie d'intrication, domaine de recherche actuellement très actif.

9.1 Opérateur densité et mélange statistique

Notre point de vue jusqu'à maintenant était de considérer que le système est dans un état associé à un vecteur d'état $|\psi\rangle$ bien déterminé. Les règles de la mécanique quantique permettent alors de calculer les distributions de probabilités associées aux différentes observables du système. Par contraste, nous souhaitons considérer des situations dans lesquelles le vecteur d'état n'est pas connu exactement.

Considérons encore une fois l'expérience de Stern–Gerlach, suivant Sakurai. Rappelons que faire passer les atomes d'argent dans une cavité contenant un champ magnétique non-homogène orienté selon un certain axe permet, grâce à l'instrication entre la fonction d'onde et le spin (voir le cours n° 6), de mesurer le spin selon cet axe.

Première expérience. On considère que le champ magnétique est orienté selon l'axe Oz. Les deux états de spin possibles selon l'axe Oz, $|\uparrow_z\rangle$ et $|\downarrow_z\rangle$, sont vecteurs propres de l'opérateur de spin S_z qui s'écrit dans cette base

$$S_z = \frac{\hbar}{2}\left(|\uparrow_z\rangle\langle\uparrow_z| - |\downarrow_z\rangle\langle\downarrow_z|\right) \mathrel{\widehat{=}} \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{9.1}$$

Les atomes d'argent sortant du four n'ont aucune raison d'avoir une direction de spin privilégiée – ont dit que le faisceau est *non-polarisé* – et on peut donc considérer que 50% d'entre eux sont dans l'état $\uparrow_z$ et les autres 50% sont dans l'état $\downarrow_z$. Nous avons donc un *mélange statistique* d'états : un atome pris au hasard a une probabilité $p = 1/2$ d'être dans l'état $\uparrow_z$ et $p = 1/2$ d'être dans l'état $\downarrow_z$.

Crucialement, cette notion de mélange statistique est complètement indépendante de la nature probabiliste de la mesure en mécanique quantique. Du point de vue expérimental, ce

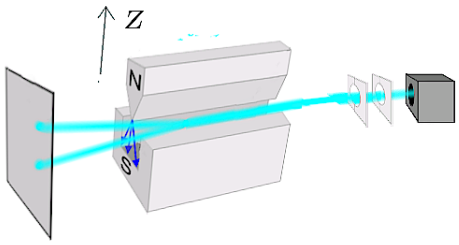

FIGURE 9.1 – Expérience de Stern et Gerlach, champ magnétique selon l'aze Oz (source : Wikipedia Commons)

mélange statistique entre les deux états de spins possibles selon Oz donne deux taches d'égale intensité sur l'écran, voir la figure 9.1

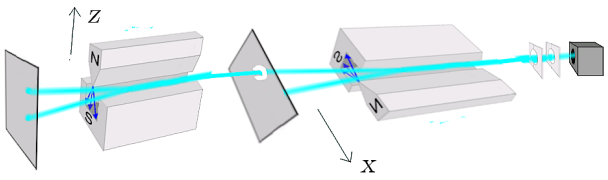

FIGURE 9.2 – Dispositifs de Stern–Gerlach successifs

Seconde expérience. Dans cette seconde expérience, le faisceau d'atomes passe dans un premier dispositif de Stern–Gerlach dont la cavité contient un champ magnétique orienté selon l'axe Ox, voir la figure 9.2.

Dans la base $\{|\uparrow_z\rangle, |\downarrow_z\rangle\}$ utilisée auparavant, l'observable associée à une mesure de spin selon l'axe Ox correspond à

$$\mathcal{S}_x = \frac{\hbar}{2}\left(|\uparrow_z\rangle\langle\downarrow_z| + |\downarrow_z\rangle\langle\uparrow_z|\right) \mathrel{\widehat{=}} \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{9.2}$$

comme nous le montrerons dans le cours n° 13. Les vecteurs propres de cet opérateur s'expriment dans la base choisie comme

$$|\uparrow_x\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_z\rangle + |\downarrow_z\rangle\right) \mathrel{\widehat{=}} \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{9.3a}$$

$$|\downarrow_x\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow_z\rangle - |\downarrow_z\rangle\right) \mathrel{\widehat{=}} \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{9.3b}$$

qui vérifient bien $\mathcal{S}_x |\uparrow_x\rangle = \frac{\hbar}{2} |\uparrow_x\rangle$ et $\mathcal{S}_x |\downarrow_x\rangle = -\frac{\hbar}{2} |\downarrow_x\rangle$.

À la sortie de cette première cavité, on ne garde que l'un des deux faisceaux ; on peut par exemple sélectionner uniquement les atomes d'argent dans l'état de spin

$$|\uparrow_x\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow_z\rangle + |\downarrow_z\rangle \right) . \tag{9.4}$$

On effectue ensuite un passage par une deuxième cavité contenant un champ magnétique selon l'axe Oz, permettant de mesurer le spin selon l'axe Oz. La probabilité de trouver l'un ou l'autre des deux états de spin possibles est :

$$p(\uparrow_z) = |\langle\uparrow_z| |\psi_1\rangle|^2 = \frac{1}{2} , \quad p(\downarrow_z) = |\langle\uparrow_z| |\psi_1\rangle|^2 = \frac{1}{2} . \tag{9.5}$$

Les deux orientations possibles sont donc équiprobables.

Pour chacune des expériences que nous venons de décrire, les probabilités d'observer les états de spin $\uparrow_z$ ou $\downarrow_z$ sont égales à 1/2 toutes les deux. Pourtant nous avons affaire physiquement à deux situations très dissemblables. Dans la première expérience nous avons un ensemble d'atomes dont statistiquement la moitié sont dans l'état $\uparrow_z$ et l'autre moitié dans l'état $\downarrow_z$. Dans la second expérience tous les atomes sont dans le *même état quantique* $(|\uparrow_z\rangle + |\downarrow_z\rangle)/\sqrt{2}$, superposition cohérente des états $\uparrow_z$ ou $\downarrow_z$.

Un moyen de se convaincre que ces deux situations sont très différentes est de remplacer la mesure finale selon Oz par une mesure selon Ox. Alors que pour la première expérience on trouverait autant d'atomes dans l'état $\uparrow_x$ que dans l'état $\downarrow_x$ (pour la même raison que l'on en trouve autant dans l'état $\uparrow_z$ que dans l'état $\downarrow_z$, car nous avons un mélange statistique des deux états) dans la deuxième experience nous sommes certains d'observer le système dans l'état $\uparrow_x$!

L'objectif de cette section est de développer un formalisme permettant de décrire les mélanges statistiques d'états, l'opérateur densité.

9.1.1 Opérateur densité pour un état pur

Soit un système quantique dans un état décrit par un vecteur $|\psi\rangle \in \mathcal{H}$. On associe à ce vecteur un opérateur appelé *opérateur densité* :

$$\boxed{\rho_\psi = |\psi\rangle \langle\psi|} \tag{9.6}$$

Cet opérateur peut être interprété comme un projecteur sur le sous-espace vectoriel de dimension 1 engendré par le vecteur $|\psi\rangle$. En effet si le vecteur $|\psi\rangle$ est normalisé nous avons :

$$\rho^2 = |\psi\rangle \langle\psi|\psi\rangle \langle\psi| = |\psi\rangle \langle\psi| = \rho . \tag{9.7}$$

Montrons qu'à l'aide de cet opérateur nous pouvons reformuler les relations importantes de la mécanique quantique. Premièrement, en représentation de Schrödinger, l'équation d'évolution de cet opérateur est donnée par :

$$\frac{d\rho}{dt} = \left(\frac{d}{dt} |\psi(t)\rangle \right) \langle\psi(t)| + |\psi(t)\rangle \left(\frac{d}{dt} \langle\psi(t)| \right) . \tag{9.8}$$

En utilisant l'équation de Schrödinger (8.70) et l'hermiticité du hamiltonien, nous avons

$$\frac{d\rho}{dt} = \left(\frac{1}{i\hbar}\mathcal{H}\,|\psi(t)\rangle\right)\langle\psi(t)| + |\psi(t)\rangle\,\langle\psi(t)|\left(-\frac{1}{i\hbar}\mathcal{H}\right) \quad\Longrightarrow\quad \boxed{i\hbar\frac{d\rho}{dt} = [\,\mathcal{H},\,\rho\,]} \tag{9.9}$$

appelée *équation de Von Neumann*.

Notons que l'opérateur densité correspondant à un état propre d'un hamiltonien indépendant du temps est indépendante du temps car

$$\forall t\ , \quad \rho(t) = e^{-iE_\ell t/\hbar}\,|\ell\rangle\left(e^{+iE_\ell t/\hbar}\,\langle\ell|\right) = |\ell\rangle\,\langle\ell|\ . \tag{9.10}$$

Deuxièmement, la valeur moyenne d'une observable $\mathcal{O}$ dans un état $|\psi\rangle$ peut s'exprimer simplement dans ce formalisme. Si on choisit une base orthonormée $\{|\ell\rangle\ ,\ \ell \in \mathcal{I}\}$ de l'espace des états nous avons :

$$\langle\mathcal{O}\rangle_\psi = \langle\psi|\,\mathcal{O}\,|\psi\rangle = \sum_{\ell\in\mathcal{I}}\langle\psi|\,\mathcal{O}\,|\ell\rangle\,\langle\ell|\psi\rangle = \sum_{\ell\in\mathcal{I}}\langle\ell|\left(\,|\psi\rangle\,\langle\psi|\,\right)\mathcal{O}\,|\ell\rangle \tag{9.11}$$

et cette dernière expression s'interprète simplement comme une trace :

$$\boxed{\langle\mathcal{O}\rangle_\psi = \mathrm{Tr}\left(\rho_\psi\mathcal{O}\right)} \tag{9.12}$$

Appliqué à l'opérateur identité, nous obtenons que

$$\mathrm{Tr}\,\rho = 1\,. \tag{9.13}$$

Nous avons vu dans le cours n° 1, éqn. (1.91), que la probabilité de mesurer une certaine valeur propre λ pour une observable était donnée par la valeur moyenne du projecteur Π_λ sur le sous-espace propre $E_\lambda \subset \mathcal{H}$ associé. Dans le formalisme de l'opérateur densité, cela donne :

$$\boxed{p(\lambda) = \mathrm{Tr}\left(\rho_\psi\Pi_\lambda\right)} \tag{9.14}$$

Finalement, on peut donner, pour une base orthonormée $\{|\ell\rangle\ ,\ \ell \in \mathcal{I}\}$ quelconque, les composantes de l'opérateur densité. Décomposons premièrement le vecteur $|\psi\rangle$ sur cette base :

$$|\psi\rangle = \sum_\ell \langle\ell|\psi\rangle\,|\ell\rangle = \sum_\ell \psi_\ell\,|\ell\rangle\ . \tag{9.15}$$

Nous avons, en insérant deux relations de fermeture,

$$\rho_\psi = |\psi\rangle\,\langle\psi| = \sum_{\ell,\ell'}|\ell\rangle\,\langle\ell|\psi\rangle\,\langle\psi|\ell'\rangle\,\langle\ell'| = \sum_{\ell,\ell'}\psi_\ell\overline{\psi}_{\ell'}\,|\ell\rangle\,\langle\ell'| \quad\Longrightarrow\quad \boxed{\rho_{\ell\ell'} = \psi_\ell\overline{\psi}_{\ell'}} \tag{9.16}$$

9.1.2 Opérateur densité pour un mélange statistique d'états

Jusqu'à maintenant, le formalisme de l'opérateur densité n'a permis qu'une reformulation de systèmes quantiques déjà étudiés. Le grand avantage de cette approche est de pouvoir décrire naturellement un mélange statistique d'états quantiques, comme l'exemple présenté dans le préambule de cette section. On se placera donc dans une situation où nous voulons décrire de manière statistique un grand nombre de copies d'un même système quantique, par exemple un grand nombre de particules, pour lesquelles nos n'avons qu'une connaissance statistique de leur état.

On choisit une base orthonormée de l'espace des états $\{|\ell\rangle\ ,\ \ell \in \mathcal{I}\}$ associé à l'une des particules, par exemple formée d'états propres du hamiltonien. On considère une répartition statistique des états décrite par l'ensemble $\{w_\ell \in [0,1], \ell \in \mathcal{I}\}$, où w_ℓ désigne la fraction de particules qui se trouvent dans l'état $|\ell\rangle$. On doit naturellement avoir

$$\sum_\ell w_\ell = 1\,. \tag{9.17}$$

On définit alors l'opérateur densité associée à ce mélange d'états par :

$$\boxed{\rho = \sum_{\ell\in\mathcal{I}} w_\ell\ |\ell\rangle\,\langle\ell|} \tag{9.18}$$

Examinons les propriétés de cette opérateur. Nous avons toujours

$$\mathrm{Tr}\,\rho = \sum_{\ell'\in\mathcal{I}} \langle\ell'|\,\rho\,|\ell'\rangle = \sum_{\ell'\in\mathcal{I}}\sum_{\ell\in\mathcal{I}} w_\ell\ \langle\ell'|\ell\rangle\,\langle\ell|\ell'\rangle = \sum_{\ell\in\mathcal{I}} w_\ell \sum_{\ell'\in\mathcal{I}} \langle\ell|\ell'\rangle\,\langle\ell'|\ell\rangle = \sum_{\ell\in\mathcal{I}} w_\ell = 1 \tag{9.19}$$

En revanche,

$$\rho^2 = \sum_{\ell',\ell} w'_\ell w_\ell\ |\ell'\rangle\,\langle\ell'|\ell\rangle\,\langle\ell| = \sum_{\ell',\ell} w'_\ell w_\ell\,\delta_{\ell,\ell'}\,|\ell'\rangle\,\langle\ell| = \sum_\ell (w_\ell)^2\,|\ell\rangle\,\langle\ell| \tag{9.20}$$

donc $\rho^2 \neq \rho$ sauf pour un état pur. Pour obtenir une caractérisation d'un état pur indépendante de la base remarquons que :

$$\mathrm{Tr}\,\rho^2 = \sum w_\ell^2\,. \tag{9.21}$$

Si le système est dans un mélange statistique d'états, nous avons

$$\forall \ell \in \mathcal{I}\ \text{ t.q. } w_\ell \neq 0\ ,\quad 0 < w_\ell < 1\ \implies\ 0 < w_\ell^2 < w_\ell\ \implies\ \sum_\ell w_\ell^2 < \sum_\ell w_\ell = 1 \tag{9.22}$$

Nous obtenons donc le critère suivant :

Propriété 1 *Soit ρ l'opérateur densité associée à un système quantique. Le système est dans un état pur si $\mathrm{Tr}\,\rho^2 = 1$, et dans un mélange statistique d'états si $\mathrm{Tr}\,\rho^2 < 1$.*

Les équations précédentes, l'équation de Von Neumann (9.9) donnant l'évolution de l'opérateur densité, l'équation (9.12) donnant l'expression de la valeur moyenne d'une observable et l'équation (9.14) donnant la distribution de probabilité associée à la mesure d'une observable restent valables pour l'opérateur densité associée à un mélange statistique d'états.

Exemple 1 (Stern–Gerlach). Revenons sur les deux expériences de Stern–Gerlach décrites au début de cette section. Pour la première expérience, nous avons un mélange statistique des deux états de spin et l'opérateur densité associée est :

$$\rho_1 = \frac{1}{2}\left(|\uparrow_z\rangle\langle\uparrow_z| + |\downarrow_z\rangle\langle\downarrow_z|\right) . \tag{9.23}$$

Nous avons alors pour la probabilité de mesurer le spin dans l'état $\uparrow_z$:

$$p(\uparrow_z) = \langle \Pi_{\uparrow_z}\rangle = \mathrm{Tr}\,\left(\rho_1 \Pi_{\uparrow_z}\right) = \mathrm{Tr}\,\left(\rho_1 |\uparrow_z\rangle\langle\uparrow_z|\right) = \mathrm{Tr}\,\left(\frac{1}{2}|\uparrow_z\rangle\langle\uparrow_z|\right) = \frac{1}{2} \tag{9.24}$$

Et pour la probabilité de mesurer le spin dans l'état $\uparrow_x$:

$$\begin{aligned} p(\uparrow_x) = \langle \Pi_{\uparrow_x}\rangle = \mathrm{Tr}\,\left(\rho_1 \Pi_{\uparrow_x}\right) = \mathrm{Tr}\,\left(\rho_1 |\uparrow_x\rangle\langle\uparrow_x|\right) &= \mathrm{Tr}\,\left(\rho_1 \frac{1}{2}\left(|\uparrow_z\rangle + |\downarrow_z\rangle\right)\left(\langle\uparrow_z| + \langle\downarrow_z|\right)\right) \\ &= \frac{1}{4}\mathrm{Tr}\,\left(|\uparrow_z\rangle\langle\uparrow_z| + |\downarrow_z\rangle\langle\downarrow_z|\right)\left(|\uparrow_z\rangle + |\downarrow_z\rangle\right)\left(\langle\uparrow_z| + \langle\downarrow_z|\right) \\ &= \frac{1}{4}\mathrm{Tr}\,\left(|\uparrow_z\rangle\langle\uparrow_z| + |\uparrow_z\rangle\langle\downarrow_z| + |\downarrow_z\rangle\langle\uparrow_z| + |\downarrow_z\rangle\langle\downarrow_z|\right) = \frac{1}{2} \end{aligned} \tag{9.25}$$

Passons à la seconde expérience où le système est dans l'état pur $\uparrow_x$. L'opérateur densité associée est donc :

$$\rho_2 = |\uparrow_x\rangle\langle\uparrow_x| = \frac{1}{2}\left(|\uparrow_z\rangle + |\downarrow_z\rangle\right)\left(\langle\uparrow_z| + \langle\downarrow_z|\right) . \tag{9.26}$$

Nous avons alors

$$p(\uparrow_z) = \langle \Pi_{\uparrow_z}\rangle = \mathrm{Tr}\,\left(\rho_2 |\uparrow_z\rangle\langle\uparrow_z|\right) = \mathrm{Tr}\,\left(\frac{1}{2}\left(|\uparrow_z\rangle + |\downarrow_z\rangle\right)\left(\langle\uparrow_z| + \langle\downarrow_z|\right)|\uparrow_z\rangle\langle\uparrow_z|\right) = \frac{1}{2} \tag{9.27}$$

et

$$p(\uparrow_x) = \langle \Pi_{\uparrow_x}\rangle = \mathrm{Tr}\,\left(\rho_2 |\uparrow_x\rangle\langle\uparrow_x|\right) = 1 . \tag{9.28}$$

9.1.3 Application : distribution canonique

On considère un système quantique, dont le hamiltonien est indépendant du temps, constitué d'un grand nombre de particules en contact avec un thermostat à la température T. [1] Le

1. On ne prendra pas ici en compte le fait que des particules indiscernables doivent suivre la statistique de Fermi–Dirac ou de Bose–Einstein. On pourra par exemple considérer un système de spins ou d'oscillateurs sur un réseau. Ou bien, comme dans l'expérience de Stern–Gerlach, on ne s'intéresse qu'aux degrés de liberté de spin d'un gaz dilué de particules dont les fonctions d'onde ont un recouvrement négligeable.

système quantique, en contact avec un système macroscopique « classique » ne peut rester dans un état pur quantique. Nous obtenons un mélange statistique d'états caractérisé par une opérateur densité. À l'équilibre nous devons avoir

$$\frac{\partial \rho}{\partial t} = 0 \implies [\,\rho,\, \mathcal{H}\,] = 0 \tag{9.29}$$

génériquement cela implique que $\rho = f(\mathcal{H})$.

Par un raisonnement développé dans le cours de physique statistique, nous savons que la distribution statistique à l'équilibre du système correspond à une distribution canonique. Si $\{|\ell\rangle\,, \ell \in \mathcal{I}\}$ est une base orthonormée de l'espace des états constituée de vecteurs propres du hamiltonien, la fraction de particules dans l'état $|\ell\rangle$ est donc donnée par le facteur de Boltzmann :

$$w_\ell = \frac{1}{Z(T)} e^{-\beta E_\ell}\ , \quad Z(T) = \sum_{\ell \in \mathcal{I}} e^{-\beta E_\ell}\ , \quad \beta = \frac{1}{k_B T}\,. \tag{9.30}$$

L'opérateur densité décrivant le système est donc :

$$\boxed{\rho(T) = \frac{1}{Z(T)} \sum_{\ell \in \mathcal{I}} e^{-\beta E_\ell}\, |\ell\rangle\, \langle \ell| = \frac{1}{Z(T)}\, e^{-\beta \mathcal{H}}} \tag{9.31}$$

Considérons maintenant la valeur moyenne d'une observable $\mathcal{O}$ en fonction de la température. Nous avons d'après (9.12) :

$$\langle \mathcal{O} \rangle(T) = \mathrm{Tr}\ (\rho(T)\mathcal{O}) \sum_{\ell'} \langle \ell'|\, (\rho(T)\mathcal{O})\, |\ell\rangle = \frac{1}{Z} \sum_{\ell',\ell} \langle \ell'|\ell\rangle\, e^{-\beta E_\ell}\, \langle \ell|\, \mathcal{O}\, |\ell'\rangle = \frac{1}{Z} \sum_{\ell} e^{-\beta E_\ell}\, \langle \ell|\, \mathcal{O}\, |\ell\rangle \tag{9.32}$$

On peut écrire cette dernière expression, indépendamment de la base choisie, comme

$$\boxed{\langle \mathcal{O} \rangle(T) = \frac{1}{Z(T)} \mathrm{Tr}\ \left(e^{-\beta \mathcal{H}} \mathcal{O}\right)} \tag{9.33}$$

Cette importante formule est le point de départ de la description quantique d'un système dans l'ensemble canonique.

Exemple 2 (spin dans un champ magnétique). Considérons un ensemble statistique de spins 1/2 dans un champ magnétique, de hamiltonien (avec $\omega > 0$) :

$$\mathcal{H} = \frac{\hbar \omega}{2}\, (|\uparrow_z\rangle\, \langle \uparrow_z| - |\downarrow_z\rangle\, \langle \downarrow_z|)\ , \tag{9.34}$$

et dont les vecteurs de base sont vecteurs propres de l'opérateur de spin S_z :

$$S_z\, |\uparrow_z\rangle = \frac{\hbar}{2}\, |\uparrow_z\rangle\ , \quad S_z\, |\downarrow_z\rangle = -\frac{\hbar}{2}\, |\downarrow_z\rangle \tag{9.35}$$

L'opérateur densité associée, dans l'ensemble canonique, est

$$\rho = \frac{1}{e^{-\beta\hbar\omega/2} + e^{\beta\hbar\omega/2}} \left(e^{-\beta\hbar\omega/2} \left|\uparrow_z\right\rangle \left\langle\uparrow_z\right| + e^{\beta\hbar\omega/2} \left|\downarrow_z\right\rangle \left\langle\downarrow_z\right| \right) . \tag{9.36}$$

La valeur moyenne du spin est alors

$$\langle \mathcal{S}_z \rangle(T) = \frac{1}{Z} \mathrm{Tr} \left(e^{-\beta\mathcal{H}} \mathcal{S}_z \right) = \frac{1}{e^{-\beta\hbar\omega/2} + e^{\beta\hbar\omega/2}} \frac{\hbar}{2} \mathrm{Tr} \left(e^{-\beta\hbar\omega/2} \left|\uparrow_z\right\rangle \left\langle\uparrow_z\right| - e^{\beta\hbar\omega/2} \left|\downarrow_z\right\rangle \left\langle\downarrow_z\right| \right) \tag{9.37}$$

soit

$$\boxed{\langle \mathcal{S}_z \rangle(T) = -\frac{\hbar}{2} \tanh \frac{\hbar\omega}{2k_B T}} \tag{9.38}$$

L'interprétation de ce résultat est simple. À très basse température ($T \ll \hbar\omega/k_B$), $\langle \mathcal{S}_z \rangle \to -\frac{\hbar}{2}$ car le système préfère se trouver dans l'état de plus basse énergie. Au contraire, à très haute température ($T \ll \hbar\omega/k_B$) on trouve $\langle \mathcal{S}_z \rangle \to 0$ en raison des fluctuations thermiques importantes.

9.2 Entropie d'un système quantique, entropie d'intrication

L'opérateur densité, qui permet de décrire un mélange statistique d'états quantiques d'un système, permet de définir une notion d'entropie pour un système quantique, appelée *entropie de Von Neumann* :

$$\boxed{S_{VN}(\rho) \overset{\text{déf.}}{=} -\mathrm{Tr} \Big(\rho \ln \rho \Big)} \tag{9.39}$$

Notons que, contrairement à l'entropie thermodynamique, l'entropie de Von Neumann est sans dimension. Pour un mélange statistique d'états, en utilisant la décomposition (9.18) de l'opérateur densité vis-à-vis d'une base de l'espace de Hilbert $\mathcal{H}$, elle s'écrit également :

$$S_{VN} = -\sum_\ell w_\ell \ln w_\ell \,, \tag{9.40}$$

faisant le lien avec l'entropie d'information ou entropie de Shannon.

Étant donné que, pour tout ℓ, $w_\ell \in [0, 1]$, l'entropie de Von Neumann est positive pour toute opérateur densité. Dans le cas d'un état pur,

$$\rho = \left|\ell\right\rangle \left\langle\ell\right| \implies S_{VN}(\rho) = 0 \,. \tag{9.41}$$

Pour un espace de Hilbert de dimension finie d, la distribution de $\{w_\ell,\, \ell \in \mathcal{I}\}$ donnant la plus grande entropie est obtenue en maximisant (9.40) avec la contrainte $\sum_\ell w_\ell = 1$. L'opérateur densité correspondante et son entropie sont données par [2] :

$$\rho = \tfrac{1}{d}\, \mathcal{I} \implies S_{VN}(\rho) = \ln d \,, \tag{9.42}$$

2. On utilise la méthode des multiplicateurs de Lagrange. Soit $F(w_\ell, \lambda) = -\sum w_\ell \ln w_\ell - \lambda(\sum w_\ell - 1)$, et on calcule sa différentielle :

$$dF = -\sum (\ln w_\ell + 1) dw_\ell - \lambda \sum dw_\ell - \big(\sum w_\ell - 1 \big) d\lambda = -\sum \big(\ln w_\ell + 1 + \lambda \big) dw_\ell - \big(\sum w_\ell - 1 \big) d\lambda \,,$$

qui s'annule pour $\lambda = -\ln w_\ell - 1$ et $w_\ell = 1/d$.

correspondant à un état de désordre maximal (tous les états sont équiprobables).

Propriété 2 *L'entropie de Von Neumann est une constante du mouvement.*

Premièrement, les valeurs propres de l'opérateur densité étant toutes dans l'intervalle $[0, 1]$, on peut faire le développement de Taylor :

$$S_{VN}(\rho) \overset{\text{déf.}}{=} -\operatorname{Tr}\Big\{\rho\big(\rho - \mathcal{I} - \tfrac{1}{2}(\rho-\mathcal{I})^2 + \cdots\big)\Big\} \tag{9.43}$$

L'évolution du système est associée à l'action d'un opérateur d'évolution unitaire,

$$\rho(t) = U^\dagger(t)\rho(0)U(t) \implies S_{VN}\big(\rho(t)\big) = \operatorname{Tr}\Big\{U^\dagger\rho U\Big(U^\dagger(\rho-\mathcal{I})U - \tfrac{1}{2}\big(U^\dagger(\rho-\mathcal{I})U\big)^2 + \cdots\Big)\Big\}, \tag{9.44}$$

puis il suffit d'utiliser la cyclicité de la trace. □

Exemple 3 (entropie canonique). Considérons l'opérateur densité associée à une distribution canonique, voir l'éqn. (9.31). L'entropie de Von Neumann associée satisfait :

$$k_B S_{VN}(\rho) = -\frac{k_B}{Z}\sum_\ell e^{-\beta E_\ell}\big(-\ln Z - \beta E_\ell\big) = k_B \ln Z + \frac{1}{T}\langle E\rangle_T\,, \tag{9.45}$$

coïncidant avec l'entropie thermodynamique canonique.

9.2.1 Entropie d'intrication

Considérons un système quantique pouvant être divisé en un sous-système A, associé à un espace de Hilbert $\mathcal{H}_A$ doté d'une base orthonormée $\mathcal{B}_A = \{|\ell_A\rangle\,,\ \ell_A \in \mathcal{I}_A\}$, et un sous-système B, associé à un sous-ensemble détats associé à un espace de Hilbert $\mathcal{H}_A$ doté d'une base orthonormée $\mathcal{B}_B = \{|\ell_B\rangle\,,\ \ell_B \in \mathcal{I}_B\}$.

L'espace des états associé au système dans son ensemble est donné par le produit tensoriel des espaces des états des sous-systèmes, $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, dont une base est

$$\mathcal{B} = \big\{\,|\ell_A\rangle \otimes |\ell_B\rangle\ ;\ \ell_A \in \mathcal{I}_A,\, \ell_B \in \mathcal{I}_B\big\}\,. \tag{9.46}$$

Un état *pur* arbitraire du système se développe dans cette base comme :

$$|\psi\rangle = \sum_{\ell_A,\,\ell_B} \psi_{\ell_A,\ell_B}\ |\ell_A\rangle \otimes |\ell_B\rangle\ . \tag{9.47}$$

L'opérateur densité associé à cet état est ainsi :

$$\rho_\psi = |\psi\rangle\,\langle\psi| = \sum_{m_A,\,m_B}\sum_{n_A,\,n_B} \overline{\psi_{m_A,\,m_B}}\,\psi_{n_A,\,n_B}\ |m_A\rangle\,\langle n_A| \otimes |m_B\rangle\,\langle n_B|\ , \tag{9.48}$$

le premier facteur de cet opérateur $\mathcal{H} \to \mathcal{H}$ étant un opérateur $\mathcal{H}_A \to \mathcal{H}_A$ et le second un opérateur $\mathcal{H}_B \to \mathcal{H}_B$. Il est possible de développer cet opérateur de manière plus utile en utilisant un résultat classique d'algèbre linéaire, que nous ne démontrerons pas ici :

Théorème 1 (décomposition en valeurs singulières). *Soit* M *une matrice complexe de dimension* $m \times n$*. Il existe une matrice unitaire* U *de dimension* $m \times m$*, une matrice unitaire* V *de dimension* $n \times n$*, et une matrice* D *de dimension* $m \times n$ *de la forme*

$$D_{ii} \geqslant 0 \ , \quad D_{ij} = 0 \, , i \neq j \, , \tag{9.49}$$

telles que

$$M = UDV^{\dagger} \, . \tag{9.50}$$

Les composantes d'un état pur (9.47) pour les bases choisies $\mathcal{B}_A$ et $\mathcal{B}_B$ peuvent être considérées comme les composantes d'une matrice M de dimension $d_A \times d_B$. On peut alors obtenir la propriété suivante :

Propriété 3 (décomposition de Schmidt). *Soit* $|\psi\rangle$ *un vecteur normé associé à un état pur appartenant à un espace de Hilbert produit tensoriel* $\mathcal{H}_A \otimes \mathcal{H}_B$*, où les deux facteurs ont pour dimensions respectives* d_A *et* d_B*. Soit* $d = min\,(d_A, d_B)$*. Il existe une famille orthonormée de vecteurs de* $\mathcal{H}_A$*,* $\{|k\rangle_A \, , \, k = 1, \ldots, d\}$*, une famille orthonormée de vecteurs de* $\mathcal{H}_B$*,* $\{|k\rangle_B \, , \, k = 1, \ldots, d\}$ *et une famille de coefficients réels* $\{\alpha_k \in [0, 1], \, k = 1, \ldots, d\}$*, tels que*

$$|\psi\rangle = \sum_{k=1}^{d} \alpha_k \, |k\rangle_A \otimes |k\rangle_B \ , \sum_k \alpha_k^2 = 1 \, . \tag{9.51}$$

Les coefficients α_k sont des réels positifs d'après le théorème précédent, et la condition $\sum \alpha_k^2 = 1$ (qui implique évidemment que $\alpha_k \leqslant 1$ pour tout k) est imposée par la normalisation de $|\psi\rangle$. □

Un état de la forme (9.47) peut être soit factorisé soit intriqué, voir la sous-section 6.1 du chapitre 6.1.5. On définit tout d'abord les matrices densité réduites associées à chacun des sous-systèmes par une trace partielle,

$$\rho_A \overset{\text{déf.}}{=} \mathrm{Tr}_{\mathcal{H}_B}(\rho) \ , \quad \rho_B \overset{\text{déf.}}{=} \mathrm{Tr}_{\mathcal{H}_A}(\rho) \, , \tag{9.52}$$

la trace partielle étant respectivement prise sur l'espace de Hilbert $\mathcal{H}_B$ et sur l'espace de Hilbert $\mathcal{H}_A$. Dans le premier cas de figure par exemple, on imagine que des mesures peuvent nous renseigner sur l'état du sous-système A mais que nous sommes dans l'ignorance concernant l'état du sous-système B.

Pour mesurer le degré d'intrication entre le sous-système A « connu » et le sous-système B « inconnu » on introduit une *entropie d'intrication* :

$$\boxed{S_A(\rho) \overset{\text{déf.}}{=} -\mathrm{Tr}_{\mathcal{H}_A}\left(\rho_A \ln \rho_A\right)} \tag{9.53}$$

On peut naturellement définir de la même manière une entropie d'intrication S_B à partir de l'opérateur densité réduite ρ_B.

Propriété 4 *Soit ρ l'opérateur densité associée à un état pur de $\mathcal{H}_A \otimes \mathcal{H}_B$. Nous avons l'égalité entre les entropies d'intrication associées aux sous-systèmes A et B, c.-à.-d.*

$$S_A(\rho) = S_B(\rho)\,. \tag{9.54}$$

En utilisant la décomposition de Schmidt pour un état pur, nous obtenons

$$\rho = |\psi\rangle\langle\psi| = \sum_{k,\ell=1}^{d} \alpha_k \alpha_\ell \, |k\rangle_A \, {}_A\langle\ell| \otimes |k\rangle_B \, {}_B\langle\ell| \tag{9.55}$$

D'où nous déduisons que

$$\rho_A = \mathrm{Tr}_{\mathcal{H}_B}(\rho) = \sum_{k=1}^{d} (\alpha_k)^2 \, |k\rangle_A \, {}_A\langle k| \;, \quad \rho_B = \mathrm{Tr}_{\mathcal{H}_A}(\rho) = \sum_{k=1}^{d} (\alpha_k)^2 \, |k\rangle_B \, {}_B\langle k| \;, \tag{9.56}$$

d'où découle trivialement l'égalité des entropies d'intrication S_A et S_B,

$$S_A(\rho) = S_B(\rho) = -\sum_{k=1}^{d} (\alpha_k)^2 \ln (\alpha_k)^2, \tag{9.57}$$

pour un état pur quelconque. Cette entropie est évidemment positive car $\alpha_k \in [0, 1]$ pour $k = 1, \ldots, d$. □

Considérons un état pur complètement factorisé entre les deux sous-systèmes A et B, c.-à.-d. sans intrication. Nous avons :

$$|\psi\rangle = |\psi\rangle_A \otimes |\psi\rangle_B \quad \Longrightarrow \quad \rho_\psi = \big(|\psi\rangle_A \, {}_A\langle\psi|\big) \otimes \big(|\psi\rangle_B \, {}_B\langle\psi|\big) = \rho_A \otimes \rho_B\,. \tag{9.58}$$

Dans cette expression les matrices densités $\rho_A = |\psi\rangle_A \, {}_A\langle\psi|$ et $\rho_B = |\psi\rangle_B \, {}_B\langle\psi|$ étant associées respectivement à un état pur de A et un état pur de B, nous avons

$$S_A(\rho) = S_B(\rho) = 0\,. \tag{9.59}$$

L'entropie d'intrication est donc nulle pour un état complètement factorisé.

Exemple 4 Considérons un état pur complètement intriqué entre deux spins,

$$|\psi\rangle = \tfrac{1}{\sqrt{2}}\big(|\uparrow\rangle_A \otimes |\downarrow\rangle_B + |\downarrow\rangle_A \otimes |\uparrow\rangle_B\big)\,, \tag{9.60}$$

dont l'opérateur densité est donnée par

$$\rho_\psi = \frac{1}{2}\big(|\uparrow\rangle_A \otimes |\downarrow\rangle_B + |\downarrow\rangle_A \otimes |\uparrow\rangle_B\big) \otimes \big({}_B\langle\downarrow| \otimes {}_A\langle\uparrow| + {}_B\langle\uparrow| \otimes {}_A\langle\downarrow|\big) \tag{9.61}$$

Nous en déduisons que

$$\rho_A = {}_B\langle\uparrow|\,\rho\,|\uparrow\rangle_B + {}_B\langle\downarrow|\,\rho\,|\downarrow\rangle_B = \frac{1}{2}\Big(|\downarrow\rangle_A \, {}_A\langle\downarrow| + |\uparrow\rangle_A \, {}_A\langle\uparrow|\Big)\,, \tag{9.62}$$

et de même

$$\rho_B = {}_A\langle\uparrow|\,\rho\,|\uparrow\rangle_A + {}_A\langle\downarrow|\,\rho\,|\downarrow\rangle_A = \frac{1}{2}\Big(\,|\downarrow\rangle_B\,{}_B\langle\downarrow| + |\uparrow\rangle_B\,{}_B\langle\uparrow|\,\Big)\,. \tag{9.63}$$

On en déduit que

$$S_A(\rho) = -\mathrm{Tr}_{\mathcal{H}_A}(\rho_A \ln \rho_A) = \ln 2\,, \tag{9.64}$$

et de la même manière on trouve que $S_B(\rho) = \ln 2$ également.

Cours n°10

Introduction à l'intégrale de chemin

Dans ce cours, nous nous penchons sur le processus de *quantification*, par lequel on associe à un système classique, par exemple une particule dans un potentiel, un système quantique. Même après avoir pratiqué la mécanique quantique, il est difficile de répondre à la question : *qu'est-ce que la mécanique quantique ?*

L'approche traditionnelle, dite de quantification canonique, formalisée en particulier par Dirac durant la première moitié du vingtième siècle [1], est celle que vous avez apprise et utilisée jusqu'à présent. Nous présenterons dans ce cours une approche alternative due au célèbre physicien Richard Feynman (1918-1988) : le formalisme de l'*intégrale de chemin* [2], inspirés de travaux antérieurs de Dirac. Cette approche est particulièrement utile dans le cadre de la théorie quantique des champs et fondée sur la description lagrangienne de la mécanique. Dans le cadre de la mécanique quantique non-relativiste, un de ses principaux mérites est de faire le lien entre mécanique classique et mécanique quantique de manière particulièrement intuitive et explicite.

Par contraste, le formalisme hamiltonien de la mécanique classique conduit, via la quantification canonique, à l'approche de Schrödinger, Heisenberg et Dirac à la mécanique quantique, en termes d'espace des états et d'opérateurs. Si ce formalisme en lui même est relativement simple à mettre en œuvre – d'avantage, à bien des égards, que l'approche de Feynman – le processus de quantification canonique est plus difficile à visualiser et à comprendre.

Ce chapitre doit se comprendre comme une introduction à un sujet très vaste, dont les développements physiques et mathématiques pourraient faire l'objet d'un cours entier.

10.1 L'intégrale de chemin

Le point de départ du formalisme de l'intégrale de chemin est la formulation lagrangienne de la mécanique classique, dont les rudiments ont été résumés dans le chapitre 3. Rappelons qu'on associe au système une fonctionnelle appelée action :

$$\mathcal{S}[q] \overset{\text{déf.}}{=} \int_0^T dt\, \mathcal{L}(q^i(t), \dot{q}^i(t), t)\,, \tag{10.1}$$

dont les extrema, déteminés par les équations d'Euler–Lagrange,

$$\frac{\mathcal{D}\mathcal{L}(q^i,\dot{q}^i,t)}{\mathcal{D}q^i}-\frac{d}{dt}\frac{\mathcal{D}\mathcal{L}(q^i,\dot{q}^i,t)}{\mathcal{D}\dot{q}^i}=0\,, \tag{10.2}$$

donnent la trajectoire classique suivie par le système.

Dans la dérivation des équations d'Euler–Lagrange (10.2), on considère une variation de la trajectoire autour de celle effectivement suivie par le système. Cependant, il ne s'agit ici que d'une étape intermédiaire de calcul ! Du point de vue physique, le système suit la trajectoire obéissant aux équations du mouvement et aux conditions aux limites en $t=0$ et $t=T$, correspondant à un jeu de fonctions $\{q^i(t), i=1,\ldots,n\}$, et n'explore pas les autres chemins qui ne minimisent pas l'action.

Principe de l'intégrale de chemin

Contrairement à un système classique qui suit sa trajectoire déterminée par les équations du mouvement, un système quantique va « explorer » l'ensemble des trajectoires disponibles (même si, formellement, celles-ci passent par la lune ou une galaxie lointaine !). Pour mettre en équations ce principe fondamental, il s'agit d'identifier quelle quantité calculer.

En mécanique classique, nous avons cherché à obtenir la trajectoire suivie par un système se propageant entre une position $q^i(0)=q^i_0$ (dans l'espace généralisé) et une position $q^i(T)=q^i_T$ pendant un intervalle de temps T. En mécanique quantique un tel concept n'a pas de sens, car nous ne pouvons déterminer avec certitude la position finale de la particule. Dit autrement, une fonction d'onde initialement localisée strictement en un point va s'étaler au cours du temps – voir plus loin éqn. (8.39).

Nous pouvons par contre revenir à la notion de *propagateur*, voir la définition 1 du cours 8. Rappelons que $U(q^i_0,q^i_T;T)$représente l'amplitude de probabilité (telle que le module carré donne la probabilité correspondante) du système d'évoluer du point de coordonnées q^i_0 à l'instant $t=0$ au point de coordonnées q^i_T à $t=T$.

Dans la conception de Feynman, pour évaluer cette quantité il faut considérer *tous* les chemins imaginables partant de q^i_0 à $t=0$ et aboutissant à q^i_T en $t=T$, voir la figure 10.1. On peut supposer, intuitivement, que tous les chemins n'ont pas une contribution égale, ou le même « poids », et que les chemins au voisinage de la trajectoire classique du système ont un poids prédominant. Ce principe peut faire penser à un poids statistique de Boltzmann mais – malgré une analogie formelle jouant un grand rôle en physique théorique – il est différent tant par nature que mathématiquement.

Soit un chemin correspondant à la donnée d'un jeu de fonctions $\{q^i(t), i=1,\ldots,n\}$ avec $t\in[0,T]$ satisfaisant aux conditions aux limites :

$$q^i(0)=q^i_0\,,\quad \forall i\in\{1,\ldots,n\}, \tag{10.3a}$$

$$q^i(T)=q^i_T\,,\quad \forall i\in\{1,\ldots,n\}. \tag{10.3b}$$

On associe à un tel chemin un poids complexe donné par :

$$e^{\frac{i}{\hbar}S[q^i]}\,,\quad S[q^i]=\int_0^T dt\,\mathcal{L}(q^i(t),\dot{q}^i(t),t)\,, \tag{10.4}$$

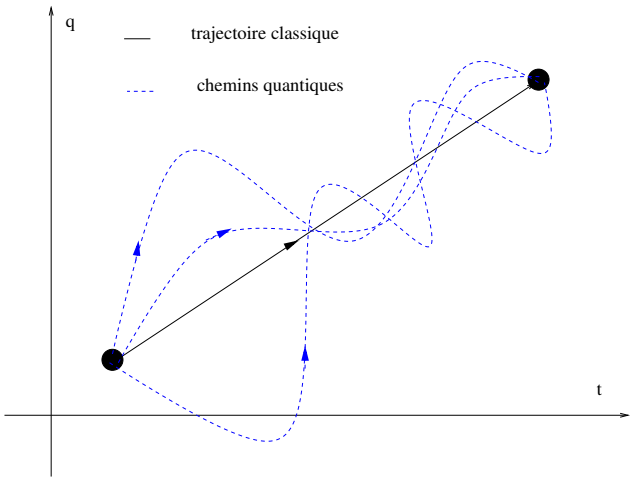

FIGURE 10.1 – Intégrale de chemin de Feynman

évalué pour une certaine trajectoire (10.3). L'action étant a priori réelle, ce poids est une pure phase, de module unité. La constante de Planck réduite $\hbar$ possède la dimension d'une énergie fois un temps, tout comme l'action.

Supposons que l'on varie le chemin par rapport à un chemin de référence. Cela va changer la valeur de l'action, et va donc induire un déphasage entre la contribution du chemin de référence et la contribution du nouveau chemin. Ces déphasages vont conduire à des interférences entre les contributions des différents chemins.

Postulat 1 (intégrale de chemin). *Le propagateur* $U(q_0^i, q_T^i; T)$ *d'un système quantique associé à un système classique d'action* $S[q_i]$ *est donné par l'expression suivante :*

$$U(q_0^i, q_T^i; T) = \int_{q^i(0)=q_0^i}^{q^i(T)=q_T^i} \mathcal{D}q^i \; e^{\frac{i}{\hbar}S[q^i]}, \tag{10.5}$$

ou l'intégrale de mesure $\mathcal{D}q^i$ *représente, formellement, l'intégration sur tous les chemins possibles entre les points extrêmes.*

Le formalisme d'intégrale de chemin permet aussi de calculer l'analogue des valeurs moyennes d'observables qui seront définies dans les cours suivants. Si on considère par exemple une certaine fonction $F(q^i)$ des coordonnées, sa valeur moyenne sera donnée, a une normalisation près, par l'intégrale $\bar{F} \propto \int \mathcal{D}q^i \, F(q^i) \exp \frac{i}{\hbar} S[q^i]$ (formule faisant là encore penser à la physique statistique).

10.2 Propagateur d'une particule libre : intégrale de chemin

On peut donner une forme explicite à l'intégrale formelle (10.5) pour des cas particuliers, l'idée générale étant de décomposer les chemins sur une base de fonctions appropriées. Pour la

plupart des systèmes néanmoins le calcul ne peut se faire de manière exacte, mais uniquement de manière approchée.

Considérons l'exemple exactement soluble d'une particule libre de masse m à une dimension, de lagrangien

$$\mathcal{L} = \frac{1}{2} m \dot{q}^2 . \tag{10.6}$$

Les équations classiques du mouvement (3.4) se réduisent naturellement à $\ddot{q} = 0$. La solution classique compatible avec des conditions aux limites données est donc

$$q_{\text{CL}}(t) = q_0 + (q_T - q_0)\frac{t}{T} . \tag{10.7}$$

Cette solution correspond à un mouvement uniforme.

Un paramétrage d'un chemin arbitraire peut alors de faire par rapport au chemin classique q_{CL} pris comme référence – sans aucune approximation à ce stade. On pose :

$$q(t) = q_{\text{CL}}(t) + f(t) , \tag{10.8a}$$
$$f(0) = f(T) = 0 . \tag{10.8b}$$

La fonction $f(t)$ est arbitraire mais doit satisfaire des conditions aux limites de Dirichlet, étant donné que la solution classique (10.7) satisfait en elle-même aux conditions aux limites (10.3).

Il est alors possible de développer $f(t)$ sur une base de fonctions satisfaisant aux conditions (10.8b), ce qui n'est rien d'autre qu'une décomposition en série de Fourier :

$$f(t) = \frac{1}{\sqrt{T}} \sum_{r \in \mathbb{N}^\star} a_r \sin \tfrac{\pi r t}{T} . \tag{10.9}$$

Les fonctions de base $t \mapsto \sin \frac{2\pi r t}{T}$ ne sont pas des solutions de l'équation du mouvement $\ddot{q} = 0$, mais n'ont aucune raison de l'être [1] ! Les degrés de liberté correspondant aux différents chemins possibles sont entièrement contenus dans le choix des coefficients réels $\{a_r,\, r \in \mathbb{N}^\star\}$. Ces coefficients prennent des valeurs arbitraires – d'un certain point de vue la particule quantique « explore » tous les chemins possibles dans l'Univers.

Si on utilise dans l'action associée au lagrangien (10.6) la décomposition (10.8) avec le développement (10.9), à l'aide de

$$\dot{q}(t) = \frac{q_T - q_0}{T} + \frac{\pi}{T^{3/2}} \sum_{r \in \mathbb{N}^\star} r\, a_r \cos \tfrac{\pi r t}{T} , \tag{10.10}$$

et

$$\int_0^T \mathrm{d}t \, \cos \frac{\pi r t}{T} \, \cos \frac{\pi r' t}{T} = \frac{T}{2} \delta_{r,r'} \tag{10.11}$$

1. Par contre, ce sont des fonctions propres de l'opérateur différentiel $-m\,\mathrm{d}^2/\mathrm{d}t^2$ apparaissant dans l'action associée au lagrangien (10.6) après intégration par parties : $\mathcal{S} = \frac{1}{2}\int_0^T \mathrm{d}t\, q(t)\,(-m\,\mathrm{d}^2/\mathrm{d}t^2)\, q(t)$.

on obtient simplement

$$\mathcal{S}[a_r] = \frac{m}{2T}\left((q_T - q_0)^2 + \frac{\pi^2}{T}\sum_{r\in\mathbb{Z}^\star} r^2 a_r^2\right). \tag{10.12}$$

Notons que notre choix de décomposition (10.9) diagonalise la forme quadratique associée à l'action de la particule libre ; il s'agit d'un principe général.

À une normalisation globale près (que l'on fixera a posteriori) on peut remplacer l'intégrale formelle sur tous les chemins par le produit d'intégrales ordinaires sur les coefficients de Fourier,

$$\int \mathcal{D}q \;\longrightarrow\; \mathcal{N}\int\prod_{r=1}^{\infty} da_r\,. \tag{10.13}$$

On obtient finalement, en remplaçant pour le propagateur (10.5) la mesure d'intégration comme (10.13) et l'action suivant (10.12) :

$$U(q_0, q_T; T) = \mathcal{N}\, e^{\frac{im(q_T-q_0)^2}{2T\hbar}} \prod_{r=1}^{\infty}\int da_r\, e^{\frac{i}{\hbar}\frac{\pi^2 r^2 m}{2T^2}a_r^2} \tag{10.14}$$

Pour effectuer l'intégrale sur chacun des coefficients de Fourier, on utilise l'intégrale de Fresnel donnée par l'éqn. (8.35) :

$$U(q_0, q_T; T) = \mathcal{N}\, e^{\frac{im(q_T-q_0)^2}{2T\hbar}} \prod_{r=1}^{\infty}\sqrt{\frac{2i\hbar T^2}{\pi m r^2}}\,. \tag{10.15}$$

Le produit infini de la dernière expression diverge, mais on peut lui donner un sens par une technique mathématique de régularisation, la *régularisation zêta* [4], que nous ne développerons pas ici (voir appendice 10.A). Contentons-nous de donner le résultat :

$$\prod_{r=1}^{\infty}\sqrt{\frac{2i\hbar T^2}{\pi m r^2}} = \left(\frac{m}{8i\pi\hbar T^2}\right)^{1/4}. \tag{10.16}$$

Il reste à évaluer le facteur de normalisation $\mathcal{N}$. Pour cela on peut faire le raisonnement suivant. Si on « coupe » un chemin arbitraire à un instant t quelconque pour le séparer en deux sous-parties, et on intègre sur toutes les positions intermédiaires possibles, on reconstitue tous les chemins possibles entre $q = q_0$ et $q = q_T$ et on ne doit donc rien changer au résultat, voir figure 10.2. Explicitement :

$$U(q_0, q_T; T) = \int dx\, U(q_0, x, t)U(x, q_T, T-t)\;, \quad \forall t \in]0, t[\,. \tag{10.17}$$

Soit, en remplaçant par (10.15) avec (10.16) et en simplifiant de part et d'autre

$$\left(\frac{m}{8i\pi\hbar}\right)^{1/4}\frac{\mathcal{N}}{\sqrt{T}}e^{\frac{im(q_T-q_0)^2}{2T\hbar}} = \left(\frac{m}{8i\pi\hbar}\right)^{1/2}\frac{\mathcal{N}^2}{\sqrt{t(T-t)}}\int dx\, e^{\frac{im(x-q_0)^2}{2t\hbar}}e^{\frac{im(x-q_T)^2}{2(T-t)\hbar}}\,, \tag{10.18}$$

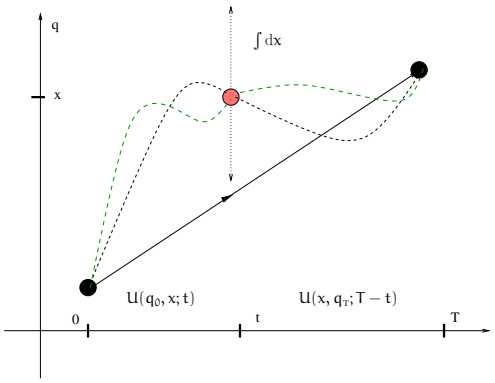

FIGURE 10.2 – Insertion d'un point intermédiaire dans le propagateur.

qui redonne une intégrale de Fresnel (8.35).

On obtient au final à partir de l'intégrale de chemin l'expression suivante du propagateur de la particule libre :

$$\boxed{U(q_0, q_T; T) = \sqrt{\frac{m}{2i\pi\hbar T}}\, e^{\frac{im(q_T - q_0)^2}{2T\hbar}}} \tag{10.19}$$

Ce résultat a déjà été obtenu au cours n° 8 à partir de l'équation de Schrödinger (dans sa version tridimensionnelle), indiquant l'équivalence entre les deux formalismes de la mécanique quantique au moins pour les particules libres.

10.3 Intégrale de chemin et limite classique

Notons la structure particulière du résultat (10.19). Le propagateur comporte un facteur de phase, $\exp \frac{i}{\hbar} S[q_{CL}]$ associé à la contribution chemin suivi par la particule dans le cadre classique, multiplié par un pré-facteur de normalisation qui peut s'interpréter comme provenant des fluctuations quantiques autour du chemin classique. On remarque que, dans cette description purement quantique, la trajectoire classique joue néanmoins un rôle privilégié.

Pour comprendre l'origine de ce résultat, qui se généralise à d'autres situations que la particule libre – par exemple à l'oscillateur harmonique – il faut examiner la limite où les fluctuations quantiques deviennent faibles. De manière imagée (mais dimensionnellement incorrecte), on souhaite considérer la limite « $\hbar \to 0$ » de l'intégrale de chemin. Une approche plus détaillée et rigoureuse de la limite classique sera donnée dans le cours n° 11.

Examinons l'expression (10.5) du propagateur défini comme une intégrale de chemin. Les « poids » associés aux différents chemins sont de la forme $\exp \frac{i}{\hbar} S$ et sont donc tous de module $\left|\exp \frac{i}{\hbar} S\right| = 1$, l'action d'une particule étant réelle. Comment peut-on comprendre que le chemin classique soit, dans un certain sens, favorisé ?

La clef du problème est de remarquer que de petites variations de la valeur de l'action entre des chemins proches l'un de l'autre vont induire un déphasage très important entre leurs contributions comme le facteur $\exp \frac{i}{\hbar} S$ va osciller très violemment. Plus précisément, si on

considère une variation infinitésimale du chemin $q(t) \mapsto q(t) + \epsilon(t)$ induisant une variation de l'action $S[q] \mapsto S[q + \epsilon] = S[q] + \delta S$, les contributions de ces deux chemins s'annuleront par interférences destructives lorsque

$$\delta S \in (2\mathbb{Z} + 1)\pi\hbar \,. \tag{10.20}$$

Génériquement, les contributions de chemins très proches – au sens du critère précédent – vont interférer de manière chaotique et leurs contributions vont, globalement, se compenser.

Pour éviter ces interférences entre des chemins infinitésimalement proches, il suffit que la variation de l'action au premier ordre soit nulle, soit

$$S[q + \epsilon] - S[q] = 0 + \mathcal{O}(\epsilon^2) \,. \tag{10.21}$$

Il faut donc que la fonction $q(t)$ corresponde à un point stationnaire de l'action, et satisfasse les équations (3.4) de la mécanique classique.

Nous en déduisons donc que *les solutions classiques des équations du mouvement d'Euler–Lagrange correspondent aux points stationnaires de l'intégrale de chemin quantique sur toutes les trajectoires.*

Si deux trajectoires $q(t)$ bien distinctes correspondent à des points stationnaires de l'action, ces deux trajectoires contribuent de manière dominante au propagateur et l'intégrale sur les trajectoires se scindera une somme de deux termes correspondant à ces points stationnaires. D'une certaine manière, la propagation d'une particule en mécanique quantique s'effectue par ces deux canaux.

Mathématiquement, l'approximation de l'intégrale de chemin au voisinage de la limite classique par les contributions des point stationnaires de l'action porte le nom de *méthode de la phase stationnaire.*

10.4 De l'intégrale de chemin à l'équation de Schrödinger

Les deux approches possibles à la mécanique quantique que nous avons évoquées, issues pour l'une de la mécanique lagrangienne et pour l'autre de la mécanique hamiltonienne, doivent être équivalentes pour tout système physique.

Au-delà de l'exemple de la particule libre, traité dans la section 10.2, il est important de montrer que les deux formalismes conduisent aux mêmes prédictions physiques. Nous donnerons ici un argumement heuristique, pouvant être présenté mathématiquement de manière plus rigoureuse. Considérons une particule massive dont le lagrangien ne dépend pas explicitement du temps, c.-à.-d. de la forme générique :

$$\mathcal{L}(q^i(t), \dot{q}^i(t), t) = \frac{m}{2} \sum_i (\dot{q}^i)^2 - V(q^i) \,. \tag{10.22}$$

Si à un instant initial $t = 0$, l'état du système est décrit par le vecteur $|\Psi(0)\rangle$, associé à une fonction d'onde $\Psi(q^i, 0)$, l'état $|\Psi(t)\rangle$ à un instant t ultérieur est associé en toute généralité à la fonction d'onde suivante :

$$\Psi(q^i, t) = \int \mathrm{d}q_0^i \, \Psi(q_0^i, 0) U(q_0^i, q^i; 0, t) \,. \tag{10.23}$$

Cette relation découle de l'équation (8.38), obtenue à partir de l'éqn. (8.38) qui, bien qu'obtenue dans ce chapitre à partir du propagateur d'une particule libre, est une conséquence directe de la définition générale du propagateur. Par la suite on posera $t = \epsilon$ et on considérera la propagation pendant un temps infiniment court.

Dans le formalisme de l'intégrale de chemin, le propagateur apparaissant dans l'équation (10.23) est donné en toute généralité par :

$$U(q_0^i, q^i; 0, \epsilon) = \mathcal{N}(\epsilon) \int_{q^i(0)=q_0^i}^{q^i(\epsilon)=q^i} \mathcal{D}q^i \; e^{\frac{i}{\hbar}\int_0^\epsilon d\tau\, \mathcal{L}(q^i(t),\dot{q}^i(\tau),\tau)} \,. \tag{10.24}$$

en termes d'un facteur de normalisation $\mathcal{N}(\epsilon)$ qui sera discuté plus tard.

Dans la limite $\epsilon \to 0^+$, la propagation de la particule, de manière intuitive, ne se fait que sur une distance infinitésimale, la trajectoire classique étant infiniment courte ; on considère ainsi dans cette limite que l'intégrale de chemin est dominée par le chemin rectiligne :

$$q^i(\tau) = q_0^i + \frac{\tau}{\epsilon}(q^i - q_0^i) \;, \quad 0 \leqslant \tau \leqslant \epsilon \,, \tag{10.25}$$

et nous pouvons également donner une expression approchée de l'action évaluée sur cette trajectoire :

$$\mathcal{S} = \epsilon \left\{ \frac{m}{2} \left(\frac{q^i - q_0^i}{\epsilon} \right)^2 - V\left(q^i\right) \right\} . \tag{10.26}$$

Le potentiel V est évalué par commodité à l'extrêmité finale du segment infinitésimal, l'erreur commise étant négligeable dans la limite $\epsilon \to 0^+$. Nous obtenons alors la relation

$$\Psi(q^i, \epsilon) = \mathcal{N}(\epsilon)\Big(1 - \frac{i\epsilon}{\hbar} V(q^i) + \mathcal{O}(\epsilon^2)\Big) \int dq_0^i \, e^{\frac{im}{2\epsilon\hbar}\left(q^i - q_0^i\right)^2} \Psi(q_0^i, 0) \,. \tag{10.27}$$

La phase présente dans l'intégrale va osciller très rapidement dans la limite $\epsilon \to 0^+$, et ainsi seules les positions initiales q_0^i proches de la position finale q_i vont contribuer significativement. On peut alors développer :

$$\Psi(q_0^i, 0) = \Psi(q^i, 0) + (q_0^i - q^i)\frac{\partial}{\partial q^i}\Psi(q^i, 0) + \frac{1}{2}(q_0^i - q^i)^2 \frac{\partial^2}{(\partial q^i)^2}\Psi(q^i, 0) + \cdots \tag{10.28}$$

et calculer les intégrales de Fresnel correspondantes, prolongement analytique d'intégrales gaussiennes, voir l'éqn. (8.35). Les puissances impaires s'annulant par parité, nous obtenons :

$$\begin{aligned} \Psi(q^i, \epsilon) &= \mathcal{N}(\epsilon) \left(\Psi(q^i, 0) + \frac{\partial^2 \Psi(q^i, 0)}{(\partial q^i)^2} \frac{\epsilon\hbar}{i} \frac{\partial}{\partial m} + \cdots \right) \sqrt{\frac{i\pi\epsilon\hbar}{m}} \Big(1 - \frac{i\epsilon}{\hbar} V(q^i) + \mathcal{O}(\epsilon^2)\Big) \\ &= \mathcal{N}(\epsilon) \left(\Psi(q^i, 0) - \frac{\partial^2 \Psi(q^i, 0)}{(\partial q^i)^2} \frac{\epsilon\hbar}{2im} + \cdots \right) \sqrt{\frac{i\pi\epsilon\hbar}{m}} \Big(1 - \frac{i\epsilon}{\hbar} V(q^i) + \mathcal{O}(\epsilon^2)\Big) \end{aligned} \tag{10.29}$$

Pour $\epsilon \to 0^+$, le premier terme du membre de droite annule bien le membre de gauche, à condition que le facteur de normalisation $\mathcal{N}(\epsilon)$ soit donné, du moins à l'ordre dominant, par :

$$\mathcal{N}(\epsilon) = \sqrt{\frac{m}{i\pi\epsilon\hbar}}, \tag{10.30}$$

en accord avec l'expression obtenue pour une particule libre dans la section précédente. Nous obtenons ensuite à l'ordre linéaire en ϵ l'équation :

$$i\hbar \frac{\Psi(q^i, \epsilon) - \Psi(q^i, 0)}{\epsilon} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(q^i, 0)}{(\partial q^i)^2} + V(q^i)\Psi(q^i, 0), \tag{10.31}$$

qui, dans la limite $\epsilon \to 0^+$, redonne comme attendu l'équation de Schrödinger évaluée à l'instant $t = 0$.

Les deux formalismes sont donc équivalents pour un système constitué d'une particule massive. Il est simple d'étendre cet argument à un ensemble de N particules.

À retenir

- la mécanique lagrangienne fournit une description alternative de la mécanique quantique, l'*intégrale de chemin* ;
- dans ce formalisme, une particule explore tous les chemins possibles d'un point à un autre ;
- chaque chemin est pondéré dans l'intégrale de chemin par un facteur de phase dépendant de la valeur de l'action correspondante ;
- dans la limite classique l'intégrale est dominée par les contributions des chemins proches des points stationnaires de l'action ;
- l'intégrale de chemin est équivalente à l'approche traditionnelle à la quantification, appelée *quantification canonique* ; dans cette dernière, les crochets de Poisson de la mécanique hamiltonienne sont remplacés par des commutateurs entre des opérateurs.

Appendices

10.A Régularisation zêta

On considère un opérateur D de valeurs propres $\{\lambda_n, n \in \mathbb{N}^\star\}$ et on cherche à calculer le produit infini $\prod_{n\in\mathbb{N}^\star} \lambda_n$ correspondant à son déterminant. En toute généralité, on peut considérer que $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n \leqslant \cdots$ et définir la fonction zêta spectrale comme :

$$\zeta_D(z) \overset{\text{déf.}}{=} \sum_{n=1}^{\infty} \lambda_n^{-z}, \tag{10.32}$$

qui converge si $\Re(z)$ est suffisamment grande. Cette fonction peut être prolongée analytiquement à tout le plan complexe sauf peut-être en un ensemble fini de points. Nous remarquons que

$$\ln\det \mathsf{D} = \sum_{n=1}^{\infty} \ln\lambda_n = -\zeta_{\mathsf{D}}'(0)\,, \tag{10.33}$$

et définissons le produit infini régularisé par

$$\widehat{\prod_{n\in\mathbb{N}^\star}} \lambda_n \overset{\text{déf.}}{=} e^{-\zeta_{\mathsf{D}}'(0)}\,. \tag{10.34}$$

Typiquement, la fonction zêta spectrale peut s'exprimer en termes de la fonction zêta de Riemann dont la présentation en série entière, donnée par

$$\zeta(z) \overset{\text{déf.}}{=} \sum_{n=1}^{\infty} \frac{1}{n^z}\,, \tag{10.35}$$

est une fonction analytique définie pour $\Re z > 1$. On peut ensuite la prolonger analytiquement sur $\mathbb{C}\backslash\{1\}$.

Dans le cas présenté ici, voir éqn. (10.16), nous devons calculer un produit infini de la forme [2] :

$$\widehat{\prod_{n\in\mathbb{N}^\star}} \frac{a}{n}\,, \quad a\in\mathbb{C}\,. \tag{10.36}$$

Nous avons premièrement

$$\zeta_{\mathsf{D}}(z) = \sum_{n=1}^{\infty} \left(\frac{a}{n}\right)^{-z} = a^{-z}\zeta(-z)\,, \tag{10.37}$$

d'où on déduit, en utilisant $\zeta(0) = -\frac{1}{2}$ et $\zeta'(0) = -\frac{1}{2}\ln 2\pi$,

$$\zeta_{\mathsf{D}}'(0) = -\ln a\,\zeta(0) - \zeta'(0) = \frac{1}{2}\ln a + \frac{1}{2}\ln 2\pi \tag{10.38}$$

et finalement

$$\widehat{\prod_{n\in\mathbb{N}^\star}} \frac{a}{n} = e^{-\zeta_{\mathsf{D}}'(0)} = \frac{1}{\sqrt{2\pi a}}\,. \tag{10.39}$$

Bibliographie

[1] P. A. M. Dirac, "The Principles of Quantum Mechanics," Oxford University Press, 1930.

[2] R. P. Feynman, "Space-time approach to nonrelativistic quantum mechanics," Rev. Mod. Phys. **20** (1948) 367. doi :10.1103/RevModPhys.20.367

2. Sur cet exemple les valeurs propres satisfont $\lambda_1 > \lambda_2 > \cdots$, en conséquence la fonction zéta spectrale est définie pour $\Re z < -1$ plutôt que pour des valeurs positives de la partie réelle.

[3] J. J. Sakurai, "Modern Quantum mechanics," Cambridge University Press, 1982

[4] S. Minakshisundaram and A. Pleijel, "Some properties of the eigenfunctions of the Laplace operator on Riemannian manifolds," Can. J. Math. **1** (1949) 242. doi :10.4153/CJM-1949-021-5

Cours n°11

Théorie des perturbations stationnaires

Pour un système quantique dont le hamiltonien $\mathcal{H}$ ne dépend pas explicitement du temps, la connaissance des valeurs propres et vecteurs propres de $\mathcal{H}$ (ou, plus généralement, du spectre de $\mathcal{H}$ et de l'opérateur unitaire lui donnant la forme d'un opérateur multiplication) permet d'en obtenir une description complète et exacte.

La diagonalisation explicite du hamiltonien est toujours possible pour un système possédant un espace des états de dimension finie, s'agissant de la diagonalisation d'une matrice hermitienne de taille finie. Dans le cas d'un espace de Hilbert de dimension infinie, la connaissance du spectre exact n'est pas toujours possible, malgré des progrès importants en physique mathématique, faisant appel a de nombreuses méthodes comme la supersymétrie.

Il est alors possible de résoudre le problème de manière approchée, si les interactions rendant le spectre non explicitement calculable sont de faible amplitude, par la méthode des *perturbations stationnaires* que nous allons développer dans ce cours.

Même lorsque le spectre exact du hamiltonien est calculable, cette méthode d'approximation est utile car elle permet de comprendre comment réagit le système lorsque les interactions sont progressivement introduites dans le système, en particulier comment les états propres du hamiltonien libre se mélangent entre eux.

11.1 Position du problème

Supposons que le hamiltonien d'un système quantique système quantique puisse se décomposer sous la forme :

$$\mathcal{H} = \mathcal{H}_0 + \epsilon \mathcal{H}_1 \,, \tag{11.1}$$

où $\mathcal{H}_0$ est un hamiltonien complètement soluble. Si $\mathcal{H}_0$ est diagonalisable on considère une base orthonormée de l'espace des états $\mathcal{H}$ constituée de vecteurs propres de $\mathcal{H}_0$, $\{\left|\ell^{(0)}\right\rangle \,,\, \ell \in \mathcal{I}\}$ tels que

$$\mathcal{H}_0 \left|\ell^{(0)}\right\rangle = \mathsf{E}_\ell^{(0)} \left|\ell^{(0)}\right\rangle \,. \tag{11.2}$$

Le paramètre sans dimension ϵ contrôle l'intensité de la perturbation. Ce paramètre peut être naturellement présent dans le modèle ; par exemple les phénomènes d'interaction entre

la matière et le rayonnement sont typiquement contrôlés par la *constante de structure fine* (voir cours n° 20) :

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \simeq \frac{1}{137} \tag{11.3}$$

Dans le cas contraire le paramètre ϵ est ajouté « à la main » et on prend la limite $\epsilon \to 1$ à la fin du calcul.

L'objectif de la méthode perturbative est de développer en puissances de ϵ tant les énergies propres que les états propres du hamiltonien total (11.1) :

$$|\ell\rangle = \left|\ell^{(0)}\right\rangle + \epsilon \left|\ell^{(1)}\right\rangle + \epsilon^2 \left|\ell^{(2)}\right\rangle + \cdots \tag{11.4a}$$

$$E_\ell = E_\ell^{(0)} + \epsilon E_\ell^{(1)} + \epsilon^2 E_\ell^{(2)} + \cdots \tag{11.4b}$$

En tronquant la série à un certain ordre en ϵ, on peut ainsi espérer obtenir une approximation de la solution exacte du problème (spectre et vecteurs propres de $\mathcal{H}_1$).

Cette méthode porte le nom de *théorie des perturbations stationnaires*. Deux questions générales peuvent se poser vis-à-vis de la validité de cette méthode :

— Pour que l'opérateur hermitien $\mathcal{H}_1$ soit considéré comme une perturbation il faut qu'il ne bouleverse pas complètement la structure de l'espace des états $\mathcal{H}$; on souhaite qu'il soit possible de développer les vecteurs propres de $\mathcal{H} = \mathcal{H}_0 + \epsilon\mathcal{H}_1$ sur une base orthonormée de $\mathcal{H}$ formée de vecteurs propres de $\mathcal{H}_0$. Par exemple, si on considère le hamiltonien d'un oscillateur harmonique,

$$\mathcal{H} = \underbrace{\frac{1}{2m}\mathcal{P}^2}_{\mathcal{H}_0} + \epsilon \underbrace{\frac{1}{2}m\omega^2\mathcal{X}^2}_{\mathcal{H}_1}, \tag{11.5}$$

il n'est pas évident qu'on puisse considérer le second terme comme une « petite » perturbation du premier. En effet le hamiltonien d'une particule libre possède n'admet pas de vecteurs propres et son spectre continu ($\sigma(\mathcal{H}_0) = \mathbb{R}^+$) alors que le spectre de l'oscillateur harmonique est discret, $\sigma(\mathcal{H}_0 + \epsilon\mathcal{H}_1) = \epsilon\hbar\omega(\mathbb{N} + 1/2)$, et que les fonctions de Hermite fournissent une base orthonormée de vecteurs propres ;

— Si on considére le développement (11.4b) comme une série entière en ϵ,

$$E_\ell(\epsilon) = \sum_{n=0}^{\infty} \epsilon^n E_\ell^{(n)} \tag{11.6}$$

se pose la question du rayon de convergence. Nous verons que typiquement le rayon de convergence est nul ! Cela ne signifie pas que le développement perturbatif est sans intérêt mais qu'il s'agit d'un *développement asymptotique* plutôt que d'une série entière. Le calcul des corrections jusqu'à un ordre donné permet tout de même d'obtenir une approximation du résultat exact.

11.1.1 Résultats généraux ⋆

Une hypothèse raisonnable pour considérer l'effet de la perturbation sous-dominant serait de demander que $\mathcal{H}_1$ soit borné (voir l'éqn. (1.31) du cours n° 1) sur le domaine de définition de $\mathcal{H}_0$. Une telle hypothèse est trop contraignante, car le hamiltonien $\mathcal{H}_0$ est en général non borné, cf. cours n° 2, et une perturbation peut être « dominée » par $\mathcal{H}_0$ sans être bornée.

Théorème 1 (théorème de Kato–Rellich). *Soit $\mathcal{O}_0$ un opérateur auto-adjoint et $\mathcal{O}_1$ un opérateur hermitien avec $D(\mathcal{O}_0) \subset D(\mathcal{O}_1)$. Si $\mathcal{O}_1$ est borné relativement à $\mathcal{O}_0$, c.-à.-d. si*

$$\exists a, b > 0\,,\ \forall |\psi\rangle \in D(\mathcal{O}_0)\ ,\quad \| \mathcal{O}_1 |\psi\rangle \| \leqslant a \| \mathcal{O}_0 |\psi\rangle \| + b \| |\psi\rangle \| \tag{11.7}$$

et si $a < 1$, alors l'opérateur $\mathcal{O}_0 + \mathcal{O}_1$ est auto-adjoint sur $D(\mathcal{O}_0)$, domaine de définition de $\mathcal{O}_0$. De plus si $\mathcal{O}_0$ est borné inférieurement alors $\mathcal{O}_0 + \mathcal{O}_1$ l'est également.

Ce théorème nous donne la condition à laquelle la perturbation est suffisamment petite pour que le hamiltonien perturbé soit auto-adjoint sur le domaine de définition du hamiltonien non-perturbé.

L'autre question est la manière dont le spectre est modifié par le hamiltonien de perturbation. On rappelle que le spectre d'un opérateur $\sigma(\mathcal{O})$ contient le *spectre ponctuel* $\sigma_p(\mathcal{O})$, qui contient les valeurs propres, c.-à.-d. les $\lambda \in \mathbb{C}$ pour lesquels on peut trouver $|\psi\rangle$ tel que $(\mathcal{O} - \lambda \mathfrak{I})\,|\psi\rangle = 0$. Le reste du spectre est appelé le *spectre essentiel* de $\mathcal{O}$:

$$\sigma_e(\mathcal{O}) \overset{\text{déf.}}{=} \sigma(\mathcal{O}) \backslash \sigma_p(\mathcal{O})\,. \tag{11.8}$$

Par exemple, pour le hamiltonien $\mathcal{H}$ de l'atome d'hydrogène (cf. cours n° 15) nous aurons :

— le spectre ponctuel est $\sigma_p(\mathcal{H}) = -E_I/(\mathbb{N}_\star)^2$ (états liés du proton et de l'électron) ;

— le spectre essentiel est $\sigma_e(\mathcal{H}) = \mathbb{R}^+$ (états où l'électron n'est pas lié au proton).

Le comportement de ces deux composantes est different sous l'effet d'une perturbation.

Théorème 2 (théorème de Weyl). *Soit $\mathcal{O}_0$ un opérateur auto-adjoint et $\mathcal{O}_1$ un opérateur hermitien . Si l'une de ces conditions est vérifée :*

— *$\mathcal{O}_1$ est un opérateur compact ;*

— *$\mathcal{O}_1$ est relaticement compact, c.-à.-d. $\exists z \in \mathbb{C}\,,\ \mathcal{O}_1(\mathcal{O}_0 - z)^{-1}$ compact*

alors $\sigma_e(\mathcal{O}_0) = \sigma_e(\mathcal{O}_0 + \mathcal{O}_1)$.

Examinons plus concrètement un cas de figure où la deuxième condition s'applique.

Propriété 1 (hamiltonien de Schrödinger). *Soit un hamiltonien d'une particule dans un potentiel V en dimension $d \leqslant 3$,*

$$\mathcal{H} = \frac{1}{2m}\vec{\mathcal{P}}^2 + V(\mathcal{X}). \tag{11.9}$$

Si on peut écrire $V(\vec{x}) = V_\infty(\vec{x}) + V_2(\vec{x})$ où $V_2(\vec{x}) \in \mathcal{L}^2(\mathbb{R}^n)$ et V_∞ bornée sur $\mathbb{R}^n$ et s'annulant à l'infini ; alors $V(\vec{\mathcal{X}})$ est relativement borné par rapport à $\mathcal{H}_0 = \frac{1}{2m}\vec{\mathcal{P}}^2$ avec $a < 1$, et relativement compact.

Dans ce cas de figure on peut appliquer premièrement le théorème de Kato–Rellich, indiquant que le hamiltonien $\mathcal{H}_0+V$ est auto-adjoint sur $D(\mathcal{H}_0)$, et borné inférieurement (le hamiltonien libre $\mathcal{H}_0$ est borné inférieurement, son spectre étant positif). Deuxièmement le théorème de Weyl implique que le spectre essentiel est invariant.

Exemple 1 (hamiltonien de l'atome d'hydrogène) L'atome d'hydrogène, qui sera étudié en détail dans le cours n° 15, est décrit par le hamiltonien

$$\mathcal{H} = \frac{1}{2m}\vec{\mathcal{P}}^2 - \frac{e^2}{4\pi\epsilon_0}\frac{1}{\|\vec{\mathcal{X}}\|}\,. \tag{11.10}$$

Soit δ_D la fonction indicatrice d'une boule de rayon r_0, c.-à.-d.

$$\delta_D(\vec{x}) = \begin{cases} 0 & \text{si} \quad \|\vec{x}\| > r_0 \\ 1 & \text{si} \quad \|\vec{x}\| \leqslant r_0 \end{cases} \tag{11.11}$$

On écrit alors

$$V(\vec{x}) = -\frac{e^2}{4\pi\epsilon_0}\Big(\underbrace{\delta_D(\vec{x})\frac{1}{\|\vec{x}\|}}_{V_2(\vec{x})} + \underbrace{\big(1-\delta_D(\vec{x})\big)\frac{1}{\|\vec{x}\|}}_{V_\infty(\vec{x})}\Big)\,. \tag{11.12}$$

Posons $r = \|\vec{x}\|$. La fonction V_2 est bien de carré sommable sur $\mathbb{R}^3$ car

$$\| V_2 \|^2 = \int_D d^3x\ |V_2(\vec{x})|^2 = 4\pi\int_0^{r_0} r^2 dr\,\frac{1}{r^2} = 4\pi r_0\,. \tag{11.13}$$

La fonction V_∞ est de son côté bornée par $1/r_0$ et tend bien vers zéro lorsque $r \to \infty$.

Les théorèmes de Kato–Rellich et de Weyl s'appliquent donc au hamiltonien de l'atome d'hydrogène (11.10). On en déduit que $\mathcal{H}$ est auto-adjoint sur $D(\mathcal{H}_0)$, voir éqn. (2.70), son spectre essentiel est le même que pour une particule libre :

$$\sigma_e(\mathcal{H}) = \sigma_e(\mathcal{H}_0) = \mathbb{R}^+\,. \tag{11.14}$$

Ce spectre correspond aux états où l'électron n'est pas lié au proton et peut, classiquement, s'échapper à l'infini. Comme annoncé plus haut, le hamiltonien de l'atome d'hydrogène possède en revanche un spectre ponctuel (états liés) que ne possède pas le hamiltonien d'une particule libre. Une autre conséquence du théorème de Kato–Rellich est que, le hamiltonien d'une particule libre état borné inférieurement (les énergies sont positives ou nulles), le hamiltonien de l'atome d'hydrogène l'est également donc ce système est stable.

Par contraste, dans le cas d'un oscillateur harmonique, disons à trois dimensions, le potentiel

$$V(\vec{x}) = \frac{1}{2}m\omega^2\vec{x}^2 \tag{11.15}$$

ne peut pas s'écrire comme somme d'une fonction bornée tendant vers zéro et d'une fonction de carré sommable. En conséquence, les théorèmes ci-dessous ne s'appliquent pas.

Même si le hamiltonien de l'oscillateur harmonique est bien auto-adjoint sur un domaine approprié, cela a pour conséquence que le spectre essentiel d'une particule libre n'est pas préservé par le potentiel. Ce résultat est certainement attendu : comme nous l'avons vu dans le cours n° 2, l'oscillateur harmonique possède une base de vecteurs propres associée à un spectre ponctuel $\sigma_p = \hbar\omega(\mathbb{N}+3/2)$, et son spectre essentiel, est vide, c.-à.-d. qu'il ne possède pas de continuum d'états contrairement à une particule libre.

11.1.2 Séries de perturbations : le cas de l'oscillateur harmonique

Avant de développer les méthodes permettant d'expliciter le développement perturbatif (11.4) des états et énergies propres en puissances du paramètre ϵ, il est utile d'illustrer la problématique de la convergence de la série (11.4b) et de la série de fonctions (11.4a) dans le cadre des perturbations au hamiltonien de l'oscillateur harmonique à une dimension,

$$\mathcal{H} = \frac{1}{2m}\mathcal{P}^2 + \frac{1}{2}m\omega^2\mathcal{X}^2 + \epsilon\mathcal{H}_1(\mathcal{X})\,. \tag{11.16}$$

Examinons au cas par cas quelques exemples simples de perturbation.

Perturbation linéaire. Considérons une perturbation de la forme

$$\mathcal{H}_1(\mathcal{X}) = -m\omega^2 x_0\mathcal{X}\,, \tag{11.17}$$

où $x_0 \in \mathbb{R}$ est une constante arbitraire. En introduisant un opérateur $\mathcal{X}_\epsilon \overset{\text{déf.}}{=} \mathcal{X} - \epsilon x_0$ il est évident que le hamiltonien (11.16) peut s'écrire

$$\mathcal{H} = \frac{1}{2m}\mathcal{P}^2 + \frac{1}{2}m\omega^2\mathcal{X}_\epsilon^2 - \frac{1}{2}m\epsilon^2\omega^2 x_0^2\,. \tag{11.18}$$

Étant donné que $[\mathcal{X}_\epsilon, \mathcal{P}] = 0$ on se réduit au problème de l'oscillateur harmonique habituel. Les énergies propres sont simplement modifiées par le terme constant dans (11.18) :

$$E_n = \hbar\omega(n + 1/2) - \frac{1}{2}m\epsilon^2\omega^2 x_0^2 \tag{11.19}$$

et, l'opérateur position étant translaté d'une constante, les solutions sont centrées en $x = \epsilon x_0$ au lieu de l'être en $x = 0$. Par exemple l'état fondamental est

$$\phi_0(x) = \Big(\frac{m\omega}{\pi\hbar}\Big)^{1/4} e^{-\frac{m\omega}{2\hbar}(x-\epsilon x_0)^2}\,. \tag{11.20}$$

Cette analyse étant valide pour tout $\epsilon \in \mathbb{R}$, on en conclut que le rayon de convergence des séries (11.4) doit être $R = +\infty$.

Perturbation quadratique. Considérons une perturbation de la forme

$$\mathcal{H}_1(X) = \frac{1}{2} m\omega^2 X^2 \,, \tag{11.21}$$

c.-à.-d. de la même forme que le potentiel initial de l'oscillateur harmonique. La théorie obtenue a donc un potentiel total $V(X) = \frac{1}{2}(1+\epsilon) m\omega^2 x^2 X$.

Le hamiltonien de l'oscillateur harmonique est auto-adjoint tant que le coefficient du potentiel est positif, assurant que le hamiltonien est borné inférieurement. Si $\epsilon < -1$ nous obtenons le hamiltonien d'un oscillateur harmonique inversé qui correspond à un système « pathologique » dont le spectre n'est pas borné inférieurement.

Nous en concluons que le rayon de convergence d'une perturbation quadratique de l'oscillateur harmonique doit être $R = 1$. Cela est confirmé par le calcul exact des énergies propres du hamiltonien. Nous avons en effet de manière évidente :

$$E_n = \hbar\omega\sqrt{1+\epsilon}(n+1/2) = \hbar\omega(n+1/2)(1+\epsilon/2+\cdots) \,, \tag{11.22}$$

et le développement en série entière de $\sqrt{1+\epsilon}$ a bien comme rayon de convergence $R = 1$.

Perturbation quartique. Considérons une perturbation part un terme quartique, c.-à.-d. de la forme

$$\mathcal{H}_1(X) = \frac{m\omega^2}{2x_0^2} X^4 \,. \tag{11.23}$$

Lorsque $x \to \infty$, le potentiel complet de la particule $W(x) = \frac{1}{2} m\omega^2 x^2 + \epsilon V(X)$ se comporte comme $W(x) \sim \epsilon x^4$.

Lorsque $\epsilon < 0$, aussi petit soit-il, nous en déduisons que le hamiltonien n'est pas borné inférieurement et nous nous retrouvons dans la situation pathologique d'un système dont le spectre n'est pas borné inférieurement. Nous en déduisons donc que, pour une perturbation de la forme (11.21), le rayon de convergence des séries perturbatives (11.4)) est égal à zéro !

Un calcul complexe qui ne sera pas reproduit ici conforte cette intuition. Les coefficients du développement (11.4b) se comportent pour un ordre k élevé comme

$$E_n^{(k)} \overset{k\gg 1}{\sim} 3^k (k-1/2)! \big(1 + \mathcal{O}(1/k)\big) \tag{11.24}$$

et le rayon de convergence déduit de cette expression est bien nul.

Séries asymptotiques. L'oscillateur harmonique perturbé par un terme quartique que nous venons d'évoquer est pourtant le cas d'école de théorie des perturbations en mécanique quantique. Il faut voir le développement (11.4b) non pas comme un développement en série mais plutôt comme un *développement asymptotique*. Une série $\sum \epsilon^k f_k$ est appelée développement asymptotique de la fonction $f(\epsilon)$ si, pour tout K,

$$f(\epsilon) - \sum_{k=1}^{K-1} \epsilon^k f_k = \mathcal{O}(\epsilon^K) \,. \tag{11.25}$$

Resommer tous les termes de la série ne redonne pas la fonction, mais, pour $|\epsilon|$ petit, la somme d'un nombre fini de termes de la série donne une bonne approximation du résultat. Considérons une série asymptotique dont les coefficients sont de la forme :

$$f_k = k! a^k k^b \,, \quad a, b > 0 \,. \tag{11.26}$$

Nous cherchons jusqu'à quel ordre K il faut pousser le développement pour minimiser l'erreur qui, d'après (11.25), est de l'ordre de $f_K \epsilon^K$. On peut dériver $f_k \epsilon^k$ par rapport à k, avec f_k donné par (11.26), en utilisant la formule de Stirling, ce qui donne :

$$f_k \sim e^{k(\ln k + \ln a - 1)} k^{b+1/2} \implies \partial_k \big(f_k \epsilon^k\big) \sim f_k (\ln k + \ln(a\epsilon) + (b+1/2)/k) \,. \tag{11.27}$$

L'ordre optimal de développement dans ce cas est donc $K \sim \frac{1}{a\epsilon}$.

Il est possible d'optimiser la convergence d'une série asymptotique en utilisant la méthode de *resommation de Borel* que nous ne développerons pas ici. De telles méthodes sont indispensables pour obtenir des prédictions d'une grande précision en physique des particules.

11.2 Perturbations dans le cas non-dégénéré

Après avoir évoqué le cadre général dans lequel s'inscrit la théorie des perturbations nous nous proposons de développer des méthodes pour obtenir explicitement les coefficients du développement perturbatif (11.4). Considérons donc un hamiltonien de la forme

$$\mathcal{H} = \mathcal{H}_0 + \epsilon \mathcal{H}_1 \,, \tag{11.28}$$

tel qu'il existe une base orthonormée de l'espace des états $\mathcal{H}$ formée de vecteurs propres de $\mathcal{H}_0$, cad une base $\{\big|\ell^{(0)}\big\rangle, \ell \in \mathcal{I}\}$ telle que

$$\forall \ell \in \mathcal{I} \,, \quad \mathcal{H}_0 = E_\ell^{(0)} \big|\ell^{(0)}\big\rangle \,. \tag{11.29}$$

On supposera également que le spectre du hamiltonien n'est pas dégénéré, c.-à.-d.

$$\ell \neq \ell' \implies E_\ell^{(0)} \neq E_{\ell'}^{(0)} \,. \tag{11.30}$$

Un vecteur propre $|\ell\rangle$ du hamiltonien $\mathcal{H}$ satisfait par définition

$$\mathcal{H} |\ell\rangle = E_\ell |\ell\rangle \,, \tag{11.31}$$

et, $\{\big|\ell^{(0)}\big\rangle\}$ étant par hypothèse une base de l'espace des états, $|\ell\rangle$ peut se développer sur cette base. L'équation (11.4a) revient à chercher à chaque ordre en ϵ le développement sur la base $\{\big|\ell^{(0)}\big\rangle\}$, en posant

$$\big|\ell^{(r)}\big\rangle = \sum_m \alpha_{m,\ell}^{(r)} \big|m^{(0)}\big\rangle \,. \tag{11.32}$$

Nous avons donc

$$|\ell\rangle = \big|\ell^{(0)}\big\rangle + \epsilon \sum_m \alpha_{m,\ell}^{(1)} \big|m^{(0)}\big\rangle + \epsilon^2 \sum_m \alpha_{m,\ell}^{(2)} \big|m^{(0)}\big\rangle + \cdots \tag{11.33}$$

Le terme d'ordre zéro en ϵ est choisi tel que, par continuité, dans la limite $\epsilon \to 0$ on retrouve un vecteur propre de $\mathcal{H}_0$. Cependant, à ce stade, le vecteur obtenu n'est pas normalisé. Nous imposerons tout d'abord la condition

$$\left\langle \ell^{(0)} \middle| \ell \right\rangle = 1 \,, \tag{11.34}$$

qui signifie à un ordre $r > 0$ donné que $\left\langle \ell^{(0)} \middle| \ell^{(r)} \right\rangle = 0$, puis nous imposerons la normalisation a posteriori du vecteur d'état obtenu en sommant les contributions des différents ordres à l'ordre en ϵ considéré.

Insérons le développement (11.33) ainsi que le développement (11.4b) des énergies propres dans l'équation aux valeur propres $\mathcal{H}\,|\ell\rangle = E_\ell\,|\ell\rangle$. Nous avons :

$$\boxed{\begin{aligned} &\Big(\mathcal{H}_0 + \epsilon\mathcal{H}_1\Big)\Big(\left|\ell^{(0)}\right\rangle + \epsilon \sum_m \alpha^{(1)}_{m,\ell} \left|m^{(0)}\right\rangle + \epsilon^2 \sum_m \alpha^{(2)}_{m,\ell} \left|m^{(0)}\right\rangle + \cdots\Big) \\ &= \Big(E^{(0)}_\ell + \epsilon E^{(1)}_\ell + \epsilon^2 E^{(2)}_\ell + \cdots\Big)\Big(\left|\ell^{(0)}\right\rangle + \epsilon \sum_m \alpha^{(1)}_{m,\ell} \left|m^{(0)}\right\rangle + \epsilon^2 \sum_m \alpha^{(2)}_{m,\ell} \left|m^{(0)}\right\rangle + \cdots\Big) \end{aligned}} \tag{11.35}$$

Le développement (11.35) doit être vérifié ordre par ordre en ϵ. Pour l'ordre zéro nous trouvons évidemment l'équation

$$\mathcal{H}_0 \left|\ell^{(0)}\right\rangle = E^{(0)}_\ell \left|\ell^{(0)}\right\rangle \,, \tag{11.36}$$

satisfaite par hypothèse.

11.2.1 Vecteur et énergie propres à l'ordre ϵ

Nous commençons la résolution systématique du problème. Les différents termes en ϵ^1 du développement (11.35) doivent satisfaire l'équation :

$$\mathcal{H}_0 \sum_m \alpha^{(1)}_{m,\ell} \left|m^{(0)}\right\rangle + \mathcal{H}_1 \left|\ell^{(0)}\right\rangle = E^{(0)}_\ell \sum_m \alpha^{(1)}_{m,\ell} \left|m^{(0)}\right\rangle + E^{(1)}_\ell \left|\ell^{(0)}\right\rangle \,. \tag{11.37}$$

Soit, $\{\left|\ell^{(0)}\right\rangle\}$ étant une base de vecteurs propres de $\mathcal{H}_0$,

$$\sum_m \alpha^{(1)}_{m,\ell} E^{(0)}_m \left|m^{(0)}\right\rangle + \mathcal{H}_1 \left|\ell^{(0)}\right\rangle = E^{(0)}_\ell \sum_m \alpha^{(1)}_{m,\ell} \left|m^{(0)}\right\rangle + E^{(1)}_\ell \left|\ell^{(0)}\right\rangle \,. \tag{11.38}$$

Multiplions cette relation à gauche par $\left\langle \ell^{(0)} \right|$. La base choisie étant orthonormée, nous avons :

$$\alpha^{(1)}_{\ell,\ell} E^{(0)}_\ell + \left\langle \ell^{(0)} \right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle = E^{(0)}_\ell \alpha^{(1)}_{\ell,\ell} + E^{(1)}_\ell \,. \tag{11.39}$$

Nous en déduisons la correction en énergie au premier ordre en ϵ :

$$\boxed{E^{(1)}_\ell = \left\langle \ell^{(0)} \right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle} \tag{11.40}$$

qui correspond à la valeur moyenne du hamiltonien de perturbation $\mathcal{H}_1$ dans l'état non perturbé $\left|\ell^{(0)}\right\rangle$.

Pour obtenir la correction au vecteur propre au premier ordre, multiplions la relation (11.41) à gauche par $\left\langle n^{(0)}\right|$ avec $n \neq \ell$. Nous avons, la base étant orthonormée :

$$E_n^{(0)}\alpha_{n,\ell}^{(1)} + \left\langle n^{(0)}\right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle = E_\ell^{(0)}\alpha_{n,\ell}^{(1)} \tag{11.41}$$

Le spectre étant non dégénéré par hypothèse, nous avons $E_\ell^{(0)} \neq E_n^{(0)}$ et nous pouvons en déduire

$$\alpha_{n,\ell}^{(1)} = \frac{1}{E_\ell^{(0)} - E_n^{(0)}} \left\langle n^{(0)}\right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle , \tag{11.42}$$

soit

$$\boxed{\left|\ell^{(1)}\right\rangle = \sum_{n\neq\ell} \frac{\left\langle n^{(0)}\right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle}{E_\ell^{(0)} - E_n^{(0)}} \left|n^{(0)}\right\rangle} \tag{11.43}$$

Cette formule indique que l'état perturbé reçoit au premier ordre en perturbations des contributions de tous les vecteurs ayant des éléments de matrice non-nuls de $\mathcal{H}_1$ avec lui.

Vérifions que le vecteur propre décrivant l'état jusqu'au premier ordre en ϵ inclus, $\left|\ell^{(0)}\right\rangle + \epsilon \left|\ell^{(1)}\right\rangle$, est bien normalisé. Nous avons

$$\begin{aligned} \left\| \left|\ell^{(0)}\right\rangle + \epsilon \left|\ell^{(1)}\right\rangle \right\|^2 &= \left(\left\langle \ell^{(0)}\right| + \epsilon \left\langle \ell^{(1)}\right| \right) \left(\left|\ell^{(0)}\right\rangle + \epsilon \left|\ell^{(1)}\right\rangle \right) \\ &= \left\langle \ell^{(0)}\middle|\ell^{(0)}\right\rangle + \epsilon \left\langle \ell^{(0)}\middle|\ell^{(1)}\right\rangle + \epsilon \left\langle \ell^{(1)}\middle|\ell^{(0)}\right\rangle + \mathcal{O}(\epsilon^2) = 1 + \mathcal{O}(\epsilon^2) , \end{aligned} \tag{11.44}$$

utilisant l'orthogonalité entre le vecteur non perturbé $\left|\ell^{(0)}\right\rangle$ et sa première correction $\left|\ell^{(1)}\right\rangle$.

11.2.2 Vecteur et énergie propres à l'ordre ϵ^2

Passons maintenant au calcul des corrections à l'ordre suivant, ϵ^2. Les termes pertinents dans le développement (11.33) conduisent à l'équation :

$$\mathcal{H}_0 \left|\ell^{(2)}\right\rangle + \mathcal{H}_1 \left|\ell^{(1)}\right\rangle = E_\ell^{(0)} \left|\ell^{(2)}\right\rangle + E_\ell^{(1)} \left|\ell^{(1)}\right\rangle + E_\ell^{(2)} \left|\ell^{(0)}\right\rangle \tag{11.45}$$

La multiplication de cette équation par $\left\langle \ell^{(0)}\right|$ donne

$$\left\langle \ell^{(0)}\right| \mathcal{H}_0 \left|\ell^{(2)}\right\rangle + \left\langle \ell^{(0)}\right| \mathcal{H}_1 \left|\ell^{(1)}\right\rangle = E_\ell^{(0)} \left\langle \ell^{(0)}\middle|\ell^{(2)}\right\rangle + E_\ell^{(1)} \left\langle \ell^{(0)}\middle|\ell^{(1)}\right\rangle + E_\ell^{(2)} \left\langle \ell^{(0)}\middle|\ell^{(0)}\right\rangle \tag{11.46}$$

Nous pouvons ensuite utiliser l'hermiticité de $\mathcal{H}_0$ et la propriété $\left\langle \ell^{(0)}\middle|\ell^{(1)}\right\rangle = 0$ pour obtenir :

$$E_\ell^{(0)} \left\langle \ell^{(0)}\middle|\ell^{(0)}\right\rangle + + \left\langle \ell^{(0)}\right| \mathcal{H}_1 \left|\ell^{(1)}\right\rangle = E_\ell^{(0)} \left\langle \ell^{(0)}\middle|\ell^{(2)}\right\rangle + E_\ell^{(2)} \tag{11.47}$$

d'où nous tirons, en utilisant le vecteur perturbé au premier ordre, éqn. (11.43) :

$$E_\ell^{(2)} = \left\langle \ell^{(0)}\right| \mathcal{H}_1 \left|\ell^{(1)}\right\rangle = \sum_{n\neq\ell} \frac{\left\langle n^{(0)}\right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle}{E_\ell^{(0)} - E_n^{(0)}} \left\langle \ell^{(0)}\right| \mathcal{H}_1 \left|n^{(0)}\right\rangle \tag{11.48}$$

Soit finalement

$$\boxed{E_\ell^{(2)} = \sum_{n\neq\ell} \frac{\left|\left\langle n^{(0)}\right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle\right|^2}{E_\ell^{(0)} - E_n^{(0)}}} \tag{11.49}$$

Pour obtenir la correction au vecteur d'état, on multiplie (11.45) par $\left\langle n^{(0)}\right|$ avec $n \neq \ell$ et on obtient :

$$\left\langle n^{(0)}\right| \mathcal{H}_0 \left|\ell^{(2)}\right\rangle + \left\langle n^{(0)}\right| \mathcal{H}_1 \left|\ell^{(1)}\right\rangle = E_\ell^{(0)} \left\langle n^{(0)}\middle|\ell^{(2)}\right\rangle + E_\ell^{(1)} \left\langle n^{(0)}\middle|\ell^{(1)}\right\rangle \tag{11.50}$$

Soit, en injectant les résultats précédents :

$$\left(E_n^{(0)} - E_\ell^{(0)}\right) \left\langle n^{(0)}\middle|\ell^{(2)}\right\rangle + \sum_{m\neq\ell} \frac{\left\langle n^{(0)}\right| \mathcal{H}_1 \left|m^{(0)}\right\rangle \left\langle m^{(0)}\right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle}{E_\ell^{(0)} - E_m^{(0)}} = \left\langle \ell^{(0)}\right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle \ \frac{\left\langle n^{(0)}\right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle}{E_\ell^{(0)} - E_n^{(0)}} \tag{11.51}$$

D'où on tire les coefficients du développement de $\left|\ell^{(2)}\right\rangle$ dans la base $\{\left|\ell^{(0)}\right\rangle\}$, avec $n \neq \ell$:

$$\alpha_{n,\ell}^{(2)} = \left\langle n^{(0)}\middle|\ell^{(2)}\right\rangle = \sum_{m\neq\ell} \frac{\left\langle n^{(0)}\right| \mathcal{H}_1 \left|m^{(0)}\right\rangle \left\langle m^{(0)}\right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle}{(E_\ell^{(0)} - E_m^{(0)})(E_\ell^{(0)} - E_n^{(0)})} - \frac{\left\langle \ell^{(0)}\right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle \left\langle n^{(0)}\right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle}{(E_\ell^{(0)} - E_n^{(0)})^2} \tag{11.52}$$

Finalement calculons la norme du résultat obtenu pour pouvoir obtenir un vecteur d'état normé. Nous avons à l'ordre ϵ^2 :

$$\begin{aligned} \left\| \left|\ell^{(0)}\right\rangle + \epsilon \left|\ell^{(1)}\right\rangle + \epsilon^2 \left|\ell^{(2)}\right\rangle \right\|^2 &= \left(\left\langle \ell^{(0)}\right| + \epsilon \left\langle \ell^{(1)}\right| + \epsilon^2 \left\langle \ell^{(2)}\right| \right) \left(\left|\ell^{(0)}\right\rangle + \epsilon \left|\ell^{(1)}\right\rangle + \epsilon^2 \left|\ell^{(2)}\right\rangle \right) \\ &= 1 + \epsilon^2 \left(\left\langle \ell^{(2)}\middle|\ell^{(0)}\right\rangle + \left\langle \ell^{(0)}\middle|\ell^{(2)}\right\rangle + \left\langle \ell^{(1)}\middle|\ell^{(1)}\right\rangle \right) + \mathcal{O}(\epsilon^3)\,, \end{aligned} \tag{11.53}$$

les termes d'ordre ϵ s'annulant en raison de la propriété $\left\langle \ell^{(0)}\middle|\ell^{(1)}\right\rangle = 0$. En utilisant de même le fait que $\left\langle \ell^{(0)}\middle|\ell^{(2)}\right\rangle = 0$ et l'équation (11.43) nous obtenons à l'ordre souhaité :

$$\left\| \left|\ell^{(0)}\right\rangle + \epsilon \left|\ell^{(1)}\right\rangle + \epsilon^2 \left|\ell^{(2)}\right\rangle \right\|^2 = 1 + \epsilon^2 \sum_{n\neq\ell} \frac{\left|\left\langle n^{(0)}\right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle\right|^2}{(E_\ell^{(0)} - E_n^{(0)})^2} \tag{11.54}$$

Suivant la terminologie de Sakurai, nous écrivons le vecteur d'état normé comme

$$\left|\ell\right\rangle_N = Z_\ell^{1/2} \left|\ell\right\rangle \ , \quad {}_N\left\langle \ell|\ell\right\rangle_N = Z_\ell \left\langle \ell|\ell\right\rangle = 1\,. \tag{11.55}$$

où le facteur Z_ℓ, appelé *renormalisation de la fonction d'onde*. Le calcul précédent nous indique que

$$Z_\ell = 1 - \epsilon^2 \sum_{n\neq\ell} \frac{\left|\left\langle n^{(0)}\right| \mathcal{H}_1 \left|\ell^{(0)}\right\rangle\right|^2}{(E_\ell^{(0)} - E_n^{(0)})^2} + \mathcal{O}(\epsilon^3)\,. \tag{11.56}$$

Étant donné que $Z_\ell^{1/2} = \left\langle \ell^{(0)}\middle|n\right\rangle_N$, le nombre Z_ℓ s'interprète comme la probabilité que le système préparé dans l'état perturbé $\left|\ell\right\rangle_N$ soit mesuré dans l'état non perturbé $\left|\ell^{(0)}\right\rangle$.

11.2.3 Exemples

Le formalisme mis en place précédemment va être illustré par deux exemples simples, un système à deux niveaux et une perturbation cubique de l'oscillateur harmonique.

Système à deux niveaux

Comme première application de la théorie des perturbation nous examinerons un système à deux niveaux. La solution explicite du problème peut naturellement être trouvée – cela revient à diagonaliser une matrice 2×2 – mais la comparaison entre le calcul exact et le calcul perturbatif est instructive.

En toute généralité, si le hamiltonien est réel dans la base choisie (et donc symétrique) il peut être mis sous la forme :

$$\mathcal{H} = \frac{\hbar\omega_0}{2}\big(|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|\big) + \frac{\hbar\omega_1}{2}\big(|\uparrow\rangle\langle\downarrow| + |\uparrow\rangle\langle\downarrow|\big) + E_0\,. \tag{11.57}$$

Nous pouvons choisir $E_0 = 0$ en redéfinissant l'origine des énergies. Utilisons la notation matricielle. Les vecteurs de base s'écrivent

$$\begin{cases} |\uparrow\rangle & \widehat{=} & \begin{pmatrix}1\\0\end{pmatrix} \\ |\downarrow\rangle & \widehat{=} & \begin{pmatrix}0\\1\end{pmatrix} \end{cases} \quad\Longrightarrow\quad \mathcal{H} \mathrel{\widehat{=}} \frac{\hbar}{2}\begin{pmatrix}\omega_0 & \omega_1\\ \omega_1 & -\omega_0\end{pmatrix}. \tag{11.58}$$

Pour obtenir simplement la solution exacte du problème (c.-à.-d. les valeurs propres et vecteurs propres de $\mathcal{H}$), introduisons un angle $\theta \in [0, \pi/4[$ tel que

$$\cos 2\theta = \frac{\omega_0}{\sqrt{\omega_0^2 + \omega_1^2}}\,, \tag{11.59a}$$

$$\sin 2\theta = \frac{\omega_1}{\sqrt{\omega_0^2 + \omega_1^2}}\,. \tag{11.59b}$$

et posons

$$\Omega \overset{\text{déf.}}{=} \sqrt{\omega_0^2 + \omega_1^2} = \omega_0\sqrt{1 + \tan^2 2\theta} = \frac{\omega_0}{\cos 2\theta}\,. \tag{11.60}$$

Il vient

$$\mathcal{H} \mathrel{\widehat{=}} \frac{\hbar\Omega}{2}\begin{pmatrix}\cos 2\theta & \sin 2\theta\\ \sin 2\theta & -\cos 2\theta\end{pmatrix}. \tag{11.61}$$

La matrice

$$M \overset{\text{déf.}}{=} \begin{pmatrix}\cos 2\theta & \sin 2\theta\\ \sin 2\theta & -\cos 2\theta\end{pmatrix} \tag{11.62}$$

a pour polynôme caractéristique $\det(\lambda\mathbb{I} - M) = \lambda^2 - 1$ donc ses valeurs propres sont $\lambda_\pm = \pm 1$. Il est aisé de vérifier qu'une base orthonormée de vecteurs propres de M, et donc de vecteurs propres de $\mathcal{H}_1$ est donnée par :

$$|+\rangle \overset{\text{déf.}}{=} \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}, \quad |-\rangle \overset{\text{déf.}}{=} \begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}. \tag{11.63}$$

Nous avons en effet

$$\begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} = \begin{pmatrix} \cos 2\theta \cos\theta + \sin 2\theta \sin\theta \\ \sin 2\theta \cos\theta - \cos 2\theta \sin\theta \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \tag{11.64}$$

et

$$\begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} = \begin{pmatrix} -\cos 2\theta \sin\theta + \sin 2\theta \cos\theta \\ -\sin 2\theta \sin\theta - \cos 2\theta \cos\theta \end{pmatrix} = -\begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} \tag{11.65}$$

Les énergies propres correspondantes sont

$$E_\pm = \pm\frac{\hbar\Omega}{2} = \pm\frac{\hbar\omega_0}{2\cos 2\theta} \tag{11.66}$$

Pour faire le lien avec la théorie des perturbations au premier ordre, développons premièrement le hamiltonien (11.61) au premier ordre en θ :

$$\mathcal{H} \mathrel{\widehat{=}} \underbrace{\frac{\hbar\omega_0}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\mathcal{H}_0} + \theta \underbrace{\hbar\omega_0 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{\mathcal{H}_1} + \mathcal{O}(\theta^2) \tag{11.67}$$

De manière équivalente le développement au premier ordre de (11.59b) donne :

$$\frac{\omega_1}{\omega_0} = 2\theta + \mathcal{O}(\theta^2) \quad \Longrightarrow \quad \theta = \frac{\omega_1}{2\omega_0} + \mathcal{O}(\theta^2)\,. \tag{11.68}$$

permettant effectivement d'identifier θ comme le paramètre de développement approprié (jouant le même rôle que ϵ dans le formalisme général). Le développement des résultats obtenus nous donne les vecteurs propres perturbés au premier ordre :

$$|+\rangle \mathrel{\widehat{=}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \theta \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \mathcal{O}(\theta^2) \tag{11.69a}$$

$$|-\rangle \mathrel{\widehat{=}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} - \theta \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mathcal{O}(\theta^2) \tag{11.69b}$$

et les énergies perturbées correspondantes sont

$$E_\pm = \pm\frac{\hbar\omega_0}{2} + \mathcal{O}(\theta^2)\,. \tag{11.70}$$

On peut voir l'angle θ dans les expressions (11.63) puis (11.69) comme un *angle de mélange*. Lorsque cet angle, donc l'intensité de la perturbation, augumente, le mélange entre les deux états propres du hamiltonien non perturbé $\mathcal{H}_0$ s'accroît. Nous avons également reproduit sur la figure 11.1 l'évolution de l'énergie des niveaux en fonction de l'angle θ d'aprés l'expression exacte (11.66). On s'aperçoit que l'écart entre le niveaux augmente avec θ, phénomène très général portant le nom de *phénomène de non-croisement des niveaux d'énergie*. Tout se passe comme si les niveaux d'énergie se repoussaient sous l'effet de la perturbation.

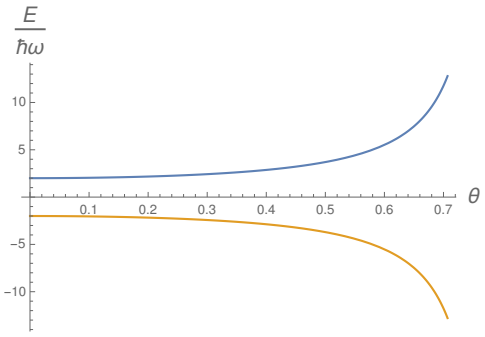

FIGURE 11.1 – Évolution de la séparation des niveaux en fonction de l'angle de mélange θ.

Reproduisons ces résultats avec la théorie des perturbations au premier ordre, voir la sous-section 11.2.1, en identifiant (à l'ordre linéaire) le paramètre de développement ϵ avec l'angle θ. Les vecteurs propres du hamiltonien non perturbé $\mathcal{H}_0$ (voir éqn. (11.67))

$$\left|+^{(0)}\right\rangle \mathrel{\widehat{=}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \qquad \left|-^{(0)}\right\rangle \mathrel{\widehat{=}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \tag{11.71}$$

La perturbation en énergie au premier ordre en θ est donnée suivant (11.40) par :

$$E_+^{(1)} = \left\langle +^{(0)}\right| \mathcal{H}_1 \left|+^{(0)}\right\rangle \mathrel{\widehat{=}} \hbar\omega_0 \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0 \tag{11.72a}$$

$$E_-^{(1)} = \left\langle -^{(0)}\right| \mathcal{H}_1 \left|-^{(0)}\right\rangle \mathrel{\widehat{=}} \hbar\omega_0 \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \tag{11.72b}$$

et s'annule en accord avec le développement (11.70) de la solution exacte. Pour les vecteurs propres en partant de la formule (11.43) nous trouvons :

$$\left|+^{(1)}\right\rangle = \frac{\left\langle -^{(0)}\right| \mathcal{H}_1 \left|+^{(0)}\right\rangle}{E_+^{(0)} - E_-^{(0)}} \left|-^{(0)}\right\rangle = \frac{\hbar\omega_0 \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}}{\hbar\omega_0} \left|-^{(0)}\right\rangle = \left|-^{(0)}\right\rangle \tag{11.73a}$$

$$\left|-^{(1)}\right\rangle = \frac{\left\langle +^{(0)}\right| \mathcal{H}_1 \left|-^{(0)}\right\rangle}{E_-^{(0)} - E_+^{(0)}} \left|+^{(0)}\right\rangle = \frac{\hbar\omega_0 \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}}{-\hbar\omega_0} \left|+^{(0)}\right\rangle = -\left|+^{(0)}\right\rangle \tag{11.73b}$$

Ces résultats sont en accord avec le développement de la solution exacte, voir éqn. (11.69).

Perturbation cubique d'un oscillateur harmonique

Considérons maintenant un oscillateur harmonique à une dimension perturbé par un terme cubique :

$$\mathcal{H} = \underbrace{\frac{1}{2m}P^2 + \frac{1}{2}m\omega^2 X^2}_{\mathcal{H}_0} + \epsilon \underbrace{\hbar\omega \left(\frac{2m\omega}{\hbar}\right)^{3/2} X^3}_{\mathcal{H}_1} \tag{11.74}$$

Nous savons déjà, par un raisonnement similaire à celui développé concernant la perturbation quartique dans la sous-section 11.1.2, que le rayon de convergence de la série de perturbations en ϵ doit être nul, car un potentiel cubique n'est pas borné inférieurement sur $\mathbb{R}$, quel que soit le signe de ϵ.

En utilisant les résultats présentés dans l'appendice 2.B du cours 2 on peut écrire ce hamiltonien sous la forme :

$$\mathcal{H} = \hbar\omega(a^\dagger a + 1/2) + \epsilon\hbar\omega(a + a^\dagger)^3 \,. \tag{11.75}$$

Pour faciliter les calculs ultérieurs on peut écrire, à l'aide de la relation de commutation (2.127), le terme de perturbation sous une forme (forme dite *ordre normal*) où l'opérateur d'annihilation a apparaît tout à droite dans tous les termes du développement de (11.75) (en utilisant de manière répétée $aa^\dagger = a^\dagger a + 1$) :

$$\begin{aligned}(a + a^\dagger)^3 &= \big(a^2 + (a^\dagger)^2 + \underbrace{aa^\dagger}_{1+a^\dagger a} + a^\dagger a\big)\big(a + a^\dagger\big) \\ &= a^3 + \underbrace{a^2 a^\dagger}_{2a+a^\dagger a^2} + (a^\dagger)^2 a + (a^\dagger)^3 + 2a^\dagger a^2 + \underbrace{2a^\dagger aa^\dagger}_{2a^\dagger(a^\dagger a+1)} + a + a^\dagger \\ &= a^3 + (a^\dagger)^3 + 3(a^\dagger)^2 a + 3a^\dagger a^2 + a + a^\dagger \,. \end{aligned} \tag{11.76}$$

Soit finalement, en introduisant l'opérateur $\mathcal{N} = a^\dagger a$ (rappelant que $\mathcal{N}|n\rangle = n|n\rangle$) :

$$(a + a^\dagger)^3 = (a^\dagger)^3 + a^3 + 3a^\dagger(\mathcal{N} + 1) + 3(\mathcal{N} + 1)a. \tag{11.77}$$

Nous pouvons ainsi obtenir facilement l'action de l'opérateur de perturbation cubique (11.75) sur un vecteur propre quelconque $\left|n^{(0)}\right\rangle$ du hamiltonien de l'oscillateur harmonique non perturbé en calculant :

$$\begin{aligned}(a + a^\dagger)^3 \left|n^{(0)}\right\rangle &= (a^\dagger)^3 \left|n^{(0)}\right\rangle + a^3 \left|n^{(0)}\right\rangle + a^\dagger(3\mathcal{N} + 1) \left|n^{(0)}\right\rangle + (3\mathcal{N} + 1)a \left|n^{(0)}\right\rangle \\ &= \sqrt{(n+1)(n+2)(n+3)} \left|(n+3)^{(0)}\right\rangle + \sqrt{n(n-1)(n-2)} \left|(n-3)^{(0)}\right\rangle + \\ &\quad + 3(n+1)^{3/2} \left|(n+1)^{(0)}\right\rangle + 3n^{3/2} \left|(n-1)^{(0)}\right\rangle \end{aligned} \tag{11.78}$$

Ce calcul n'est en principe valable que pour $n \geqslant 2$ mais on peut l'appliquer pour $n = 0, 1$ en remarquant que les termes problématiques (avec un niveau négatif) ont un coefficient nul. On remarque également que les termes diagonaux de l'opérateur de perturbation sont tous nuls.

En conséquence, la correction en énergie au premier ordre en ϵ d'un vecteur propre $\left|n^{(0)}\right\rangle$ de l'oscillateur harmonique non perturbé est égale à zéro :

$$E_n^{(1)} = \left\langle n^{(0)}\right| \mathcal{H}_1 \left|n^{(0)}\right\rangle = \hbar\omega \left\langle n^{(0)}\right| (a + a^\dagger)^3 \left|n^{(0)}\right\rangle = 0 \,, \tag{11.79}$$

utilisant l'absence de termes diagonaux pour la perturbation $\mathcal{H}_1$, voir l'éqn. (11.76).

Passons finalement au calcul de la correction au vecteurs propres au premier ordre dans le paramètre de perturbation ϵ. La formule générale (11.43) donne, pour $n > 2$:

$$\begin{aligned}\left|n^{(1)}\right\rangle &= \sum_{m\neq n} \frac{\left\langle m^{(0)}\right| \mathcal{H}_1 \left|n^{(0)}\right\rangle}{E_n^{(0)} - E_m^{(0)}} \left|m^{(0)}\right\rangle = \sum_{m\neq n} \frac{\left\langle m^{(0)}\right| (a+a^\dagger)^3 \left|n^{(0)}\right\rangle}{n-m} \left|m^{(0)}\right\rangle \\ = \sum_{m\neq n} \frac{\left|m^{(0)}\right\rangle}{n-m} \left\langle m^{(0)}\right| &\left(\sqrt{(n+1)(n+2)(n+3)} \left|(n+3)^{(0)}\right\rangle + \sqrt{n(n-1)(n-2)} \left|(n-3)^{(0)}\right\rangle \right. \\ &\left. + 3(n+1)^{3/2} \left|(n+1)^{(0)}\right\rangle + 3n^{3/2} \left|(n-1)^{(0)}\right\rangle \right) \quad (11.80)\end{aligned}$$

Pour finir le calcul, on utilise l'orthonormalité de la base, $\left\langle m^{(0)}\middle|n^{(0)}\right\rangle = \delta_{m,n}$ impliquant qu'un seul terme de la somme est non nul pour chacun des termes entre parenthèses. On obtient ainsi :

$$\begin{aligned}\left|n^{(1)}\right\rangle = -\frac{1}{3}\sqrt{(n+1)(n+2)(n+3)} \left|(n+3)^{(0)}\right\rangle + \frac{1}{3}\sqrt{n(n-1)(n-2)} \left|(n-3)^{(0)}\right\rangle \\ - 3(n+1)^{3/2} \left|(n+1)^{(0)}\right\rangle + 3n^{3/2} \left|(n-1)^{(0)}\right\rangle \quad (11.81)\end{aligned}$$

Nous avons donc finalement au premier ordre de la théorie des perturbations :

$$\begin{aligned}|n\rangle = \left|n^{(0)}\right\rangle + \epsilon \left(-\frac{1}{3}\sqrt{(n+1)(n+2)(n+3)} \left|(n+3)^{(0)}\right\rangle + \frac{1}{3}\sqrt{n(n-1)(n-2)} \left|(n-3)^{(0)}\right\rangle \right. \\ \left. - 3(n+1)^{3/2} \left|(n+1)^{(0)}\right\rangle + 3n^{3/2} \left|(n-1)^{(0)}\right\rangle \right) + \mathcal{O}(\epsilon^2)\,. \quad (11.82)\end{aligned}$$

Comme précédemment, nous pouvons utiliser ce résultat pour tout $n \in \mathbb{N}$ étant donné que les coefficients des termes problématiques pour $n = 0$ ou $n = 1$ sont nuls.

En présence de la perturbation cubique, nous pouvons par exemple calculer la probabilité que le système dans l'état perturbé $|n\rangle$ soit dans l'état $\left|(n+1)^{(0)}\right\rangle$ de l'oscillateur non perturbé :

$$p(n \to n+1) = \left|\left\langle (n+1)^{(0)}\middle|n\right\rangle\right|^2 = 9\epsilon^2(n+1)^3\,. \quad (11.83)$$

Il faut naturellement que $|\epsilon|$ soit suffisament petit pour que $p(n) \leqslant 1$.

L'analyse se poursuit sans difficulté au second ordre dans la théorie des perturbations, voir la sous-section 11.2.2, même si les calculs deviennent un peu rébarbatifs.

11.3 Perturbations dans le cas dégénéré

Dans la section 11.2 nous avons étudié l'effet d'une perturbation sur le spectre et les vecteurs propres d'un hamiltonien $\mathcal{H}_0$ dont le spectre n'était pas dégénéré, c.-à.-d. tel que deux

vecteurs propres linéairement indépendants correspondent à deux valeurs propres distinctes de $\mathcal{H}_0$. Dans ce cas chacun des vecteurs propres $|\ell\rangle$ du hamiltonien perturbé était relié sans ambiguïté à un vecteur propre $\left|\ell^{(0)}\right\rangle$ de $\mathcal{H}_0$ dans la limite $\epsilon \to 0$, par continuité.

Si une énergie $E_\ell^{(0)}$ du hamiltonien non perturbé $\mathcal{H}_0$ correspond à un sous-espace propre $\mathcal{H}_\ell \overset{\text{déf.}}{=} \ker(\mathcal{H}_0 - E_\ell^{(0)}\mathcal{I})$ de dimension $d_\ell \geqslant 2$, un problème potentiel se pose, voir la figure 11.2.

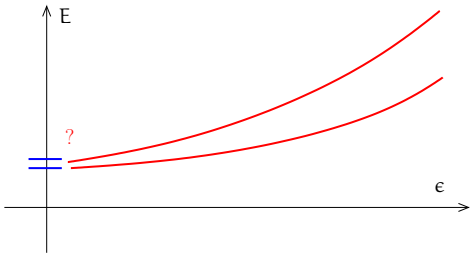

FIGURE 11.2 – Perturbation d'un niveau doublement dégénéré.

Dans le cas typique où ce niveau d'énergie dégénéré de $\mathcal{H}_0$ se scinde en plusieurs sous-niveaux sous l'effet de la perturbation, il n'est pas évident a priori comment, dans la limite $\epsilon \to 0$, relier chacun de ces sous-niveaux à l'un des niveaux dégénérés d'origine.

Considérons une base orthonormée de $\mathcal{H}_\ell$, c.-à.-d. une famille orthonormée de vecteurs $\{\left|\ell^{(0)}; i\right\rangle \ , \ i = 1, \dots d_\ell\}$ tels que :

$$\forall i \in \{1, \dots, d_\ell\} \ , \quad \mathcal{H}_0 \left|\ell^{(0)}; i\right\rangle = E_\ell^{(0)} \left|\ell^{(0)}; i\right\rangle \ . \tag{11.84}$$

On souhaite développer en puissances de ϵ l'équation aux valeurs propres pour le hamiltonien perturbé $\mathcal{H}_0 + \epsilon\mathcal{H}_1$ comme précédemment, voir les équations (11.4). On souhaiterait écrire :

$$(\mathcal{H}_0 + \epsilon\mathcal{H}_1)\left(\left|\ell^{(0)}; i\right\rangle + \epsilon\left|\ell^{(1)}; i\right\rangle + \cdots\right) \overset{?}{=} \left(E_\ell^{(0)} + \epsilon E_\ell^{(1)} + + \cdots\right)\left(\left|\ell^{(0)}; i\right\rangle + \epsilon\left|\ell^{(1)}; i\right\rangle + \cdots\right).$$

Cela pose un problème car, si les vecteurs $\left|\ell^{(0)}; i\right\rangle$ ont même énergie à l'ordre zéro, rien ne dit qu'ils aient même énergie à l'ordre suivant, ni même qu'ils soient vecteurs propres du hamiltonien perturbé.

Pour trouver la solution du problème, introduisons Π_ℓ, le projecteur sur le sous-espace propre $\mathcal{H}_\ell$:

$$\Pi_\ell = \sum_{i=1}^{d_\ell} \left|\ell^{(0)}; i\right\rangle \left\langle \ell^{(0)}; i\right| \ . \tag{11.85}$$

et $\Pi_\ell^\perp$ la projection sur son complément orthogonal. Décomposons l'équations au valeurs propres comme :

$$0 = (\mathcal{H}_0 + \epsilon\mathcal{H}_1 - E_\ell\mathcal{I})\,|\psi\rangle = \left((E_\ell^{(0)} - E_\ell)\mathcal{I} + \epsilon\mathcal{H}_1\right)\Pi_\ell\,|\psi\rangle + (\mathcal{H}_0 + \epsilon\mathcal{H}_1 - E_\ell\mathcal{I})\,\Pi_\ell^\perp\,|\psi\rangle \tag{11.86}$$

où nous avons utilisé le fait que la restriction de $\mathcal{H}_0$ au sous-espace propre $\mathcal{H}_\ell$ est, par définition,

$$\mathcal{H}_0\big|_{\mathcal{H}_\ell} = \mathsf{E}_\ell^{(0)}\mathfrak{I}_{\mathcal{H}_\ell}\,. \tag{11.87}$$

Si on multiplie à gauche l'éqn. (11.86) par Π_ℓ, on obtient une première relation concernant la projection de l'équation aux vecteurs propres dans le sous-espace $\mathcal{H}_\ell$:

$$\Pi_\ell\left((\mathsf{E}_\ell^{(0)}-\mathsf{E}_\ell)\mathfrak{I}+\epsilon\mathcal{H}_1\right)\Pi_\ell\,|\psi\rangle+\Pi_\ell\,(\mathcal{H}_0+\epsilon\mathcal{H}_1)\,\Pi_\ell^\perp\,|\psi\rangle\;. \tag{11.88}$$

Cette équation se simplifie en utilisant $\Pi_\ell^2=\Pi_\ell$, $[\,\Pi_\ell,\,\mathcal{H}_0\,]=0$ et $\Pi_\ell\Pi_\ell^\perp=0$:

$$(\mathsf{E}_\ell^{(0)}-\mathsf{E}_\ell)\Pi_\ell\,|\ell\rangle+\epsilon\Pi_\ell\mathcal{H}_1\Pi_\ell\,|\psi\rangle+\epsilon\Pi_\ell\mathcal{H}_1\Pi_\ell^\perp\,|\psi\rangle=0 \tag{11.89}$$

De même en multipliant à gauche l'éqn. (11.86) par $\Pi_\ell^\perp$ on obtient :

$$\Pi_\ell^\perp\epsilon\mathcal{H}_1\Pi_\ell\,|\psi\rangle+\left(\mathcal{H}_0+\Pi_\ell^\perp\epsilon\mathcal{H}_1-\mathsf{E}_\ell\mathfrak{I}\right)\Pi_\ell^\perp\,|\psi\rangle=0 \tag{11.90}$$

Contentons-nous dans le cadre de ce cours de calculer la première correction aux niveaux d'énergie du système. L'équation (11.90) donne à l'ordre ϵ^0 :

$$(\mathcal{H}_0-\mathsf{E}_\ell\mathfrak{I})\,\Pi_\ell^\perp\,|\psi\rangle=0\,. \tag{11.91}$$

Étant donné que, par construction, $\Pi_\ell^\perp\,|\psi\rangle\notin\mathcal{E}_\ell$ et que $\mathsf{E}_\ell^{(0)}-\mathsf{E}_\ell=\mathcal{O}(\epsilon)$, nous avons nécessairement

$$\Pi_\ell^\perp\,|\psi\rangle=\mathcal{O}(\epsilon) \tag{11.92}$$

Nous en déduisons donc, en gardant les termes d'ordre ϵ dans l'éqn. (11.89) :

$$\left((\mathsf{E}_\ell-\mathsf{E}_\ell^{(0)})\mathfrak{I}+\epsilon\Pi_\ell\mathcal{H}_1\Pi_\ell\right)\Pi_\ell\,|\psi\rangle=0 \tag{11.93}$$

Cette équation n'est rien d'autre qu'une équation aux vecteurs propres pour l'opérateur $\epsilon\Pi_\ell\mathcal{H}_1\Pi_\ell$ dans le sous-espace $\mathcal{H}_\ell$. On obtient donc les corrections à l'énergie au premier ordre en théorie des perturbations en calculant les valeurs propres de l'opérateur $\Pi_\ell\mathcal{H}_1\Pi_\ell$.

L'opérateur $\Pi_\ell\mathcal{H}_1\Pi_\ell$: $\mathcal{H}_\ell\to\mathcal{H}_\ell$, troncation de $\mathcal{H}_1$ au sous-espace propre considéré, correspond à une matrice $\mathsf{d}_\ell\times\mathsf{d}_\ell$ hermitienne dont les composantes sont, dans la base choisie :

$$\mathsf{V}_{ij}\overset{\text{déf.}}{=}\left\langle\ell^{(0)};i\right|\mathcal{H}_1\left|\ell^{(0)};j\right\rangle \tag{11.94}$$

On diagonalise la matrice V par un changement de base unitaire T, c.-à.-d.

$$\boxed{\mathsf{V}=\mathsf{T}\,\mathrm{diag}\big(\mathsf{E}_{\ell;1}^{(1)},\ldots,\mathsf{E}_{\ell;\mathsf{d}_\ell}^{(1)}\big)\,\mathsf{T}^{-1}}\,, \tag{11.95}$$

et, à l'ordre ϵ :

Le niveau $\mathsf{E}_\ell^{(0)}$ se scinde en niveaux d'énergies $\mathsf{E}_\ell^{(0)}+\epsilon\mathsf{E}_{\ell;k}^{(1)}+\mathcal{O}(\epsilon^2)$, $k=1,\ldots,\mathsf{d}_\ell$

Chaque combinaison linéaire

$$\left|\ell_k^{(1)}\right\rangle \overset{\text{déf.}}{=} \sum_{i=1}^{d_j} T_{ik} \left|\ell^{(0)}; i\right\rangle \tag{11.96}$$

est donc à la fois un vecteur propre de $\mathcal{H}_0$ et de la perturbation $\mathcal{H}_1$. Le calcul du vecteur corrigé à l'ordre ϵ est un peu plus complexe, et ne sera pas nécessaire pour les applications considérées dans la suite de ce cours.

Exemple 2 *(oscillateur harmonique 2d).* Considérons un oscillateur harmonique isotrope à deux dimensions, de hamiltonien :

$$\mathcal{H}_0 = \hbar\omega \left(a_x^\dagger a_x + a_y^\dagger a_y + 1\right) . \tag{11.97}$$

Une base d'états propres est naturellement donnée par un produit tensoriel.

$$\{|n_x; n_y\rangle \overset{\text{déf.}}{=} |n_x\rangle \otimes |n_y\rangle \ , \quad n_x, n_y \in \mathbb{N}\} \tag{11.98}$$

Les énergies propres sont donc

$$\mathcal{H}_0 |n_x; n_y\rangle = \hbar\omega(n+1) |n_x; n_y\rangle \ , \quad n \overset{\text{déf.}}{=} n_x + n_y \tag{11.99}$$

et la dégénérescence d'un niveau n est donnée par $n+1$. Par exemple pour $n = 1$ (premier niveau excité) les deux vecteurs propres d'énergie $2\hbar\omega$ sont $|1;0\rangle$ et $|0;1\rangle$. Considérons une perturbation de la forme :

$$\mathcal{H}_1 = 2m\omega^2 X \otimes Y = \hbar\omega(a_x + a_x^\dagger) \otimes (a_y + a_y^\dagger) \tag{11.100}$$

La matrice associée à la perturbation dans le sous-espace propre de $\mathcal{H}_0$ de valeur propre $2\hbar\omega$ s'obtient en calculant

$$\mathcal{H}_1 |1;0\rangle = \hbar\omega \left(a_x + a_x^\dagger\right) |1_x\rangle \otimes \left(a_y + a_y^\dagger\right) |0_y\rangle = \hbar\omega \left(|0_x\rangle + \sqrt{2}\,|2_x\rangle\right) \otimes |1_y\rangle \tag{11.101}$$

et de même

$$\mathcal{H}_1 |0;1\rangle = \hbar\omega \left(a_x + a_x^\dagger\right) |0_x\rangle \otimes \left(a_y + a_y^\dagger\right) |1_y\rangle = \hbar\omega\, |1_x\rangle \otimes \left(|0_y\rangle + \sqrt{2}\,|2_y\rangle\right) \tag{11.102}$$

Notons que le membre de droite de (11.101,11.102) contient des termes qui ne sont pas dans le sous-espace propre considéré, néanmoins nous avons vu dans le traitement général que seule la troncation au sous-espace propre du vecteur obtenu par l'action de V contribue au premier ordre. Calculons donc les éléments de matrice

$$\begin{cases} V_{10;10} &= \langle 1;0|\, \mathcal{H}_1\, |1;0\rangle = 0 \\ V_{10;01} &= \langle 1;0|\, \mathcal{H}_1\, |0;1\rangle = \hbar\omega \\ V_{01;10} &= \langle 0;1|\, \mathcal{H}_1\, |1;0\rangle = \hbar\omega \\ V_{01;01} &= \langle 0;1|\, \mathcal{H}_1\, |0;1\rangle = 0 \end{cases} \implies V = \hbar\omega \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{11.103}$$

Les vecteurs et valeurs propres de V sont naturellement

$$|\pm\rangle = \frac{1}{\sqrt{2}} \left(|1;0\rangle \pm |0;1\rangle\right) \ , \quad V|\pm\rangle = \pm\hbar\omega\, |\pm\rangle \ . \tag{11.104}$$

Les énergies des états considérés aux premier ordre inclus en théorie des perturbations sont donc

$$E = \hbar\omega(2 \pm \epsilon) . \tag{11.105}$$

À retenir

— principe de la méthode des perturbations stationnaires ;
— notion de série asymptotique ;
— perturbation d'un niveau non-dégénéré ;
— perturbation d'un niveau dégénéré.

Cours n°12

Limite semi-classique : approximation WKB

L'objectif de ce cours est de mieux comprendre le lien entre physique classique et physique quantique dans la limite semi-classique, valable lorsque les effets quantique sont « faibles », suivant un critère que nous préciserons plus loin.

Dans le cours n° 11, nous avons étudié un oscillateur harmonique perturbé par un terme cubique, de la forme :

$$\mathcal{H} = \frac{1}{2m}\mathcal{P}^2 + \frac{1}{2}m\omega^2\left(\mathcal{X}^2 - \epsilon\,\mathcal{X}^3/x_0\right)\,, \tag{12.1}$$

Grâce à la théorie des perturbations stationnaires, nous avions pu, ordre par ordre en puissances de ϵ, voir l'effet de la perturbation sur états propres et les énergies propres du système. Sur le principe, ce calcul peut être mené à un ordre arbitraire en ϵ même si, comme nous l'avions déjà remarqué, il ne peut rendre compte d'un aspect essentiel de la physique du problème.

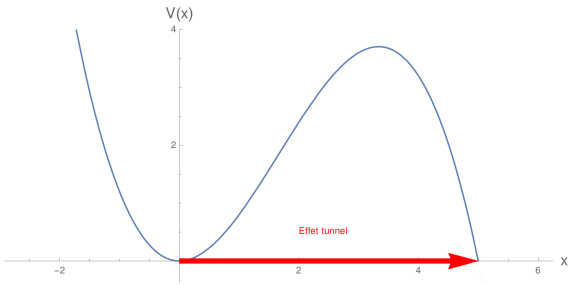

FIGURE 12.1 – Effet tunnel pour un oscillateur harmonique avec une perturbation cubique.

En effet on s'attend à ce que le système dans l'état fondamental puisse passer par effet tunnel de $x = 0$ à

$$x_c = \frac{x_0}{\epsilon} \tag{12.2}$$

cf. figure 12.1. Cela indique que le système quantique associé au hamiltonien (12.1) est instable, car son potentiel n'est pas borné inférieurement dans la région $x > x_0$ et que le système

passera inévitablement la barrière par effet tunnel quel que soit son état initial.

On supposera, pour éviter ce problème, que le potentiel est tronqué pour $x > x_c$, c.-à.-d. que $V(x > x_c) = 0$ (cela ne change rien à l'analyse en théorie des perturbations autour du minimum local $x = 0$ effectuée auparavant).

Nous verrons que le taux de transition vers la région $x > x_c$ où le potentiel est nul va en $\Gamma(\epsilon) \sim e^{-1/\epsilon^2}$, fonction qui n'est pas développable en puissances de ϵ autour de $\epsilon = 0$; cet effet est donc invisible en théorie des perturbations.

De tels effets, qui décroissent exponentiellement lorsque le paramètre de perturbation ϵ tend verz zéro, sont appelés effets *non-perturbatifs*. Il est possible de prendre en compte ces effets dans un régime d'approximation valide si le potentiel varie lentement à l'échelle de la longueur d'onde de de Broglie $\lambda = h/p$ de la particule. L'objectif de cette section est d'introduire une méthode permettant d'analyser un système quantique dans ce régime dit *semi-classique*.

12.1 Équation de Hamilton–Jacobi

Pour bien comprendre le sens physique ce cette méthode, commençons par quelques rappels de mécanique. Considérons l'action d'une particule à une dimension, évaluée sur une trajectoire satisfaisant aux équations d'Euler–Lagrange 3.4 entre les instants t_0 et t_1, du point de coordonnée q_0 au point de coordonnée q_1 à l'instant t :

$$S = \int_{t_0}^{t_1} dt\, \mathcal{L}(q, \dot{q}, t) \ , \quad q(0) = q_0 \ , \quad q(t) = q_1 \tag{12.3}$$

Considérons l'action comme fonction de la coordonnée $q_1(t)$ du point d'arrivée de la trajectoire et de t_1. Sous une variation $q \mapsto q + \delta q$ de la trajectoire, avec $q(t_0) = q_0$ fixé, la variation de l'action est donnée par

$$\delta S = \left[\frac{\partial \mathcal{L}}{\partial \dot{q}} \delta q\right]_0^t + \int_0^t dt \, \left(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}}\right) . \tag{12.4}$$

L'action étant évaluée sur une solution des équations d'Euler–Lagrange, le second terme s'annule et nous trouvons

$$\delta S = p_1 \delta q_1 \quad \Longrightarrow \quad \frac{\partial S}{\partial q_1} = p_1 \, . \tag{12.5}$$

Considérons maintenant la dépendance de l'action dans l'instant d'arrivée t_1 de la particule en $q = q_1$. Nous avons d'une part par définition $\frac{d}{dt_1} S = \mathcal{L}$ et d'autre part

$$\frac{d}{dt_1} S = \frac{\partial}{\partial t_1} S + \frac{\partial S}{\partial q_1} \dot{q}_1 \quad \Longrightarrow \quad \frac{\partial}{\partial t_1} S = \mathcal{L} - p_1 \dot{q}_1 = -\mathcal{H}(q_1, p_1, t_1) \, . \tag{12.6}$$

La différentielle totale de l'action S, vue comme fonction de l'instant d'arrivée et de la position d'arrivée, est donc

$$dS = p_1 dq_1 - H dt_1 \, . \tag{12.7}$$

En utilisant la relation (12.5), nous obtenons une équation différentielle pour l'action $\mathcal{S}(q,t)$:

$$\frac{\partial \mathcal{S}}{\partial t_1} + \mathcal{H}\left(q_1, \frac{\partial \mathcal{S}}{\partial q_1}, t_1\right) = 0\,, \tag{12.8}$$

appelée *équation de Hamilton–Jacobi*, point de départ pour étudier le mouvement de la particule équivalent aux autres formulations de la mécanique (Euler–Lagrange ou Hamilton).

Lorsque le lagrangien ne dépend pas explicitement du temps, l'énergie est conservée et nous avons

$$d\mathcal{S} = p_1 dq_1 - E dt_1 \implies \mathcal{S} = \int_{t_0}^{t_1} p dq - E(t_1 - t_0)\,. \tag{12.9}$$

On introduit l'*action réduite*

$$\tilde{\mathcal{S}} = \mathcal{S} + Et = \int_{t_0}^{t_1} p dq\,, \tag{12.10}$$

et le principe de moindre action appliqué à $\tilde{\mathcal{S}}$, ou *principe de Maupertuis*, correspond à la recherche d'un extremum sur les trajectoires à énergie fixée.

12.2 Fonction d'onde WKB

La méthode présentée dans cette section, développée principalement par Gregor Wentzel [1], Hendrik Kramers [2] et Léon Brillouin [3], porte le nom d'*approximation WKB*.

Considérons l'équation de Schrödinger pour la fonction d'onde d'une particule à une dimension dans un potentiel :

$$i\hbar \frac{\partial}{\partial t}\psi(x,t) = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right)\psi(x,t) \tag{12.11}$$

Cherchons en toute généralité une solution de l'équation d'onde de la forme :

$$\psi(x,t) = \rho(x,t) e^{\frac{i}{\hbar}S(x,t)} \tag{12.12}$$

où ρ et S sont des fonctions réelles. Nous avons

$$\frac{\partial}{\partial t}\psi(x,t) = \left(\frac{\partial \rho}{\partial t} + \frac{i\rho}{\hbar}\frac{\partial S}{\partial t}\right) e^{\frac{i}{\hbar}S(x,t)} = \left(\frac{1}{\rho}\frac{\partial \rho}{\partial t} + \frac{i}{\hbar}\frac{\partial S}{\partial t}\right)\psi(x,t) \tag{12.13a}$$

$$\begin{aligned}\frac{\partial^2}{\partial x^2}\psi(x,t) &= \left(\frac{\partial^2 \rho}{\partial x^2} + \frac{2i}{\hbar}\frac{\partial \rho}{\partial x}\frac{\partial S}{\partial x} + \frac{i\rho}{\hbar}\frac{\partial^2 S}{\partial x^2} - \frac{\rho}{\hbar^2}\left(\frac{\partial S}{\partial x}\right)^2\right) e^{\frac{i}{\hbar}S(x,t)} \\ &= \left(\frac{1}{\rho}\frac{\partial^2 \rho}{\partial x^2} + \frac{2i}{\hbar}\frac{1}{\rho}\frac{\partial \rho}{\partial x}\frac{\partial S}{\partial x} + \frac{i}{\hbar}\frac{\partial^2 S}{\partial x^2} - \frac{1}{\hbar^2}\left(\frac{\partial S}{\partial x}\right)^2\right)\psi(x,t)\end{aligned} \tag{12.13b}$$

L'équation de Schrödinger (12.11) peut alors s'écrire :

$$i\hbar\left(\frac{1}{\rho}\frac{\partial \rho}{\partial t} + \frac{i}{\hbar}\frac{\partial S}{\partial t}\right) = -\frac{\hbar^2}{2m}\left(\frac{1}{\rho}\frac{\partial^2 \rho}{\partial x^2} + \frac{2i}{\hbar}\frac{1}{\rho}\frac{\partial \rho}{\partial x}\frac{\partial S}{\partial x} + \frac{i}{\hbar}\frac{\partial^2 S}{\partial x^2} - \frac{1}{\hbar^2}\left(\frac{\partial S}{\partial x}\right)^2\right) + V(x) \tag{12.14}$$

dont on tire la partie réelle :

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 + V(x) = \frac{\hbar^2}{2m}\frac{1}{\rho}\frac{\partial^2 \rho}{\partial x^2}\,. \tag{12.15}$$

Intuitivement le dernier terme est lié aux effets quantiques, puisque c'est le seul comportant une dépendance en $\hbar$. Si on néglige ce terme, nous obtenons l'équation de Hamilton–Jacobi (12.8) en identifiant S avec l'action évaluée le long d'une trajectoire classique. Il s'agit de l'*approximation WKB*.

Pour mieux comprendre la signification de cette approximation et donner une solution explicite correspondante à l'équation de Schrödinger, nous pouvons poser, le potentiel $V(x)$ ne dépendant pas explicitement du temps :

$$\psi(x,t) = u(x)e^{-\frac{iEt}{\hbar}}\,, \quad u(x) = \rho(x)e^{\frac{i}{\hbar}(S(x,t)+Et)} = \rho(x)e^{\frac{i}{\hbar}\tilde{S}(x)}\,. \tag{12.16}$$

où nous avons introduit, en analogie avec l'action réduite (voir éqn. (12.10)), la fonction $\tilde{S}(x)$, indépendante du temps par construction :

$$\tilde{S}(x) = S(x,t) + Et\,. \tag{12.17}$$

De l'équation (12.15) nous tirons alors :

$$\left(\frac{\partial \tilde{S}}{\partial x}\right)^2 = 2m\big(E - V(x)\big) + \hbar^2 \cancel{\frac{1}{\rho}\frac{d^2\rho}{dx^2}}\,, \tag{12.18}$$

et, dans l'approximation WKB, le dernier terme est négligé. De la partie imaginaire de l'équation de Schrödinger (12.14) nous obtenons, avec $\dot{\rho} = 0$:

$$\frac{2}{\rho}\frac{d\rho}{dx}\frac{d\tilde{S}}{dx} + \frac{d^2\tilde{S}}{dx^2} = 0 \;\Leftrightarrow \frac{d}{dx}\left(\ln\frac{d\tilde{S}}{dx} + 2\ln\rho\right) = 0\,. \tag{12.19}$$

Dont la solution est

$$\rho(x) = c\left(\frac{d\tilde{S}}{dx}\right)^{-1/2}\,,\; c \in \mathbb{R}_{>0}\,. \tag{12.20}$$

Pour résoudre (12.18) dans l'approximation WKB on introduit la fonction :

$$p(x) = \begin{cases} \sqrt{2m\big(E - V(x)\big)} & \text{si} \quad E > V(x) \\ i\sqrt{2m\big(V(x) - E\big)} & \text{si} \quad E < V(x) \end{cases} \tag{12.21}$$

associée, lorsque $E > V$, à l'impulsion classique. Nous avons alors

$$\tilde{S}(x) = \pm\int_{x_0}^{x} p(x')dx' \tag{12.22}$$

d'où on tire la *fonction d'onde WKB :*

$$\boxed{u(x) = \frac{C_+}{\sqrt{p(x)}} e^{+\frac{i}{\hbar}\int^x p(x')dx'} + \frac{C_-}{\sqrt{p(x)}} e^{-\frac{i}{\hbar}\int^x p(x')dx'}} \tag{12.23}$$

Lorsque $E > V(X)$, ce qui correspond au régime autorisant des trajectoires classiques, cette solution possède localement la forme de la superposition d'une onde plane se propageant vers la gauche et d'une onde plane se propageant vers la droite. Au contraire, si $E < V(x)$ la solution est la somme d'une exponentielle décroissante, associée à un phénomène d'effet tunnel, et une exponentielle croissante qui n'est a priori pas physique.

Revenons sur la condition de validité de l'approximation WKB. Nous avons résolu l'équation de Schrödinger de manière approchée dans l'approximation :

$$\left|\hbar^2 \frac{1}{\rho}\frac{d^2\rho}{dx^2}\right| \ll \left(\frac{d\tilde{S}}{dx}\right)^2 . \tag{12.24}$$

Dans le régime où la fonction d'onde est oscillante ($p^2(x) > 0$) cette condition peut s'écrire, en utilisant l'éqn. (12.20) comme :

$$\left|\frac{3}{4}\frac{(dp/dx)^2}{p^2} - \frac{1}{2}\frac{d^2p/dx^2}{p}\right| \ll \frac{p^2(x)}{\hbar^2} \tag{12.25}$$

On prendra la condition simplifiée suivante :

$$\boxed{\left|\frac{1}{p}\frac{dp(x)}{dx}\right| \ll \frac{p}{\hbar} = \frac{2\pi}{\lambda}} \tag{12.26}$$

où $\lambda = h/p$ est la longueur d'onde de Broglie associée à la particule. Cela signifie que le potentiel varie peu sur des échelles de longueur de l'ordre de λ.

Pour mieux comprendre comment interpréter l'approximation WKB, regardons premièrement la généralisation des calculs précédents à trois dimensions. Nous avons premièrement, à la place de l'éqn. (12.18), dans l'approximation WKB,

$$\left(\vec{\nabla}\tilde{S}(\vec{x})\right)^2 = 2m\big(E - V(\vec{x})\big) \tag{12.27}$$

et nous introduisons le vecteur impulsion $\vec{p}(\vec{x})$ tel que

$$\vec{p}(\vec{x}) = \vec{\nabla}\tilde{S}(\vec{x})\,, \tag{12.28}$$

analogue à la relation de mécanique classique entre l'action réduite et l'impulsion, voir la sous-section 12.1. Au lieu de l'éqn. (12.19), nous avons :

$$\frac{2}{\rho}\vec{\nabla}\rho\cdot\vec{\nabla}\tilde{S} + \Delta\tilde{S} = 0 \implies \vec{\nabla}\cdot\left(\rho^2\vec{\nabla}\tilde{S}\right) = 0 \implies \rho(\vec{x}) = c\left(\|\vec{\nabla}\tilde{S}\|\right)^{-1/2} \tag{12.29}$$

et la fonction d'onde WKB à trois dimensions est de la forme

$$u(\vec{x}) = \frac{C_+}{\sqrt{\|\vec{p}(\vec{x})\|}} e^{+\frac{i}{\hbar}\int^{\vec{x}}\vec{p}(\vec{x}')\cdot d\vec{x}'} + \frac{C_-}{\sqrt{\|\vec{p}(\vec{x})\|}} e^{-\frac{i}{\hbar}\int^{\vec{x}}\vec{p}(\vec{x}')\cdot d\vec{x}'} \tag{12.30}$$

12.3 Approximation WKB et intégrale de chemin

Dans la phase de chacun des termes de la fonction d'onde WKB donnée par l'éqn. (12.30), le facteur $\int^{\vec{x}} \vec{p}(\vec{x}') \cdot d\vec{x}'$ correspond, tant que $E > V(\vec{x})$, à une intégrale sur la trajectoire classique du système obtenue par principe de moindre action appliqué à l'action réduite, voir l'équation (12.10).

Cette remarque nous permet de faire le lien entre l'approximation WKB et le formalisme d'intégrale de chemin présenté dans le cours n° 10. En effet, lorsque les effets quantiques sont faibles (c.-à.-d. dans la limite « $\hbar \to 0$ ») nous avons vu, en appliquant la méthode de la phase stationnaire, que l'intégrale est dominée par les extrema de l'action, correspondant aux trajectoires classiques. La contribution d'une trajectoire $\vec{x}_c(t)$ à l'intégrale de chemin est alors de la forme $\int \mathcal{D}x \exp \frac{i}{\hbar} S \sim \lambda \exp \frac{i}{\hbar} S[x_c]$, où le pré-facteur λ représente l'intégrale sur les fluctuations quantiques quadratiques autour de la trajectoire classique.

Établir précisément la correspondance entre l'approximation WKB et la limite semi-classique de l'intégrale de chemin est cependant assez technique et ne sera pas présenté ici. Brièvement, on décompose tout d'abord le propagateur sur la base des états propres du hamiltonien,

$$\begin{aligned} U(x_0, x_1; t) = \langle x_0, 0 | x_1, t \rangle &= \int dE \, \langle x_0, 0 | E \rangle \, \langle E | x_1, t \rangle = \int dE \, e^{i \frac{Et}{\hbar}} \, \langle x_0, 0 | E \rangle \, \langle E | x_1, 0 \rangle \\ &= \int dE \, e^{i \frac{Et}{\hbar}} \, \overline{u_E(x_0)} u_E(x_1) \, , \end{aligned} \tag{12.31}$$

où $u_E(x)$ est la fonction d'onde indépendante du temps associée à un état stationnaire d'énergie E. On évalue ensuite cette fonction d'onde avec la méthode WKB exposée ci-dessus, et enfin l'intégrale sur l'énergie est calculée de manière approchée par la méthode de la phase stationnaire. Le résultat est identique à celui obtenu par la contribution dominante l'intégrale de chemin dans la limite semi-classique, généralisation du calcul effectué en travaux dirigés pour l'oscillateur harmonique à un potentiel $V(x)$ arbitraire.

12.4 Points de rebroussement et effet tunnel

L'approximation WKB est clairement invalide lorsque $p(x) = 0$, c.-à.-d. pour $V(x) = E$. Pour une particule classique, d'énergie mécanique $E = \frac{p^2}{2m} + V(x)$, en un tel point l'impulsion de la particule s'annule et, si $V'(x)$ ne s'annule pas en ce point, la particule fait demi-tour. Pour cette raison un tel point est appelé *point de rebroussement*.

Considérons la situation de la figure 12.2. Une particule venant de la droite d'énergie E fera demi-tour en $x = r$ tel que $V(r) = E$. Dans la limite semi-classique, l'approximation WKB, voir l'éqn. (12.23), donne :

— une fonction d'onde oscillante dans la région $x \gg r$, avec une longueur d'onde « locale » $\lambda = h/p(x)$ augmentant à l'approche du point de rebroussement ;

— une fonction d'onde décroissant exponentiellement dans la région $x \ll r$, région interdite classiquement mais où la pénétration est possible par effet tunnel.

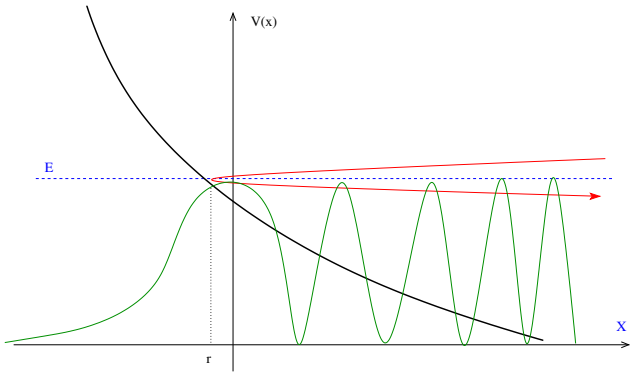

FIGURE 12.2 – Comparaison de la trajectoire classique (en rouge) et de la fonction d'onde (en vert) au voisinage d'un point de rebroussement.

Le raccord de ces deux solutions pour $x \sim r$ ne peut se faire avec l'approximation WKB qui n'est pas valide dans cette région. Nous allons plutôt résoudre exactement l'équation de Schrödinger, mais avec une approximation locale du potentiel :

$$V(x) = E + (x-r)V'(r) + \mathcal{O}\big((x-r)^2\big)\,. \tag{12.32}$$

L'équation de Schrödinger indépendante du temps dans ce potentiel linéaire est :

$$-\frac{\hbar^2}{2m}\psi''(x) + (x-r)V'(r)\psi(x) = 0\,. \tag{12.33}$$

Faisons le changement de variable (avec ici $V'(r) < 0$) :

$$y = a(r-x)\ , \quad a = \left(-\frac{2mV'(r)}{\hbar^2}\right)^{1/3} > 0\,, \tag{12.34}$$

où la dimension de a est l'inverse d'une longueur, et posons $f(y) = \psi(x)$. L'équation différentielle à résoudre est donc de la forme :

$$f''(y) - y\,f(y) = 0\,. \tag{12.35}$$

Il s'agit d'une équation d'Airy, dont la solution régulière pour $y \to -\infty$ est donnée par

$$f(y) = \lambda A_i(y)\ , \quad A_i(y) = \frac{1}{\pi}\int_0^\infty dt\, \cos(t^3/3 + yt)\,. \tag{12.36}$$

Le comportement asymptotique de cette fonction pour $y \to -\infty$ est :

$$A_i(y) \overset{y\to-\infty}{\sim} \frac{1}{\sqrt{\pi}y^{1/4}}\cos(2y^{3/2}/3 - \pi/4)\,, \tag{12.37}$$

correspondant d'après le changement de variable (12.34) à la région à droite sur la figure 12.2, où la fonction d'onde est oscillante ; ce comportement indique que le déphasage entre les deux

ondes planes dans l'équation (12.23) doit être $\Delta\phi = \pi/4$. Nous avons donc dans cette région, à une phase globale près,

$$u(x) = \frac{c}{\sqrt{p(x)}} \cos\left(\frac{1}{\hbar}\int_r^x p(x')dx' - \frac{\pi}{4}\right) . \tag{12.38}$$

ce déphasage peut s'interpréter comme une conséquence de la pénétration de la fonction d'onde dans la barrière de potentiel sur une distance effective d'un huitième de longueur d'onde avant d'être réfléchie. On peut vérifier que les fonctions (12.37) et (12.38) coïncident dans une région intermédiaire ou les deux approximations peuvent être considérées valables, moyennant un choix approprié de la constante λ en fonction de la constante c. Nous avons à droite de la barrière (c.-à.-d. pour $x > r$) :

$$p(x) = \sqrt{2m(E - V(x))} \overset{x-r\to 0}{\sim} \sqrt{-2mV'(r)(r-x)} + \mathcal{O}(r-x)^2 , \tag{12.39}$$

Donnant bien dans l'éq. (12.38) :

$$\frac{1}{\hbar}\int_r^x p(x')dx = 2\sqrt{\frac{-2mV'(r)}{\hbar^2}}(r-x)^{3/2} = 2y^{3/2} \tag{12.40}$$

et conduisant à l'identification des facteurs devant les exponentielles et donc au raccordement de la fonction d'onde WKB oscillante dans la région asymptotique et la fonction d'onde associée au potentiel linéaire au voisinage du point de rebroussement :

$$\frac{c}{(-2mV'(r))^{1/4}a^{-1/4}y^{1/4}} = \frac{\lambda}{\sqrt{\pi}y^{1/4}} . \tag{12.41}$$

Considérons le comportement asymptotique de la fonction de Airy pour $y \to +\infty$, c.-à.-d. de l'autre côté de la barrière de potentiel :

$$A_i(y) \overset{y\to+\infty}{\sim} \frac{1}{2\sqrt{\pi}y^{1/4}} e^{-\frac{2y^{3/2}}{3}} . \tag{12.42}$$

La fonction d'onde WKB donne de son côté, pour $x < r$, en ne gardant que l'exponentielle décroissante pour $x \to -\infty$:

$$u(x) = \frac{\tilde{c}}{\big(2m\big(V(x) - E\big)\big)^{1/4}} e^{\frac{\sqrt{2m}}{\hbar}\int_x^r (V(x')-E)^{1/2}dx'} \tag{12.43}$$

Lorsque $x \to r^-$, on peut remplacer $V(X) - E$ par $(x-r)V'(r)$ et on obtient :

$$u(x) \overset{x\to r^-}{\sim} \frac{c}{\big(2m(x-r)V'(r)\big)^{1/4}} e^{\frac{\sqrt{-2mV'(r)}}{\hbar}\int_x^r (r-x')^{1/2}dx'} \sim e^{-\frac{2}{3}\frac{\sqrt{-2mV'(r)}}{\hbar}(r-x)^{3/2}} \tag{12.44}$$

dont le comportement asymptotique coïncide avec celui de la fonction de Airy, voir l'éqn. (12.42), une fois effectué le changement de variable (12.34). Pour obtenir un accord exact nous imposons ici également l'égalité des facteurs devant les exponentielles, soit

$$\frac{\lambda}{2\sqrt{\pi}y^{1/4}} = \frac{\tilde{c}}{\left(-2mV'(r)\right)^{1/4}} \frac{1}{\left(-\frac{2mV'(r)}{\hbar^2}\right)^{-1/12} y^{1/4}}, \tag{12.45}$$

d'où nous déduisons la relation entre λ et $\tilde{c}$. Nous avons pu ainsi raccorder les solutions approchées fournies par l'approximation WKB des deux côtés de la barrière de potentiel.

En traversant une barrière de largeur finie, située entre x_a et x_b, la fonction d'onde WKB décroît d'un facteur (en négligeant la variation du préfacteur dans l'éqn. (12.43), qui est sous-dominante) :

$$\frac{u(x_b)}{u(x_a)} \sim e^{-\int_{x_a}^{x_b} dx \sqrt{\frac{2m}{\hbar^2}(V(x)-E)}}, \tag{12.46}$$

Dont le carré donne le coefficient de transmission de la barrière de potentiel :

$$\boxed{\mathcal{T} = e^{-2\int_{x_a}^{x_b} dx \sqrt{\frac{2m}{\hbar^2}(V(x)-E)}}} \tag{12.47}$$

Pour revenir au problème initial de l'oscillateur harmonique perturbé par un potentiel cubique, voir la figure 12.1 :

— les calculs perturbatifs effectués dans le cours n° 11 donnent les corrections aux niveaux d'énergie au voisinage du minimum local en $x = 0$, dans le développement perturbatif en puissances de ϵ

— On peut calculer la probabilité par unité de temps de passer de l'autre côté de la barrière par effet tunnel ; cette probabilité est proportionnelle au taux de transmission de la barrière. Dans le cas limite $E \to 0$ on trouve par exemple un taux de transmission :

$$e^{-\frac{2m\omega}{\hbar}\int_0^{x_c}\sqrt{x^2(1-x/x_c)}} = e^{-\frac{8m\omega x_0^2}{15\hbar}\frac{1}{\epsilon^2}}. \tag{12.48}$$

qui n'est pas développable en puissances de ϵ. Il s'agit donc d'un effet non-perturbatif.

À retenir

— signification de l'approximation semi-classique ;

— fonction d'onde WKB ;

— analyse de l'effet tunnel.

Bibliographie

[1] G. Wentzel, "Eine Verallgemeinerung der Quantenbedingungen für die Zwecke der Wellenmechanik," Z. Phys. **38** (1926) no.6, 518. doi :10.1007/BF01397171

[2] H. A. Kramers, "Wellenmechanik und halbzahlige Quantisierung," Z. Phys. **39** (1926) no.10, 828. doi :10.1007/BF01451751

[3] L. Brillouin, "La mécanique ondulatoire de Schrödinger ; une méthode générale de résolution par approximations successives," Compt. Rend. Hebd. Seances Acad. Sci. **183** (1926) no.1, 24.

Cours n°13

Moment cinétique I : rotations et spin

La notion de spin est apparue en mécanique quantique suite à différentes expériences menées au début du XXe siècle, comme les expériences de Stern–Gerlach décrites en détail dans les chapitres précédents ou l'analyse de l'effet Zeeman dit « anormal ».

Rappelons quelques faits élémentaires. En mécanique classique, une particule chargée en rotation porte un moment magnétique μ donné par :

$$\vec{\mu} = \frac{q}{2}\,\vec{r}\wedge\vec{v} = \frac{q}{2m}\vec{L}, \tag{13.1}$$

tel que cette particule, dans un champ magnétique uniforme, subit un couple $\vec{\Gamma} = \vec{\mu}\wedge\vec{B}$, responsable du mouvement de précession de Larmor, à la pulsation $\omega = \gamma|B|$ où $\gamma = q/2m$ est le rapport gyromagnétique de cette particule. Plongé dans un champ magnétique quelconque, un moment magnétique subit une force dérivant de l'énergie potentielle

$$U = -\vec{\mu}\cdot\vec{B}\,. \tag{13.2}$$

Dans l'expérience de Stern et Gerlach, les atomes d'argent étant électriquement neutres, leur mouvement ne va pas en lui-même leur faire subir une force dans la cavité contenant un champ magnétique inhomogène. L'expérience montre, comme nous l'avons déjà vu, que :

— les atomes d'argent sont déviés, suggérant l'existence d'un moment magnétique intrisèque, existant même pour un atome au repos ;

— tout se passe comme si seulement deux valeurs, opposées, de la projection du moment cinétique sur l'axe portant le champ magnétique étaient possibles.

Comme nous le savons aujourd'hui, ce moment magnétique est porté par le 47e et dernier électron de l'atome d'argent, le moment magnétique nucléaire étant trop faible en comparaison pour jouer un rôle significatif dans l'expérience.

Historiquement, l'expérience de Stern–Gerlach n'a pas conduit à la découverte du spin de l'électron, même si ce dernier en a donné retrospectivement l'explication. L'*effet Zeeman* [1], la séparation des raies spectrales d'un atome sous l'action d'un champ magnétique externe, a conduit Uhlenbeck et Goudsmit à proposer en 1925 l'hypothèse d'un moment cinétique

intrinsèque à l'électron [2], nommé *spin*, responsable de cet effet que nous décrirons dans le cours n° 17.

Nous avons montré dans le cours n° 3 comment le moment cinétique $\vec{L}$ – que nous appelerons dorénavant « orbital » pour le distinguer du spin $\vec{S}$ – générait les rotations dans l'espace. Le spin est aussi lié, quoique de manière plus subtile, aux rotations, ce que nous découvrirons apprès une analyse plus approfondie du groupe des rotations.

Le spin, comme la masse et la charge, est une caractéristique fondamentale des particules élémentaires ; les particules composites, par exemple les atomes, sont aussi caractérisés dans leur état fondamental par un spin obtenu par composition des spins des particules le composant, comme nous l'étudierons dans le cours n° 16.

13.1 Représentations du groupe des rotations

Appliquons la notion de représentation d'un groupe exposée dans le cours n° 3, au groupe des rotations à trois dimensions. Nous avons appris dans le cours n° 4 que le groupe des rotations est un *groupe de Lie*, groupe dépendant continuement de trois paramètres (θ, ϕ, ψ) indiquant pour les deux premiers la direction du vecteur $\vec{n}$ orthogonal au plan de rotation et pour le dernier l'angle de rotation dans le plan. Le groupe proprement dit est défini par les règles de composition de rotations successives dans l'espace,

$$g(\theta, \phi, \psi) = g(\theta_1, \phi_1, \psi_1) \cdot g(\theta_2, \phi_2, \psi_2)\,. \tag{13.3}$$

Ces règles de composition sont induites des règles de transformations des vecteurs, c.-à.-d. de la *représentation vectorielle* du groupe des rotations, qui est une représentation réelle de dimensions 3. En effet un vecteur de l'espace se transforme par des matrices 3×3 réelles :

$$\vec{v} \;\mapsto\; R[g(\theta, \phi, \psi)]\,\vec{v}\,, \quad R \in SO(3) \subset GL(3, \mathbb{R})\,, \tag{13.4}$$

telles que

$$R[g(\theta_1, \phi_1, \psi_1) \cdot g(\theta_2, \phi_2, \psi_2)] = R[g(\theta_1, \phi_1, \psi_1)]R[g(\theta_2, \phi_2, \psi_2)]\,. \tag{13.5}$$

Il est évident que cette représentation est irreductible, puisqu'il n'existe aucun sous-espace vectoriel de $\mathbb{R}^3$ qui soit stable sous toutes les rotations de l'espace.

Les mêmes règles (13.3) définissant le groupe peuvent être réalisées dans des représentations irreductibles et unitaires de dimension différente de trois.

La représentation la plus simple du groupe des rotations, comme pour tous les groupes, est la *représentation triviale* de dimension réelle un. Dans le cas présent il s'agit de quantités invariantes par toutes les rotations de l'espace, soit de quantités scalaires comme le produit scalaire $\vec{v} \cdot \vec{w}$ entre deux vecteurs quelconques de l'espace.

Nous chercherons dans la suite de ce cours à déterminer toutes les autres représentations irreductibles et unitaires inéquivalentes.

13.1.1 Exemple : représentation irreductible de dimension 5

À l'aide de deux vecteurs de composantes v^i et w^j il est possible de construire (par produit tensoriel), un objet à deux indices qui se transforme par rotation de la manière suivante :

$$v^i w^j \;\mapsto\; \sum_{k,l} R(\theta,\phi,\psi)^i{}_k \, R(\theta,\phi,\psi)^j{}_l \, v^k w^l \,. \tag{13.6}$$

Les paires d'indices (i,j) d'une part et (k,l) d'autre part pouvant prendre chacune 9 valeurs, on peut représenter les objets $R^i{}_k \, R^j{}_l$ comme des matrices réelles 9×9 et on a ainsi une représentation réelle de dimension 9. Nous pouvons en effet représenter (13.6) sous la forme :

$$\begin{pmatrix} u^1 = v^1 w^1 \\ u^2 = v^1 w^2 \\ u^3 = v^1 w^3 \\ u^4 = v^2 w^1 \\ u^5 = v^2 w^2 \\ u^6 = v^2 w^3 \\ u^7 = v^3 w^1 \\ u^8 = v^3 w^2 \\ u^9 = v^3 w^3 \end{pmatrix} \;\mapsto\; \begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} u^1 = v^1 w^1 \\ u^2 = v^1 w^2 \\ u^3 = v^1 w^3 \\ u^4 = v^2 w^1 \\ u^5 = v^2 w^2 \\ u^6 = v^2 w^3 \\ u^7 = v^3 w^1 \\ u^8 = v^3 w^2 \\ u^9 = v^3 w^3 \end{pmatrix} \tag{13.7}$$

Cependant cette représentation n'est pas irreductible car on réalise immédiatement qu'on peut obtenir une combinaison des composantes $v^i w^j$ stable par l'action des rotations : le produit scalaire

$$\vec{v}\cdot\vec{w} = \sum_i v^i w_i = \mathrm{Tr}\left[(vw)^i{}_j\right] \,. \tag{13.8}$$

Les 8 combinaisons linéaires indépendantes restantes ne forment pas non plus une représentation irreductible ; pour le voir on peut remarquer que le produit vectoriel de deux vecteurs se transforme par rapport aux rotations comme un vecteur [1] :

$$(\vec{v}\wedge\vec{w})^i = \sum_{kl} \epsilon^{ikl} v^k w^l \;\mapsto R^i{}_j \, (\vec{v}\wedge\vec{w})^j \,. \tag{13.9}$$

Comme nous l'avons vu dans le cours n° 3, éqn. (3.26), on peut associer le produit vectoriel à une matrice 3×3 antisymétrique, qui est la partie antisymétrique de la matrice $v^i w^j$:

$$A^{ij} = v^{[i} w^{j]} \overset{\text{déf.}}{=} \frac{1}{2}\left(v^i w^j - v^j w^i\right) = \frac{1}{2}\begin{pmatrix} 0 & v^1 w^2 - v^2 w^1 & v^1 w^3 - v^3 w^1 \\ -v^1 w^2 + v^2 w^1 & 0 & v^2 w^3 - v^3 w^2 \\ v^3 w^1 - v^1 w^3 & v^3 w^2 - v^2 w^3 & 0 \end{pmatrix} \tag{13.10}$$

Il reste cinq composantes indépendantes, formant une matrice symétrique de trace nulle :

$$S^{ij} \overset{\text{déf.}}{=} \frac{1}{2}\left(v^i w^j + v^j w^i\right) - \frac{1}{3}\mathrm{Tr}\left(v^k w^l\right)\delta^{ij} \,. \tag{13.11}$$

1. Cependant, par rapport à l'inversion le produit vectoriel de deux vecteurs change de signe contrairement à un vecteur, d'où le nom de *pseudo-vecteur*. Cette distinction est pertinente si on considère le groupe $O(3)$ qui contient l'inversion, mais pas le groupe $SO(3)$; voir la section 4.3 du cours n° 4.

se transformant dans une représentation irreductible réelle de dimension $9-1-3=5$. La décomposition en représentations irreductibles que nous venons de détailler :

$$v^i w^j = A^{ij} + S^{ij} + \frac{1}{3}\mathrm{Tr}\left(v^k w^l\right)\delta^{ij} \tag{13.12}$$

se résume sous la forme symbolique :

$$\mathbf{3}\otimes\mathbf{3} = \mathbf{1}\oplus\mathbf{3}\oplus\mathbf{5}\,, \tag{13.13}$$

où chaque nombre correspond à la dimension d'une représentation irreductible et unitaire. Ce type de décomposition sera développé de manière systématique à partir du chapitre n° 13.

Nous avons en particulier construit une représentation irreductible de dimension 5, agissant sur les tenseurs symétriques de trace nulle. On peut explicitement déterminer les matrices 5×5 correspondantes à partir des matrices 3×3 de rotation $R(\theta,\phi,\psi)$ et de la loi de transformation (13.6).

13.1.2 Représentations de l'algèbre de Lie

La notion de sous-groupe à un paramètre, voir l'éqn. (3.65) du cours 3, permet d'associer à un plan de rotation donné, spécifié par les angles (θ,ϕ) ou de manière équivalente par l'axe de rotation $\vec{n}$, un générateur infinitésimal des rotations dans ce plan que nous noterons $T_{\vec{n}}$:

$$\frac{\mathrm{d}}{\mathrm{d}\psi}g(\vec{n},\psi)\Big|_{\psi=0} = T_{\vec{n}}\,. \tag{13.14}$$

Ce concept se transpose à toute représentation linéaire du groupe des rotations. Nous avons vu précédemment que, pour la représentation de dimension 3,

$$\frac{\mathrm{d}}{\mathrm{d}\theta}R(\vec{n},\psi)\Big|_{\psi=0} = \vec{n}\cdot\vec{j}, \tag{13.15}$$

où le triplet de matrices (j_x,j_y,j_z), qui sont anti-hermitiennes et de dimension 3×3, est donné par l'équation (4.7). Nous avons constaté que ces matrices satisfont aux relations de commutation

$$[\,j_a,\,j_b\,] = \sum_c \epsilon_{ab}{}^c\, j_c \tag{13.16}$$

qui expriment pour des rotations proches de l'identité, les règles de composition de rotations successives d'axes différents appliquées à des vecteurs.

Le même concept peut être appliqué à une représentation D^r de dimension r quelconque, c.-à.-d. de matrices $r\times r$ vérifiant

$$D^r\big(g(\theta_1,\phi_1,\psi_1)\cdot g(\theta_2,\phi_2,\psi_2)\big) = D^r\big(g(\theta_1,\phi_1,\psi_1)\big)\,D^r\big(g(\theta_2,\phi_2,\psi_2)\big)\,. \tag{13.17}$$

Nous définissions alors les générateurs

$$\frac{\mathrm{d}}{\mathrm{d}\psi}D^r\big(g(\vec{n},\psi)\big)\Big|_{\psi=0} = t^r_{\vec{n}}\,, \tag{13.18}$$

où les $t^r_{\vec{n}}$ sont des matrices $r \times r$, également anti-hermitiennes. Prenons une base (t^r_x, t^r_y, t^r_z) de cet espace de matrices, où les trois matrices associées respectivement aux rotations d'axes (x, y, z).

Les matrices $D^r(g)$ formant une représentation de $SO(3)$, les matrices (t^r_x, t^r_y, t^r_z) vérifient des relations de commutation analogues aux relations (13.19) :

$$[\, t^r_a,\, t^r_b \,] = \sum_c \epsilon_{ab}{}^c\, t^r_c\,, \tag{13.19}$$

car ces relations de commutation comportent l'information permettant de reconstruire le produit dans le groupe. De cette manière nous avons fait correspondre à un représentation du groupe des rotations $SO(3)$ une représentation de son *algèbre de Lie*, notée $\mathfrak{so}_3$.

La notion de représentations reductibles et irreductibles des groupes permet de définir une notion équivalente pour les algèbres de Lie :

Définition 1 (représentation reductible de $\mathfrak{so}_3$) *Une représentation de dimension r de $\mathfrak{so}_3$ est reductible s'il existe un sous-espace invariant par l'action des matrices $t^r_{\vec{n}}$ définies par l'éqn. (13.18).*

On pourrait imaginer que, par exponentiation des matrices $t^r_{\vec{n}}$, une représentation irreductible de $\mathfrak{so}_3$ donne une représentation irreductible du groupe des rotations $SO(3)$. Il existe cependant une subtilité, d'une portée considérable, que nous allons rencontrer à la prochaine section.

13.2 Groupe $SU(2)$ et spineurs

À partir de la notion de représentation d'une algèbre de Lie nous introduirons dans cette section des objets fondamentaux de la mécanique quantique, les spineurs.

13.2.1 Représentation de dimension 2 de $\mathfrak{so}_3$

Cherchons explicitement la représentation non-triviale de plus petite dimension de l'algèbre de Lie $\mathfrak{so}_3$ du groupe des rotations. Introduisons les *matrices de Pauli* (qui forment une base de l'espace vectoriel des matrices 2×2 hermitiennes de trace nulle) :

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{13.20a}$$

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{13.20b}$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{13.20c}$$

Nous remarquons ces matrices satisfont :

$$\left[-\frac{i\sigma_1}{2}, -\frac{i\sigma_2}{2}\right] = -\frac{1}{4}\left(\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix} - \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\right) = -\frac{i\sigma_3}{2} \tag{13.21a}$$

$$\left[-\frac{i\sigma_2}{2}, -\frac{i\sigma_3}{2}\right] = -\frac{1}{4}\left(\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} - \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}\right) = -\frac{i\sigma_1}{2} \tag{13.21b}$$

$$\left[-\frac{i\sigma_3}{2}, -\frac{i\sigma_1}{2}\right] = -\frac{1}{4}\left(\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} - \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right) = -\frac{i\sigma_2}{2} \tag{13.21c}$$

Nous avons ainsi obtenu une représentation complexe de dimension égale à 2 de l'algèbre (13.19) en identifiant

$$\boxed{t_x = -\frac{i\sigma_1}{2}\ , \quad t_y = -\frac{i\sigma_2}{2}\ , \quad t_z = -\frac{i\sigma_3}{2}} \tag{13.22}$$

que l'on regroupera dans un vecteur $\vec{t}$ de composantes (t_x, t_y, t_z), qui sont toutes les trois des matrices anti-hermitiennes. On considère également le vecteur $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ composé de matrices hermitiennes.

13.2.2 Groupe SU(2)

Examinons les propriétés des matrices complexes 2×2 obtenues par exponentiation des générateurs (13.22),

$$M(\vec{n}, \psi) \overset{\text{déf.}}{=} e^{\psi \vec{t}\cdot\vec{n}} = e^{-\frac{i}{2}\psi\vec{\sigma}\cdot\vec{n}}\,. \tag{13.23}$$

Il est utile de remarquer que les matrices de Pauli (13.20) satisfont

$$\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{jk}\mathbb{I}\,, \tag{13.24}$$

d'où on déduit que

$$\left(\sum_k \sigma_k v_k\right)^2 = \sum_{k,l} v_k v_l \sigma_k\sigma_l = \frac{1}{2}\sum_{k,l} v_k v_l\big(\sigma_k\sigma_l + \sigma_k\sigma_k\big) = \sum_k (v_k)^2 \tag{13.25}$$

Cela permet d'obtenir, pour tout $h \in \mathbb{Z}_{\geqslant 0}$:

$$\left(\sum_k \sigma_k v_k\right)^{2h} = \left(\sum_k v_k^2\right)^h \mathbb{I} \tag{13.26a}$$

$$\left(\sum_k \sigma_k v_k\right)^{2h+1} = \left(\sum_k v_k^2\right)^h \left(\sum_k \sigma_k v_k\right) \tag{13.26b}$$

Et on peut exprimer les matrices M comme :

$$\begin{aligned} M(\vec{n},\psi) &= \sum_{l=0}^{\infty} \frac{(-i\psi/2)^l}{l!}(\sigma_1 n_x + \sigma_2 n_y + \sigma_3 n_3)^l \\ &= \sum_{h=0}^{\infty} \frac{(-1)^h}{(2h)!}\left(\psi^2 \frac{n_x^2+n_y^2+n_z^2}{4}\right)^{2h} \\ &\quad - \sum_{h=0}^{\infty} \frac{(-1)^h}{(2h+1)!}\left(\psi^2 \frac{n_x^2+n_y^2+n_z^2}{4}\right)^{2h} \frac{i\psi}{2}(\sigma_1 n_x + \sigma_2 n_y + \sigma_3 n_3) \end{aligned} \quad (13.27)$$

Étant donné que, par hypothèse, $\vec{n}^2 = 1$, nous trouvons

$$\boxed{M(\vec{n},\psi) = \mathbb{I}\cos\frac{\psi}{2} - i\vec{n}\cdot\vec{\sigma}\,\sin\frac{\psi}{2}} \quad (13.28)$$

Ces matrices complexes sont unitaires. Nous avons en effet, les matrices de Pauli étant hermitiennes :

$$\begin{aligned} M(\vec{n},\psi)^{-1} = M(\vec{n},-\psi) &= \mathbb{I}\cos\frac{\psi}{2} + i\vec{n}\cdot\vec{\sigma}\,\sin\frac{\psi}{2} \\ &= \left(\mathbb{I}\cos\frac{\psi}{2} - i\vec{n}\cdot\vec{\sigma}\sin\frac{\psi}{2}\right)^{\dagger} = M(\vec{n},\psi)^{\dagger}\,. \end{aligned} \quad (13.29)$$

Pour montrer la propriété suivante, explicitons la matrice M. Nous avons

$$M(\vec{n},\psi) = \mathbb{I}\cos\frac{\psi}{2} - i(\sigma_1 n_x + \sigma_2 n_y + \sigma_3 n_3)\sin\frac{\psi}{2} = \begin{pmatrix} \cos\frac{\psi}{2} - in_3\sin\frac{\psi}{2} & -i(n_x - in_y)\sin\frac{\psi}{2} \\ -i(n_x + in_y)\sin\frac{\psi}{2} & \cos\frac{\psi}{2} + in_3\sin\frac{\psi}{2} \end{pmatrix} \quad (13.30)$$

d'où

$$\det M(\vec{n},\psi) = \cos^2\frac{\psi}{2} + n_3^2\sin^2\frac{\psi}{2} + (n_x^2+n_y^2)\sin^2\frac{\psi}{2} = 1\,. \quad (13.31)$$

Les matrices $M(\vec{n},\psi)$ forment donc le groupe matriciel des matrices 2×2 unitaires de déterminant unité qui est naturellement un sous-groupe de $GL(2,\mathbb{C})$.

Les matrices (13.28) définissent-elles une représentation du groupe des rotations de dimension 2 ? On remarque la propriété :

$$M(\vec{n},\psi+2\pi) = \mathbb{I}\cos\left(\frac{\psi}{2}+\pi\right) + 2\vec{n}\cdot\vec{t}\,\sin\left(\frac{\psi}{2}+\pi\right) = -M(\vec{n},\psi)\,, \quad (13.32)$$

Si on veut interpréter les matrices M comme les matrices d'une représentation du groupe des rotations $SO(3)$, on peut constater que :

$$M\Big[\underbrace{g(\vec{n},\pi)\cdot g(\vec{n},\pi)}_{g(n,2\pi)=e}\Big] = -\underbrace{M\left[g(\vec{n},\pi)\right]}_{-i\vec{n}\cdot\sigma}\underbrace{M\left[g(\vec{n},\pi)\right]}_{-i\vec{n}\cdot\sigma} \quad (13.33)$$

Le signe apparaissant dans le membre de droite de (13.33) signifie que nous n'avons pas une représentation usuelle du groupe des rotations $SO(3)$, mais une *représentation projective*, voir la section 3.4 du cours n° 3, c.-à.-d. une représentation à une phase près. Nous avions vu dans ce cours que les représentations projectives sont autorisées en mécanique quantique, au contraire de la mécanique classique.

L'ensemble des matrices (13.28), qui correspond donc pas à une représentation du groupe $SO(3)$ proprement dite, définit un groupe appelé $SU(2)$, associé aux matrices unitaires de déterminant unité. Son algèbre de Lie, qui est isomorphe à celle de $SO(3)$, est notée $\mathfrak{su}_2$.

13.2.3 Spineurs

Les matrices $M(\vec{n}, \psi)$ agissent sur un espace vectoriel complexe à deux dimensions. Les éléments de cet espace sont appelés des *spineurs* :

$$\chi = \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} , \tag{13.34}$$

et ils se transforment par rotation selon

$$\chi \;\mapsto\; M(\vec{n}, \psi)\, \chi \,, \tag{13.35}$$

où les matrices 2×2 complexes de transformation sont données par l'éqn. (13.28).

La propriété étonnante de ces objets est leur comportement sous une rotation de l'espace de 2π. Nous avons :

$$\chi \mapsto M(\vec{n}, 2\pi)\chi = -\chi \tag{13.36}$$

Ils leur faut donc effectuer « deux tours sur eux-mêmes » pour revenir dans leur état initial ! Les particules de spin 1/2, comme les électrons, se transforment sous les rotations comme des spineurs et possèdent donc cette propriété qui n'a pas d'analogue classique.

Les particules se transformant comme des spineurs sont décrites par un couple de fonctions d'ondes, une pour chaque composante de (13.34) :

$$\chi(\vec{x}) = \begin{pmatrix} \chi_1(\vec{x}) \\ \chi_2(\vec{x}) \end{pmatrix} , \tag{13.37}$$

qui se transforme sous une rotation comme :

$$\begin{pmatrix} \chi_1(\vec{x}) \\ \chi_2(\vec{x}) \end{pmatrix} \xmapsto{g(\vec{n},\psi)} M(\vec{n}, \psi) \begin{pmatrix} \chi_1\big(R^{-1}(\vec{n}, \psi)\vec{x}\big) \\ \chi_2\big(R^{-1}(\vec{n}, \psi)\vec{x}\big) \end{pmatrix} . \tag{13.38}$$

Ce type de transformation, combinant l'action sur le spineur et sur l'argument des fonctions d'onde χ_1 et χ_2 – c.-à.-d. l'action du moment cinétique de spin et du moment cinétique orbital – sera étudiée en détail dans le cours n° 16.

L'espace des états associé à une particule de spin 1/2 possède ainsi la structure d'un produit tensoriel :

$$\mathcal{H} \cong \mathcal{L}^2(\mathbb{R}^3) \otimes \mathbb{C}^2 \,. \tag{13.39}$$

En physique l'usage est d'utiliser, plutot que les générateurs anti-hermitiens (13.22), un triplet d'*opérateurs de spin* hermitiens,

$$\boxed{\mathcal{S}_a \overset{\text{déf.}}{=} \hbar\frac{\sigma_a}{2}\,,} \tag{13.40}$$

satisfaisant aux relations de commutation

$$\boxed{[\,\mathcal{S}_a,\,\mathcal{S}_b\,] = i\hbar\sum_c \epsilon^{abc}\mathcal{S}_c\,.} \tag{13.41}$$

En raison du facteur $\hbar$ ces opérateurs ont la dimension physique d'un moment cinétique et sont appelés pour cette raison *moment cinétique de spin*. Cependant, de notre point de vue, il est préférable de voir le spin comme une propriété géométrique de transformation sous les rotations plutôt que comme un « moment cinétique intrinsèque ».

Moment magnétique de spin

Du point de vue expérimental, le spin 1/2 de l'électron se manifeste par l'existence d'un moment magnétique intrinsèque donné par l'opérateur

$$\boxed{\vec{\mu} = -g_s\mu_B\,\frac{1}{\hbar}\,\vec{\mathcal{S}}} \tag{13.42}$$

où la constante μ_B, appelée *magneton de Bohr*,[2] donne la valeur d'un quanta de moment magnétique de l'électron (q_e et m_e étant respectivement la charge et la masse de l'électron) :

$$\mu_B = \frac{|q_e|\hbar}{2m_e}\,, \quad \mu_B \simeq 0,9\times 10^{-23}\mathrm{J}\cdot\mathrm{T}^{-1} \tag{13.43}$$

et où la constante sans dimension g_s est appelée *facteur de Landé*.

La description relativiste de l'électron, via l'équation de Dirac, indique que le couplage d'une particule de spin 1/2 au champ électromagnétique induit l'existence d'un moment magnétique intrinsèque, avec un facteur de Landé $g_s = 2$. Cette valeur est faiblement corrigée lorsque la quantification du champ électromagnétique est prise en compte :

$$g_s \simeq 2,002319 \tag{13.44}$$

La quantité $(g_s - 2)/2$ mesurant cette déviation, appelée *moment magnétique anormal* peut se calculer avec une très grande précision en théorie quantique des champs et correspond parfaitement à l'expérience, fournisssant un test très sévere de l'électrodynamique quantique.

2. Certains auteurs définissent un magnéton de Bohr négatif pour l'électron, auquel cas il n'y a pas de signe négatif dans l'éqn. (13.42).

Notation

En mécanique quantique tout triplet d'opérateurs hermitiens satisfaisant aux relations de commutation (13.41) est appelé moment cinétique. Nous utiliserons suivant l'usage courant les notation et terminologie suivantes :

— le moment cinétique de spin est associé au triplet d'opérateurs $\mathcal{S}_x$, $\mathcal{S}_y$ et $\mathcal{S}_z$ correspondant à l'action géométrique des rotations, sur les spineurs pour les particules de spin 1/2, ou plus généralement sur les vecteurs à $2s+1$ composantes de l'espace vectoriel associé à une représentation de spin s ;

— le moment cinétique orbital est associé au triplet d'opérateurs $\mathcal{L}_x$, $\mathcal{L}_y$ et $\mathcal{L}_z$. Comme nous l'avons vu dans le cours n° 4, ce sont des générateurs de l'action des rotations sur l'argument des fonctions d'ondes, indépendamment du spin, c.-à.-d.

$$\psi(R^{-1}\vec{x}) = e^{-\frac{i\alpha}{\hbar}\,\vec{n}\cdot\vec{\mathcal{L}}}\,\psi(\vec{x})\,. \tag{13.45}$$

— un moment cinétique général associé à tout triplet d'opérateurs $\mathcal{J}_x$, $\mathcal{J}_y$ et $\mathcal{J}_z$ hermitiens satisfaisant (13.41) ; cela inclut en particulier la somme des moments cinétique de spin et orbital, qui sera étudiée dans le cours n° 16.

Du point de vue de la théorie des groupes, étant donné que chacun de ces triplets d'opérateurs satisfait aux relations de commutation :

$$[\,\mathcal{J}_a,\,\mathcal{J}_b\,] = i\hbar\sum_c \epsilon^{abc}\mathcal{J}_c\,, \tag{13.46}$$

nous avons dans tous les cas une représentation de l'algèbre de Lie $\mathfrak{su}(2)$.

Groupe de recouvrement $\star$

Le groupe $SU(2)$ est le groupe des matrices unitaires 2×2 de déterminant unité. Un paramétrage possible de telles matrices est

$$M = \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix} \tag{13.47a}$$

$$|\alpha|^2 + |\beta|^2 = 1\,. \tag{13.47b}$$

Si on pose

$$\alpha = w + ix\,,\quad \beta = y + iz\,,\quad w,x,y,z\in\mathbb{R} \tag{13.48}$$

La condition (13.47b) sur le déterminant donne l'équation d'une hypersphère de rayon unité dans $\mathbb{R}^4$:

$$w^2 + x^2 + y^2 + z^2 = 1\,. \tag{13.49}$$

Cette hyper-sphère est la variété de groupe correspondant à $SU(2)$. Si on compare les parametrages (13.31) et (13.47a) des éléments du groupe, on se rend compte qu'une rotation de 2π autour de l'axe $\vec{n}$, c.-à.-d. $\psi\mapsto\psi+2\pi$, donne

$$M(\vec{n},\psi+2\pi) = -M(\vec{n},\psi+2\pi)\,, \tag{13.50}$$

et revient donc à l'inversion ou parité dans $\mathbb{R}^4$:

$$P_4 : (w, x, y, z) \mapsto (-w, -x, -y, -z) , \tag{13.51}$$

Dans le groupe $SO(3)$, une rotation de 2π autour d'un axe quelconque envoie un élément du groupe sur lui-même. La relation entre les deux groupes de Lie est donnée par le quotient

$$SO(3) \cong \frac{SU(2)}{P_4} , \tag{13.52}$$

où P_4, du point de vue géométrique, identifie les points antipodaux sur l'hypersphère. Le groupe $SU(2)$ est appelé *groupe de recouvrement* de $SO(3)$.

Rappelons maintenant qu'en mécanique quantique, on admet pour un groupe de symétrie non seulement des représentations « ordinaires » mais également des représentations projectives, c.-à.-d. à une phase près :

$$D(g)D(g') = \epsilon(g, g')D(g \cdot g') , \quad |\epsilon(g, g')| = 1 . \tag{13.53}$$

De manière générale les représentations projectives d'un groupe G sont associées à des représentations « ordinaires » d'un groupe plus grand au sens de l'inclusion, son groupe de recouvrement.

Dans le cas du groupe $SO(3)$, le groupe de recouvrement est le groupe $SU(2)$ nous montrerons dans la section suivante que :

— les représentations de $SU(2)$ de spin entier « descendent » vers des représentations ordinaires de $SO(3)$

— les représentations de $SU(2)$ de spin demi-entier « descendent » vers des représentations projectives de $SO(3)$, telles que

$$D[g(\vec{n}, \pi) \cdot g(\vec{n}, \pi)] = -D[g(\vec{n}, 2\pi)] . \tag{13.54}$$

L'existence de particules de spin demi-entier comme les électrons indique que le « véritable » groupe de symétrie associé aux rotations en mécanique quantique est $SU(2)$ et non pas $SO(3)$.

13.3 Représentations unitaires irreductibles de $SU(2)$

Le groupe des rotations $SO(3)$ possède une représentation fondamentale, la représentation vectorielle de dimension réelle 3, à partir de laquelle les propriétés du groupes ont été déduites. Cette représentation vectorielle est également une représentation du groupe de recouvrement $SU(2)$, qui pour sa part a pour représentation fondamentale la représentation spinorielle, de dimension complexe 2.

L'objectif de cette section est de déterminer les représentations unitaires et irreductibles de $SU(2)$, qui interviendront dans tous les problèmes de mécanique quantique où jouent un

rôle le spin – par exemple en présence d'un champ magnétique – ou bien la symétrie de rotation – par exemple en physique atomique.

Nous utiliserons le langage de la mécanique quantique pour formuler le problème, même si les résultats que nous obtiendrons sont des résultats purement mathématiques existant indépendemment de la physique.

D'après le théorème de Peter–Weyl (voir théorème 3, cours n° 3) le groupe $\mathsf{SU}(2)$ étant compact (ayant la topologie d'une hypersphère), toute représentation unitaire se décompose en somme directe de représentations irreductibles unitaires et finies. Nous considérerons donc uniquement des espaces hermitiens de dimension finie dans ce qui suit.

13.3.1 Opérateur de Casimir

Soit $\mathcal{H}$ un espace hermitien de dimension $\mathfrak{d}$ associé à un espace des états en mécanique quantique sur lequel agissent trois opérateurs hermitiens $\mathcal{J}_x$, $\mathcal{J}_y$ et $\mathcal{J}_z$. Ces opérateurs correspondent à des opérateurs de moment cinétique s'ils satisfont aux relations de commuation

$$\boxed{[\mathcal{J}_a, \mathcal{J}_b] = i\hbar \sum_c \epsilon^{abc} \mathcal{J}_c} \tag{13.55}$$

Mathématiquement, nous dirions que l'action de ces opérateurs sur $\mathcal{H}$ définit une représentation de dimension $\mathfrak{d}$ de l'algèbre de Lie $\mathfrak{su}_2$.

Dans la « convention des physiciens », une transformation infinitésimale du groupe $\mathsf{SU}(2)$ correspondra dans la représentation considérée à un opérateur

$$\mathcal{R}(\vec{n}, \epsilon) = \mathfrak{I} - \frac{i\epsilon}{\hbar} \vec{n} \cdot \vec{\mathcal{J}}, \tag{13.56}$$

et, si la représentation est unitaire, les opérateurs $\mathcal{J}^a$ sont hermitiens. En effet

$$\mathcal{R}(\vec{n}, \epsilon)^\dagger \mathcal{R}(\vec{n}, \epsilon) = \mathfrak{I} - \frac{i\epsilon}{\hbar} \vec{n} \cdot \left(\vec{\mathcal{J}} - \vec{\mathcal{J}}^\dagger \right) + \mathcal{O}(\epsilon^2) = \mathcal{O}(\epsilon^2) \implies \vec{\mathcal{J}} = \vec{\mathcal{J}}^\dagger \tag{13.57}$$

Si cette représentation de dimension $\mathfrak{d}$ est une *représentation irreductible*, le lemme de Schur (voir lemme 1 du cours n° 3) nous apprend qu'un opérateur qui commute avec $\mathcal{J}_x$, $\mathcal{J}_y$ et $\mathcal{J}_z$ est proportionnel à l'identité. Nous cherchons donc à trouver un tel opérateur afin de caractériser les représentations irreductibles, il est appelé *opérateur de Casimir*.

Un tel opérateur pour l'algèbre $\mathfrak{su}_2$ (plus précisément son opérateur de Casimir quadratique) est donné par :

$$\boxed{\mathcal{J}^2 \overset{\text{déf.}}{=} \mathcal{J}_x^2 + \mathcal{J}_y^2 + \mathcal{J}_z^2} \tag{13.58}$$

En effet, calculons tout d'abord

$$\left[\mathcal{J}_a, \mathcal{J}_b^2 \right] = [\mathcal{J}_a, \mathcal{J}_b] \mathcal{J}_b + \mathcal{J}_b [\mathcal{J}_a, \mathcal{J}_b] = i\hbar \sum_c \epsilon^{abc} (\mathcal{J}_c \mathcal{J}_b + \mathcal{J}_b \mathcal{J}_c) \tag{13.59}$$

Nous avons donc

$$\left[\mathcal{J}_a, \mathcal{J}^2 \right] = i\hbar \sum_{b,c} \epsilon^{abc} (\mathcal{J}_c \mathcal{J}_b + \mathcal{J}_b \mathcal{J}_c) = 0, \tag{13.60}$$

où nous avons utilisé le fait que le terme entre parenthèses, symétrique par échange de b et c, est contracté avec le symbole de Levi–Civita, antisymétrique par échange de b et c (analogue de la multiplication d'une matrice symétrique par une matrice antisymétrique).

Supposons que $\mathcal{H}$ corresponde à une représentation irreductible de $\mathfrak{su}_2$. L'opérateur $\mathcal{J}^2$ est proportionnel à l'identité d'après le lemme de Schur. Les opérateurs $\mathcal{J}_a$, étant hermitiens,

$$\forall\,|\alpha\rangle \in \mathcal{H}\ , \quad \imath(|\alpha\rangle\,, \mathcal{J}_a^2\,|\alpha\rangle) = \imath(\mathcal{J}_a\,|\alpha\rangle\,, \mathcal{J}_a\,|\alpha\rangle) = \|\,\mathcal{J}_a\,|\alpha\rangle\,\|^2 \geqslant 0\,. \tag{13.61}$$

Les opérateurs $\mathcal{J}_a^\dagger \mathcal{J}_a = \mathcal{J}_a^2$, et en en conséquence $\mathcal{J}^2$, sont donc des opérateurs *positifs*. Nous en déduisons que la constante de proportionnalité est positive et nous poserons, en toute généralité :

$$\mathcal{J}^2 = \hbar^2 j(j+1)\,\mathfrak{I}\ , \quad j \in \mathbb{R}^+\,, \tag{13.62}$$

où le nombre j est sans dimension.

13.3.2 Construction des représentations

La construction des représentations irreductibles de $\mathfrak{su}_2$ ressemble fortement à la construction algébrique du spectre de l'oscillateur harmonique. Premièrerement, pour caractériser complètement les vecteurs de la représentations, nous pouvons considérer un ensemble maximal d'observables commutantes.

Sachant que les trois opérateurs hermitiens $\mathcal{J}_a$ ne commutent pas entre eux mais commutent tous avec $\mathcal{J}^2$, nous pouvons considérer une base orthonormée commune de vecteurs propres de $\mathcal{J}^2$ et de l'un des trois autres opérateurs, par convention $\mathcal{J}_z$:

$$\boxed{\begin{aligned} \mathcal{J}^2\,|j,m\rangle &= \hbar^2 j(j+1)\,|j,m\rangle && j \in \mathbb{R}^+\,, \qquad (13.63a)\\ \mathcal{J}_z\,|j,m\rangle &= \hbar m\,|j,m\rangle && m \in \mathbb{R}\,. \qquad (13.63b) \end{aligned}}$$

avec

$$\langle j, m|j', m'\rangle = \delta_{j,j'}\delta_{m,m'}\,. \tag{13.64}$$

Notons que les formules mathématiques correspondantes sont identiques, les facteur de $\hbar$ en moins.

Pour contraindre les valeurs possibles de j et m et donc des valeurs propres nous introduisons les opérateurs d'échelle :

$$\boxed{\mathcal{J}_\pm = \mathcal{J}_x \pm i\mathcal{J}_y\,,} \tag{13.65}$$

qui sont conjugués hermitiens l'un de l'autre. La propriété essentielle de ces deux opérateurs, outre qu'ils commutent évidemment avec $\mathcal{J}^2$ et la suivante :

$$[\,\mathcal{J}_z,\,\mathcal{J}_\pm\,] = [\,\mathcal{J}_z,\,\mathcal{J}_x\,] \pm i\,[\,\mathcal{J}_z,\,\mathcal{J}_y\,] = i\hbar\big(\mathcal{J}_y \mp i\mathcal{J}_x\big) \implies \boxed{[\,\mathcal{J}_z,\,\mathcal{J}_\pm\,] = \pm\hbar\,\mathcal{J}_\pm} \tag{13.66}$$

Nous avons également

$$\mathcal{J}_\pm\mathcal{J}_\mp = \big(\mathcal{J}_x \pm i\mathcal{J}_y\big)\big(\mathcal{J}_x \mp i\mathcal{J}_y\big) = \mathcal{J}_x^2 + \mathcal{J}_y^2 \mp i\,[\,\mathcal{J}_x,\,\mathcal{J}_y\,] = \mathcal{J}_x^2 + \mathcal{J}_y^2 \pm \hbar\mathcal{J}_z = \mathcal{J}^2 - \mathcal{J}_z^2 \pm \hbar J_z\,. \tag{13.67}$$

Considérons un certain vecteur $|j,m\rangle$, vecteur propre commun de $\mathcal{J}^2$ et $\mathcal{J}_z$. Les vecteurs $\mathcal{J}_\pm|j,m\rangle$ satisfont donc :

$$\mathcal{J}^2\big(\mathcal{J}_\pm|j,m\rangle\big) = \hbar^2 j(j+1)\mathcal{J}_\pm|j,m\rangle \tag{13.68a}$$
$$\mathcal{J}_z\big(\mathcal{J}_\pm|j,m\rangle\big) = \hbar(m\pm 1)\mathcal{J}_\pm|j,m\rangle \tag{13.68b}$$

La relation (13.68b) indique que les vecteurs $\mathcal{J}_\pm|j,m\rangle$, s'ils sont non nuls, sont proportionnels à des vecteurs propres $|j,m\pm 1\rangle$:

$$\mathcal{J}_\pm|j,m\rangle = \lambda_{j,m,\pm}|j,m\pm 1\rangle \ , \quad \lambda_{j,m,\pm}\in\mathbb{C}^\star\,. \tag{13.69}$$

La valeur absolue des coefficients $\lambda_{j,m,\pm}$ s'obtient en utilisant la propriété (13.67) :

$$\begin{aligned}\|\,\mathcal{J}_\pm|j,m\rangle\,\|^2 = \langle j,m|\,\mathcal{J}_\mp\mathcal{J}_\pm|j,m\rangle &= \langle j,m|\left(\mathcal{J}^2-\mathcal{J}_z^2\mp\hbar J_z\right)|j,m\rangle\\ &= \hbar^2\left(j(j+1)-m^2\mp m\right)\langle j,m|j,m\rangle\end{aligned} \tag{13.70}$$

Nous avons donc, en choisisant de manière appopriée la phase arbitraire pour les vecteurs $|j,m\pm 1\rangle$ par rapport au vecteur $|j,m\rangle$, la relation :

$$\boxed{\mathcal{J}_\pm|j,m\rangle = \hbar\sqrt{j(j+1)-m(m\pm 1)}\,|j,m\pm 1\rangle} \tag{13.71}$$

Valeurs possibles de j et m. Les représentations étudiées étant unitaires, il existe des contraintes importantes sur les valeurs possibles de j et m. En effet les normes carrées des vecteurs $\mathcal{J}_\pm|j,m\rangle$, voir éqn. (13.70), doivent toutes être positives.

Considérons premièrement l'application répétée de $\mathcal{J}_+$ sur un certain vecteur $|j,m\rangle$. Le coefficient de la norme carrée,

$$|\lambda_{j,m,+}|^2 = \hbar^2\big(j(j+1)-m(m+1)\big)\,, \tag{13.72}$$

est, j étant fixé, une fonction positive entre $m=-1-j<0$ et $m=j\geqslant 0$ s'annulant pour ces deux valeurs.

Pour éviter que ce coefficient devienne négatif par actions successives de $\mathcal{J}_+$, on doit choisir m comme étant de la forme

$$m = j-r\ , \quad r\in\mathbb{N} \tag{13.73}$$

impliquant que la séquence stoppe pour

$$\boxed{\mathcal{J}_+|j,j\rangle = 0\,.} \tag{13.74}$$

On peut alors, partant par le vecteur $|j,j\rangle$, appelé *vecteur de plus haut poids*, construire d'autres vecteurs propres normés de $\mathcal{J}^2$ et $\mathcal{J}_z$ par application répétée de $\mathcal{J}_-$. D'après (13.71) nous avons (si $j>0$) :

$$|j,j-1\rangle = \frac{1}{\hbar\sqrt{j(j+1)-j(j-1)}}\,\mathcal{J}_-\,|j,j\rangle = \frac{1}{\hbar\sqrt{2j}}\,\mathcal{J}_-\,|j,j\rangle\ , \tag{13.75}$$

et ainsi de suite :

$$|j, m-1\rangle = \frac{1}{\hbar\sqrt{j(j+1)-m(m-1)}}\, \mathcal{J}_- \, |j, m\rangle \ . \tag{13.76}$$

Au contraire de l'oscillateur harmonique cette séquence ne peut se poursuivre indéfiniment. L'action de $\mathcal{J}_-$ sur un vecteur $|j, m\rangle$ donne un vecteur de norme carrée

$$|\lambda_{j,m,-}|^2 = \hbar^2\big(j(j+1)-m(m-1)\big)\,, \tag{13.77}$$

Cette fonction est positive entre $m = -j \leqslant 0$ et $m = 1+j > 0$ s'annulant pour ces deux valeurs. Pour éviter que ce coefficient devienne négatif par actions successives de $\mathcal{J}_-$, on doit choisir m comme étant de la forme

$$m = -j + r'\,, \tag{13.78}$$

et nous atteignons par application répétée de $\mathcal{J}_-$:

$$\boxed{\mathcal{J}_- \, |j, -j\rangle = 0\,.} \tag{13.79}$$

La combinaison des conditions (13.73) et (13.73) nous donne toutes les valeurs possibles de j :

$$m = j - r = -j + r' \quad \Longrightarrow \quad 2j = r - r' \in \mathbb{Z} \cap \mathbb{R}^+\,. \tag{13.80}$$

Finalement, remarquons que nous avons supposé implicitement que, pour j et m donnés, il n'existait qu'un vecteur propre commun de $\mathcal{J}^2$ et $\mathcal{J}_z$, noté $|j, m\rangle$, dans la représentation considérée. Si on avait supposé l'existence d'un autre vecteur propre $|j, m\rangle_2$, linéairement indépendant de $|j, m\rangle$, alors on aurait pu construire par action de $\mathcal{J}_\pm$ une autre représentation de $\mathfrak{su}_2$, contredisant l'hypothèse d'irreducibilité.

Représentations irreductibles de $\mathfrak{su}_2$

Les représentations de l'algèbre $\mathfrak{su}_2$ sont étiquetées par un nombre j,

$$j \in \frac{1}{2}\mathbb{N}\,. \tag{13.81}$$

Une base orthonormée d'une telle représentation, de dimension

$$d_j = 2j + 1\,, \tag{13.82}$$

est donnée par les vecteurs $|j, m\rangle$, avec

$$m = -j, -j+1, \ldots, j-1, j\,. \tag{13.83}$$

Si j est entier (resp. demi-entier), toutes les valeurs de m sont entières (resp. demi-entières).

Toute représentation de l'algèbre $\mathfrak{su}_2$ de spin j donne, par exponentiation, une représentation du groupe $SU(2)$ de dimension $d_j = 2j+1$:

$$D^j(\vec{n},\psi) = e^{-\frac{i\psi}{\hbar}\vec{n}\cdot\vec{J}}. \tag{13.84}$$

Ces représentations correspondent à des représentations de $SO(3)$, le groupe des rotations, qui peuvent être projectives ou non car le groupe $SU(2)$ est le groupe de recouvrement du groupe $SO(3)$.

Pour élucider cet aspect, considérons une rotation autour de l'axe Oz, d'un angle ψ. L'élément de la représentation de spin j du groupe $SU(2)$ correspondant est donné par l'opérateur :

$$D^j(\vec{e}_z,\psi) = e^{-\frac{i}{\hbar}\psi J_z}. \tag{13.85}$$

Nous en déduisons, pour une rotation de 2π,

$$D^j(\vec{e}_z,2\pi)\,|j,m\rangle = e^{-\frac{2i\pi}{\hbar}J_z}\,|j,m\rangle = e^{-2i\pi m}\,|j,m\rangle = (-1)^{2j}\,|j,m\rangle\ , \tag{13.86}$$

car m est de la forme $m = -j+r$ avec $r \in \mathbb{N}$.

Nous arrivons donc à la conclusion suivante :

- les représentations de $\mathfrak{su}_2$ avec $j \in \mathbb{N}$ correspondent à des représentations ordinaires du groupe des rotations $SO(3)$;
- les représentations de $\mathfrak{su}_2$ avec $j \in \mathbb{N}+\frac{1}{2}$ correspondent à des représentations *projectives* du groupe des rotations $SO(3)$.

Pour une valeur de j donnée, les vecteurs $|j,m\rangle$ sont donnent la base sur laquelle on peut développer un « vecteur » se transformant dans la représentation correspondante. Par exemple, le spineur (13.34) peut se représenter comme :

$$|\chi\rangle = \chi_+\,|1/2,1/2\rangle + \chi_-\,|1/2,-1/2\rangle\ . \tag{13.87}$$

Considérons également la représentation $j=1$, de dimension $d_1 = 3$. Elle donne une représentation ordinaire de $SO(3)$, correspondant aux transformations des vecteurs de $\mathbb{R}^3$ par rotation. Les trois états contenus dans cette représentation sont donc les 3 composantes d'un vecteur ordinaire dans une certaine base :

$$\vec{v} \mathrel{\widehat{=}} \begin{pmatrix} v_1 \\ v_0 \\ v_{-1} \end{pmatrix} \ \leftrightarrow\ |\vec{v}\rangle = v_1\,|1,1\rangle + v_0\,|1,0\rangle + v_{-1}\,|1,-1\rangle\ . \tag{13.88}$$

Pour trouver la base de $\mathbb{R}^3$ correspondante, appelée *base sphérique*, on utilise la relation (13.85) :

$$D^1(\vec{e}_z,\psi)\,|1,m\rangle = e^{-im\psi}\,|1,m\rangle\ \ ,\quad m \in \{-1,0,1\}. \tag{13.89}$$

d'où, comme les vecteurs transforment par rotation d'axe Oz selon

$$\vec{e}_x \mapsto \cos\psi\,\vec{e}_x + \sin\psi\,\vec{e}_y\ ,\quad \vec{e}_y \mapsto -\sin\psi\,\vec{e}_x + \cos\psi\,\vec{e}_y \tag{13.90}$$

on obtient par identification (à une convention de signe près qui sera justifiée plus tard) :

$$\left\{\begin{array}{lcl} |1,1\rangle & \leftrightarrow & -\frac{1}{\sqrt{2}}\left(\vec{e}_x + i\vec{e_y}\right) \\ |1,0\rangle & \leftrightarrow & \vec{e}_z \\ |1,-1\rangle & \leftrightarrow & \frac{1}{\sqrt{2}}\left(\vec{e}_x - i\vec{e_y}\right) \end{array}\right. \longrightarrow \left\{\begin{array}{lcl} v_x & = & -\frac{1}{\sqrt{2}}\left(v_1 - v_{-1}\right) \\ v_y & = & -\frac{i}{\sqrt{2}}\left(v_1 + v_{-1}\right) \\ v_z & = & v_0 \end{array}\right. \tag{13.91}$$

À retenir

— Représentations de l'algèbre $\mathfrak{su}_2$ et du groupe $\mathrm{SU}(2)$;
— représentation de spin 1/2 et spineurs ;
— comportement d'un spineur par une rotation de 2π ;
— notion de recouvrement du groupe des rotations ;
— moment magnétique de spin ;
— représentation de $\mathfrak{su}_2$:
  — opérateur de Casimir ;
  — opérateurs d'échelle ;
  — valeurs de j et m permises ;
  — dimension d'une représentation ;
  — spins entier et demi-entier.

Bibliographie

[1] P. Zeeman, "Over de invloed eener magnetisatie op den aard van het door een stof uitgezonden licht," Koninklijk Akademie van Wetenschappen te Amsterdam, 1896

[2] G.E. Uhlenbeck et S. Goudsmit, "Ersetzung der Hypothese vom unmechanischen Zwang durch eine Forderung bezüglich des inneren Verhaltens jedes einzelnen Elektrons," Naturwissenschaften, 1925

Cours n°14

Moment cinétique II : harmoniques sphériques

Dans le cours n° 13 nous avons vu que les particules (élémentaires ou composites) sont caractérisés, outre leur masse et leur charge, par leur *spin*, associé à leur transformation par rotation dans une certaine représentation du groupe $SU(2)$ mélangeant leurs différentes composantes, par exemple les 2 composantes d'un spineur ou les 3 composantes d'un vecteur.

Il existe en mécanique quantique un autre type d'action du groupe des rotations, lié à la transformation de l'argument des fonctions d'onde, que la particule possède un spin ou non. Nous avons vu dans le cours n° 4 qu'une fonction d'onde d'une particule sans spin se transforme selon :

$$\psi(\vec{x}) \xmapsto{R(\vec{n},\psi)} \psi\big(R^{-1}(\vec{n},\psi)\vec{x}\big) = e^{-\frac{i\psi}{\hbar}\vec{n}\cdot\vec{\mathcal{L}}}\,\psi(\vec{x})\,, \tag{14.1}$$

où $\vec{\mathcal{L}} = \vec{X} \wedge \vec{\mathcal{P}}$ est appelé *moment cinétique orbital*, pour le distinguer du « moment cinétique de spin ».

Si on utilise les coordonnées sphériques (r, θ, ϕ) dans $\mathbb{R}^3$, le groupe des rotations autour de l'origine va uniquement agir sur les angles (θ, ϕ), coordonnées sur la sphère de rayon r. L'action du groupe que nous devons considérer est donc une action de $SO(3)$ sur les fonctions définies sur une sphère.

L'objectif de ce cours est d'analyser cette action du groupe des rotations sur l'espace des fonctions sur la sphère $\mathcal{L}^2(S^2)$, qui correspond une représentation de dimension infinie, en décomposant toute fonction sur la sphère sur une base de fonctions se transformant dans des representations irreductibles de $SO(3)$, appelées *harmoniques sphériques*. Ces harmoniques sphériques apparaîtront dans les fonctions d'ondes associées à des hamiltoniens invariants par rotation, par exemple pour les orbitales atomiques qui seront étudiées dans le cours n° 15.

14.1 Fonctions sur la sphère

En physique, nous sommes souvent amenés à considérer des fonctions sur S^2, la sphère usuelle à deux dimensions dans l'espace euclidien à trois dimensions. Par exemple, nous étudierons dans le cours n° 15 une particule dans un potentiel ne dépendant que de la coordonnée

radiale r ; dans ce cas nous verrons qu'il est possible de trouver une base de solutions avec séparation de variables, correspondant à des fonctions d'onde $\psi(\vec{x}) = \psi_r(r)\psi_a(\vec{n})$, où la deuxième fonction dépend uniquement du vecteur sur la sphère de rayon unité :

$$\vec{n} = \begin{pmatrix} \cos\theta \\ \sin\theta\cos\phi \\ \sin\theta\sin\phi \end{pmatrix} . \tag{14.2}$$

correspondant au vecteur de base radial $\vec{e}_r$ des coordonnées sphériques (r, θ, ϕ) de $\mathbb{R}^3$, c.-à.-d. que la fonction $\psi_a(\vec{n})$ est une fonction de la sphère S^2 vers $\mathbb{C}$.

La sphère admettant une action naturelle du groupe des rotations, l'analyse des fonctions sur la sphère découle de l'analyse des fonctions sur le groupe $SO(3)$ que nous aborderons en premier.

14.1.1 Fonctions sur un groupe

Considérons une fonction f à valeurs complexes définie sur un groupe de Lie G, ici le groupe des rotations :

$$f : \left\{ \begin{array}{ccc} G & \to & \mathbb{C} \\ g & \mapsto & f(g) \end{array} \right. . \tag{14.3}$$

Rappelons qu'un groupe de Lie admet une action du groupe sur lui-même. Pour intégrer une fonction sur le groupe, on introduit une mesure dite *mesure de Haar*, qui s'exprime pour le groupe $SO(3)$ comme :

$$d\mu(g) = \tfrac{1}{2}\sin^2\tfrac{\psi}{2}\sin\theta d\psi d\theta d\phi . \tag{14.4}$$

On peut montrer explicitement que, pour un système de coordonnées donné sur le groupe, cette mesure est l'unique mesure invariante par l'action (à gauche) du groupe, c.-à.-d. telle que :

$$\forall h \in G , \quad \int_G d\mu(g)\, f(h\cdot g) = \int_G d\mu(g)\, f(g) . \tag{14.5}$$

Si on prend comme exemple le groupe $SO(2)$ des rotations à deux dimensions, voir éqn. (3.59), la mesure de Haar est naturellement $d\theta$, invariante par rotation, c.-à.-d. sous $\theta \mapsto \theta + \theta_0$.

On peut alors définir l'espace de Hilbert $\mathcal{L}^2(G)$, espace des fonctions de carré sommable sur le groupe G, avec le produit hermitien

$$\langle f_1, f_2\rangle \overset{\text{déf.}}{=} \int_G d\mu(g)\, \overline{f_1(g)} f_2(g) . \tag{14.6}$$

Un groupe de Lie G admettant une action sur lui même, on peut associer à $\mathcal{L}^2(G)$ une représentation du groupe G. Nous définissons une action du groupe sur la fonction f par :

$$\big[\rho(h)\circ f\big](g) \overset{\text{déf.}}{=} f(h^{-1}\cdot g) \tag{14.7}$$

et cette action définit bien une représentation du groupe :

$$\begin{aligned}\big[\rho(h_1\cdot h_2)\circ f\big](g) &= f\big((h_1\cdot h_2)^{-1}\cdot g\big) = f(h_2^{-1}\cdot h_1^{-1}\cdot g) = \big[\rho(h_2)\circ f\big](h_1^{-1}g)\\ &= \big[\rho(h_1)\circ\rho(h_2)\circ f\big](g)\end{aligned} \tag{14.8}$$

Cette représentation est unitaire, grâce à la propriété (14.5) :

$$\forall h\in G,\ \langle\rho(h)\circ f_1,\rho(h)\circ f_2\rangle = \int_G d\mu(g)\,\overline{f_1(h^{-1}g)}f_2(h^{-1}g) = \int_G d\mu(g)\,\overline{f_1(g)}f_2(g) = \langle f_1,f_2\rangle \tag{14.9}$$

D'après le théorème de Peter–Weyl, voir le théorème 3 du cours n° 3, cette représentation unitaire (de dimension infinie) se décompose en somme directe de représentations unitaires irreductibles de dimension finie.

Fonctions sur $SO(3)$

Considérons la décomposition d'une fonction $f(\vec{n},\psi)$ sur $SO(3)$ en termes de ses représentations unitaires irreductibles ; une telle décomposition est la généralisation de la décomposition de Fourier sur un cercle.

Une représentation unitaire irreductible du groupe $SO(3)$ étant caractérisée par $j\in\mathbb{N}$, nous avons

$$f(\vec{n},\psi) = \sum_{j=0}^{\infty}\sum_{m=-j}^{j}\sum_{m'=-j}^{j} f^j_{mm'}D^j_{mm'}(\vec{n},\psi)\,, \tag{14.10}$$

en termes des éléments de matrice

$$D^j_{mm'}(\vec{n},\psi) = \langle j,m|\,D^j(\vec{n},\psi)\,|j,m'\rangle\ , \tag{14.11}$$

appelée éléments de la *matrice D de Wigner.*[1] Les sommes sur m et m' dans (14.10) sont sur tous les entiers entre $-j$ et j.

Les éléments de la matrice D de Wigner satisfont à des relations d'orthogonalité que l'on admettra :

$$\int_{SO(3)} d\mu(g)\,\overline{D^{j_1}_{m_1m_1'}(g)}D^{j_2}_{m_2m_2'} = \frac{8\pi^2}{2j+1}\delta_{j_1,j_2}\delta_{m_1,m_2}\delta_{m_1',m_2'}\,, \tag{14.12}$$

et constituent une base de $\mathcal{L}^2\big(SO(3)\big)$.

Ces relations d'orthogonalité permettent d'obtenir les coefficients de Fourier du développement (14.10) suivant :

$$f^j_{mm'} = \tfrac{2j+1}{8\pi^2}\int_{SO(3)} d\mu(g)\,\overline{D^j_{mm'}(g)}f(g)\,. \tag{14.13}$$

et la série (14.10) converge vers f au sens de la norme $\mathcal{L}^2\big(SO(3)\big)$, de manière analogue à la convergence des séries de Fourier usuelles sur le cercle.

1. Strictement parlant la matrice D est définie en paramétrant les éléments du groupe des rotations par les angles d'Euler définis plus bas.

14.1.2 Développement des fonctions sur la sphère

La sphère S^2 dans $\mathbb{R}^3$ admet une action naturelle du groupe des rotations de l'espace, qui est transitive car deux points quelconques de la sphère sont reliés par une rotation. Rappelons que le groupe $SU(2)$ est équivalent à une sphère à trois dimensions dans $\mathbb{R}^4$:

$$M = \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix} , \quad |\alpha|^2 + |\beta|^2 = 1 . \tag{14.14}$$

avec $\alpha = w + ix$ et $\beta = y + iz$, et que le groupe $SO(3)$ est obtenu en identifiant $M \sim -M$.

Considérons le quotient $SU(2)/U(1)$, c.-à.-d. l'identification des éléments du groupe (14.14) par rotation selon l'axe Oz :

$$\begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix} \sim \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix} \begin{pmatrix} e^{-i\psi} & 0 \\ 0 & e^{i\psi} \end{pmatrix} = \begin{pmatrix} \alpha e^{-i\psi} & -\bar{\beta} e^{i\psi} \\ \beta e^{-i\psi} & \bar{\alpha} e^{i\psi} \end{pmatrix} . \tag{14.15}$$

Étant donné qu'une rotation d'angle $\psi = \pi$ revient à $M \mapsto -M$, on peut tout aussi bien décrire l'espace obtenu comme le quotient de $SO(3)$ par le sous-groupe des rotations selon Oz.

On peut choisir un représentant du quotient avec $\alpha \in \mathbb{R}$, en absorbant la phase de α avec ψ. La condition (13.31) sur le déterminant devient alors

$$\alpha^2 + y^2 + z^2 = 1 , \tag{14.16}$$

décrivant bien une sphère à deux dimensions S^2 de rayon unité.

On peut ainsi obtenir des fonctions sur la sphère S^2 en partant de fonctions sur le groupe $SO(3)$ invariantes par l'action (à droite) d'une rotation selon l'axe Oz :

$$\forall g \in SO(3) , \ \forall \psi \in [0, 2\pi[, \ f(g) = f\big(g \cdot g(\vec{e}_z, \psi)\big) . \tag{14.17}$$

On peut imposer cette condition sur la décomposition (14.10) et nous trouvons que les éléments de matrice de coefficient $f^j_{mm'}$ non-nul doivent satisfaire la condition :

$$D^j_{mm'}\big(g\big) = D^j_{mm'}\big(g \cdot g(\vec{e}_z, \psi)\big) \ \Leftrightarrow \langle j, m| \, D^j(g) \, |j, m'\rangle = \langle j, m| \, D^j(g) D^j\big(g(\vec{e}_z, \psi)\big) \, |j, m'\rangle . \tag{14.18}$$

Par construction $|j, m'\rangle$ est un vecteur propre de $D^j\big(g(\vec{e}_z, \psi)\big)$ donc nous avons

$$\langle j, m| \, D^j(g) D^j\big(g(\vec{e}_z, \psi)\big) \, |j, m'\rangle = e^{-im'\psi} \langle j, m| \, D^j(g) \, |j, m'\rangle . \tag{14.19}$$

La condition (14.17) impose donc que seuls les termes avec $m' = 0$ interviennent dans la décomposition (14.10). On utilisera conventionnellement la lettre l au lieu de la lettre j dans le présent contexte et nous obtenons

$$f(g) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f^l_{m0} D^l_{m0}(g) . \tag{14.20}$$

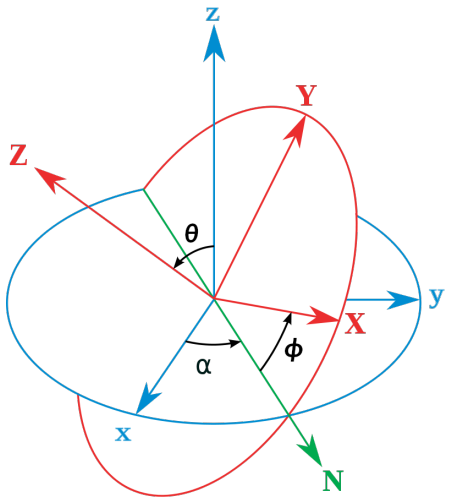

FIGURE 14.1 – Angles d'Euler (source : Wikicommons)

Les fonctions $D^l_{m0}(g)$ sont essentiellement les *harmoniques sphériques* que nous déterminerons explicitement plus bas.

Pour mieux comprendre la dépendance angulaire de ces éléments de matrice devons utiliser une autre représentation des matrices de rotation, utilisant les *angles d'Euler*, voir la figure 14.1.

Une rotation quelconque peut se décomposer comme une composition de trois rotations, dans l'ordre :

— une rotation d'axe Oz et d'angle α ;

— une rotation d'axe Oy et d'angle θ, envoyant l'axe Oz sur l'axe Oz' dans le plan (Ox, Oz) ;

— une rotation d'axe Oz et d'angle ϕ, envoyant l'axe Oz' sur l'axe OZ.

La matrice de rotation s'écrit donc

$$\begin{aligned} R(\phi,\theta,\alpha) &= R_z(\phi)R_y(\theta)R_z(\alpha) \\ &= \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \tag{14.21}$$

L'axe et l'angle de cette rotation sont donnés par des fonctions compliquées de (θ, ϕ, α) qui ne seront pas nécessaires ici. On remarquera cependant que, avec ce paramétrage,

$$R(\phi,\theta,\alpha)\vec{e}_z = \vec{n}(\theta,\phi)\,, \tag{14.22}$$

indépendamment de α. Nous avons ensuite

$$\begin{aligned}\langle l, m|\, D^l(\phi,\theta,\alpha)\,|l,0\rangle &= \langle l, m|\, D^l\left[R_z(\phi)R_y(\theta)R_z(\alpha)\right]|l,0\rangle \\ &= \underbrace{\langle l, m|\, D^l\left[R_z(\phi)\right]}_{e^{-im\phi}\langle l,m|} D^l\left[R_y(\theta)\right] \underbrace{D^l\left[R_z(\alpha)\right] D^l\,|l,0\rangle}_{|l,0\rangle} \\ &= e^{-im\phi}\,\langle l, m|\, D^l\left[R_y(\theta)\right]|l,0\rangle\ . \end{aligned} \tag{14.23}$$

Les éléments de matrice sont donc de la forme

$$\langle l, m|\, D^l(\phi,\theta,\alpha)\,|l,0\rangle = e^{-im\phi} d^l_m(\theta)\,, \tag{14.24}$$

où les *éléments de matrice réduits* d^l_m ne dépendent que de l'angle θ. On peut réécrire le développement (14.20) sous la forme :

$$\boxed{f(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} f^l_m\, e^{-im\phi}\, d^l_m(\theta)}\ , \quad d^l_m(\theta) = \langle l, m|\, D^l\left[R_y(\theta)\right]|l,0\rangle\ . \tag{14.25}$$

14.2 Harmoniques sphériques

L'objectif de cette section est de déterminer explicitement l'expression des éléments de matrice réduits $d^l_m(\theta)$, afin d'obtenir l'expression des fonctions propres communes des opérateurs de moment cinétique orbital $\vec{\mathcal{L}}^{\,2}$ et $\mathcal{L}_z$.

Les harmoniques sphériques proprement dites sont les fonctions sur la sphère S^2 associées aux vecteurs $|l, m\rangle$ de base d'une représentation irreductible l, c.-à.-d. :

$$\boxed{Y^m_l(\theta,\phi) \overset{\text{déf.}}{=} \langle \vec{n}|l, m\rangle\ .} \tag{14.26}$$

Nous pouvons remarquer, voir l'éqn. (14.22), que

$$\forall\alpha\ ,\quad \langle\vec{n}| = \langle\vec{e}_z|\,\mathcal{R}(\phi,\theta,\alpha)^\dagger\,, \tag{14.27}$$

où $\mathcal{R}(\phi,\theta,\alpha)$ est l'opérateur de rotation qui, restreint au sous-espace associé à la représentation D^l, correspond à la matrice $D^l[R(\phi,\theta,\alpha)]$.

Étant donné que les éléments de matrice de $D^l[R(\phi,\theta,\alpha)]$ ne mélangent pas différentes représentations irreductibles, c.-à.-d. les différentes valeurs de l, nous avons la relation de fermeture

$$\sum_{m'} |l, m'\rangle\,\langle l, m'| = \mathfrak{I}_{2l+1}\,, \tag{14.28}$$

où $\mathfrak{I}_{2l+1}$ est l'opérateur identité dans l'espace vectoriel de dimension $2l+1$ sur lequel agit la représentation D^l. Nous écrivons ensuite :

$$\begin{aligned}\langle\vec{n}|l, m\rangle &= \langle\vec{e}_z|\, D^l[R(\phi,\theta,\alpha)]^\dagger\,|l, m\rangle \\ &= \sum_{m'}\langle\vec{e}_z|l, m'\rangle\,\langle l, m'|\, e^{\frac{i\alpha}{\hbar}\mathcal{L}_z} e^{\frac{i\theta}{\hbar}\mathcal{L}_y} e^{\frac{i\phi}{\hbar}\mathcal{L}_z}\,|l, m\rangle \\ &= \sum_{m'}\langle\vec{e}_z|l, m'\rangle\, e^{im'\alpha+im\phi}\,\langle l, m'|\, e^{\frac{i\theta}{\hbar}\mathcal{L}_y}\,|l, m\rangle \end{aligned} \tag{14.29}$$

Le résultat devant être indépendant de l'angle α, nous concluons comme auparavant que seul $m' = 0$ contribue à la somme – excluant au passage les valeurs j demi-entières – et nous avons

$$\begin{aligned} Y_l^m(\theta,\phi) &= e^{im\phi} \langle \vec{e}_z | l, 0\rangle \langle l, 0 | e^{\frac{i\theta}{\hbar}\mathcal{L}_y} | l, m\rangle = e^{im\phi} \langle \vec{e}_z | l, 0\rangle \overline{d_m^l(\theta)} \\ &= Y_0^l(0,\phi) e^{im\phi} \overline{d_m^l(\theta)}\,, \end{aligned} \tag{14.30}$$

où $Y_l^0(0,\phi)$ est en réalité une constante (l'angle ϕ des coordonnées sphériques n'étant pas défini en $\theta = 0$) que nous déterminerons plus loin, voir éqn. (14.60). De manière équivalente, nous pouvons écrire finalement :

$$\boxed{Y_l^m(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}}\ \overline{\langle l, m | D^l(\phi,\theta,\alpha) | l, 0\rangle}} \tag{14.31}$$

Expression explicite des harmoniques sphériques

Pour trouver explicitement la forme des harmoniques sphériques, on projette sur $\langle \vec{n}|$ les équations aux vecteurs propres :

$$\boxed{\begin{aligned} \langle \vec{n} | \mathcal{L}^2 | l, m\rangle &= \hbar^2 l(l+1) \langle \vec{n} | l, m\rangle \qquad (14.32a) \\ \langle \vec{n} | \mathcal{L}_z | l, m\rangle &= \hbar m \langle \vec{n} | l, m\rangle \qquad (14.32b) \end{aligned}}$$

et on exprime ensuite les composantes du moment cinétique comme opérateurs différentiels sur $\mathcal{L}^2(S^2)$. Nous avons, pour une fonction $f(\theta,\phi)$ sur la sphère,

$$\begin{aligned} \langle \vec{n} | \vec{\mathcal{L}} | f\rangle &= -i\hbar r \vec{e}_r \wedge \left(\vec{e}_r \partial_r + \frac{1}{r}\vec{e}_\theta \partial_\theta + \frac{1}{r\sin\theta}\vec{e}_\phi \partial_\phi \right) f(\theta,\phi) \\ &= -i\hbar \left(\vec{e}_\phi \partial_\theta - \frac{1}{\sin\theta}\vec{e}_\theta \partial_\phi \right) f(\theta,\phi) \end{aligned} \tag{14.33}$$

d'où nous tirons premièrement :

$$\begin{aligned} \langle \vec{n} | \vec{\mathcal{L}}^2 | f\rangle &= -\hbar^2 \left(\vec{e}_\phi \partial_\theta - \frac{1}{\sin\theta}\vec{e}_\theta \partial_\phi \right)^2 f(\theta,\phi) \\ &= -\hbar^2 \left(\partial_\theta^2 + \operatorname{cotan}\theta \partial_\theta + \frac{1}{\sin^2\theta}\partial_\phi^2 \right) f(\theta,\phi) \end{aligned} \tag{14.34}$$

et deuxièmement :

$$\langle \vec{n} | \mathcal{L}_z | f\rangle = \vec{e}_z \cdot \langle \vec{n} | \vec{\mathcal{L}} | f\rangle = -i\hbar \partial_\phi f(\theta,\phi)\,.$$

Il est utile pour plus tard de noter que les opérateurs d'échelle $\mathcal{L}_\pm = \mathcal{L}_x \pm i\mathcal{L}_y$ correspondent aux opérateurs différentiels :

$$\begin{aligned} \langle \vec{n} | \mathcal{L}_\pm | f\rangle &= -i\hbar(\vec{e}_x \pm i\vec{e}_y) \cdot \langle \vec{n} | \vec{\mathcal{L}} | f\rangle \\ &= \hbar e^{\pm i\phi} \big(\pm \partial_\theta + i \operatorname{cotan}\theta\, \partial_\phi \big) f(\theta,\phi)\,, \end{aligned} \tag{14.35}$$

où nous avons utilisé

$$\vec{e}_\theta = \cos\theta\cos\phi\vec{e}_x + \cos\theta\sin\phi\vec{e}_y - \sin\theta\vec{e}_z \tag{14.36a}$$

$$\vec{e}_\phi = -\sin\phi\vec{e}_x + \cos\phi\vec{e}_y \tag{14.36b}$$

Les harmoniques sphériques doivent donc satisfaire les équations aux dérivées partielles suivantes :

$$\left(\partial_\theta^2 + \operatorname{cotan}\theta\partial_\theta + \frac{1}{\sin^2\theta}\partial_\phi^2\right) Y_l^m(\theta,\phi) = -l(l+1)Y_l^m(\theta,\phi) \tag{14.37a}$$

$$\partial_\phi Y_l^m(\theta,\phi) = im Y_l^m(\theta,\phi) \tag{14.37b}$$

Nous savons, voir éqn. (14.30), que les harmoniques sphériques sont le produit d'une fonction de ϕ et de θ :

$$Y_l^m(\theta,\phi) = e^{im\phi}F_l^m(\theta)\,, \tag{14.38}$$

satisfaisant naturellement à l'équation (14.37b). L'équation (14.37a) devient alors une équation différentielle du second ordre :

$$\left(\partial_\theta^2 + \operatorname{cotan}\theta\partial_\theta - \frac{m^2}{\sin^2\theta} + l(l+1)\right) F_l^m(\theta) = 0\,. \tag{14.39}$$

Au lieu de résoudre cette équation du second ordre, nous préférerons, comme pour l'oscillateur harmonique, ramener le problème à la résolution d'une équation du premier ordre. Nous savons que l'état $|l, m = -l\rangle$ satisfait

$$\mathcal{L}_- |l, -l\rangle = 0 \tag{14.40}$$

donnant à l'aide des équations (14.33,14.35) :

$$\boxed{e^{-i\phi}\left(-\partial_\theta + i\operatorname{cotan}\theta\,\partial_\phi\right) Y_l^{-l}(\theta,\phi) = 0} \tag{14.41}$$

Soit finalement

$$\frac{d}{d\theta}F_l^{-l}(\theta) = l\operatorname{cotan}(\theta)\,F_l^{-l}(\theta) \tag{14.42}$$

dont la solution est

$$F_l^{-l}(\theta) = c_{-l}\sin^l\theta \tag{14.43}$$

La constante de normalisation c_{-l} s'obtient, à une phase près, en posant

$$1 = |c_{-l}|^2\int_0^{2\pi}d\phi\int_0^\pi \sin\theta d\theta\left|Y_l^l(\theta,\phi)\right|^2 = 2\pi|c_{-l}|^2\int_0^\pi d\theta\,(\sin\theta)^{2l+1}\ . \tag{14.44}$$

L'intégrale dite de *Wallis* peut se calculer en utilisant la récurrence (immédiate par intégration par parties) [2]

$$\int_0^\pi \sin^n\theta d\theta = \frac{n-1}{n}\int_0^\pi \sin^{n-2}\theta d\theta \tag{14.45}$$

2. $\int_0^\pi \sin^n\theta\,d\theta = \int_0^\pi \sin\theta\sin^{n-1}\theta\,d\theta = (n-1)\int_0^\pi \cos^2\theta\sin^{n-2}\theta\,d\theta = (n-1)\int_0^\pi(1-\sin^2\theta)\sin^{n-2}\theta\,d\theta$ d'où le résultat.

d'où nous tirons

$$\begin{aligned}\int_0^\pi \sin^{2l+1}\theta d\theta &= \frac{2l}{2l+1}\int_0^\pi \sin^{2l-1}\theta d\theta = \frac{(2l)(2l-2)}{(2l+1)(2l-1)}\int_0^\pi \sin^{2l-3}\theta d\theta \\ &= \frac{(2l)(2l-2)(2l-4)\cdots}{(2l+1)(2l-1)(2l-3)\cdots}\times\int_0^\pi \sin\theta d\theta \\ &= \frac{2^l l!}{(2l+1)(2l-1)(2l-3)\cdots}\times 2 = \frac{\left(2^l l!\right)^2}{(2l+1)!}\times 2 \quad (14.46)\end{aligned}$$

Soit, avec un choix de phase conventionnel,

$$\boxed{Y_l^{-l}(\theta,\phi) = \sqrt{\frac{(2l+1)!}{4\pi}}\frac{1}{2^l l!}e^{-il\phi}\sin^l\theta} \quad (14.47)$$

On peut ensuite construire les autres harmoniques sphériques par application répétée de $\mathcal{L}_+$. On obtient (voir éqn. (13.75) du cours n° 13) :

$$|l,-l+1\rangle = \frac{1}{\hbar\sqrt{2l}}\mathcal{L}_+|l,-l\rangle \implies Y_l^{-l+1}(\theta,\phi) = \frac{1}{\sqrt{2l}}e^{i\phi}\left(\partial_\theta + i\cot an\,\theta\partial_\phi\right)Y_l^{-l}(\theta,\phi) \quad (14.48)$$

Soit

$$\begin{aligned}Y_l^{-l+1}(\theta,\phi) &= \frac{1}{\sqrt{2l}}\sqrt{\frac{(2l+1)!}{4\pi}}\frac{1}{2^l l!}e^{i(-l+1)\phi}\left(\partial_\theta + l\cot an\,\theta\right)\sin^l\theta\,. \\ &= \sqrt{\frac{2l(2l+1)!}{4\pi}}\frac{1}{2^l l!}e^{i(-l+1)\phi}\cos\theta\sin^{l-1}\theta\,. \quad (14.49)\end{aligned}$$

On a ensuite pour tout $m\in\{-l,-l+1,\cdots,l\}$:

$$\begin{aligned}&Y_l^m(\theta,\phi) \\ &= \frac{1}{\sqrt{(l+m)(l-m+1)}}\frac{1}{\sqrt{(l+m-1)(l-m+2)}}\cdots\frac{1}{\sqrt{2(2l-1)}}\frac{1}{\sqrt{2l}} \\ &\qquad\times\left(e^{i\phi}\left(\partial_\theta + i\cot an\,\theta\partial_\phi\right)\right)^{l+m}Y_l^{-l}(\theta,\phi) \\ &= \sqrt{\frac{(l-m)!}{(2l)!(l+m)!}}\left(e^{i\phi}\left(\partial_\theta + i\cot an\,\theta\partial_\phi\right)\right)^{l+m}Y_l^{-l}(\theta,\phi) \quad (14.50)\end{aligned}$$

Cela donne :

$$\begin{aligned}Y_l^m(\theta,\phi) &= \sqrt{\frac{(l-m)!}{(2l)!(l+m)!}}\sqrt{\frac{(2l+1)!}{4\pi}}\frac{1}{2^l l!}e^{im\phi} \\ &\times\left(\frac{d}{d\theta} + (-m+1)\cot an\,\theta\right)\cdots\left(\frac{d}{d\theta} + (l-1)\cot an\,\theta\right)\left(\frac{d}{d\theta} + l\cot an\,\theta\right)\sin^l\theta \quad (14.51)\end{aligned}$$

Pour finir le calcul on utilise l'identité

$$\left(\frac{\mathrm{d}}{\mathrm{d}\theta} + n\cot\theta\right) f(\theta) = \frac{1}{\sin^n\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\big(\sin^n\theta\, f(\theta)\big) \tag{14.52}$$

d'où

$$\begin{aligned} Y_l^m(\theta,\phi) &= \frac{1}{2^l l!}\sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}e^{im\phi} \\ &\times \left(\frac{1}{\sin^{-m+1}\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\sin^{-m+1}\theta\right)\cdots\left(\frac{1}{\sin^{l-1}\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\sin^{l-1}\theta\right)\left(\frac{1}{\sin^{l}\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\sin^{l}\theta\right)\sin^l\theta \end{aligned} \tag{14.53}$$

Et on obtient au final :

$$Y_l^m(\theta,\phi) = \frac{1}{2^l l!}\sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}e^{im\phi}\sin^m\theta\left(\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\right)^{l+m}\sin^{2l}\theta \tag{14.54}$$

En posant $x = \cos\theta$, nous avons en effet

$$\sin^m\theta\left(\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\right)^{l+m}\sin^{2l}\theta = (-1)^{l+m}(1-x^2)^{m/2}\frac{\mathrm{d}^{l+m}}{\mathrm{d}x^{l+m}}(1-x^2)^l \tag{14.55}$$

Nous avons fait apparaître les *polynômes de Legendre associés* – qui ne sont en réalité des polynômes que pour m pair ! – définis par

$$\boxed{P_l^m(x) = \frac{(-1)^l}{2^l l!}(1-x^2)^{m/2}\frac{\mathrm{d}^{l+m}}{\mathrm{d}x^{l+m}}(1-x^2)^l} \tag{14.56}$$

Permettant d'écrire les harmoniques sphériques dans leur forme finale :

$$\boxed{Y_l^m(\theta,\phi) = (-1)^m\sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}e^{im\phi}P_l^m(\cos\theta)} \tag{14.57}$$

Donnons une propriété importante des harmoniques sphériques que nous démontrerons plus loin :

$$\boxed{Y_l^{-m}(\theta,\phi) = (-1)^m\overline{Y_l^m(\theta,\phi)}} \tag{14.58}$$

reliant les harmoniques sphériques pour $m > 0$ à celles pour $m < 0$.

Finalement, nous pouvons expliciter le lien donné par l'éqn. (14.30) entre les harmoniques sphériques et les éléments de matrice des représentations de $SU(2)$. Nous premièrement

$$Y_l^0(\theta) = \sqrt{\frac{(2l+1)}{4\pi}}\frac{(-1)^l}{2^l l!}\frac{\mathrm{d}^l}{\mathrm{d}x^l}(1-x^2)^l\Big|_{x=\cos\theta} \tag{14.59}$$

Remarquons que, pour $\theta = 0$, cette expression se simplifie grandement car dans $\frac{d^l}{dx^l}(1-x^2)^l = \frac{d^l}{dx^l}(1-x)^l(1+x)^l$ seuls les termes où toutes les dérivées agissent sur les facteurs $(1-x)$ donne un résultat non-nul pour $x = 1$. Nous en déduisons que

$$Y_l^0(0) = \sqrt{\frac{(2l+1)}{4\pi}}\frac{(-1)^l}{2^l l!}(-1)^l l(l-1)\cdots 1(1+1)^l = \sqrt{\frac{2l+1}{4\pi}} \tag{14.60}$$

donc

$$Y_l^m(\theta,\phi) = \sqrt{\frac{(2l+1)}{4\pi}}e^{-im\phi}\overline{d_m^l(\theta)} = \sqrt{\frac{2l+1}{4\pi}}e^{-im\phi}\overline{d_m^l(\theta)} \tag{14.61}$$

et comme annoncé plus haut

$$\langle l, m|\, D^l(\theta,\phi,\phi)\,|l,0\rangle = \sqrt{\frac{4\pi}{2l+1}}\,\overline{Y_l^m(\theta,\phi)} \tag{14.62}$$

14.3 Propriétés des harmoniques sphériques

Plusieurs propriétés des harmoniques sphériques, bien que pouvant être prouvées directement, proviennent de leur relation avec les éléments de matrice des représentations de $SU(2)$ développée au début de ce cours.

Premièrement, les vecteurs $|l,m\rangle$ formant par construction une base orthonormée de la représentation D^l nous avons

$$\delta_{l,l'}\delta_{m,m'} = \langle lm|l'm'\rangle = \int_0^{2\pi} d\phi \int_0^{\pi} d\theta\, \sin\theta\, \langle lm|\vec{n}\rangle\, \langle \vec{n}|l'm'\rangle \tag{14.63}$$

soit la relation d'orthonormalité :

$$\boxed{\int_0^{2\pi} d\phi \int_0^{\pi} d\theta\, \sin\theta\, \overline{Y_l^m(\theta,\phi)}Y_{l'}^{m'}(\theta,\phi) = \delta_{l,l'}\delta_{m,m'}} \tag{14.64}$$

Deuxièmement, d'après le théorème de Peter–Weyl, on peut décomposer une fonction de carré sommable sur la sphère en harmoniques sphériques, voir éqn. (14.20) :

$$\begin{aligned} f(\theta,\phi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} f_{m0}^l D_{m0}^l(\theta,\phi) &= \sum_{l=0}^{\infty}\sqrt{\frac{4\pi}{2l+1}}\sum_{m=-l}^{l} f_{m0}^l \overline{Y_l^m(\theta,\phi)} \\ &= \sum_{l=0}^{\infty}\sqrt{\frac{4\pi}{2l+1}}\sum_{m=-l}^{l} f_{m0}^l (-)^m Y_l^{-m}(\theta,\phi) \end{aligned} \tag{14.65}$$

où nous avons utilisé l'éqn. (14.58). Remplaçant m par $-m$ dans la somme et en définissant

$$f_l^m = \sqrt{\frac{4\pi}{2l+1}}(-)^m f_{-m0}^l \tag{14.66}$$

nous avons finalement :

$$\boxed{f(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_l^m \, Y_l^m(\theta,\phi)} \tag{14.67}$$

avec

$$\boxed{f_l^m = \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \, \sin\theta \, \overline{Y_l^m(\theta,\phi)} f(\theta,\phi)} \tag{14.68}$$

Ainsi, les harmoniques sphériques forment une base de Hilbert de $\mathcal{L}^2(S^2)$, espace de Hilbert des fonctions de carré sommable sur la sphère.

Parité et renversement du temps

L'opérateur parité $\mathfrak{P}$, associé à la symétrie de parité $P : \vec{x} \mapsto -\vec{x}$, agit en coordonnées cartésiennes comme

$$\langle \vec{x} | \, \mathfrak{P} = \langle -\vec{x} | \ , \tag{14.69}$$

donc son action en coordonnées sphériques est

$$\langle \theta, \phi | \, \mathfrak{P} = \langle \pi - \theta, \phi + \pi | \ . \tag{14.70}$$

Nous avons donc

$$Y_l^m(\theta,\phi) = \langle \theta, \phi | l, m \rangle \ \mapsto \ \langle \theta, \phi | \, \mathfrak{P} \, | l, m \rangle = \langle \pi - \theta, \phi + \pi | l, m \rangle = Y_l^m(\pi - \theta, \phi + \pi) \tag{14.71}$$

et, d'après la définition (14.56),

$$\begin{aligned} Y_l^m(\pi - \theta, \phi + \pi) &= (-1)^m \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} e^{im(\phi+\pi)} P_l^m(-\cos\theta) \\ &= (-1)^m \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} e^{im(\phi+\pi)} (-1)^{l+m} P_l^m(\cos\theta) \end{aligned} \tag{14.72}$$

Soit

$$\boxed{P \ : \ Y_l^m(\theta,\phi) \ \mapsto \ (-1)^l Y_l^m(\theta,\phi)} \tag{14.73}$$

Examinons le cas plus délicat du renversement du temps T, associé à un opérateur anti-unitaire χ. Nous avons obtenu dans le cours n° 8, éqn. (8.90), l'action de χ sur les observables position et impulsion. Nous en déduisons pour le moment cinétique orbital :

$$\chi^{-1} \mathcal{L}_i \chi = \chi^{-1} \sum_{j,k} \epsilon^{ijk} \mathcal{X}_j \mathcal{P}_k \chi = \sum_{j,k} \epsilon^{ijk} \chi^{-1} \mathcal{X}_j \chi \chi^{-1} \mathcal{P}_k \chi = -\mathcal{L}_i \, . \tag{14.74}$$

Nous avons donc

$$\mathcal{L}_z \chi \, | l, m \rangle = -\chi \mathcal{L}_z \, | l, m \rangle = -\hbar m \chi \, | l, m \rangle \tag{14.75}$$

donc $\chi\,|l,m\rangle$ est un vecteur propre de $\mathcal{L}_z$ de valeur propre $-\hbar m$ et

$$\exists \alpha_m \in \mathbb{C}\,,\ \chi\,|l,m\rangle = \alpha_m\,|l,-m\rangle \tag{14.76}$$

Pour trouver la constante de proportionnalité on procède de la manière suivante. L'opérateur χ étant anti-unitaire nous avons

$$\big(\mathcal{L}_x + i\mathcal{L}_y\big)\chi = -\chi\big(\mathcal{L}_x - i\mathcal{L}_y\big) \tag{14.77}$$

Donc d'une part

$$\mathcal{L}_+\chi\,|l,m\rangle = \alpha_m\mathcal{L}_+\,|l,-m\rangle = \alpha_m\hbar\sqrt{l(l+1)-m(m-1)}\,|l,-m+1\rangle \tag{14.78}$$

et d'autre part

$$\begin{aligned}\mathcal{L}_+\chi\,|l,m\rangle = -\chi\mathcal{L}_-\,|l,m\rangle &= -\chi\hbar\sqrt{l(l+1)-m(m-1)}\,|l,m-1\rangle \\ &= -\hbar\sqrt{l(l+1)-m(m-1)}\alpha_{m-1}\,|l,-m+1\rangle\end{aligned} \tag{14.79}$$

donc

$$\alpha_m = -\alpha_{m-1} \tag{14.80}$$

et on peut choisir

$$\alpha_m = (-1)^m\,, \tag{14.81}$$

compatible avec l'invariance de Y_l^0. Le renversement du temps agit donc sur les harmoniques sphériques comme

$$\boxed{\mathsf{T}\ :\ Y_l^m(\theta,\phi)\ \mapsto\ (-1)^m Y_l^{-m}(\theta,\phi)} \tag{14.82}$$

Nous savons par ailleurs (voir éqn. (8.93) du cours n° 8) que l'effet du renversement du temps sur une fonction d'onde est de prendre son complexe conjugué. On obtient donc l'importante relation :

$$\boxed{Y_l^{-m}(\theta,\phi) = (-1)^m\overline{Y_l^m(\theta,\phi)}} \tag{14.83}$$

qui peut également s'obtenir directement à partir de l'expression (14.56) des polynômes de Legendre associés.

Rotations

L'action des rotations sur les harmoniques sphériques est aisée à dériver en utilisant leur relation avec les éléments de matrice de $\mathsf{SU}(2)$, voir éqn. (14.62), alors qu'une preuve directe à partir des polynômes de Legendre associés est fort complexe.

Considérons une action du groupe $\mathsf{SU}(2)$ sur lui-même, envoyant un point (θ,ϕ,ψ) de l'hypersphère sur (θ',ϕ',ψ') :

$$g(\theta,\phi,\psi) \mapsto g(\theta_0,\phi_0,\psi_0)\cdot g(\theta,\phi,\psi) = g(\theta',\phi',\psi') \tag{14.84}$$

Si nous nous plaçons dans une certaine représentation irreductible D^l nous obtenons pour les éléments de matrice de $SU(2)$ correspondants :

$$\langle l, m| D^l(\theta, \phi, \psi) |l, 0\rangle \mapsto \langle l, m| D^l(\theta', \phi', \psi') |l, 0\rangle = \langle l, m| D^l(\theta_0, \phi_0, \psi_0) D^l(\theta, \phi, \psi) |l, 0\rangle \tag{14.85}$$

Étant donné que $\{|l, m'\rangle, m' = -l, \cdots l\}$ forme une base orthonormée de la représentation, nous avons, en utilisant la relation de fermeture associée :

$$\begin{aligned}&\langle l, m| D^l(\theta_0, \phi_0, \psi_0) D^l(\theta, \phi, \psi) |l, 0\rangle \\ &\qquad = \sum_{m'=-l}^{l} \langle l, m| D^l(\theta_0, \phi_0, \psi_0) |l, m'\rangle \langle l, m'| D^l(\theta, \phi, \psi) |l, 0\rangle\end{aligned} \tag{14.86}$$

faisant apparaître les éléments de la matrice D de Wigner, voir l'éqn. (14.11). En utilisant l'éqn. (14.62) cela implique que

$$\overline{Y_l^m(\theta, \phi)} \mapsto \sum_{m'=-l}^{l} D^l_{m,m'}(\theta_0, \phi_0, \psi_0) \overline{Y_l^{m'}(\theta, \phi)} \tag{14.87}$$

Soit, en prenant le complexe conjugué

$$\boxed{Y_l^m(\theta, \phi) \xmapsto{R(\theta_0, \phi_0, \psi_0)} Y_l^m(\theta', \phi') = \sum_{m'=-l}^{l} \overline{D^l_{m,m'}(\theta_0, \phi_0, \psi_0)} Y_l^{m'}(\theta, \phi)} \tag{14.88}$$

14.4 Exemples d'harmoniques sphériques

Pour clore ce cours nous donnons quelques exemples simples d'harmoniques sphériques, en introduisant également la notation spectroscopique associée, sur laquelle nous reviendrons lors de l'étude de l'atome d'hydrogène.

Représentation $l = 0$ (orbitale s) Pour $l = 0$ il n'existe qu'une seule harmonique sphérique,

$$Y_0^0(\theta, \phi) = \frac{1}{\sqrt{4\pi}}, \tag{14.89}$$

correspondant à une fonction constante sur la sphère.

Représentation $l = 1$ (orbitales p) Pour $l = 1$ il existe trois harmoniques sphériques,

$$Y_1^{\pm 1}(\theta, \phi) = \mp \sqrt{\frac{3}{8\pi}} \sin\theta e^{\pm i\phi}, \tag{14.90a}$$

$$Y_1^0(\theta, \phi) = \sqrt{\frac{3}{4\pi}} \cos\theta \tag{14.90b}$$

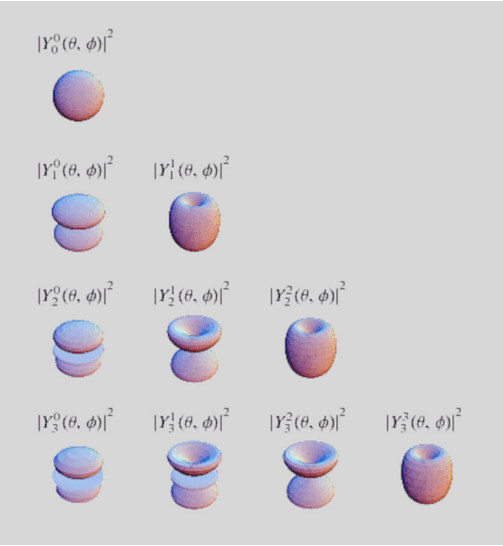

FIGURE 14.2 – Représentation de $|Y_l^m(\theta,\phi)|^2$ (source : Wolfram)

Exprimons ces harmoniques sphériques en termes des coordonnées cartésiennes. Nous avons :

$$Y_1^{\pm1}(x,y,z) = \mp\frac{1}{r}\sqrt{\frac{3}{4\pi}}\,\frac{x \pm iy}{\sqrt{2}}\,, \tag{14.91a}$$

$$Y_1^0(x,y,z) = \frac{1}{r}\sqrt{\frac{3}{4\pi}}\,z \tag{14.91b}$$

qui s'identifie avec la base sphérique définie par l'équation (13.91), justifiant a posteriori les conventions de signe que nous avions utilisées.

Représentation $l = 2$ (orbitales d) Pour $l = 2$ il existe cinq harmoniques sphériques,

$$Y_2^{\pm2}(\theta,\phi) = \sqrt{\frac{15}{32\pi}}\sin^2\theta e^{\pm2i\phi}\,, \tag{14.92a}$$

$$Y_2^{\pm1}(\theta,\phi) = \mp\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{\pm i\phi}\,, \tag{14.92b}$$

$$Y_2^0(\theta,\phi) = \sqrt{\frac{5}{16\pi}}(3\cos\theta^2 - 1) \tag{14.92c}$$

On peut représenter graphiquement les harmoniques sphériques de différentes manières, par exemple en traçant $f(\theta,\phi) = |Y_l^m(\theta,\phi)|^2$ en fonction des angles (θ,ϕ), donnant la probabilité de présence en fonction de la direction anglulaire, voir la figure 14.2.

À retenir

— Éléments de matrice de $SO(3)$ et harmoniques sphériques ;

— équations différentielles satisfaites par les harmoniques sphériques ;
— transformation sous parité et sous renversement du temps ;
— rotation d'une harmonique sphérique.

Cours n°15

Potentiel central et atome d'hydrogène

L'objectif de ce cours est d'étudier le spectre en énergie de systèmes quantiques associés à des particules dans un potentiel à symétrie sphérique en nous servant des outils développés dans les cours précédents, en particulier l'étude du moment cinétique. Nous appliquerons ensuite ce formalisme à l'exemple le plus important de tel système, l'atome d'hydrogène.

15.1 Hamiltonien à potentiel central

Considérons une particule massive sans spin à trois dimensions dont le hamiltonien est de la forme :

$$\mathcal{H} = \frac{1}{2m}\vec{\mathcal{P}}^2 + V\big(\|\vec{\mathcal{X}}\|\big)\,, \tag{15.1}$$

Le potentiel, ne dépendant que de la distance par rapport à l'origine des coordonnées, est appelé *potentiel central*.

Un tel système est possède une symétrie de rotation, voir le cours n° 3, éqn. (3.31), pour une démonstration dans le cadre classique. Cette symétrie nous permettra d'exploiter notre connaissance des représentations du groupe des rotations $SO(3)$ pour obtenir les états propres du système.

15.1.1 Séparation de variables

Pour exploiter l'invariance par rotation, il est nécessaire de transformer l'opérateur $\vec{\mathcal{P}}^2$ afin de faire apparaître explicitement le moment cinétique orbital $\vec{\mathcal{L}}$. On note premièrement que :

$$\vec{\mathcal{L}}^2 = \sum_i \sum_{j,k} \epsilon^{ijk}\mathcal{X}_j\mathcal{P}_k \sum_{j',k'} \epsilon^{ij'k'}\mathcal{X}_{j'}\mathcal{P}_{k'} = \sum_{j,k,j',k'} \big(\sum_i \epsilon^{ijk}\epsilon^{ij'k'}\big)\mathcal{X}_j\mathcal{P}_k\mathcal{X}_{j'}\mathcal{P}_{k'} \tag{15.2}$$

La quantité $\sum_i \epsilon^{ijk}\epsilon^{ij'k'}$ étant antisymétrique en (j,k) d'une part et en (j',k') d'autre part, et symétrique par échange $(j,k) \leftrightarrow (j',k')$ nous avons

$$\sum_i \epsilon^{ijk}\epsilon^{ij'k'} = \delta_{jj'}\delta_{k,k'} - \delta_{kj'}\delta_{k'j} \tag{15.3}$$

d'où

$$\begin{aligned}\vec{\mathcal{L}}^2 = \sum_{j,k,j',k'} \big(\delta_{jj'}\delta_{k,k'} - \delta_{kj'}\delta_{k'j}\big)\mathcal{X}_j\mathcal{P}_k\mathcal{X}_{j'}\mathcal{P}_{k'} &= \sum_{j,k} \mathcal{X}_j\mathcal{P}_k\mathcal{X}_j\mathcal{P}_k - \sum_{j,k} \mathcal{X}_j\mathcal{P}_k\mathcal{X}_k\mathcal{P}_j \\ &= \sum_{j,k} \mathcal{X}_j\big(\mathcal{X}_j\mathcal{P}_k - i\hbar\delta_{jk}\big)\mathcal{P}_k - \sum_{j,k} \mathcal{X}_j\mathcal{P}_k\big(\mathcal{P}_j\mathcal{X}_k + i\hbar\delta_{jk}\big) \\ &= \vec{\mathcal{X}}^2\vec{\mathcal{P}}^2 - i\hbar\vec{\mathcal{X}}\cdot\vec{\mathcal{P}} - \sum_{j,k} \mathcal{X}_j\mathcal{P}_j\mathcal{P}_k\mathcal{X}_k - i\hbar\vec{\mathcal{X}}\cdot\vec{\mathcal{P}} \\ &= \vec{\mathcal{X}}^2\vec{\mathcal{P}}^2 - 2i\hbar\vec{\mathcal{X}}\cdot\vec{\mathcal{P}} - \sum_{j,k} \mathcal{X}_j\mathcal{P}_j\big(\mathcal{X}_k\mathcal{P}_k - i\hbar\delta_{k,k}\big) \\ &\Longrightarrow \quad \boxed{\vec{\mathcal{L}}^2 = \vec{\mathcal{X}}^2\vec{\mathcal{P}}^2 - \Big(\vec{\mathcal{X}}\cdot\vec{\mathcal{P}}\Big)^2 + i\hbar\vec{\mathcal{X}}\cdot\vec{\mathcal{P}}} \end{aligned} \tag{15.4}$$

Étant donné que

$$\langle\vec{x}|\,\vec{\mathcal{X}}\cdot\vec{\mathcal{P}}\,|\psi\rangle = -i\hbar\vec{x}\cdot\vec{\nabla}\psi(\vec{x}) = -i\hbar r\frac{\partial}{\partial r}\psi(r,\theta,\phi)\,, \tag{15.5}$$

en passant aux coordonnées sphériques (r,θ,ϕ), nous obtenons :

$$\langle\vec{x}|\,\vec{\mathcal{L}}^2\,|\psi\rangle = r^2\,\langle\vec{x}|\,\vec{\mathcal{P}}^2\,|\psi\rangle + \hbar^2 r\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\psi(r,\theta,\phi)\right) + \hbar^2 r\frac{\partial}{\partial r}\psi(r,\theta,\phi) \tag{15.6}$$

Nous obtenons donc la relation entre l'action de $\vec{\mathcal{P}}^2$ et l'action de $\vec{\mathcal{L}}^2$:

$$\langle\vec{x}|\,\vec{\mathcal{P}}^2\,|\psi\rangle = \frac{1}{r^2}\,\langle\vec{x}|\,\vec{\mathcal{L}}^2\,|\psi\rangle - \hbar^2\frac{\partial^2}{\partial r^2}\psi - \frac{2\hbar^2}{r}\frac{\partial}{\partial r}\psi \tag{15.7}$$

Et, partant du hamiltonien (15.1),

$$\boxed{\langle r,\theta,\phi|\,\mathcal{H}\,|\psi\rangle = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right)\psi(r,\theta,\phi) + \frac{1}{2mr^2}\,\langle\vec{x}|\,\vec{\mathcal{L}}^2\,|\psi\rangle + V(r)\psi(r,\theta,\phi)} \tag{15.8}$$

Étant donné que l'opérateur $\vec{\mathcal{L}}^2$ agit sur les fonctions d'onde comme un opérateur différentiel en les coordonnées (θ,ϕ) sur la sphère, voir éqn. (14.34), il est naturel de chercher des fonctions propres du hamiltonien sous une forme factorisée.

L'opérateur $\vec{\mathcal{L}}^2$ étant proportionnel à l'identité dans une représentation irreductible de $SO(3)$, cherchons des solutions se transformant dans une représentation D^l donnée. Une base de fonctions d'ondes sur la sphère associées à une telle représentation étant fournie par les harmoniques sphériques Y_l^m, nous posons :

$$\boxed{\psi_{\alpha,l,m}(r,\theta,\phi) = \rho_{\alpha,l}(r)Y_l^m(\theta,\phi)} \tag{15.9}$$

où α désigne d'autres nombres quantiques éventuels. [1] Nous savons, voir l'éqn. (14.37a), que

$$\langle\theta,\phi|\,\vec{\mathcal{L}}^2\,|l,m\rangle = \hbar^2 l(l+1)Y_l^m(\theta,\phi) \tag{15.10}$$

L'équation aux valeurs propres est donnée en représentation position par :

$$\Big[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)+\frac{\hbar^2 l(l+1)}{2mr^2}+V(r)\Big]\rho_{\alpha,l}(r)Y_l^m(\theta,\phi) = E\rho_{\alpha,l}(r)Y_l^m(\theta,\phi) \tag{15.11}$$

La résolution du problème, dans la représentation D^l du moment cinétique orbital se ramène donc à la résolution de l'*équation radiale* :

$$\boxed{-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2}+\frac{2}{r}\frac{\partial}{\partial r}\right)\rho_{\alpha,l}(r)+\Big(\frac{\hbar^2 l(l+1)}{2mr^2}+V(r)\Big)\rho_{\alpha,l}(r) = E\rho_{\alpha,l}(r)} \tag{15.12}$$

Cette équation se simplifie en posant

$$\rho_{\alpha,l}(r) = \frac{R_{\alpha,l}(r)}{r} \tag{15.13}$$

et l'équation (15.12) devient :

$$\boxed{-\frac{\hbar^2}{2m}R''_{\alpha,l}(r)+\underbrace{\left(V(r)+\frac{\hbar^2 l(l+1)}{2mr^2}\right)}_{V_{\text{EFF}}(r)}R_{\alpha,l}(r) = ER_{\alpha,l}(r)} \tag{15.14}$$

Nous nous sommes donc ramenés à l'équation aux énergies propres pour une particule à une dimension dans un potentiel effectif

$$V_{\text{EFF}}(r) = V(r)+\frac{\hbar^2 l(l+1)}{2mr^2} \tag{15.15}$$

qui comporte un terme répulsif à courte distance pour les représentations avec $l>0$.

Les harmoniques sphériques étant normalisées, voir l'éqn. (14.64), nous avons la condition de normalisation suivante :

$$\int d^3r|\psi_{l,m}(r,\theta,\phi)|^2 = \Big(\int \sin\theta d\theta d\phi|Y_l^m(\theta,\phi)|^2\Big)\int_0^\infty r^2dr\,|\rho_{\alpha,l}(r)|^2 \implies \boxed{\int_0^\infty dr\,|R_{\alpha,l}(r)|^2 = 1} \tag{15.16}$$

1. Dit autrement, l'espace des états se scinde en somme directe sur les sous-espaces $\mathcal{H}_l$ de dimension $(2l+1)$ associés aux différentes représentations irreductibles D^l du groupe des rotations,

$$\mathcal{H} \cong \bigoplus_{l=0}^{\infty}\mathcal{H}_l$$

et chaque sous-espace $\mathcal{H}_l$ se factorise comme un produit tensoriel

$$\mathcal{H}_l \cong \mathcal{L}^2(\mathbb{R}^+, r^2\,dr)\otimes\mathbb{C}^{2l+1}\,,$$

où $\mathcal{L}^2(\mathbb{R}^+, r^2\,dr)$ désigne l'espace des fonctions de carré sommable sur $\mathbb{R}^+$ pour la mesure $r^2\,dr$ associée à l'intégrale radiale en coordonnées sphériques, c.-à.-d. telles que $\int_0^\infty r^2dr\,|f(r)|^2<\infty$.

15.1.2 Comportement asymptotique

Examinons le comportement de l'équation radiale (15.14) dans la limite $r \to 0^+$, pour $l > 0$. En supposant que $V(r)$ diverge moins vite que $1/r^2$ lorsque $r \to 0^+$, nous avons, en ne gardant que les termes les plus singuliers,

$$R''_{\alpha,l}(r) \overset{r\to 0^+}{\sim} \frac{l(l+1)}{r^2} R_{\alpha,l}(r) \tag{15.17}$$

Cette relation indique un comportement asymptotique de la forme

$$R_{\alpha,l}(r) \overset{r\to 0^+}{\sim} Ar^{l+1} + Br^{-l}. \tag{15.18}$$

Pour que la condition de normalisation (15.16) soit satisfaite, il faut que R_l soit intégrable au voisinage de zéro, ce qui impose $B = 0$, donc nous avons

$$R_l(r) \overset{r\to 0^+}{\sim} r^{l+1} \tag{15.19}$$

au voisinage de l'origine. La probabilité de présence dans une sphère de rayon r_0 autour de l'origine se comporte donc pour $r_0 \to 0$ comme

$$p(r_0) \sim \int_0^{r_0} |Ar^{l+1}|^2 \propto r_0^{2l+3} \tag{15.20}$$

qui tend vers zéro lorsque $r_0 \to 0$.

Examinons maintenant le comportement de l'équation radiale (15.14) à grande distance, dans la limite $r \to +\infty$, pour $l > 0$. En supposant que $V(r)$ tend vers zéro à l'infini nous avons

$$R''_{\alpha,l}(r) \overset{r\to +\infty}{\sim} = \frac{-2mE}{\hbar^2} R_{\alpha,l}(r) \tag{15.21}$$

Il faut distinguer deux types de comportement, suivant le signe de E.

États liés S'il existe des solutions avec $E < 0$, la particule se trouve dans un état lié (leur existence est liée à la forme du potentiel). Si nous posons

$$\kappa = \sqrt{\frac{-2mE}{\hbar^2}} \in \mathbb{R}^+ \tag{15.22}$$

Nous trouvons la solution physiquement acceptable :

$$R_{\alpha,l}(r) \overset{r\to +\infty}{\sim} \lambda e^{-\kappa r} \tag{15.23}$$

États asymptotiques S'il existe des solutions avec $E > 0$, la particule se trouve dans un état asymptotique. Si nous posons

$$k = \sqrt{\frac{2mE}{\hbar^2}} \in \mathbb{R}^+ \tag{15.24}$$

Nous trouvons des solutions de la forme :

$$R_{\alpha,l}(r) \overset{r\to +\infty}{\sim} \lambda e^{ikr} + \mu e^{-ikr} \tag{15.25}$$

15.1.3 Exemple : particule libre

Le formalisme développé précédemment est valable pour tout potentiel central, en particulier lorsque le potentiel est nul.

Nous avons déjà obtenu le spectre du hamiltonien d'une particule libre par transformation de Fourier, voir le cours n° 2, correspondant à la décomposition d'une fonction d'onde quelconque comme une superposition d'ondes planes (qui ne constituent pas, rappelons-le, une base de $\mathcal{L}^2(\mathbb{R}^3)$). Cette décomposition diagonalisait simultanément le hamiltonien et les translations spatiales, ces dernières étant associées à l'opérateur impulsion.

Nous nous proposons ici de considérer une décomposition diagonalisant simultanément le hamiltonien et les rotation spatiales. Nous savons déjà, voir l'éqn. (2.71), que le spectre est donné par

$$E = \frac{\hbar^2 k^2}{2m} \tag{15.26}$$

ce qui conduit à l'équation radiale :

$$R''_{k,l}(r) + \left(k^2 - \frac{l(l+1)}{r^2}\right) R_l(r) = 0 \tag{15.27}$$

Si on pose $u = kr$, et $f_{k,l}(u) = R_{k,l}(r)$ on obtient

$$f''_{k,l}(u) + \left(1 - \frac{l(l+1)}{u^2}\right) f_{k,l}(u) = 0 \tag{15.28}$$

En revenant à la fonction radiale d'origine, posant $f_{k,l}(u) = u g_{k,l}(u)$, nous arrivons à

$$\boxed{g''_{k,l}(u) + \frac{2}{u} g'_{k,l}(u) + \left(1 - \frac{l(l+1)}{u^2}\right) g_{k,l}(u) = 0} \tag{15.29}$$

appelée *équation de Bessel sphérique*. Les solutions de cette équation sont, naturellement, les fonctions de Bessel sphérique. En ne considérant que les solutions régulières à l'origine, notées $j_l(u)$, nous avons pour les premières d'entre elles :

$$j_0(u) = \frac{\sin u}{u} \tag{15.30a}$$

$$j_1(u) = \frac{\sin u}{u^2} - \frac{\cos u}{u} \tag{15.30b}$$

$$j_2(u) = \left(\frac{3}{u^3} - \frac{1}{u}\right) \sin u - \frac{3\cos u}{u^2} \tag{15.30c}$$

qui possèdent bien le comportement asymptotique attendu pour $r \to 0$ et pour $r \to \infty$, associé dans le second cas à des états asymptotiques.

15.2 Potentiel coulombien et atome d'hydrogène

L'exemple le plus important de potentiel central en mécanique quantique, en particulier d'un point de vue historique, est celui du potentiel coulombien, particulièrement dans l'exemple de l'atome d'hydrogène.

Rappelons que, dans une description classique et non relativiste, atome d'hydrogène est un état lié d'un proton (de masse m_p et de charge $q > 0$) et d'un électron (de masse m_e et de charge $-q < 0$), obtenu par interaction électrostatique.

L'opérateur hamiltonien correspondant, obtenu par quantification canonique, est de la forme :

$$\mathcal{H} = \frac{\vec{\mathcal{P}}_p^2}{2m_p} + \frac{\vec{\mathcal{P}}_e^2}{2m_e} - \frac{q^2}{4\pi\epsilon_0}\frac{1}{||\vec{\mathcal{X}}_e - \vec{\mathcal{X}}_p||} \tag{15.31}$$

où $\vec{\mathcal{P}}_p$ (resp. $\vec{\mathcal{P}}_e$) désigne l'opérateur impulsion agissant sur l'espace des états du proton (resp. de l'électron) ; de même $\vec{\mathcal{X}}_p$ (resp. $\vec{\mathcal{X}}_e$) désigne l'opérateur position agissant sur l'espace des états du proton (resp. de l'électron).

15.2.1 Réduction d'un problème à deux corps

Le hamiltonien (15.31) contient deux types de degrés de liberté ; ceux du centre de masse, dont la dynamique est triviale, et ceux du mouvement relatif qui nous intéressent.

Pour séparer ces deux types de degrés de liberté, nous introduisons tout d'abord les opérateurs associés au centre de masse :

$$\vec{\mathcal{X}}_c \overset{\text{déf.}}{=} \frac{m_e\vec{\mathcal{X}}_e + m_p\vec{\mathcal{X}}_p}{m_e + m_p}\,, \quad \vec{\mathcal{P}}_c \overset{\text{déf.}}{=} \vec{\mathcal{P}}_p + \vec{\mathcal{P}}_e\,. \tag{15.32}$$

ainsi que ceux associés au mouvement relatif :

$$\vec{\mathcal{X}}_{rel} \overset{\text{déf.}}{=} \vec{\mathcal{X}}_e - \vec{\mathcal{X}}_p\,, \quad \vec{\mathcal{P}}_{rel} \overset{\text{déf.}}{=} \frac{m_p\vec{\mathcal{P}}_e - m_e\vec{\mathcal{P}}_p}{m_e + m_p}\,. \tag{15.33}$$

On peut vérifier aisément que les impulsions sont bien conjuguées aux positions correspondantes. Nous avons premièrement :

$$\begin{aligned}\left[\,\mathcal{X}_c^i,\,\mathcal{P}_c^j\,\right] &= \frac{m_e}{m_e+m_p}\left[\,\mathcal{X}_e^i,\,\mathcal{P}_e^j\,\right] + \frac{m_p}{m_e+m_p}\left[\,\mathcal{X}_p^i,\,\mathcal{P}_p^j\,\right] \\ &= \frac{m_e}{m_e+m_p}i\hbar\delta_{ij} + \frac{m_p}{m_e+m_p}i\hbar\delta_{ij} = i\hbar\delta_{ij}\end{aligned} \tag{15.34}$$

Et deuxièmement :

$$\begin{aligned}\left[\,\mathcal{X}_{rel}^i,\,\mathcal{P}_{rel}^j\,\right] &= \frac{m_p}{m_e+m_p}\left[\,\mathcal{X}_e^i,\,\mathcal{P}_e^j\,\right] + \frac{m_e}{m_e+m_p}\left[\,\mathcal{X}_p^i,\,\mathcal{P}_p^j\,\right] \\ &= \frac{m_p}{m_e+m_p}i\hbar\delta_{ij} + \frac{m_e}{m_e+m_p}i\hbar\delta_{ij} = i\hbar\delta_{ij}\end{aligned} \tag{15.35}$$

On vérifie ensuite que ces deux jeux de degrés de liberté sont indépendants. Nous avons tout d'abord :

$$\left[X^i_{rel}, P^j_c \right] = \left[X^i_e, P^j_e \right] - \left[X^i_p, P^j_p \right] = 0 \tag{15.36}$$

Puis :

$$\left[X^i_c, P^j_{rel} \right] = \frac{m_e}{m_e + m_p} \frac{m_p}{m_e + m_p} \left[X^i_e, P^j_e \right] - \frac{m_p}{m_e + m_p} \frac{m_e}{m_e + m_p} \left[X^i_p, P^j_p \right] = 0 \tag{15.37}$$

Le hamiltonien (15.31) prend alors la forme :

$$\mathcal{H} = \underbrace{\frac{1}{2M} \vec{P}_c^{\,2}}_{\mathcal{H}_c} + \underbrace{\frac{1}{2\mu} \vec{P}_{rel}^{\,2} - \frac{q^2}{4\pi\epsilon_0} \frac{1}{\|\vec{X}_{rel}\|}}_{\mathcal{H}_{rel}} \tag{15.38}$$

en fonction de la masse totale et de la masse réduite, respectivement

$$M \overset{\text{déf.}}{=} m_e + m_p \ , \quad \mu \overset{\text{déf.}}{=} \frac{m_e m_p}{m_e + m_p} \tag{15.39}$$

En pratique, comme les masses de l'électron et du proton sont respectivement

$$m_e \simeq 9,109 \times 10^{-31} \,\text{kg} \ , \quad m_p \simeq 1,672 \times 10^{-27} \,\text{kg} \tag{15.40}$$

on peut considérer que $M \simeq m_p$ et $\mu \simeq m_e$. D'un point de vue de la mécanique classique, on peut ainsi considérer un électron tournant autour d'un proton immobile dans le référentiel du centre de masse.

La factorisation du hamiltonien que nous avons obtenue entre hamiltonien du centre de masse $\mathcal{H}_c$ et hamiltonien du mouvement relative $\mathcal{H}_{rel}$, signifie que l'espace des états possède la structure d'un produit tensoriel $\mathcal{L}^2(\mathbb{R}^3) \otimes \mathcal{L}^2(\mathbb{R}^3)$ sur lequel agit le hamiltonien :

$$\mathcal{H} = \mathcal{H}_c \otimes \mathcal{I} + \mathcal{I} \otimes \mathcal{H}_{rel} \,. \tag{15.41}$$

Le hamiltonien du centre de masse correspond à une particule fictive de masse M, dont nous avons déjà étudié le spectre dans le cours n° 2.

15.2.2 Résolution de l'équation radiale

À l'aide du formalisme développé dans la section 15.1, nous pouvons analyser le spectre du hamiltonien relatif $\mathcal{H}_{rel}$ qui possède bien un potentiel central (nous avons retiré l'indice rel pour plus de clarté) :

$$\boxed{\mathcal{H}_{rel} = \frac{1}{2\mu} \vec{P}^{\,2} - \frac{q^2}{4\pi\epsilon_0} \frac{1}{\|\vec{X}\|}} \tag{15.42}$$

Nous cherchons donc, pour une représentation D^l, un vecteur d'état $|\alpha, l, m\rangle$ associé à une fonction d'onde de la forme :

$$\psi_{\alpha,l,m}(r, \theta, \phi) = \frac{1}{r} R_{\alpha,l}(r) Y^m_l(\theta, \phi) \tag{15.43}$$

où la fonction radiale $R_{\alpha,l}$ satisfait à l'équation :

$$-\frac{\hbar^2}{2\mu}R''_{\alpha,l}(r) - \frac{q^2}{4\pi\epsilon_0}\frac{1}{r} + \frac{\hbar^2 l(l+1)}{2\mu r^2}R_{\alpha,l}(r) = ER_{\alpha,l}(r) \tag{15.44}$$

Soit

$$R''_{\alpha,l} - \frac{l(l+1)}{r^2}R_{\alpha,l} + \frac{q^2}{4\pi\epsilon_0}\frac{2\mu}{\hbar^2}\frac{1}{r}R_{\alpha,l} = -\frac{2\mu}{\hbar^2}ER_{\alpha,l} \tag{15.45}$$

On voit apparaître une échelle de longueur caractéristique, le *rayon de Bohr* :

$$a_0 \overset{\text{déf.}}{=} \frac{4\pi\epsilon_0\hbar^2}{\mu q^2} \tag{15.46}$$

À partir de laquelle on peut obtenir une échelle d'énergie, l'*énergie de ionisation*

$$E_I \overset{\text{déf.}}{=} \frac{\hbar^2}{2\mu a_0^2} = \frac{\mu q^4}{2\hbar^2(4\pi\epsilon_0)^2} \tag{15.47}$$

On peut poser en toute généralité pour un état lié (énergie négative) :

$$E = -E_I/n^2\,, \quad n \in \mathbb{R}_{>0} \tag{15.48}$$

Nous avons

$$R''_{\alpha,l} - \frac{l(l+1)}{r^2}R_{\alpha,l} + \frac{2}{a_0 r}R_{\alpha,l} = \frac{1}{(na_0)^2}R_{\alpha,l} \tag{15.49}$$

Effectuons ensuite le changement de variable

$$u = \frac{r}{na_0}\,, \quad f_{\alpha,l}(u) = R_{\alpha,l}(r) \tag{15.50}$$

Nous arrivons enfin à l'équation radiale sous sa forme adimensionnée,

$$\boxed{f''_{\alpha,l} - \frac{l(l+1)}{u^2}f_{\alpha,l} + \frac{2n}{u}f_{\alpha,l} = f_{\alpha,l}} \tag{15.51}$$

Cette équation correspond à un problème de Schrödinger pour un potentiel effectif adimensionné

$$V_{\text{EFF}} = \frac{l(l+1)}{u^2} - \frac{2n}{u}\,, \tag{15.52}$$

représenté sur la figure 15.1. Nous voyons que, pour $l \neq 0$, le terme associé au moment cinétique est un terme répulsif qui, pour prendre une image classique, « repousse » une particule s'approchant de l'origine. Nous nous attendons à ce que ce potentiel admette un spectre ponctuel d'états liés pour $E < 0$ et un spectre continu pour $E \geqslant 0$.

Revenons au calcul. Pour prendre en compte les comportements asymptotiques analysés dans la sous-section 15.1.2, nous posons

$$f_{\alpha,l}(u) = u^{l+1}e^{-u}g_{\alpha,l}(u)\,. \tag{15.53}$$

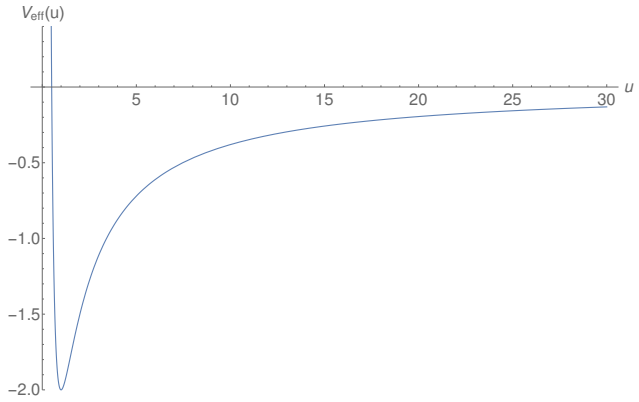

FIGURE 15.1 – Potentiel effectif adimensionné pour l'atome d'hydrogène ($n = 2$ et $l = 1$).

En injectant cette décomposition dans l'équation (15.51) nous trouvons une équation pour $g_{\alpha,l}$:

$$\boxed{u g''_{\alpha,l} + 2(l+1-u) g'_{\alpha,l} + 2\big(n-(l+1)\big) g_{\alpha,l} = 0} \tag{15.54}$$

appelée *équation de Kummer*. Posons

$$a = l + 1 - n \,, \quad b = 2(l+1) \,. \tag{15.55}$$

Les solutions de cette équation sont appelées *fonctions hypergéométriques confluentes* et bien connues dans la littérature mathématique. Elles admettent un développement en série :

$${}_1F_1(a,b;2u) = 1 + \frac{a}{b}\frac{2u}{1!} + \frac{a(a+1)}{b(b+1)}\frac{(2u)^2}{2!} + \frac{a(a+1)(a+2)}{b(b+1)(b+2)}\frac{(2u)^3}{3!} + \cdots \tag{15.56}$$

que l'on peut trouver directement en cherchant une solution de l'équation (15.54) de la forme d'une série entière. [2] Si aucun des facteurs $(a+k)$, $k \in \mathbb{N}$ apparaissant dans le développement ne s'annule, le comportement asymptotique de cette function est donné par

$${}_1F_1(a,b;2u) \overset{u\to\infty}{\sim} \frac{\Gamma(a)}{\Gamma(b)} e^{2u} (2u)^{a-b} \tag{15.57}$$

Un tel comportement à l'infini n'est pas admissible, car il modifie le comportement asymptotique de la fonction d'onde que nous avions déjà pris en compte (voir la sous-section 15.1.2). Cela signifie donc que :

$$\exists k \in \mathbb{N} \,, \quad a + k = 0 \;\Leftrightarrow\; n = (l+1) + k \tag{15.58}$$

2. En effet, si on pose $x = 2u$ et $f(x) = g_{\alpha,l}(u)$, la série $f(x) = \sum_{k=0}^{\infty} c_k x^k$ est solution de l'"équation

$$x f'' + (b - x) f' - a f = 0$$

si

$$c_k = \frac{a+n-1}{n(n+b-1)} c_{k-1}$$

Lorsque cette condition est satisfaite, la fonction hypergéométrique ${}_1F_1(a,b;2u)$ se réduit à un polynôme de degré k.

Ces polynômes sont appelés *polynômes de Laguerre associés*, notés $L_k^\beta(x)$ et solutions de l'équation différentielle (avec $\beta = 2l+1$, $x = 2u$ et $y(x) = g(u)$) :

$$xy'' + (\beta + 1 - x)y' + ky = 0\,, \quad k \in \mathbb{N}\,. \tag{15.59}$$

Ils constituent une famille orthogonale sur $\mathbb{R}^+$, au sens de :

$$\int_0^\infty x^\beta e^{-x} L_n^\beta(x) L_m^\beta(x) = \frac{(n+\beta)!}{n!}\delta_{n,m} \tag{15.60}$$

et forment une base de Hilbert de $\mathcal{L}^2(\mathbb{R}^+, x^\beta e^{-x}dx)$, c.-à.-d. des fonctions de carré sommable sur $\mathbb{R}^+$ par rapport à la mesure d'intégration $x^\beta e^{-x}dx$. Le premiers de ces polynômes sont :

$$L_0^\beta(x) = 1 \tag{15.61a}$$
$$L_1^\beta(x) = 1 + \beta - x \tag{15.61b}$$
$$L_2^\beta(x) = \frac{(\beta+1)(\beta+2)}{2} - (2+\beta)x + \frac{1}{2}x^2 \tag{15.61c}$$
$$L_3^\beta(x) = \frac{(\beta+1)(\beta+2)(\beta+3)}{6} - \frac{(\beta+2)(\beta+3)}{2}x + \frac{\beta+3}{2}x^2 - \frac{1}{6}x^3 \tag{15.61d}$$

Solution de l'équation radiale. En reprenant à rebours tous les changements de variables effectués précédemment, nous trouvons que les fonctions d'onde radiales sont de la forme

$$\rho_{n,l}(r) = c_{n,l}\left(\frac{2r}{na_0}\right)^l e^{-\frac{r}{na_0}} L_{n-l-1}^{2l+1}\left(\frac{2r}{a_0}\right)\,, \tag{15.62}$$

avec la condition

$$n - l - 1 \geqslant 0 \;\leftrightarrow\; l = 0, 1, \ldots, n-1. \tag{15.63}$$

La constante de normalisation $c_{n,l}$ est déterminée par la condition

$$\int_0^\infty dr\, r^2\, |\rho_l(r)|^2 = 1 \tag{15.64}$$

Un calcul un peu fastidieux donne :

$$|c_{n,l}|^2 2n \left(\frac{na_0}{2}\right)^3 \frac{(n+l)!}{(n-l-1)!} = 1 \tag{15.65}$$

D'où nous tirons, avec un choix de phase approprié,

$$\boxed{\rho_{n,l}(r) = \sqrt{\left(\frac{2}{na_0}\right)^3 \frac{(n-l-1)!}{2n(n+l)!}}\, e^{-\frac{r}{na_0}} \left(\frac{2r}{na_0}\right)^l L_{n-l-1}^{2l+1}\left(\frac{2r}{na_0}\right)} \tag{15.66}$$

15.3 Analyse physique des solutions

Le spectre du hamiltonien de l'atome d'hydrogène se compose de deux composantes. Premièrement, le hamiltonien (15.42) possède un spectre essentiel,

$$\sigma_e(\mathcal{H}) = \mathbb{R}^+ , \tag{15.67}$$

comme nous l'avions indiqué dans le cours n° 11, example 1. Ce spectre est associé aux états asymptotiques pour lesquels, dans une vision classique, l'électron peut « s'échapper » du potentiel électrostatique du proton.

Le spectre contient aussi un spectre ponctuel, associé aux états liés du système dont les énergies discrètes sont données par :

$$\boxed{E = -\frac{E_I}{n^2} , \quad n = 1, 2, \dots} \tag{15.68}$$

en termes de l'énergie d'ionisation donnée par (dans l'approximation $\mu \simeq m_e$) :

$$\boxed{E_I = \frac{m_e q^4}{2\hbar^2 (4\pi\epsilon_0)^2} \simeq 13,6 \ eV} \tag{15.69}$$

Cette énergie correspond, comme son nom l'indique, à l'énergie minimale à fournir au système pour le faire passer de son état fondamental ($n = 1$) vers un état asymptotique (avec $E \geqslant 0$).

Pour une valeur donnée de n, appelé *nombre quantique principal*, les représentations du groupe des rotations apparaissant dans le spectre sont données par

$$l = 0, 1, \dots n-1 . \tag{15.70}$$

Une représentation donnée (c.-à.-d. une valeur de l donnée) correspond à un sous-espace vectoriel de dimension $2l+1$, dont les vecteurs de base sont caractérisés par $m = -l, \cdots l$. Une base des sous-espaces propres d'énergies (15.68) est donc donnée par :

$$\boxed{E = -\frac{E_I}{n^2} \ \leftrightarrow \ \left\{ |n, l, m\rangle \ , \ l = 0, \dots, n-1 \, , \ m = -l, \dots, l \right\}} \tag{15.71}$$

Les fonctions d'onde associées sont données par :

$$\boxed{\psi_{n,l,m}(r, \theta, \phi) = \langle r, \theta, \phi | n, l, m \rangle = \rho_{n,l}(r) Y_l^m(\theta, \phi)} \tag{15.72}$$

en termes des fonctions radiales

$$\rho_{n,l}(r) = \sqrt{\left(\frac{2}{n a_0}\right)^3 \frac{(n-l-1)!}{2n(n+l)!}} \ e^{-\frac{r}{n a_0}} \left(\frac{2r}{n a_0}\right)^l L_{n-l-1}^{2l+1}\left(\frac{2r}{n a_0}\right) \tag{15.73}$$

et des harmoniques sphériques

$$Y_l^m(\theta, \phi) = Y_l^m(\theta, \phi) = (-1)^m \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} e^{im\phi} P_l^m(\cos\theta) \tag{15.74}$$

Les fonctions radiales dépendent d'un paramètre ayant la dimension d'une longueur, le rayon de Bohr, donné par (dans l'approximation $\mu \simeq m_e$) :

$$\boxed{a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e q^2} \simeq 5,29 \times 10^{-11} \text{ m}} \tag{15.75}$$

La dégénérescence totale, pour des états liés d'énergie $E = -E_I/n^2$ avec n donné, se calcule de la manière suivante. Toutes les valeurs entières de l entre 0 et $n-1$ sont réalisées une fois, et la représentation D^l est de dimension $(2l+1)$, correspondant aux différentes valeurs de m possibles. La dégénérescence totale, que nous noterons d_n, est donnée par :

$$d_n = \sum_{l=0}^{n-1}(2l+1) = 2\frac{(n-1)n}{2} + n \quad \Longrightarrow \quad \boxed{d_n = n^2} \tag{15.76}$$

Notons pour finir l'existence de deux types de dégénerescences de nature différente dans le spectre ponctuel :

— la dégénérescence en énergie des états propres $|n, l, m\rangle$ pour les différentes valeurs de m, $m = -l, \cdots, l$, avec n et l fixé, est une conséquence de la symétrie de rotation du système, comme nous l'avons montré en toute généralité dans le cours n° 3, corollaire 1. En effet les vecteurs $|l, m\rangle$ forment une base de la représentation D^l du groupe des rotations et le hamiltonien est constant dans le sous-espace vectoriel associé. On trouverait un résultat analogue pour tout hamiltonien invariant par rotation ;

— la dégénérescence en énergie des états propres $|n, l, m\rangle$ pour les différentes valeurs de l, $l = 0, \ldots, n-1$ avec n fixé est une *dégénérescence accidentelle* pour le potentiel coulombien, qui disparaît dès que le hamiltonien (15.42) est modifié, par exemple par les corrections relativistes que nous étudierons plus loin. Cette dégénérescence accidentelle est due à l'apparition d'une symétrie accidentelle (cf. vecteur de Runge–Lenz).

Orbitale (1s)

Analysons plus en détail l'état fondamental du système $|n = 1, l = 0, m = 0\rangle$, ou (1s) dans la notation spectroscopique. La fonction d'onde associée est très simple, car en particulier la dépendance angulaire est triviale :

$$\boxed{\psi_{1,0,0}(r, \theta, \phi) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0}} \tag{15.77}$$

L'absence de dépendance angulaire signifie que la distribution de probabilité de présence est isotrope. La probabilité de trouver l'électron dans une coquille d'épaisseur dr à une distance r du proton est donnée par :

$$dp(r) = 4\pi|\psi_{1,0,0}(r)|^2 r^2 dr = 4\left(\frac{r}{a_0}\right)^2 e^{-\frac{2r}{a_0}} \frac{dr}{a_0}. \tag{15.78}$$

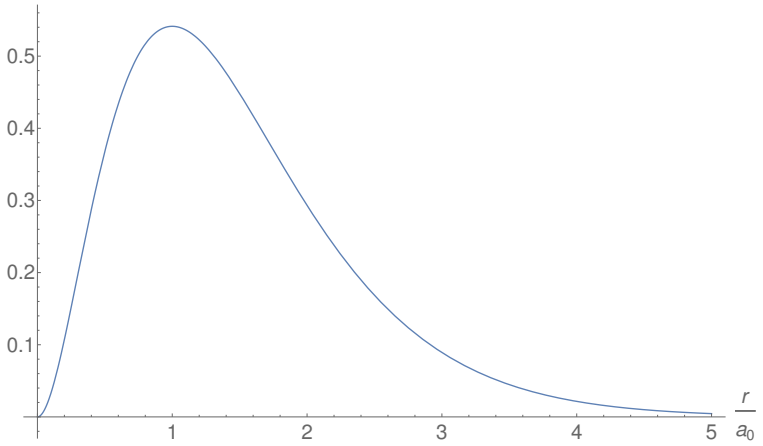

FIGURE 15.2 – Probabilité de présence dans l'état fondamental

Cette distribution est maximale à la distance $r = a_0$ correspondant au rayon de Bohr et s'annule à l'origine, voir la figure 15.2

Nous pouvons aussi considérer la distance moyenne de l'électron au proton, donnée par le calcul suivant :

$$\langle \mathcal{R} \rangle_{1,0,0} = 4\pi \int dr\, r^2 \, |\psi_{1,0,0}(r)|^2 \, r = \frac{4}{a_0^3} \int_0^\infty r^3 e^{-\frac{2r}{a_0}} dr = 4a_0 \int_0^\infty u^3 e^{-2u} du = \frac{3}{2} a_0 \tag{15.79}$$

Nous voyons que a_0 donne l'ordre de grandeur du rayon de l'atome d'hydrogène dans son état fondamental.

Orbitales (2s) et (2p)

Pour $n = 1$ nous avons premièrement un état $|n = 1, l = 0, m = 0\rangle$, ou (2s) dans la notation spectroscopique. La fonction d'onde associée est :

$$\boxed{\psi_{2,0,0}(r, \theta, \phi) = \frac{1}{4\sqrt{2\pi a_0^3}} \left(2 - \frac{r}{a_0}\right) e^{-\frac{r}{2a_0}}} \tag{15.80}$$

possédant également une symétrie sphérique. La probabilité de présence est donnée par :

$$dp(r) = 4\pi |\psi_{2,0,0}(r)|^2 r^2 dr = \frac{1}{8} \left(\frac{r}{a_0}\right)^2 \left(2 - \frac{r}{a_0}\right)^2 e^{-\frac{r}{a_0}} \frac{dr}{a_0} \tag{15.81}$$

représentée sur la figure 15.3.

Nous avons deuxièmement les orbitales (2p), états se transformant dans la représentation $l = 1$ (c.-à.-d. la représentation vectorielle), soit les états $|2, 1, 1\rangle$, $|2, 1, -1\rangle$ et $|2, 1, 0\rangle$.

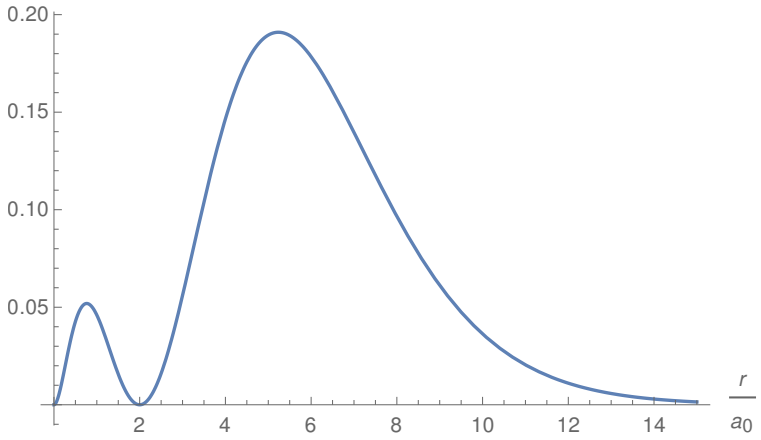

FIGURE 15.3 – Probabilité de présence dans l'état (2s)

Explicitement, les fonctions d'onde associées sont :

$$\psi_{2,1,\pm1} = \mp\frac{1}{8\sqrt{\pi a_0^3}}\,\frac{r}{a_0}\,e^{-\frac{r}{2a_0}}\,e^{\pm i\phi}\sin\theta \tag{15.82a}$$

$$\psi_{2,1,0} = \frac{1}{4\sqrt{2\pi a_0^3}}\,\frac{r}{a_0}\,e^{-\frac{r}{2a_0}}\cos\theta \tag{15.82b}$$

Atomes hydrogénoïdes

L'analyse conduite ici reste valable pour des atomes hydrogénoïdes, c.-à.-d. des atomes de numéro atomique Z dans un état ionisé ne comportant qu'un seul électron dans un état lié. Le hamiltonien d'un tel système est :

$$\mathcal{H} = \frac{1}{2\mu}\vec{\mathcal{P}}^2 - \frac{Zq^2}{4\pi\epsilon_0}\frac{1}{||\vec{\mathcal{X}}||} \tag{15.83}$$

Ainsi toute l'analyse que nous avons conduite dans ce cours reste valable, avec la substitution $q^2 \mapsto Zq^2$ dans toutes les formules précédentes. Nous voyons en particulier que le rayon moyen de l'atome dans l'état fondamental est :

$$\langle\mathcal{R}\rangle_{1,0,0} = \frac{3}{2}\frac{a_0}{Z} \tag{15.84}$$

Ce rayon est plus petit d'un facteur $1/Z$, en accord avec notre intuition car la force qui s'exerce entre le noyau et l'électron est plus forte d'un facteur Z en comparaison avec l'atome d'hydrogène.

Nous pouvons analyser de la même manière des systèmes plus exotiques comme le *positronium*, constitué d'un électron et d'un positron, dont la durée de vie est évidemment extrêmement brève en raison de l'annihilation entre les deux particules (de l'ordre de la nanoseconde).

À retenir

— hamiltonien à potentiel central : séparation de variables
— comportements asymptotiques ;
— développement en ondes sphériques pour une particule libre ;
— mouvement relatif pour un problème à deux corps ;
— résolution de l'équation radiale pour l'atome d'hydrogène ;
— états propres $|\mathfrak{n}, \mathfrak{l}, \mathfrak{m}\rangle$, énergie et dégénérescence ;
— atomes hydrogénoïdes.

Cours n°16

Moment cinétique III : composition

Les cours n° 13 et n° 14 ont été consacrés à deux concepts distincts mais reliés par un formalisme commun, la théorie des représentations des groupes $SU(2)$ et $SO(3)$. Dans le cours n° 13 nous avons vu que le spin s est une caractéristique intrinsèque des particules, au même titre que leur masse, caractérisant leur comportement possible sous une rotation dans l'espace, associé à la représentation irreductible de $SU(2)$ de dimension $2s+1$. Le cours n° 14 fut consacré à l'étude de l'opérateur moment cinétique en mécanique quantique, dont nous avons étudié les fonctions propres, ou harmoniques sphériques, en termes des éléments de matrice des représentations irreductibles de $SO(3)$.

16.1 Composition de moments cinétiques : exemples

Il arrive fréquemment que la description physique naturelle d'un système se fasse en termes de représentations *réductibles* du groupe des rotations, par exemple un système de plusieurs spins ou le vecteur d'état « complet » associé à une particule de spin non-nul. La description du système dans la base associée à cette représentation réductible est parfaitement valable, cependant il peut être utile de décomposer cette représentation en représentations irreductibles, par exemple en présence d'une perturbation du hamiltonien s'exprimant en fonction du moment cinétique total.

Un exemple de décomposition d'une représentation de $SO(3)$ en représentations irreductibles a déjà été présenté dans le cours n° 16, sous-section 13.1.1.

16.1.1 Transformation d'un spineur

Reprenons plus en détail la transformation sous une rotation d'un spineur, ou plutôt d'un « spineur de fonctions d'ondes », que nous avons déjà étudiée dans le cours n° 13, sous-section 13.2.3. Le vecteur d'état $|\chi\rangle$ d'une particule de spin 1/2 est associé à une paire de fonctions d'onde :

$$|\chi\rangle = |\chi_1\rangle \otimes |\uparrow_z\rangle + |\chi_2\rangle \otimes |\downarrow_z\rangle \ , \quad \chi_1(\vec{x}) = \big(\langle\uparrow_z| \otimes \langle\vec{x}| \big) |\chi\rangle \ , \quad \chi_2(\vec{x}) = \big(\langle\downarrow_z| \otimes \langle\vec{x}| \big) |\chi\rangle \ , \tag{16.1}$$

et appartient donc à l'espace de Hilbert

$$\mathcal{H} \cong \mathcal{L}^2(\mathbb{R}^3) \otimes \mathbb{C}^2 \,. \tag{16.2}$$

Appelons par la suite $|1/2, m_s\rangle$ les états propres communs de $\mathcal{S}^2$ et de $\mathcal{S}_z$, de valeurs propres respectives $3\hbar^2/4$ et $\hbar m_s$.

Si le hamiltonien de la particule possède une symétrie de rotation (potentiel central), il est naturel de décomposer l'espace $\mathcal{L}^2(\mathbb{R}^3)$ des fonctions d'onde se scinde comme

$$\mathcal{L}^2(\mathbb{R}^3) \cong \bigoplus_{l=0}^{\infty} \mathcal{H}_l \tag{16.3}$$

et chaque sous-espace $\mathcal{H}_l$ se factorise comme un produit tensoriel

$$\mathcal{H}_l \cong \mathcal{L}^2(\mathbb{R}^+, r^2\,dr) \otimes \mathbb{C}^{2l+1} \,, \tag{16.4}$$

où $\mathcal{L}^2(\mathbb{R}^+, r^2\,dr)$ désigne l'espace des fonctions de carré sommable sur $\mathbb{R}^+$ pour la mesure $r^2\,dr$ associée à l'intégrale radiale en coordonnées sphériques et $\mathbb{C}^{2l+1}$ est l'espace vectoriel associé à la représentation de moment cinétique orbital l.

En d'autres termes, les états propres du hamiltonien de la particule, pour une valeur de l donnée, se factorisent comme

$$|\alpha; l, m; m_s\rangle \overset{\text{déf.}}{=} |\alpha, l\rangle \otimes |l, m\rangle \otimes |1/2, m_s\rangle \in \mathcal{L}^2(\mathbb{R}^+, r^2\,dr) \otimes \mathbb{C}^{2l+1} \otimes \mathbb{C}^2 \tag{16.5}$$

et les fonctions d'onde associées sont sous la forme factorisée

$$\psi_{\alpha,l,m;m_s}(\vec{x}) = \big(\langle r| \otimes \langle \theta, \phi| \otimes |1/2, m_s\rangle \big)\, |\alpha; l, m; m_s\rangle = \rho_{\alpha,l,m_s}(r) Y_l^m(\theta, \phi) \,. \tag{16.6}$$

Elles peuvent naturellement se regrouper en un spineur de fonctions d'onde, c.-à.-d.

$$\chi_{\alpha,l,m} \overset{\text{déf.}}{=} Y_l^m(\theta, \phi) \begin{pmatrix} \rho_{\alpha,l,+}(r) \\ \rho_{\alpha,l,-}(r) \end{pmatrix} \tag{16.7}$$

Examinons la transformation de spineurs de fonctions d'onde sous une rotation. Nous avons vu dans le cours n° 13 la loi de transformation générale :

$$\begin{pmatrix} \chi_1(\vec{x}) \\ \chi_2(\vec{x}) \end{pmatrix} \xmapsto{g(\vec{n},\psi)} M(\vec{n}, \psi) \begin{pmatrix} \chi_1\big(R^{-1}(\vec{n}, \psi)\vec{x}\big) \\ \chi_2\big(R^{-1}(\vec{n}, \psi)\vec{x}\big) \end{pmatrix} , \tag{16.8}$$

combinant la transformation d'un spineur (spin 1/2) avec l'action sur les fonctions d'onde associée à une représentation réductible de $SO(3)$ de dimension infinie.

La décomposition 16.3 de l'espace de Hilbert $\mathcal{L}^2(\mathbb{R}^3)$ est associée à la décomposition de cette représentation de dimension infinie en somme directe de représentations de dimension $2l+1$ et nous obtenons pour les vecteurs d'état :

$$\begin{aligned} |\alpha; l, m; m_s\rangle &\xmapsto{g(\vec{n},\psi)} \Big(D^l(\vec{n}, \psi) \otimes D^{1/2}(\vec{n}, \psi) \Big)\, |\alpha; l, m; m_s\rangle \\ &= |\alpha, l\rangle \otimes \Big(D^l(\vec{n}, \psi)\, |l, m\rangle \Big) \otimes \Big(D^{1/2}(\vec{n}, \psi)\, |1/2, m_s\rangle \Big) \\ &= |\alpha, l\rangle \otimes \Big(e^{-\frac{i}{\hbar}\vec{n}\cdot\vec{\mathcal{L}}}\, |l, m\rangle \Big) \otimes \Big(e^{-\frac{i}{\hbar}\vec{n}\cdot\vec{\mathcal{S}}}\, |1/2, m_s\rangle \Big) \,. \end{aligned}$$

Ainsi, les états $|\alpha; l, m; m_s\rangle$ pour α et l fixés se transforment, du point de vue de la théorie des groupes, dans la représentation *réductible*

$$D = D^l \otimes D^{1/2} \tag{16.9}$$

de dimension $d = (2l+1)\times 2$. Nous omettrons désormais l'étiquette α pour ces états, associée aux états radiaux, qui n'intervient pas dans le raisonnement.

Pour analyser cette représentation et la décomposer en représentations irreductibles, introduisons le *moment cinétique total* :

$$\vec{J} \overset{\text{déf.}}{=} \vec{\mathcal{L}} + \vec{S}\,. \tag{16.10}$$

Étant donné que les opérateurs $\mathcal{L}_i$ commutent avec les opérateurs S_i, agissant sur des espaces différents, les opérateurs J_i satisfont bien aux relations de commutation d'un moment cinétique – ou du point de vue de la théorie des groupes aux relations de commutation de l'algèbre $\mathfrak{su}_2$:

$$\begin{aligned}[J_i, J_j] = [\mathcal{L}_i + S_i, \mathcal{L}_j + S_j] = [\mathcal{L}_i, \mathcal{L}_j] + [S_i, S_j] = i\hbar \sum_k \epsilon^{ijk}\mathcal{L}_k + i\hbar \sum_k \epsilon^{ijk} S_k \\ \Longrightarrow \quad [J_i, J_j] = i\hbar \sum_k \epsilon^{ijk} J_k\,. \end{aligned} \tag{16.11}$$

Nous pouvons contruire l'opérateur de Casimir associé :

$$J^2 \overset{\text{déf.}}{=} \vec{J}^2 = J_x^2 + J_y^2 + J_z^2\,. \tag{16.12}$$

Cette opérateur commute, par construction, avec les opérateurs J_i mais également avec les opérateurs de Casimir associés au moment cinétique orbital et au spin :

$$\left[J^2, \mathcal{L}^2\right] = 0\,, \quad \left[J^2, S^2\right] = 0\,. \tag{16.13}$$

En revanche, J^2 ne *commute pas* avec les opérateurs $\mathcal{L}_i$ et les opérateurs S_i !

Il est donc possible de décrire les vecteurs propres de moment cinétique et de spin de deux manières différentes :

— en termes des vecteurs de base associés à la représentation réductible $D^l \otimes D^{1/2}$,

$$\left\{ |l, m; 1/2, m_s\rangle\ ,\ m = l, \ldots, l\ ,\ m_s = -1/2, 1/2\right\} \tag{16.14}$$

vecteurs propres communs de $\mathcal{L}^2$, $\mathcal{L}_z$, S^2 et S_z ;

— en termes des représentations irreductibles D^j du moment cinétique total :

$$\left\{ |j, M\rangle\ ,\ M = -j, \ldots, j\right\}, \tag{16.15}$$

pour des valeurs de j qui restent à déterminer (vecteurs propres communs de $\mathcal{L}^2$, S^2, J^2 et J_z).

Le passage d'une base orthonormée à l'autre s'effectue par un opérateur unitaire.

Décomposition en représentations irreductibles

Il reste à déterminer les valeurs de j possibles, c.-à.-d. les représentations de $SU(2)$ apparaissant dans la décomposition de $D^l \otimes D^{1/2}$ en représentations irreductibles pour une valeur de l donnée.

La première étape, él'ementaire, consisiste à remarquer la relation suivante entre les valeurs propres de $\mathcal{L}_z$, $\mathcal{S}_z$ et $\mathcal{J}_z$:

$$\mathcal{J}_z = \mathcal{L}_z + \mathcal{S}_z \quad \Longrightarrow \quad M = m + m_s \,. \tag{16.16}$$

Les valeurs possibles de m et m_s étant connues (respectivement $\{-l, -l+1, \cdots, l-1, l\}$ et $\{-1/2, 1/2\}$) nous pouvons immédiatement obtenir les valeurs de M réalisées :

m_s \ m	$-l$	$-l+1$	$\cdots$	-1	0	1	$\cdots$	$l-1$	l
$-1/2$	$-l-1/2$	$-l+1/2$	$\cdots$	$-3/2$	$-1/2$	$1/2$	$\cdots$	$l-3/2$	$l-1/2$
$1/2$	$-l+1/2$	$-l+3/2$	$\cdots$	$-1/2$	$1/2$	$3/2$	$\cdots$	$l-1/2$	$l+1/2$

(16.17)

Nous observons que nous obtenons les valeurs $M \in \{-l-1/2, -l+1/2, \cdots, l-1/2, l+1/2\}$ et que toutes ces valeurs sont réalisées deux fois, *à l'exception* des valeurs extrêmes $M = l+1/2$ et $M = l-1/2$ qui ne sont réalisées qu'une fois.

Nous savons (voir le cours n° 13, section 13.3) que les représentations irreductibles de $\mathfrak{su}_2$ ne sont pas dégénérées, c.-à.-d. que les sous-espaces propres (j, M) sont tous de dimension 1. Les dégénérescences obtenues peuvent donc s'expliquer par une décomposition en représentations irreductibles de la forme :

$$D^l \otimes D^{1/2} \cong D^{j_1} \oplus D^{j_2} \,, \tag{16.18}$$

soit la décomposition du produit tensoriel des représentations l et $1/2$ en somme directe de deux représentations irreductibles, avec j_1 et j_2 à determiner.

La remarque cruciale permettant de résoudre le problème est la suivante. Le vecteur $|l, m = l; 1/2, m_s = 1/2\rangle$ est *nécessairement* identifié (à une phase près) au vecteur de plus haut poids $|j = l+1/2, M = l+1/2\rangle$. En effet nous avons :

$$\begin{cases} \mathcal{J}_z\,|l, l; 1/2, 1/2\rangle = \big(\mathcal{L}_z \otimes \mathcal{I} + \mathcal{I} \otimes \mathcal{S}_z\big)\,|l, l; 1/2, 1/2\rangle = \hbar(l+1/2)\,|l, m = l; 1/2, 1/2\rangle \\ \mathcal{J}_+\,|l, l; 1/2, 1/2\rangle = \big(\mathcal{L}_+ \otimes \mathcal{I} + \mathcal{I} \otimes \mathcal{S}_+\big)\,|l, l; 1/2, 1/2\rangle = 0 \end{cases}$$

$$\Longrightarrow \quad \boxed{|l+1/2, l+1/2\rangle = |l, l; 1/2, 1/2\rangle} \tag{16.19}$$

Les autres vecteurs de base pour la représentation $D^{l+1/2}$ se contruisent alors de la manière habituelle avec les opérateurs d'échelle, c.-à.-d. en utilisant

$$\mathcal{J}_\pm |J, M\rangle = \hbar\sqrt{J(J+1) - M(M\pm 1)}\, |J, M\pm 1\rangle \ , \tag{16.20}$$

valable pour tout moment cinétique (c.-à.-d. $\vec{\mathcal{J}}$, $\vec{\mathcal{L}}$ ou $\vec{\mathcal{J}}$). Nous avons ainsi l'expression du vecteur $|J = l + 1/2, M = l - 1/2\rangle$ dans la base produit tensoriel :

$$\begin{aligned} |l+1/2, l-1/2\rangle &= \frac{1}{\hbar\sqrt{2l+1}}\mathcal{J}_- |l+1/2, l+1/2\rangle \\ &= \frac{1}{\hbar\sqrt{2l+1}}\left(\mathcal{L}_- \otimes \mathcal{I} + \mathcal{I} \otimes \mathcal{S}_-\right) |l, l; 1/2, 1/2\rangle \\ &= \sqrt{\frac{2l}{2l+1}}\, |l, l-1; 1/2, 1/2\rangle + \frac{1}{\sqrt{2l+1}}\, |l, l; 1/2, -1/2\rangle \end{aligned} \tag{16.21}$$

et ainsi de suite, jusqu'à atteindre par sauts entiers

$$|l+1/2, -l-1/2\rangle = \frac{1}{\sqrt{2l+1}}\mathcal{J}_- |l+1/2, -l+1/2\rangle = |l, -l; 1/2, -1/2\rangle \ , \tag{16.22}$$

la dernière égalité pouvant s'obtenir, outre par l'action répétée de $\mathcal{J}_-$, par un raisonnement analogue à celui de l'équation (16.50) en utilisant $\mathcal{J}_-$ au lieu de $\mathcal{J}_+$.

D'après la table (16.17), les sous-espaces propres $\mathcal{E}_M$ pour $\mathcal{J}_z$ sont de dimension 2 pour toutes les valeurs de M dans l'intervalle $-l+1/2 \leqslant M \leqslant l-1/2$. La construction précédente de la représentation $D^{l+1/2}$ nous donnant un vecteur pour chaque valeur de M dans cet intervalle, les vecteurs restant doivent s'interpréter comme vecteurs de base d'une représentation D^{j_2}.

Dans le sous-espace $\mathcal{E}_{M=l-1/2}$, un tel vecteur normé doit être orthogonal à $|l+1/2, l-1/2\rangle$ par orthogonalité des représentations irreductibles. Ce vecteur est unique à une phase près :

$$|\psi\rangle = \frac{1}{\sqrt{2l+1}}\, |l, l-1; 1/2, 1/2\rangle - \sqrt{\frac{2l}{l+1}}\, |l, l; 1/2, -1/2\rangle \tag{16.23}$$

et correspond au vecteur de plus haut poids $|l-1/2, l-1/2\rangle$ de la représentation $D^{l-1/2}$. En effet :

$$\mathcal{J}_+ |\psi\rangle = \left(\mathcal{L}_+ \otimes \mathcal{I} + \mathcal{I} \otimes \mathcal{S}_+\right) |\psi\rangle = \frac{\hbar\sqrt{2l}}{\sqrt{2l+1}}\, |l, l; 1/2, 1/2\rangle - \hbar\sqrt{\frac{2l}{2l+1}}\, |l, l; 1/2, 1/2\rangle = 0$$

$$\Longrightarrow \boxed{|l-1/2, l-1/2\rangle = \frac{1}{\sqrt{l+1}}\, |l, l-1; 1/2, 1/2\rangle - \sqrt{\frac{l}{l+1}}\, |l, l; 1/2, -1/2\rangle} \tag{16.24}$$

et tous les vecteurs $|l-1/2, M\rangle$ se contruisent par application répétée de $\mathcal{J}_-$ comme précédemment.

En conclusion, nous avons pu obtenir la décomposition en représentations irreductibles de $\mathfrak{su}_2$ suivante :

$$\boxed{D^l \otimes D^{1/2} \cong D^{l+1/2} \oplus D^{l-1/2}} \tag{16.25}$$

et nous pouvons vérifier que les dimensions des espaces vectoriels associés aux membres de gauche et de droite de cette relation coïncident :

$$(2l+1) \times 2 = \big(2(l-1/2)+1\big) + \big(2(l+1/2)+1\big)\,. \tag{16.26}$$

Naturellement la décomposition (16.25) ne s'applique pas au cas $l = 0$ (représentation triviale), et dans ce cas nous avons évidemment $D^0 \otimes D^{1/2} = D^{1/2}$.

Le changement de base associé à cette décomposition, depuis la base produit tensoriel $|l, m\rangle \otimes |1/2, m_s\rangle$ vers la base $|j, M\rangle$, est utile pour traiter de certaines interactions qui commutent avec le moment cinétique total mais pas avec $\vec{\mathcal{L}}$ et $\vec{\mathcal{S}}$. Par exemple le couplage spin-orbite que nous étudierons dans le cours n° 17 est une perturbation du hamiltonien de l'atome d'hydrogène proportionnelle à

$$2\,\vec{\mathcal{L}} \cdot \vec{\mathcal{S}} = \vec{\mathcal{J}}^2 - \vec{\mathcal{L}}^2 - \vec{\mathcal{S}}^2\,. \tag{16.27}$$

16.1.2 Composition de deux spins 1/2

Considérons un système composé de deux particules de spin 1/2, par exemple les noyaux d'une molécule de dihydrogène dans que nous évoquerons plus bas, en nous focalisant sur les degrés de liberté de spin uniquement.

Nous noterons $\vec{\mathcal{S}}_1$ (resp. $\vec{\mathcal{S}}_2$) les opérateurs de spin correspondant à la première (resp. à la seconde) particule. Ne prenant pas en compte provisoirement le postulat d'antisymétrisation pour des fermions identiques, nous pouvons décrire l'espace des états du système comme le produit tensoriel $D^{1/2} \otimes D^{1/2}$, dont une base est donnée par :

$$|m_1; m_2\rangle \overset{\text{déf.}}{=} |1/2, m_1\rangle \otimes |1/2, m_2\rangle \quad , \quad m_1, m_2 \in \{-1/2,\, 1/2\}\,. \tag{16.28}$$

Comme précédemment nous souhaitons décomposer cette représentation réductible de $\mathfrak{su}_2$ en représentations irréductibles. Nous introduisons le spin total

$$\vec{\mathcal{S}} \overset{\text{déf.}}{=} \vec{\mathcal{S}}_1 + \vec{\mathcal{S}}_2 \tag{16.29}$$

et nous chercherons une base de vecteurs propres communs de $(\mathcal{S}^2, \mathcal{S}_1^2, \mathcal{S}_2^2, \mathcal{S}_z)$. Nous noterons également $M = m_1 + m_2$ les valeurs propres de $\mathcal{S}_z/\hbar$. Le tableau analogue à (16.17) prend une forme nettement plus simple :

m_2 \ m_1	$-1/2$	$1/2$
$-1/2$	-1	(0)
$1/2$	(0)	1

(16.30)

et l'analyse s'effectue selon les mêmes principes que l'exemple précédent. Nous trouvons premièrement un vecteur de plus haut poids $|1, M = 1\rangle$ associé à une représentation D^1 de spin $s = 1$:

$$\left\{ \begin{array}{l} \mathcal{S}_z\,|1/2;1/2\rangle = \big(\mathcal{S}_{1,z}\otimes\mathcal{I} + \mathcal{I}\otimes\mathcal{S}_{2,z}\big)\,|1/2;1/2\rangle = \hbar(1/2+1/2)\,|1/2;1/2\rangle \\ \mathcal{S}_+\,|1/2;1/2\rangle = \big(\mathcal{S}_{1,+}\otimes\mathcal{I} + \mathcal{I}\otimes\mathcal{S}_{2,+}\big)\,|1/2;1/2\rangle = 0 \end{array} \right.$$
$$\Longrightarrow \boxed{|1,1\rangle = |1/2;1/2\rangle} \quad (16.31)$$

et nous pouvons construire les deux autres vecteurs de base de l'espace de la représentation D^1 aisément. Nous avons deuxièmement :

$$|1,0\rangle = \frac{1}{\hbar\sqrt{2}}\mathcal{S}_-\,|1,1\rangle = \frac{1}{\hbar\sqrt{2}}\big(\mathcal{S}_{1,-}\otimes\mathcal{I} + \mathcal{I}\otimes\mathcal{S}_2\big)\,|1/2;1,2\rangle = \frac{1}{\sqrt{2}}\big(\,|-1/2;1/2\rangle + |1/2,-1/2\rangle\,\big)\,, \quad (16.32)$$

et finalement :

$$|1,-1\rangle = \frac{1}{\hbar\sqrt{2}}\mathcal{S}_-\,|1,0\rangle = \frac{1}{\hbar\sqrt{2}}\big(\mathcal{S}_{1,-}\otimes\mathcal{I} + \mathcal{I}\otimes\mathcal{S}_2\big)\frac{|-1/2;1/2\rangle + |1/2,-1/2\rangle}{\sqrt{2}} = |-1/2;-1/2\rangle \quad (16.33)$$

La table (16.30) indique que le sous-espace propre de $\mathcal{S}_z$ de valeur propre $M = 0$, pour l'espace vectoriel associé à la représentation réductible $D^{1/2}\otimes D^{1/2}$, est de dimension 2. Une fois soustraits les trois vecteurs de base $|1, M\rangle$ de la représentation irreductible D^1, il reste un unique vecteur, vecteur de base de la représentation triviale D^0. Il s'obtient par orthogonalité des représentations irreductibles à partir de l'éqn. (16.32), à un choix de phase arbitraire près :

$$\langle 1,0|0,0\rangle = 0 \;\;\Longrightarrow\;\; |0,0\rangle = \frac{1}{\sqrt{2}}\big(\,|-1/2;1/2\rangle - |1/2,-1/2\rangle\,\big)\,. \quad (16.34)$$

Nous avons ainsi obtenu l'importante décomposition en représentation irreductibles

$$\boxed{D^{1/2}\otimes D^{1/2} = D^1 \oplus D^0}\,. \quad (16.35)$$

Les 3 états $|1, M\rangle$ sont usuellement appelés *états triplet* alors que l'état $|0, 0\rangle$ est appelé *état singulet*.

Notons le comportement différent des deux représentations sous la symétrie de permutation des deux spins, qui agit comme :

$$\mathcal{T}_{12}\;|1/2, m_1\rangle\otimes|1/2, m_2\rangle \;\mapsto\; |1/2, m_2\rangle\otimes|1/2, m_1\rangle\;. \quad (16.36)$$

On obtient immédiatement que :

— les états triplet sont symétriques par permutation, $\mathcal{T}_{12}\,|1, M\rangle = |1, M\rangle$;

— l'états singulet est antisymétrique par permutation, $\mathcal{T}_{12}\,|0, 0\rangle = -\,|0, 0\rangle$.

Application : ortho- et para-hydrogène

Considérons une molécule de dihydrogène, dans l'état électronique fondamental. Nous nous placerons dans l'approximation dite de *Born–Oppenheimer*, dans laquelle les électrons suivent adiabatiquement le mouvement des noyaux, ces derniers étant beaucoup plus massifs. Pour des énergies telles que les modes de vibration de la molécule ne sont pas excités nous modéliserons ce système par un *rotateur rigide*, c.-à.-d. que le hamiltonien relatif du système (voir la sous-section 15.2.1 du cours n° 15) est donné par

$$\mathcal{H}_r = \frac{\mathcal{L}^2}{2\mu r_0^2}\,, \tag{16.37}$$

où $\mu = m/2$ est la masse réduite et r_0 la distance moyenne entre les deux noyaux dans l'état fondamental. Les états propres de ce hamiltonien sont naturellement les états propres $|l, m\rangle$ du moment cinétique orbital, correspondant aux états de rotation de la molécule, et les énergies propres associées sont :

$$E_l = \frac{\hbar^2 l(l+1)}{m_e r_0^2}\,. \tag{16.38}$$

Si nous prenons en compte le spin des noyaux, l'espace des états de la molécule a pour base :

— les états triplet $|l, m\rangle \otimes |1, M\rangle$;

— les états singulet $|l, m\rangle \otimes |0, 0\rangle$.

Les protons étant des particules de spin demi-entier, ils suivent la statistique de Fermi–Dirac et les vecteurs d'état physiques doivent être antisymétriques par permutation des noyaux. Dans ce système, une permutation des noyaux revient à une opération de parité (ou d'inversion spatiale) par rapport au centre de masse de la particule, qui échange effectivement leur deux positions. Nous avons vu dans le cours n° 14, voir l'éqn. (14.73), que l'action de la parité sur un état propre du moment cinétique est

$$\mathfrak{P}\,|l, m\rangle = (-1)^l\,|l, m\rangle \tag{16.39}$$

Le postulat d'antisymétrisation pour les fermions concerne le vecteur d'état « complet » du système, c.-à.-d.

$$|l, m; s, M\rangle \overset{\text{déf.}}{=} |l, m\rangle \otimes |s, M\rangle\,. \tag{16.40}$$

Nous obtenons donc le résultat suivant :

— les états de rotation avec l impair sont dans un état de spin nucléaire triplet ($s = 1$) ;

— les états de rotation avec l pair sont dans un état de spin nucléaire singulet ($s = 0$).

Cela implique en particulier que l'état fondamental du système ($l = 0$) est un singulet ; il s'agit du para-hydrogène. Lorsque les spins nucléaires sont dans l'état triplet (ortho-hydrogène) l'état d'énergie la plus basse correspond à $l = 1$. Il s'agit d'un état métastable, car la relaxation des spins nucléaires est très longue.

16.2 Théorie générale : coefficients de Clebsch–Gordan

Après avoir détaillé l'addition de moments cinétiques sur deux exemples, l'objectif de cette section est de donner la théorie générale la décomposition en représentations irreductibles de $\mathfrak{su}_2$ du produit tensoriel de représentations, c.-à.-d. exprimer $D^{j_1} \otimes D^{j_2}$ comme une somme directe de représentations irreductibles. Notons que chacune des représentations D^{j_1} et D^{j_2} peut être associée physiquement soit à un spin, soit à un moment cinétique orbital, soit à une combinaison des deux types de moments cinétiques. En toute généralité nous avons

$$D^{j_1} \otimes D^{j_2} \cong \bigoplus_{2J=0}^{\infty} \rho_J D^J \ , \quad \rho_J \in \mathbb{N} \tag{16.41}$$

où ρ_J désigne la multiplicité (éventuellement nulle) de la représentation D^J dans la décomposition en représentations irreductibles. Comme nous le verrons plus bas, chaque représentation n'apparaît au plus qu'une fois, c.-à.-d. que $\rho_J \in \{0, 1\}$.

Comme dans les exemples précédents nous introduisons l'opérateur de moment cinétique total,

$$\boxed{\vec{\mathcal{J}} \overset{\text{déf.}}{=} \vec{\mathcal{J}}_1 + \vec{\mathcal{J}}_2} \tag{16.42}$$

et nous cherchons dans toute représentation D^J apparaissant dans la décomposition (16.41) une base de vecteurs propres communs à $\mathcal{J}_1^2$, $\mathcal{J}_2^2$, $\mathcal{J}^2$ et $\mathcal{J}_z$, que nous noterons $|J, M\rangle$ avec

$$\frac{\mathcal{J}_1^2}{\hbar^2} |J, M\rangle = j_1(j_1 + 1) |J, M\rangle \tag{16.43a}$$

$$\frac{\mathcal{J}_2^2}{\hbar^2} |J, M\rangle = j_2(j_2 + 1) |J, M\rangle \tag{16.43b}$$

$$\frac{\mathcal{J}^2}{\hbar^2} |J, M\rangle = J(J + 1) |J, M\rangle \tag{16.43c}$$

$$\frac{\mathcal{J}_z}{\hbar} |J, M\rangle = M |J, M\rangle \ . \tag{16.43d}$$

Nous remarquons, comme nous l'avions déjà fait, que l'opérateur de Casimir $\mathcal{J}^2$ ne commute pas avec les composantes $\mathcal{J}_{1,i}$ et $\mathcal{J}_{2,i}$ des moments cinétiques $\vec{\mathcal{J}}_1$ et $\vec{\mathcal{J}}_2$. Remarquons également que :

$$\mathcal{J}_z = \mathcal{J}_{1,z} \otimes \mathfrak{I} + \mathfrak{I} \otimes \mathcal{J}_{2,z} \quad \Longrightarrow \quad \boxed{M = m_1 + m_2} \tag{16.44}$$

Pour décrire l'espace hermitien sur lequel agit la représentation réductible $D^{j_1} \otimes D^{j_2}$ de $\mathfrak{su}_2$, qui est de dimension

$$d = (2j_1 + 1)(2j_2 + 2) \, , \tag{16.45}$$

nous disposons donc de deux bases orthonormées distinctes :

— la base associée au produit tensoriel de représentations $D^{j_1} \otimes D^{j_2}$:

$$\Big\{ |j_1, m_1; j_2, m_2\rangle \overset{\text{déf.}}{=} |j_1, m_1\rangle \otimes |j_2, m_2\rangle \ , \quad m_1 = -j_1, \dots, j_1 \ , \quad m_2 = -j_2, \dots, j_2 \Big\} . \tag{16.46}$$

— pour toute valeur de J donnant une multiplicité non-nulle dans la décomposition (16.41), les vecteurs $|J, M\rangle$ forment une base orthonormée de l'espace hermitien de dimension $2J+1$ associé à la représentation D^J. Nous avons donc au total pour la somme directe de représentations la base orthonormée :

$$\left\{ |J, M\rangle \ , \quad M = -J, \dots, J \ , \quad J \in \tfrac{1}{2}\mathbb{N} \ \text{ t.q. } \ \rho_J = 1 \right\} \tag{16.47}$$

Définition 1 (coefficients de Clebsch–Gordan). *Soit* $D^{j_1} \otimes D^{j_2}$ *un produit tensoriel de représentations irreductibles de* $\mathfrak{su}_2$ *et* $\sum_J \rho_J D^J$ *sa décomposition en représentations irreductibles. Les coefficients de Clebsch–Gordan, définis par*

$$\boxed{C^{J\,M}_{j_1\,m_1\,,\,j_2 m_2} \overset{\text{déf.}}{=} \langle j_1, m_1; j_2, m_2 | J, M\rangle} \tag{16.48}$$

correspondent aux éléments de la matrice de passage d'une base orthonormée de $D^{j_1} \otimes D^{j_2}$ *vers une base orthonormée de* $\sum_J \rho_J D^J$ *:*

$$|j_1, j_2; J, M\rangle = \sum_{m_1, m_2} |j_1, m_1; j_2, m_2\rangle \, \langle j_1, m_1; j_2, m_2 | J, M\rangle = \sum_{m_1, m_2} C^{J\,M}_{j_1\,m_1\,,\,j_2 m_2} \, |j_1, m_1; j_2, m_2\rangle \tag{16.49}$$

Les deux bases étant orthonormées, la matrice de passage est unitaire (cf. cours n° 1). Afin de déterminer les coefficients de Clebsch–Gordan, nous pouvons tout d'abord obtenir des *règles de sélection* indiquant quels éléments de matrice sont obligatoirement nuls.

Valeurs de J de multiplicité non nulle

Les valeurs de J apparaissant dans la décomposition (16.41) s'obtiennent de la manière suivante. Comme dans les deux exemples étudiés au début de ce cours, nous considérons l'unique vecteur avec $M = j_1 + j_2$, c.-à.-d. le vecteur $|j_1, j_1; j_2, j_2\rangle$, que nous pouvons identifier, à un choix de phase près, avec le vecteur de plus haut poids $|j_1 + j_2, j_1 + j_2\rangle$ de la représentation de spin $J = j_1 + j_2$. Nous avons en effet :

$$\begin{cases} \mathcal{J}_z \, |j_1, j_1; j_2, j_2\rangle = \left(\mathcal{J}_{1,z} \otimes \mathcal{I} + \mathcal{I} \otimes \mathcal{J}_{2,z}\right) |j_1, j_1; j_2, j_2\rangle = \hbar(j_1 + j_2) \, |j_1, j_1; j_2, j_2\rangle \\ \mathcal{J}_+ \, |l, l; 1/2, 1/2\rangle = \left(\mathcal{J}_{1,+} \otimes \mathcal{I} + \mathcal{I} \otimes \mathcal{J}_{2,+}\right) |j_1, j_1; j_2, j_2\rangle = 0 \end{cases}$$
$$\Longrightarrow \ |j_1 + j_2, j_1 + j_2\rangle = |j_1, j_1; j_2, j_2\rangle \tag{16.50}$$

Une fois ce vecteur de plus haut poids identifié, nous en déduisons l'expression de tous les $(2J+1)$ vecteurs de la représentation D^J en termes de la base produit tensoriel $|j_1, m_1; j_2, m_2\rangle$

par applications successives de $\mathcal{J}_-$, en commençant par :

$$\begin{aligned}|j_1+j_2, j_1+j_2-1\rangle &= \frac{1}{\hbar\sqrt{2(j_1+j_2)}}\mathcal{J}_-\,|j_1+j_2, j_1+j_2\rangle \\ &= \frac{1}{\hbar\sqrt{2(j_1+j_2)}}\left(\mathcal{J}_{1,-}\otimes\mathfrak{I}+\mathfrak{I}\otimes\mathcal{J}_{2,-}\right)|j_1,j_1;j_2,j_2\rangle\,\big) \\ &= \sqrt{\frac{j_1}{j_1+j_2}}\,|j_1,j_1-1;j_2,j_2\rangle+\sqrt{\frac{j_2}{j_1+j_2}}\,|j_1,j_1;j_2,j_2-1\rangle \qquad (16.51)\end{aligned}$$

Considérons ensuite les deux vecteurs avec $M = j_1+j_2-1$ de la base produit tensoriel, c.-à.-d. $|j_1,j_1-1;j_2,j_2\rangle$ et $|j_1,j_1;j_2,j_2-1\rangle$. D'après le paragraphe précédent, une combinaison linéaire de ces deux vecteurs correspond au vecteur $|j_1+j_2,j_1+j_2-1\rangle$, voir éqn. (16.51). Un vecteur normé orthogonal à $|j_1+j_2,j_1+j_2-1\rangle$ est identifié, à une phase près, au vecteur de de plus haut poids $|j_1+j_2-1,j_1+j_2-1\rangle$ de la représentation $D^{j_1+j_2-1}$. Posons donc, avec un choix de phase approprié

$$|j_1+j_2-1,j_1+j_2-1\rangle = \sqrt{\frac{j_2}{j_1+j_2}}\,|j_1,j_1-1;j_2,j_2\rangle-\sqrt{\frac{j_1}{j_1+j_2}}\,|j_1,j_1;j_2,j_2-1\rangle \qquad (16.52)$$

En effet nous avons premièrement

$$\begin{aligned}&\mathcal{J}_+\,|j_1+j_2-1,j_1+j_2-1\rangle \\ &\quad=\left(\mathcal{J}_{1,+}\otimes\mathfrak{I}+\mathfrak{I}\otimes\mathcal{J}_{2,+}\right)\left(\sqrt{\frac{j_2}{j_1+j_2}}\,|j_1,j_1-1;j_2,j_2\rangle-\sqrt{\frac{j_2}{j_1+j_2}}\,|j_1,j_1;j_2,j_2-1\rangle\right) \\ &\qquad=\hbar\sqrt{2j_1}\sqrt{\frac{j_2}{j_1+j_2}}\,|j_1,j_1;j_2,j_2\rangle-\hbar\sqrt{2j_2}\sqrt{\frac{j_1}{j_1+j_2}}\,|j_1,j_1;j_2,j_2\rangle=0 \qquad (16.53)\end{aligned}$$

et deuxièmement :

$$\begin{aligned}&\mathcal{J}_z\,|j_1+j_2-1,j_1+j_2-1\rangle \\ &\quad=\left(\mathcal{J}_{1,z}\otimes\mathfrak{I}+\mathfrak{I}\otimes\mathcal{J}_{2,z}\right)\left(\sqrt{\frac{j_2}{j_1+j_2}}\,|j_1,j_1-1;j_2,j_2\rangle-\sqrt{\frac{j_1}{j_1+j_2}}\,|j_1,j_1;j_2,j_2-1\rangle\right) \\ &\qquad=\hbar(j_1-1)\sqrt{\frac{j_2}{j_1+j_2}}\,|j_1,j_1-1;j_2,j_2\rangle-\hbar j_1\sqrt{\frac{j_1}{j_1+j_2}}\,|j_1,j_1;j_2,j_2-1\rangle \\ &\qquad+\hbar j_2\sqrt{\frac{j_2}{j_1+j_2}}\,|j_1,j_1-1;j_2,j_2\rangle-\hbar(j_2-1)\sqrt{\frac{j_1}{j_1+j_2}}\,|j_1,j_1;j_2,j_2-1\rangle \\ &\qquad=\hbar(j_1+j_2-1)\,|j_1+j_2-1,j_1+j_2-1\rangle \qquad (16.54)\end{aligned}$$

Une fois construites les bases des représentations irreductibles $D^{j_1+j_2}$ et $D^{j_1+j_2-1}$, on considère le vecteur normé pour la valeur $M = j_1+j_2-2$ qui est orthogonal aux vecteurs de base

des représentations $D^{j_1+j_2}$ et $D^{j_1+j_2-1}$ (qui est unique à une phase près) ; on l'identifie avec le vecteur de plus haut poids de la représentation $D^{j_1+j_2-2}$, et ainsi de suite.

Le processus s'arrête lorsque nous avons épuisé les vecteurs disponibles dans l'espace vectoriel de dimension $(2j_1+1)(2j_2+1)$. Il est facile de voir que ce raisonnement conduit à la décomposition en représentations irreductibles de $\mathfrak{su}_2$:

$$\boxed{D^{j_1} \otimes D^{j_2} \cong \bigoplus_{\substack{J \in j_1+j_2-\mathbb{N} \\ |j_1-j_2| \leqslant J \leqslant j_1+j_2}} D^J} \tag{16.55}$$

En effet, la somme des dimensions des représentations irreductibles apparaissant dans le membre de droite est, par exemple pour j_1+j_2 entier

$$\begin{aligned}\sum_{J=|j_1-j_2|}^{j_1+j_2} (2J+1) &= (j_1+j_2)(j_1+j_2+1) - (|j_1-j_2|-1)|j_1-j_2| + \big(j_1+j_2-|j_1-j_2|\big) \\ &= 4j_1j_2 + 2(j_1+j_2) = (2j_1+1)(2j_2+1)\,, \end{aligned} \tag{16.56}$$

et nous obtenons le même résultat pour j_1+j_2 demi-entier. Nous en déduisons donc que

$$\boxed{C^{J\,M}_{j_1\,m_1\,,\,j_2 m_2} = 0 \quad \text{si} \quad J \notin \big[|j_1-j_2|,\, j_1+j_2\big]} \tag{16.57}$$

Ensuite, nous obtenons très simplement un règle de sélection reliant les valeurs propres de $\mathcal{J}_z$, $\mathcal{J}_{1,z}$ et $\mathcal{J}_{2,z}$, en utilisant l'éqn. (16.44) :

$$\begin{aligned} 0 &= \langle j_1, m_1; j_2, m_2| \left(\mathcal{J}_z - \mathcal{J}_{1,z} \otimes \mathcal{I} - \mathcal{I} \otimes \mathcal{J}_{2,z}\right) |J, M\rangle \\ &= (M - m_1 - m_2)\, \langle j_1, m_1; j_2, m_2 | J, M\rangle = (M - m_1 - m_2) C^{J\,M}_{j_1\,m_1\,,\,j_2 m_2}\,, \end{aligned} \tag{16.58}$$

où nous avons appliqué $\mathcal{J}_z$ à droite et les autres opérateurs à gauche, soit

$$\boxed{C^{J\,M}_{j_1\,m_1\,,\,j_2 m_2} = 0 \quad \text{si} \quad M \neq m_1 + m_2} \tag{16.59}$$

En conclusion, les coefficients de Clebsch–Gordan $C^{J\,M}_{j_1\,m_1\,,\,j_2 m_2}$ sont a priori non-nuls si les conditions suivantes sont satisfaites :

$$m_1 \in \{-j_1, -j_1+1, \cdots, j_1-1, j_1\} \tag{16.60a}$$
$$m_2 \in \{-j_2, -j_2+1, \cdots, j_2-1, j_2\} \tag{16.60b}$$
$$J \in \big(j_1+j_2-\mathbb{N}\big) \tag{16.60c}$$
$$|j_1-j_2| \leqslant J \leqslant j_1+j_2 \tag{16.60d}$$
$$M = m_1 + m_2 \tag{16.60e}$$
$$M \in \{-J, -J+1, \cdots, J-1, J\} \tag{16.60f}$$

16.2.1 Calcul des coefficients de Clebsch–Gordan

Les coefficients de Clebsch–Gordan non-nuls peuvent se calculer à partir de relations générales que nous derivons dans cette sous-section, relations de récursion d'une part et relations d'orthonormalité d'autre part.

Relations de récursion

Pour obtenir les coefficients de Clebsch–Gordan dans le domaine précédemment défini par les relations (16.60), nous pouvons dériver des relations de récursion reliant un coefficient aux coefficients voisins. Nous écrivons premièrement (action à gauche) :

$$\begin{aligned}
&\frac{1}{\hbar}\langle j_1, m_1; j_2 m_2|\left(\mathcal{J}_{1,\pm}\otimes\mathfrak{I}+\mathfrak{I}\otimes\mathcal{J}_{2,\pm}\right)|J,M\rangle \\
&\qquad=\Big(\sqrt{j_1(j_1+1)-m_1(m_1\mp1)}\,\langle j_1, m_1\mp1; j_2 m_2| \\
&\qquad+\sqrt{j_2(j_2+1)-m_2(m_2\mp1)}\,\langle j_1, m_1; j_2 m_2\mp2|\Big)|J,M\rangle\,\Big) \\
&=\sqrt{j_1(j_1+1)-m_1(m_1\mp1)}C^{J\,M}_{j_1\,m_1\mp1\,,\,j_2 m_2}+\sqrt{j_2(j_2+1)-m_2(m_2\mp1)}C^{J\,M}_{j_1\,m_1\,,\,j_2 m_2\mp1}
\end{aligned} \tag{16.61}$$

Deuxièmement, étant donné que $\mathcal{J}_\pm=\mathcal{J}_{1,\pm}\otimes\mathfrak{I}+\mathfrak{I}\otimes\mathcal{J}_{2,\pm}$ nous avons (action à droite) :

$$\begin{aligned}
&\frac{1}{\hbar}\langle j_1, m_1; j_2 m_2|\left(\mathcal{J}_{1,\pm}\otimes\mathfrak{I}+\mathfrak{I}\otimes\mathcal{J}_{2,\pm}\right)|J,M\rangle=\frac{1}{\hbar}\langle j_1, m_1; j_2 m_2|\,\mathcal{J}_\pm\,|J,M\rangle \\
&=\sqrt{J(J+1)-M(M\pm1)}\,\langle j_1, m_1; j_2 m_2|\,|J,M\pm1\rangle=\sqrt{J(J+1)-M(M\pm1)}C^{J\,M\pm1}_{j_1\,m_1\,,\,j_2 m_2}
\end{aligned} \tag{16.62}$$

Nous obtenons donc les relations de récursion suivantes :

$$\boxed{\begin{aligned}
&\sqrt{J(J+1)-M(M\pm1)}C^{J\,M\pm1}_{j_1\,m_1\,,\,j_2 m_2} \\
&\quad=\sqrt{j_1(j_1+1)-m_1(m_1\mp1)}C^{J\,M}_{j_1\,m_1\mp1\,,\,j_2 m_2}+\sqrt{j_2(j_2+1)-m_2(m_2\mp1)}C^{J\,M}_{j_1\,m_1\,,\,j_2 m_2\mp1}
\end{aligned}} \tag{16.63}$$

Relations d'orthonormalité.

L'autre type de contrainte importante sur les coefficients de Clebsch–Gordan provient de leur interprétation comme éléments de matrice de la matrice unitaire associée au changement de base de la base produit tensoriel $|j_1, m_1; j_2, m_2\rangle$ vers la base $|J, M\rangle$ associée à la décomposition en représentations irreductibles de $\mathfrak{su}_2$.

Il est possible de choisir les conventions de phase associées aux choix de vecteurs de base des représentations irreductibles $|J, M\rangle$ pour que cette matrice soit réelle, comme nous

l'avions fait dans les deux exemples traités dans la section 16.1. Une matrice unitaire réelle étant une matrice orthogonale, nous avons, en sommant sur les vecteurs de base de toutes les représentations irreductibles apparaissant dans la décomposition (16.55) :

$$\sum_{J,M} \langle j_1, m_1; j_2, m_2 | J, M\rangle \langle J, M | j_1, m_1'; j_2, m_2'\rangle = \delta_{m_1,m_1'}\delta_{m_2,m_2'}$$

$$\Longrightarrow \boxed{\sum_{J,M} C^{J\,M}_{j_1\,m_1\,,\,j_2 m_2} C^{J\,M}_{j_1\,m_1'\,,\,j_2 m_2'} = \delta_{m_1,m_1'}\delta_{m_2,m_2'}} \quad (16.64)$$

Nous trouvons de même, en sommant sur m_1 et m_2 dans la base produit tensoriel :

$$\sum_{m_1,m_2} \langle J, M | j_1, m_1; j_2, m_2\rangle \langle j_1, m_1; j_2, m_2 | J', M'\rangle = \delta_{J,J'}\delta_{M,M'}$$

$$\Longrightarrow \boxed{\sum_{m_1,m_2} C^{J\,M}_{j_1\,m_1\,,\,j_2 m_2} C^{J'\,M'}_{j_1\,m_1\,,\,j_2 m_2} = \delta_{J,J'}\delta_{M,M'}} \quad (16.65)$$

Les relations d'orthonormalité (16.65,16.65) et les relations de récursion (16.63) déterminent complètement les coefficients de Clebsch–Gordan, à des conventions de signe près.

En pratique : les coefficients de Clebsch–Gordan, qui sont objets mathématiques associés à la théorie des représentations de l'algèbre $\mathfrak{su}_2$, sont indépendants du système physique considéré et ont été calculés une fois pour toutes depuis longtemps. Pour résoudre un problème physique donné il suffit de consulter des tables de coefficients ou de les calculer par ordinateur (voir par exemple ce lien pour un code Python).

Composition de deux spins 1/2

Vérifions que les relations entre les coefficients de Clebsch–Gordan sont compatibles les résultats de la section 16.1, en étudiant par exemple la décomposition

$$D^{1/2} \otimes D^{1/2} \cong D^1 \oplus D^0\,. \quad (16.66)$$

Les coefficients de Clebsch–Gordan non-nuls que nous avions trouvés sont :

$$C^{0\,0}_{1/2\,-1/2\,,\,1/2\,1/2} = -C^{0\,0}_{1/2\,1/2\,,\,1/2\,-1/2} = \frac{1}{\sqrt{2}} \quad (16.67a)$$

$$C^{1\,0}_{1/2\,-1/2\,,\,1/2\,1/2} = C^{0\,0}_{1/2\,1/2\,,\,1/2\,-1/2} = \frac{1}{\sqrt{2}} \quad (16.67b)$$

$$C^{1\,\pm1}_{1/2\,\pm1/2\,,\,1/2\,\pm1/2} = 1 \quad (16.67c)$$

Vérifions premièrement les relations de récursion. Nous avons par exemple

$$\sqrt{2} C^{1\,0}_{1/2\,\mp1/2\,,\,1/2\,\pm1/2} = C^{1\,1}_{1/2\,1/2\,,\,1/2\,1/2} \quad (16.68)$$

ou bien

$$\sqrt{2}C^{1\,-1}_{1/2\,-1/2,\,1/2\,-1/2} = C^{1\,0}_{1/2\,1/2,\,1/2\,-1/2} + C^{1\,0}_{1/2\,-1/2,\,1/2\,1/2} \tag{16.69}$$

qui sont satisfaites explicitement par les coefficients (16.67). Les relations d'orthogonalité donnent par exemple (somme sur J et M) :

$$\begin{aligned}
&\left(C^{0\,0}_{1/2\,-1/2,\,1/2\,1/2}\right)^2 + \left(C^{1\,0}_{1/2\,-1/2,\,1/2\,1/2}\right)^2 = \left(C^{0\,0}_{1/2\,1/2,\,1/2\,-1/2}\right)^2 + \left(C^{1\,0}_{1/2\,1/2,\,1/2\,-1/2}\right)^2 = 1\\
&C^{0\,0}_{1/2\,-1/2,\,1/2\,1/2}\, C^{0\,0}_{1/2\,1/2,\,1/2\,-1/2} + C^{1\,0}_{1/2\,-1/2,\,1/2\,1/2} C^{1\,0}_{1/2\,1/2,\,1/2\,-1/2} = 0
\end{aligned} \tag{16.70}$$

Nous laissons en exercice la confirmation expliicte que les coefficients (16.67) vérifient les différentes relations ci-dessus, ainsi que les autres relations analogues que nous n'avons pas écrites.

À retenir

— la notion de composition de moment cinétique, ou décomposition d'un produit tensoriel de représentations en représentations irreductibles ;

— la transformation d'un spineur de fonctions d'onde par rotation ;

— les états singulet et triplet, l'application au dihydrogène ;

— la décomposition du produit tensoriel en représentations irreductibles, en particulier les valeurs $J = |j_1 - j_2|$, $J = |j_1 - j_2| + 1, \cdots, j_1 + j_2$ apparaissant dans la décomposition avec une multiplicité non-nulle ;

— la définition des coefficients de Clebsch–Gordan et leurs propriétés principales.

Appendices

16.A Coefficients de Clebsch–Gordan et matrices de rotation

Les coefficients de Clebsch–Gordan permettent d'obtenir la relation entre les matrices de la représentation réductible $D(g) \cong D^{j_1}(g) \otimes D^{j_2}(g)$ du groupe $SU(2)$ dans la base produit tensoriel avec les matrices associées aux différentes représentations irreductibles de la décomposition (16.55).

Considérons premièrement l'expression des éléments de la matrice D de Wigner dans la base produit tensoriel. Nous avons par définition du produit tensoriel :

$$\begin{aligned}
&\langle j_1, m_1; j_2, m_2 |\, D(g)\, | j_1, m_1'; j_2, m_2' \rangle\\
&\qquad = \langle j_1, m_1; j_2, m_2 |\, D^{j_1}(g) \otimes D^{j_2}(g)\, | j_1, m_1'; j_2, m_2' \rangle\\
&\qquad = \langle j_1, m_1 |\, D^{j_1}(g)\, | j_1, m_1' \rangle \, \langle j_2, m_2 |\, D^{j_2}(g)\, | j_2, m_2' \rangle\\
&\qquad = D^{j_1}_{m_1 m_1'}(g) D^{j_2}_{m_2 m_2'}(g)
\end{aligned} \tag{16.71}$$

Deuxièmement, nous pouvons utiliser la base $|J, M\rangle$ associée à la décomposition en représentations irreductibles. Dans cette base les matrices de la représentation de $SU(2)$ prennent une forme diagonale par blocs :

$$D(g) = \mathrm{diag}\left(D^{j_1+j_2}(g), D^{j_1+j_2-1}(g), \ldots, D^{|j_1-j_2|+1}(g), D^{|j_1-j_2|}(g)\right). \tag{16.72}$$

Nous avons donc en insérant deux relations de fermeture :

$$\begin{aligned}
&\langle j_1, m_1; j_2, m_2| D(g) |j_1, m_1'; j_2, m_2'\rangle \\
&\quad = \sum_{\substack{J\in j_1+j_2-\mathbb{N}\\ J\geqslant|j_1-j_2|}} \sum_{\substack{M,M'\in J-\mathbb{N}\\ M\geqslant -J}} \langle j_1, m_1; j_2, m_2|J, M\rangle \langle J, M| D^J(g) |J, M'\rangle \langle J, M'|j_1, m_1'; j_2, m_2'\rangle \\
&\qquad\qquad = \sum_{\substack{J\in j_1+j_2-\mathbb{N}\\ J\geqslant|j_1-j_2|}} \sum_{\substack{M,M'\in J-\mathbb{N}\\ M\geqslant -J}} C^{J\,M}_{j_1\,m_1\,,\,j_2 m_2} C^{J\,M'}_{j_1\,m_1'\,,\,j_2 m_2'} \langle J, M| D^J(g) |J, M'\rangle
\end{aligned} \tag{16.73}$$

où nous avons utilisé à la dernière étape du calcul la réalité des coefficients de Clebsch–Gordan.

Nous obtenons finalement en comparant les expressions (16.71) et (16.71) la relation annoncée entre éléments de matrice de la représentation produit tensoriel et de la somme directe de représentations irreductibles :

$$\boxed{D^{j_1}_{m_1 m_1'}(g) D^{j_2}_{m_2 m_2'}(g) = \sum_{\substack{J\in j_1+j_2-\mathbb{N}\\ J\geqslant|j_1-j_2|}} \sum_{\substack{M,M'\in J-\mathbb{N}\\ M\geqslant -J}} C^{J\,M}_{j_1\,m_1\,,\,j_2 m_2} C^{J\,M'}_{j_1\,m_1'\,,\,j_2 m_2'} D^J_{MM'}(g)} \tag{16.74}$$

connue sous le nom de *série de Clebsch–Gordan*.

Une application immédiate de ce résultat est la décomposition d'un produit d'harmoniques sphériques en somme d'harmoniques sphériques. En évaluant la relation précédente pour $m_1' = m_2' = 0$ (impliquant $M' = 0$) :

$$D^{j_1}_{m_1 0}(g) D^{j_2}_{m_2 0}(g) = \sum_{\substack{J\in j_1+j_2-\mathbb{N}\\ J\geqslant|j_1-j_2|}} \sum_{\substack{M\in J-\mathbb{N}\\ M\geqslant -J}} C^{J\,M}_{j_1\,m_1\,,\,j_2 m_2} C^{J\,0}_{j_1\,0\,,\,j_2 0} D^J_{M0}(g) \tag{16.75}$$

et en utilisant la relation (14.62) entre harmoniques sphériques et éléments de matrice des représentations de $SU(2)$ nous obtenons après conjugaison complexe (sans effet sur les coefficients de Clebsch–Gordan, réels) :

$$\begin{aligned}
&\frac{4\pi}{\sqrt{(2l_1+1)(2l_2+1)}} Y^{m_1}_{l_1}(\theta,\phi) Y^{m_2}_{l_2}(\theta,\phi) \\
&\qquad = \sum_{L=|l_1-l_2|}^{l_1+l_2} \sum_{M=-L}^{L} C^{L\,M}_{l_1\,m_1\,,\,l_2 m_2} C^{L\,0}_{l_1\,0\,,\,l_2 0} \sqrt{\frac{4\pi}{2L+1}} Y^L_M(\theta,\phi).
\end{aligned} \tag{16.76}$$

Cours n°17

Structure fine et effet Zeeman

La résolution exacte de l'équation de Schrödinger pour un potentiel coulombien décrite dans le cours n° 16 est un tour de force mathématique et la description obtenue des atomes hydrogénoïdes constitue le fondement de la physique atomique et de la physique moléculaire en mécanique quantique.

Cependant, si les valeurs propres et vecteurs propres associés au états liés du hamiltonien (15.42) ont été obtenus exactement, le hamiltonien lui-même ne fournit pas une description exacte du problème. Il existe trois sources notables de corrections aux résultats obtenus précédemmment :

- la description relativiste de l'électron dans le potentiel central du proton, via l'*équation de Dirac*, donne un ensemble de corrections regroupées sous le nom de *structure fine* du spectre de l'atome d'hydrogène, faisant en particulier intervenir le spin de l'électron ;
- la prise en compte du spin du proton, en particulier de l'interaction entre les moments magnétiques de spin du proton est de l'électron, donne la *structure hyperfine* du spectre de l'atome d'hydrogène ;
- les *corrections radiatives*, obtenues en théorie quantique des champs en prenant en compte le caractère quantique du champ éléctromagnétique. [1]

Dans le cadre de ce cours, nous nous restreindrons à l'étude de la structure fine, en particulier du couplage spin-orbite en tirant profit de la composition de moments cinétiques développée dans le cours n° 16. Nous étudierons ensuite le déplacement des niveaux sous l'effet d'un champ magnétique, l'*effet Zeeman*, qui a joué un rôle historique majeur dans la découverte du spin de l'électron (voir l'introduction du cours n° 13).

17.1 Structure fine de l'atome d'hydrogène

La description de l'atome d'hydrogène que nous avons donnée dans le cours n° 17 e st fondée, comme le reste de ce cours, sur une formulation non-relativiste de la mécanique

1. Plus généralement il existe des corrections quantiques dues aux champs associés aux interactions fondamentales électrofaible et forte.

quantique. Il est judicieux de se demander a posteriori si cette approximation est valide.

Pour évaluer l'ordre de grandeur des corrections relativistes aux solutions obtenues, calculons tout d'abord, dans l'état fondamental de l'atome d'hydrogène (voir éqn. (15.77)) :

$$\begin{aligned}\langle\vec{\mathcal{P}}^2\rangle_{100} &= \int r^2 dr\, \sin\theta d\theta d\phi\, \overline{\Psi_{100}(r)}\big(-\hbar^2\Delta\big)\Psi_{100}(r)\\ &= -4\pi\frac{\hbar^2}{\pi a_0^3}\int_0^\infty r^2 dr\, e^{-r/a_0}\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}e^{-r/a_0}\right) = \frac{4\hbar^2}{a_0^4}\int_0^\infty dr\, e^{-2r/a_0}\big(2r - r^2/a_0\big) = \frac{\hbar^2}{a_0^2}\end{aligned} \tag{17.1}$$

D'où nous déduisons la valeur moyenne du paramètre relativiste ($\vec{\mathcal{V}}$ étant l'opérateur associé à la vitesse de l'électron) :

$$\frac{\langle\vec{\mathcal{V}}^2\rangle_{100}}{c^2} = \frac{\langle\vec{\mathcal{P}}^2\rangle_{100}}{m_e^2c^2} = \frac{\hbar^2}{m_e^2a_0^2c^2} = \frac{q^4}{(4\pi\epsilon_0)^2\hbar^2c^2} = \alpha^2\,, \tag{17.2}$$

soit le carré de la *constante de structure fine*

$$\alpha = \frac{q^2}{4\pi\epsilon_0\hbar c} \simeq \frac{1}{137} \ll 1\,. \tag{17.3}$$

En conclusion, la description non-relativiste du système est appropriée mais il est possible de prendre en compte, en théorie des perturbations, les corrections relativistes qui s'exprimeront comme un développement en puissances de α^2. Ce développement peut s'obtenir systématiquement à partir de l'équation de Dirac que nous n'arborderons pas ici. [2] Nous nous contenterons de justifier, par des moyens heuristiques, les différents termes à l'ordre le plus bas en α^2 non-trivial de ce développement.

17.1.1 Correction au terme cinétique et terme de Darwin

Le premier type de correction, de nature cinématique, provient du développement de l'énergie d'une particule libre relativiste en puissances de $\frac{p}{m_ec}$. Nous avons :

$$E = c\sqrt{\vec{p}^{\,2} + m_e^2c^2} = m_ec^2 + \frac{\vec{p}^{\,2}}{2m_e} - \frac{\vec{p}^{\,4}}{8m_e^3c^2} + \mathcal{O}\left[\left(\frac{p}{m_ec}\right)^6\right]. \tag{17.4}$$

Laissant de côté l'énergie au repos de la particule (terme constant) nous obtenons une première correction au hamiltonien (15.42) de l'atome d'hydrogène :

$$\boxed{\mathcal{H}_1^c \overset{\text{déf.}}{=} -\frac{\vec{\mathcal{P}}^4}{8m_e^3c^2}} \tag{17.5}$$

2. Plus précisément, ce calcul s'effectue dans l'approximation où le proton est infiniment lourd, en considérant un électron relativiste dans un champ coulombien, au lieu de considérer le mouvement relatif des deux particules dans le référentiel du centre de masse.

Cette perturbation, ne faisant pas intervenir le moment cinétique, est déjà diagonale dans la base $|n, l, m\rangle$ est il on peut utiliser le formalisme de la théorie des perturbations stationnaires sans prendre en compte la dégénérescence des niveaux d'énergie non perturbés.

Le calcul de la correction aux niveaux d'énergie peut se faire explicitement. Donnons par exemple, par un calcul proche de l'éqn. (17.1), la correction à l'état fondamental. L'opérateur $\vec{\mathcal{P}}^2$ étant hermitien, nous avons

$$\begin{aligned}\langle \mathcal{H}_1^c\rangle_{100} &= -\frac{1}{8m_e^3c^2}\langle \vec{\mathcal{P}}^4\rangle_{100} = -\frac{1}{8m_e^3c^2}\left\| \vec{\mathcal{P}}^2 |100\rangle \right\|^2 = -\frac{1}{8m_e^3c^2}4\pi\int r^2 dr\Big(-\hbar^2\Delta\Psi_{100}(r)\Big)^2 \\ &= -\frac{\hbar^4}{2m_e^3c^2}\frac{1}{a_0^3}\int dr\, r^2\left(\frac{1}{r^2}\frac{d}{dr}r^2\frac{d}{dr}e^{-r/a_0}\right)^2 = -\frac{\hbar^4}{2m_e^3c^2}\frac{1}{a_0^5}\underbrace{\int dr\,(2-r/a_0)^2 e^{-2r/a_0}}_{=5a_0/4} \\ &= \frac{-5\hbar^4}{8m_e^3c^2}\frac{1}{a_0^4} \qquad (17.6)\end{aligned}$$

En utilisant l'identité

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e q^2} = \frac{\hbar}{m_e c\alpha}\,. \qquad (17.7)$$

nous trouvons finalement

$$\langle \mathcal{H}_1^c\rangle_{100} = -\frac{5}{8}m_e c^2\alpha^4\,. \qquad (17.8)$$

On montre de manière générale (voir par exemple [1], page 322) que la correction à un niveau quelconque au premier ordre en théorie des perturbations est donnée par

$$\delta E_{nl}^c = \langle \mathcal{H}_1^c\rangle_{nlm} = -\frac{1}{2}m_e c^2\alpha^4\left(-\frac{3}{4n^4} + \frac{1}{n^3(l+1/2)}\right)\,, \qquad (17.9)$$

brisant comme attendu la dégénérescence accidentelle entre les niveaux d'énergie ayant des valeurs de l différentes à n fixé tout en étant indépendant de m en raison de la symétrie de rotation.

Le second type de correction au même ordre, est le *terme de Darwin*, provenant lui aussi du développement de l'équation de Dirac, est donné pour un électron dans potentiel électrique $\phi(\vec{x})$ quelconque par :

$$\boxed{\mathcal{H}^d \overset{\text{déf.}}{=} -\frac{\hbar^2 q}{8m_e^2c^2}\Delta\phi(\vec{x})\,.} \qquad (17.10)$$

Son interprétation est moins intuitive et ne sera pas discutée ici. La correction aux niveaux d'énergie pour le potentiel coulombien $\phi(r) = \frac{q}{4\pi\epsilon_0 r}$ est donnée par :

$$\langle \mathcal{H}^d\rangle_{lmn} = -\frac{\hbar^2q^2}{32\pi\epsilon_0 m_e^2c^2}\int d^3x|\Psi_{lmn}(r,\theta,\phi)|^2\underbrace{\Delta\Big(\frac{1}{r}\Big)}_{=-4\pi\delta(\vec{x})} = \frac{\hbar^2q^2}{8\epsilon_0 m_e^2c^2}|\Psi_{lmn}(\vec{x}=\vec{0})|^2 \qquad (17.11)$$

Seuls les fonctions d'onde avec $l = 0$ ayant une valeur non-nulle à l'origine des coordonnées – il suffit de constater que le potentiel effectif (15.52) ne possède pas de terme répulsif pour

$l = 0$, ou de manière équivalente d'examiner le comportement asymptotique (15.19) – nous avons, en utilisant l'éqn. (15.73) :

$$\delta E^{d}_{nl} = \langle \mathcal{H}^{d} \rangle_{lmn} = \delta_{l,0}\delta_{m,0} \frac{\hbar^2 q^2}{8\epsilon_0 m_e^2 c^2} \big(Y_0^0\big)^2 \frac{1}{2n^2} \left(\frac{2}{n a_0}\right)^3 \big(\underbrace{L^1_{n-1}(0)}_{=n} \big)^2 = \delta_{l,0}\delta_{m,0} \, \frac{1}{2n^3} \, m_e c^2 \alpha^4 \, . \tag{17.12}$$

17.1.2 Couplage spin-orbite

Le dernier terme de structure fine apportant des corrections d'ordre α^4 au spectre ponctuel de l'atome d'hydrogène est le couplage spin-orbite. Obtenu également par un développement de l'équation de Dirac, il peut se comprendre de la manière suivante. Dans le référentiel du proton au repos (toujours dans la limite $m_p \gg m_e$ où on peut considérer classiquement l'électron en orbite autour du proton infiniment massif) nous avons le champ électrique :

$$\vec{E} = -\vec{\nabla}\Phi(r) = \frac{q}{4\pi\epsilon_0 r^2} \vec{e}_r \, . \tag{17.13}$$

Dans un autre référentiel d'inertie, ayant une vitesse $\vec{v}$ par rapport au référentiel du proton, une transformation de Lorentz du champ électromagnétique va donner en particulier le champ magnétique :

$$\vec{B} = -\frac{\gamma}{c^2} \, \vec{v} \wedge \vec{E} \, , \quad \gamma = 1/\sqrt{1 - v^2/c^2} \, , \tag{17.14}$$

soit à l'ordre le plus bas

$$\vec{B} = -\frac{q}{4\pi\epsilon_0 r^2} \frac{\vec{p}}{m_e c^2} \wedge \vec{e}_r = \frac{q}{4\pi\epsilon_0 m_e c^2} \frac{1}{r^3} \, \vec{L} \, . \tag{17.15}$$

L'électron possède un moment magnétique de spin (là encore conséquence de l'équation de Dirac), voir l'éqn. (13.42), donnant un terme d'énergie potentielle dans le hamiltonien en présence du champ magnétique :

$$\mathcal{U} = -\vec{\mu} \cdot \vec{B} = g_s \mu_B \, \frac{1}{\hbar} \, \vec{S} \cdot \vec{B} \simeq \frac{q}{m_e} \vec{S} \cdot \vec{B} \, . \tag{17.16}$$

Naturellement, dans une analyse classique, l'électron ne suit pas une trajectoire rectiligne et uniforme dans le champ électrique créé par le proton. Le mouvement accéléré de l'électron sur son orbite va induire un phénomène de précession du spin, appelé *précession de Thomas*. Il en résulte, une fois les calculs faits, un terme supplémentaire de même forme que le couplage du moment magnétique au champ magnétique (17.15) mais avec un facteur -1/2. La correction au hamiltonien de l'atome d'hydrogène, appelée *couplage spin-orbite*, prend finalement la forme :

$$\boxed{\mathcal{H}^{so}_1 = \frac{q^2}{8\pi\epsilon_0 m_e^2 c^2} \frac{1}{\mathcal{R}^3} \, \vec{\mathcal{L}} \cdot \vec{S}} \tag{17.17}$$

L'analyse de la perturbation dans la base produit tensoriel

$$|n,l,m,s\rangle \overset{\text{déf.}}{=} |n,l\rangle \otimes |l,m\rangle \otimes |1/2,m_s\rangle \;\in\; \mathbb{L}^2(\mathbb{R}^+,r^2dr)\otimes \mathbb{C}^{2l+1}\otimes \mathbb{C}^2 \tag{17.18}$$

où les trois termes sont associés respectivement à une fonction d'onde radiale, à une harmonique sphérique et aux degrés de liberté de spin n'est pas aisée car la perturbation n'est pas diagonale dans ces variables.

Il est utile de considérer la décomposition de la représentation $D^l\otimes D^{1/2}$ en représentations irreductibles (voir le cours n° 16) section 16.1), valable pour $l\neq 0$:

$$D^l\otimes D^{1/2} = D^{l-1/2}\oplus D^{l+1/2}\,, \tag{17.19}$$

en introduisant le moment cinétique total

$$\vec{\mathcal{J}} = \vec{\mathcal{L}} + \vec{\mathcal{S}}\,, \tag{17.20}$$

En considérant les vecteurs propres communs de $\mathcal{J}^2$ et $\mathcal{J}_z$:

$$|l,1/2;J,M\rangle \;,\quad J\in\{l-1/2,l+1/2\}\,,\quad M\in\{-J,-J+1,\cdots,J-1,J\}\,, \tag{17.21}$$

correspondant à une base de l'espace des états constituée des vecteurs

$$|n,l;J,M\rangle \overset{\text{déf.}}{=} |n,l\rangle\otimes|l,1/2;J,M\rangle \;\in\; \mathcal{L}^2(\mathbb{R}^+,r^2dr)\otimes\mathbb{C}^{2J+1}\,. \tag{17.22}$$

Nous utilisons ensuite l'identité (déjà mentionnée dans le cours n° 16) :

$$2\vec{\mathcal{L}}\cdot\vec{\mathcal{S}} = \left(\vec{\mathcal{L}}+\vec{\mathcal{S}}\right)^2 - \vec{\mathcal{L}}^{\,2} - \vec{\mathcal{S}}^{\,2} = \mathcal{J}^2-\mathcal{L}^2-\mathcal{S}^2 \tag{17.23}$$

pour réexprimer le couplage spin-orbite comme :

$$\boxed{\mathcal{H}_1^{so} = \frac{q^2}{16\pi\epsilon_0 m_e c^2}\frac{1}{\mathcal{R}^3}\left(\mathcal{J}^2-\mathcal{L}^2-\mathcal{S}^2\right).} \tag{17.24}$$

Ce couplage est manifestement diagonal dans la base (17.22) et la correction aux niveaux d'énergie est donnée, au premier ordre en théorie des perturbations stationnaires, par

$$\begin{aligned}\delta E^{so}_{n,l,J} = \langle n,l;J,M|\,\mathcal{H}_1^{so}\,|n,l,J,M\rangle &= \frac{q^2}{16\pi\epsilon_0 m_e^2c^2}\,\langle n,l|\,\frac{1}{\mathcal{R}^3}\,|n,l\rangle \\ &\quad\times\langle l,1/2;J,M|\left(\mathcal{J}^2-\mathcal{L}^2-\mathcal{S}^2\right)|l,1/2;J,M\rangle\end{aligned} \tag{17.25}$$

Intégrale radiale. Supposons pour l'instant que $l\neq 0$. Le premier terme à calculer dans l'éqn. (17.28) correspond à une intégrale radiale :

$$\langle n,l|\,\frac{1}{\mathcal{R}^3}\,|n,l\rangle = \int_0^\infty r^2dr\;|\rho_{n,l}(r)|^2\,\frac{1}{r^3} \tag{17.26}$$

Le calcul de cette intégrale, un peu technique, est présenté dans l'appendice 17.A. Nous obtenons :

$$\langle n,l|\,\frac{1}{\mathcal{R}^3}\,|n,l\rangle = \frac{1}{n^3a_0^3l(l+1/2)(l+1)}\,. \tag{17.27}$$

Cette expression est divergente si on la prolonge formellement pour $l=0$.

Couplage entre moments cinétiques. Reste à calculer le second terme de l'éqn. (17.28), dépendant uniquement du moment cinétique, qui s'exprime très simplement dans la base associée à la décomposition en représentations irreductibles comme :

$$\frac{1}{\hbar^2}\langle l, 1/2; J, M|\left(\mathcal{J}^2 - \mathcal{L}^2 - \mathcal{S}^2\right)|l, 1/2; J, M\rangle = J(J+1) - l(l+1) - 3/4$$
$$= \begin{cases} l & \text{si} \quad J = l + 1/2 \\ -l-1 & \text{si} \quad J = l - 1/2 \end{cases} \tag{17.28}$$

pour $l \neq 0$. Introduisons la notation :

$$l_\pm \overset{\text{déf.}}{=} \begin{cases} l & \text{si} \quad J = l + 1/2 \\ -l-1 & \text{si} \quad J = l - 1/2 \end{cases} \tag{17.29}$$

Dans le cas particulier $l = 0$ nous avons nécessairement $J = 1/2$ impliquant $\langle J, M|\left(\mathcal{J}^2 - \mathcal{L}^2 - \mathcal{S}^2\right)|l, 1/2; J, M\rangle = 0$.

En assemblant les deux parties du calcul, nous obtenons la correction suivante :

$$\delta E^{so}_{n,l,J} = \frac{\hbar^2 q^2}{16\pi\epsilon_0 m_e^2 c^2}\frac{l_\pm}{n^3 a_0^3 l(l+1/2)(l+1)} = \frac{1}{2}\alpha^2 E_I \frac{l_\pm}{n^3 l(l+1/2)(l+1)} \tag{17.30}$$

En utilisant

$$E_I = \frac{1}{2} m_e c^2 \alpha^2 , \tag{17.31}$$

nous obtenons l'expression :

$$\boxed{\delta E^{so}_{n,l,J} = -\frac{l_\pm}{2nl(l+1/2)(l+1)}\alpha^2 E_n = \frac{l_\pm}{n^3 l(l+1/2)(l+1)}\alpha^4 m_e c^2} \tag{17.32}$$

Dans le cas $l = 0$ on s'attend à ce que $\delta E^{so}_{n,l,J} = 0$ car l'électron ne porte pas de moment cinétique orbital.

Néanmoins on peut formellement prolonger la formule ci-dessus pour $l = 0$, en sélectionnant naturellement l_+ et en prenant la limite $l \to 0$. On obtient par cette procédure formelle une expression finie

$$\delta E^{so}_{n,l,J} \overset{?}{\longrightarrow} \frac{2}{n^3}\alpha^4 m_e c^2 \tag{17.33}$$

qui coïncide exactement avec le terme de Darwin, voir éqn. (17.10) ; nous tirerons profit de cette observation dans la sous-section suivante.

17.1.3 Correction complète à l'ordre α^4

La correction complète au spectre du hamiltonien est donnée par la somme des trois contributions que nous venons de considérer, étant toutes les trois du même ordre. Nous étions donc le hamiltonien

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1^f , \quad \mathcal{H}_1^f \overset{\text{déf.}}{=} \mathcal{H}_1^c + \mathcal{H}^d + \mathcal{H}_1^{so} . \tag{17.34}$$

Au premier ordre en théorie des perturbations, l'effet de $\mathcal{H}_f$ sur le spectre s'obtient en sommant les trois contributions précédemment calculées. Nous obtenons, pour $l \neq 0$:

$$\begin{aligned}\delta E_{nlm} = \delta E^{c}_{nl} + \delta E^{so}_{n,l,J} &= \Big\{ -\frac{1}{2}\left(-\frac{3}{4n^4} + \frac{1}{n^3(l+1/2)}\right) + \frac{l_\pm}{n^3 l(l+1/2)(l+1)}\Big\}\alpha^4 m_e c^2 \\ &= -\frac{\alpha^4 m_e c^2}{2n^4}\left\{-\frac{3}{4} - \frac{n}{l+1/2}\left(\frac{l_\pm}{2l(l+1)} - 1\right)\right\} \end{aligned} \tag{17.35}$$

On remarque que, pour $J = l + 1/2$ ou pour $J = l - 1/2$, le dernier terme vaut toujours $n/(J+1/2)$ et on obtient ainsi le résultat final :

$$\boxed{\delta E_{nJ} = -\frac{\alpha^4 m_e c^2}{2n^4}\left(\frac{n}{J+1/2} - \frac{3}{4}\right)} \tag{17.36}$$

Si $l = 0$, on doit en principe ajouter le terme de Darwin (17.12) mais, comme nous l'avions expliqué auparavant (cf. éqn. (17.33)) le prolongement du couplage spin-orbite pour $l = 0$ donne la même contribution et on peut donc appliquer la formule (17.36) pour tous les états propres de l'atome d'hydrogène.

On remarque finalement que cette formule ne dépend que de l'opérateur de Casimir du moment cinétique total $\vec{J}$, qui est la seule quantité conservée associée aux rotations (expliquant au passage pourquoi cette formule est valide pour $l = 0$ car elle ne peut dépendre de la valeur individuelle du moment cinétique orbital).

Exemple 1 (structure fine du niveau $n = 2$). Avant d'inclure les corrections de structure fine, le niveau d'énergie $n = 2$ possède une dégénérescence totale $d_2 = 8$ en prenant en compte les degrés de liberté de spin. Ce sous-espace des états possède comme base :

— les états de base $|n, l, m, m_s\rangle = |2, 0, 0, \pm 1/2\rangle$ pour l'orbitale $2s$;

— les états de base $|2, 1, 1, \pm 1/2\rangle$, $|2, 1, 0, \pm 1/2\rangle$ et $|2, 1, -1, \pm 1/2\rangle$ pour l'orbitale $2p$.

Pour utiliser l'expression (17.36) il faut passer de la base produit tensoriel $D^l \otimes D^{1/2}$ (avec $l = 0, 1$) à la base des représentations irreductibles.

— Pour l'orbitale $2s$ nous avons $D^0 \otimes D^{1/2} = D^{1/2}$ et donc $J = 1/2$. L'énergie des deux états $|2, 0, 0, \pm 1/2\rangle$ est, à l'ordre α^4,

$$E_{2s_{1/2}} = -\frac{1}{8} m_e c^2 \alpha^2 - \frac{5}{128} m_e c^2 \alpha^4 \tag{17.37}$$

— Pour les orbitales $2p$ nous avons $D^1 \otimes D^{1/2} = D^{1/2} \oplus D^{3/2}$ et donc $J \in \{1/2, 3/2\}$. Nous obtenons les énergies

$$E_{2p_{1/2}} = -\frac{1}{8} m_e c^2 \alpha^2 - \frac{5}{128} m_e c^2 \alpha^4 \tag{17.38a}$$

$$E_{2p_{3/2}} = -\frac{1}{8} m_e c^2 \alpha^2 + \frac{1}{128} m_e c^2 \alpha^4 , \tag{17.38b}$$

avec respectivement 2 et quatre états. Remarquons qu'on ne peut assigner à ces états une valeur propre m_s bien définie, car ils ne sont pas états propres de l'opérateur S_z. En revanche les états $|n = 2, l = 0; J = 1/2, M\rangle$ du niveau $2s_{1/2}$ (provenant de l'orbitale $2s$) et les états $|n = 2, l = 1; J = 1/2, M\rangle$ du niveau $2p_{1/2}$ (provenant de l'orbitale $2p$) ont même énergie peuvent se distinguer par la valeur propre de $\mathcal{L}^2$, qui commute avec $\mathcal{J}^2$ et $\mathcal{J}_z$.

Les premiers niveaux d'énergie de structure fine sont représentés sur la figure 17.1. Le nombre en indice indique le moment cinétique total J correspondant ; par exemple le niveau $2p_{3/2}$ est caractérisé par $n = 2$, $l = 1$ et $J = 3/2$.

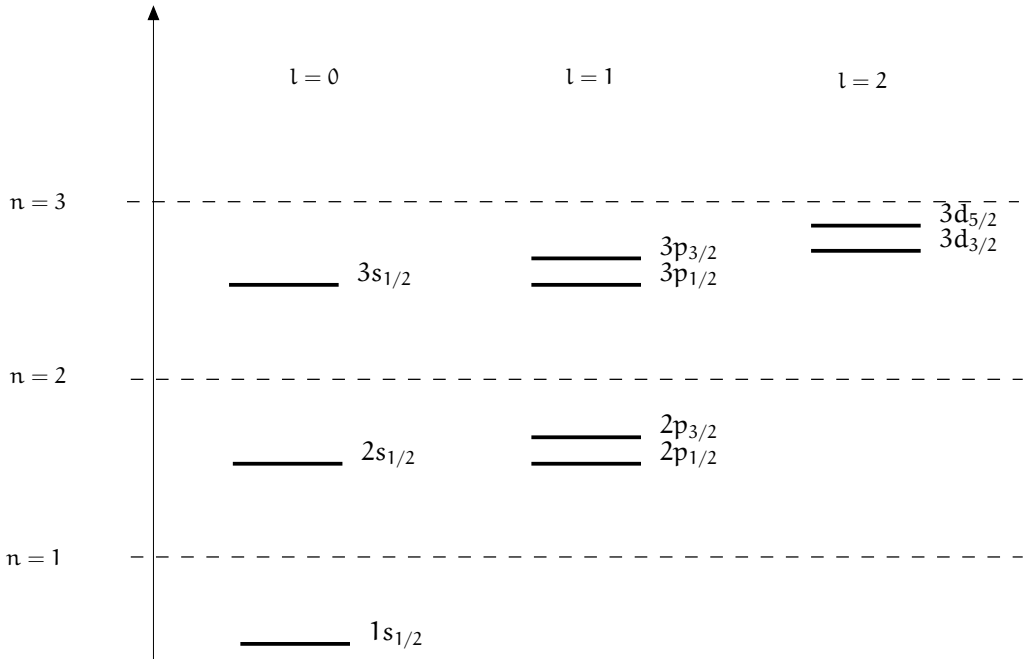

FIGURE 17.1 – Niveaux d'énergie de structure fine de l'atome d'hydrogène.

17.2 Le théorème de Wigner–Eckart

Avant d'analyser la manière dont un champ magnétique extérieur affecte la structure fine de l'atome d'hydrogène que nous venons d'étudier, il est utile de présenter des résultats très généraux permettant d'effectuer de manière efficace, en exploitant la symétrie de rotation sous-jacente dans tous les problèmes de physique atomique, le type de calcul de perturbations que nous y rencontrerons.

17.2.1 Opérateurs tensoriels

Rappelons tout d'abord la définition d'un opérateur vectoriel, déjà étudié dans le cours n° 4, sous-section 4.1.3. Un triplet d'opérateurs $(\mathcal{V}_x, \mathcal{V}_y, \mathcal{V}_z)$ est appelé opérateur vectoriel s'il se

transforme par rotation comme :

$$\mathcal{V}_i \;\mapsto\; \mathcal{R}(\vec{n},\psi)^\dagger \mathcal{V}_i \mathcal{R}(\vec{n},\psi)^\dagger = e^{\frac{i}{\hbar}\vec{n}\cdot\vec{\mathcal{J}}}\mathcal{V}_i e^{-\frac{i}{\hbar}\vec{n}\cdot\vec{\mathcal{J}}} \overset{!}{=} \sum_j R(\vec{n},\psi)_{ij}\mathcal{V}_j\,, \tag{17.39}$$

où les opérateurs $\mathcal{J}_i$ sont les générateurs des rotations agissant sur l'espace de Hilbert $\mathcal{H}$ et où $R(\vec{n},\psi)$ la matrice de rotation usuelle agissant sur les vecteurs de l'espace à 3 dimensions. Nous avons pour une transformation infinitésimale d'angle ε

$$\mathcal{V}_i \;\mapsto\; \mathcal{V}_i + \frac{i}{\hbar}\Big[\,\vec{n}\cdot\vec{\mathcal{J}},\,\mathcal{V}_i\,\Big]\,\varepsilon + \mathcal{O}(\varepsilon^2) \tag{17.40}$$

et les opérateurs $\mathcal{V}_i$ doivent satisfaire aux relations de commutation (provenant du développement des deux membres de l'égalité 17.39 au premier ordre en ε) :

$$[\,\mathcal{J}_i,\,\mathcal{V}_j\,] = i\hbar\sum_k \epsilon_{ij}{}^k\,\mathcal{V}_k\,. \tag{17.41}$$

Il est utile d'introduire la base sphérique d'états propres de $\mathcal{J}_z$ (voir éqn. (13.91) du cours n° 13 et éqn. (14.91) du cours n° 14) pour ces opérateurs :

$$\begin{cases} \mathcal{V}_1^1 & \overset{\text{déf.}}{=} & -\frac{1}{\sqrt{2}}\big(\mathcal{V}_x + i\mathcal{V}_y\big) \\ \mathcal{V}_1^0 & \overset{\text{déf.}}{=} & \mathcal{V}_z \\ \mathcal{V}_1^{-1} & \overset{\text{déf.}}{=} & \frac{1}{\sqrt{2}}\big(\mathcal{V}_x - i\mathcal{V}_y\big) \end{cases} \tag{17.42}$$

De telle sorte que les relations de commutation (17.41) peuvent se reformuler comme (avec $\mathcal{J}_\pm = \mathcal{J}_x \pm i\mathcal{J}_y$) :

$$[\,\mathcal{J}_z,\,\mathcal{V}_1^r\,] = \hbar r\mathcal{V}_1^r \tag{17.43a}$$

$$[\,\mathcal{J}_\pm,\,\mathcal{V}_1^r\,] = \hbar\sqrt{1(1+1)-r(r\pm1)}\,\mathcal{V}_1^{r\pm1} \tag{17.43b}$$

Ces relations de commutation rappellent naturellement l'action des générateurs d'une représentation irreductible D^J sur des états de base $|J,M\rangle$, avec $J=1$, c.-à.-d. la représentation vectorielle. Cela suggère une généralisation naturelle pour des valeurs de J différentes.

Définition 1 (Opérateur tensoriel sphérique). *Soit* $\{\mathcal{T}_k^r,\ r=-k,\dots,k\}$ *une famille d'opérateurs, avec* $k\in\mathbb{N}$. *Ces opérateurs forment les composantes d'un opérateurs tensoriel sphérique de poids* k *si*

$$\boxed{[\,\mathcal{J}_z,\,\mathcal{T}_k^r\,] = \hbar r\mathcal{T}_k^r} \tag{17.44a}$$

$$\boxed{[\,\mathcal{J}_\pm,\,\mathcal{T}_k^r\,] = \hbar\sqrt{k(k+1)-r(r\pm1)}\,\mathcal{T}_k^{r\pm1}} \tag{17.44b}$$

c.-à.-d. que ce tenseur se transforme dans la représentation irreductible D^k *de* $\mathfrak{so}_3$.

Les tenseurs sphériques de poids k apparaissent par exemple dans le développement multipolaire du potentiel électrique créé par une distributions de charges, au-delà de l'approximation dipolaire habituellement utilisée.

Exemple 2 (tenseur sphérique de poids 2). Nous avions déjà vu, du point de vue géométrique, comment construire un tenseur se transformant dans une représentation irreductible de $SO(3)$ avec $J = 2$, voir la sous-section 13.1.1 du cours n° 13. Partant d'un opérateur vectoriel de composantes $\mathcal{V}^j$ on peut ainsi construire un opérateur tensoriel de poids 2 :

$$S^{ij} \overset{\text{déf.}}{=} \frac{1}{2}\left(\mathcal{V}^i\mathcal{V}^j + \mathcal{V}^j\mathcal{V}^i\right) - \frac{1}{3}\mathrm{Tr}\left(\mathcal{V}^k\mathcal{V}^l\right)\delta^{ij}\,. \tag{17.45}$$

Il faut ensuite convertir les composantes cartésiennes de ce tenseur dans une base adaptée, la base sphérique, pour obtenir les relations de commutation (17.44). Pour trouver l'expression du tenseur dans la base appropriée, nous pouvons remarquer que ces relations sont identiques à l'action des générateurs du moment cinétique orbital $\mathcal{L}_j$, vus comme opérateurs différentiels, agissant sur les harmoniques sphériques. On peut alors (comme nous l'avons fait au-dessus pour un opérateur vectoriel) obtenir la décomposition voulue en remplaçant dans les harmoniques sphériques $Y_2^m(\theta,\phi)$ le vecteur unitaire $\vec{n} = \vec{x}/r$ par l'opérateur vectoriel $\vec{\mathcal{V}}$. Explicitement nous avons

$$Y_2^{\pm 2} = \sqrt{\frac{15}{32\pi}}\sin^2\theta e^{\pm 2i\phi} = \sqrt{\frac{15}{32\pi}}\frac{(x \pm iy)^2}{r^2} \quad \to\ \mathcal{T}_2^{\pm 2} = \sqrt{\frac{15}{32\pi}}\left(\mathcal{V}^1 \pm i\mathcal{V}^2\right)^2 \tag{17.46a}$$

$$Y_2^{\pm 1} = \mp\sqrt{\frac{15}{8\pi}}\sin\theta\cos\theta e^{\pm i\phi} = \mp\sqrt{\frac{15}{8\pi}}\frac{z(x \pm iy)}{r^2} \quad \to\ \mathcal{T}_2^{\pm 1} = \mp\sqrt{\frac{15}{8\pi}}\mathcal{V}^3\left(\mathcal{V}^1 \pm i\mathcal{V}^2\right) \tag{17.46b}$$

$$Y_2^0 = \sqrt{\frac{5}{16\pi}}(3\cos\theta^2 - 1) = \sqrt{\frac{5}{16\pi}}(3\cos\theta^2 - 1) \quad \to\ \mathcal{T}_2^0 = \sqrt{\frac{5}{16\pi}}\left(3\left(\mathcal{V}^3\right)^2 - 1\right) \tag{17.46c}$$

17.2.2 Théorème de Wigner–Eckart

L'analyse des perturbations au hamiltonien (15.42) de l'atome d'hydrogène, en particulier le couplage à un champ magnétique extérieur, necessite souvent de calculer les valeurs moyennes, ou plus généralement les élements de matrices, d'opérateurs tensoriels. Ce type de calcul se simplifie considérablement en considérant :

— une décomposition de l'espace de Hilbert en représentations irreductibles de $SU(2)$, c.-à.-d. le choix d'une base $|j, m\rangle$ d'états propres du moment cinétique ;

— la décomposition de l'opérateur tensoriel en représentation irreductibles, c.-à.-d. en opérateurs tensoriels sphériques $\mathcal{T}_k^r$.

Nous arrivons donc au problème suivant. Comme auparavant l'espace de Hilbert du système est décomposé en représentation irreductibles de $SU(2)$, c.-à.-d. comme une somme

directe de sous-espaces de représentations :

$$\mathcal{H} \cong \bigoplus_{j} \mathcal{H}_j \,, \tag{17.47}$$

où une base de $\mathcal{H}_j$ est fournie par des vecteurs $|\alpha; j, m\rangle$, états propres de $(\mathcal{J}^2, \mathcal{J}_z)$ de valeurs propres respectives $\hbar^2 j(j+1)$ et $\hbar m$, et α désignant l'ensemble des autres nombres quantiques permettant de spécifier un état.

Le calcul d'éléments de matrice d'un opérateur tensoriel peut se ramener, par la double décomposition de l'espace de Hilbert $\mathcal{H}$ en sous-espaces $\mathcal{H}_l$ associés à des représentations irreductibles de $SU(2)$ et de l'opérateur tensoriel en tenseurs sphériques $\mathcal{T}^r_k$, associés également à des représentations irreductibles de $SU(2)$, au calcul d'éléments de matrice de la forme :

$$\langle \alpha', j', m' | \, \mathcal{T}^r_k \, | \alpha, j, m \rangle \ . \tag{17.48}$$

Un premier résultat très simple concernant le calcul de tels éléments de matrice est une règle de sélection portant sur les valeurs propres de $\mathcal{J}_z$.

Propriété 1 (première règle de sélection pour les tenseurs sphériques). *Un élément de matrice* $\langle \alpha', j', m' | \, \mathcal{T}^r_k \, | \alpha, j, m \rangle$ *pour une composante d'un opérateur tensoriel sphérique ne peut être non-nul que si* $\boxed{m' - m = r}$.

La preuve de cette propriété est évidente. Nous avons d'une part

$$\begin{aligned} \langle \alpha', j', m' | \, [\, \mathcal{J}_z, \, \mathcal{T}^r_k \,] \, | \alpha, j, m \rangle &= \Big(\langle \alpha', j', m' | \, \mathcal{J}_z \Big) \mathcal{T}^r_k \, | \alpha, j, m \rangle - \langle \alpha', j', m' | \, \mathcal{T}^r_k \Big(\mathcal{J}_z \, | \alpha, j, m \rangle \Big) \\ &= \hbar (m' - m) \, \langle \alpha', j', m' | \, \mathcal{T}^r_k \, | \alpha, j, m \rangle \end{aligned} \tag{17.49}$$

et d'autre part, d'après la définition (17.44a) :

$$\langle \alpha', j', m' | \, [\, \mathcal{J}_z, \, \mathcal{T}^r_k \,] \, | \alpha, j, m \rangle = \hbar r \, \langle \alpha', j', m' | \, \mathcal{T}^r_k \, | \alpha, j, m \rangle \tag{17.50}$$

Ces deux quantités ne peuvent être égales que si $m' - m = r$. □

Nous en arrivons maintenant au résultat le plus important de cette section, le théorème de Wigner–Eckart portant sur le calcul d'éléments de matrice de la forme (17.48).

Théorème 1 (théorème de Wigner–Eckart). *Soit* $\mathcal{T}^r_k$ *une composante d'un tenseur sphérique de poids* k. *Les éléments de matrice de ce tenseur entre un vecteur* $|\alpha, j, m\rangle \in \mathcal{H}_j$ *et un vecteur* $|\alpha', j', m'\rangle \in \mathcal{H}_{j'}$ *peuvent se mettre sous la forme :*

$$\boxed{\langle \alpha', j', m' | \, \mathcal{T}^r_k \, | \alpha, j, m \rangle = C^{j' \, m'}_{j \, m, \, kr} \, \langle \alpha', j' \| \mathcal{T}_k \| \alpha, j \rangle} \tag{17.51}$$

où $C^{j' \, m'}_{j \, m, \, kr} = \langle j, m; k, r | j, m \rangle$ *sont les coefficients de Clebsch–Gordan de la décomposition de* $D^j \otimes D^k$ *pour la représentation* $D^{j'}$ *et où les éléments de matrices réduits*

$$\langle \alpha', j' \| \mathcal{T}_k \| \alpha, j \rangle$$

sont indépendants de m, m' *et* r.

Pour démontrer ce théorème, on utilise un raisonnement analogue à la régle de sélection précédente. Nous avons d'une part

$$\begin{aligned}\langle \alpha', j', m'| \, [\mathcal{J}_\pm, \mathcal{T}^r_k] \, |\alpha, j, m\rangle &= \Big(\langle \alpha', j', m'| \, \mathcal{J}_\pm \Big) \mathcal{T}^r_k \, |\alpha, j, m\rangle - \langle \alpha', j', m'| \, \mathcal{T}^r_k \Big(\mathcal{J}_\pm \, |\alpha, j, m\rangle \Big) \\ &= \hbar \sqrt{j'(j'+1) - m'(m' \mp 1)} \, \langle \alpha', j', m' \mp 1| \, \mathcal{T}^r_k \, |\alpha, j, m\rangle \\ &\quad - \hbar \sqrt{j'(j'+1) - m'(m' \pm 1)} \, \langle \alpha', j', m'| \, \mathcal{T}^r_k \, |\alpha, j, m \pm 1\rangle \end{aligned} \quad (17.52)$$

et d'autre part, d'après (17.44b) :

$$\langle \alpha', j', m'| \, [\mathcal{J}_\pm, \mathcal{T}^r_k] \, |\alpha, j, m\rangle = \hbar \sqrt{k(k+1) - r(r \pm 1)} \, \langle \alpha', j', m'| \, \mathcal{T}^{r-1}_k \, |\alpha, j, m \pm 1\rangle \quad (17.53)$$

L'égalité entre ces deux expressions montre que les éléments de matrice $\langle \alpha', j', m'| \, \mathcal{T}^r_k \, |\alpha, j, m\rangle$ satisfont exactement les mêmes relations de récursion que les coefficients de Clebsch–Gordan $C^{j'\,m'}_{j\,m,\,kr}$, voir l'éqn. (16.63) du cours n° 16. Ces équations linéaires homogènes entre les différents coefficients $C^{j'\,m'}_{j\,m,\,kr}$, à j, j' et k fixés, déterminent complètement leur expression à une constante multiplicative près, ces équations étant homogènes, constante pouvant dépendre de j, j' et k.

Les éléments de matrice $\langle \alpha', j', m'| \, \mathcal{T}^r_k \, |\alpha, j, m\rangle$ satisfaisant aux même relations de récursion doivent être proportionnels au coefficients de Clebsch–Gordan $C^{j'\,m'}_{j\,m,\,kr}$ à une constante muiltiplicative près, pouvant donc dépendre de j, j' et k. □

En pratique : ce théorème nous indique qu'il suffit de savoir calculer un seul élément de matrice $\langle \alpha', j', m'| \, \mathcal{T}^r_k \, |\alpha, j, m\rangle$ pour obtenir tous les autres associés aux différentes valeurs possibles de m, m' et r. Le passage d'un élément de matrice à l'autre nécessite seulement de connaître les coefficients de Clebsch–Gordan associés, qui sont, rappelons-le, déjà connus et calculés.

Une conséquence immédiate du théorème de Wigner–Eckart est une règle de sélection concernant les éléments de matrices entre sous-espaces associés à des représentations de $\mathfrak{su}_2$ différentes.

Propriété 2 (seconde règle de sélection pour les tenseurs sphériques). *Un élément de matrice* $\langle \alpha', j', m'| \, \mathcal{T}^r_k \, |\alpha, j, m\rangle$ *pour une composante d'un opérateur tensoriel sphérique ne peut être non-nul que si* $\boxed{|j - k| \leqslant j' \leqslant j + k}$.

Nous savons en effet que les coefficients de Clebsch–Gordan sont nuls en dehors de ce domaine, voir l'éqn. (16.55) du cours n° 16.

Exemple 3 (éléments de matrice d'un opérateur scalaire). Pour un opérateur scalaire le théorème prend une forme particulièrement simple car les coefficients de Clebsch–Gordan correspondent au produit tensoriel avec la représentation triviale : $D^j \otimes D^0 = D^j$. En conséquence,

$$\boxed{\langle \alpha', j', m'| \, \mathcal{T}^0_0 \, |\alpha, j, m\rangle = \langle \alpha', j' \| \mathcal{T}_0 \| \alpha, j\rangle \, \delta_{j,j'} \delta_{m,m'}} \quad (17.54)$$

c.-à.-d. qu'un opérateur scalaire ne peut changer les valeurs de j et m.

Exemple 4 (éléments de matrice d'un opérateur vectoriel). Soit un opérateur vectoriel $\vec{\mathcal{V}}$ de composantes cartésiennes $\mathcal{V}_x$, $\mathcal{V}_y$ et $\mathcal{V}_z$. On construit tout d'abord les composantes du vecteur dans la base sphérique :

$$\mathcal{V}_1^{\pm 1} = \mp \frac{\mathcal{V}_x \pm i\mathcal{V}_y}{\sqrt{2}} \tag{17.55a}$$

$$\mathcal{V}_1^0 = \mathcal{V}_z\,. \tag{17.55b}$$

Les éléments de matrice de cet opérateur vectoriel sont

$$\boxed{\langle \alpha', j', m' | \, \mathcal{V}_1^r \, | \alpha, j, m \rangle = C^{j'\,m'}_{j\,m,\,1r} \, \langle \alpha', j' \| \mathcal{V}_1 \| \alpha, j \rangle} \tag{17.56}$$

Notons tout d'abord qu'une transition $j = 0 \to j' = 0$ est impossible (car $j' = 0$ n'apparaît pas dans la décomposition de $D^0 \otimes D^1$). Nous avons également les règles de sélection $m' - m \in \{-1, 0, 1\}$ et $j' - j \in \{-1, 0, 1\}$.

Un cas particulier de calcul d'élément de matrice par le théorème de Wigner–Eckart revêt une importante particulière. Pour formuler ce théorème il est utile de transformer les composantes du moment cinétique de la base cartésienne $\mathcal{J}_x$, $\mathcal{J}_y$ et $\mathcal{J}_z$ à la base sphérique :

$$\mathcal{J}_1^{\pm 1} = \mp \frac{\mathcal{J}_x \pm i\mathcal{J}_y}{\sqrt{2}} = \mp \frac{\mathcal{J}_\pm}{\sqrt{2}} \tag{17.57a}$$

$$\mathcal{J}_1^0 = \mathcal{J}_z\,. \tag{17.57b}$$

Théorème 2 (théorème de projection) *Soit $\mathcal{V}_1^r$ un opérateur vectoriel sphérique. Nous avons*

$$\boxed{\langle \alpha'; j, m' | \, \mathcal{V}_1^r \, | \alpha; j, m \rangle = \frac{\langle \alpha'; j, m | \, \vec{\mathcal{J}} \cdot \vec{\mathcal{V}} \, | \alpha; j, m \rangle}{\hbar^2 j(j+1)} \, \langle j, m' | \, \mathcal{J}_1^r \, | j, m \rangle} \tag{17.58}$$

En comparaison de l'équation (17.51), nous avons ici $j = j'$. Ce théorème permet de contourner le calcul des coefficients de Clebsch–Gordan.

Pour démontrer le théorème de projection, calculons premièrement

$$\begin{aligned}
\langle \alpha'; j, m | \, \vec{\mathcal{J}} \cdot \vec{\mathcal{V}} \, | \alpha; j, m \rangle &= \langle \alpha'; j, m | \, \left(\mathcal{J}_x \mathcal{V}_x + \mathcal{J}_y \mathcal{V}_y + \mathcal{J}_z \mathcal{V}_z \right) | \alpha; j, m \rangle \\
&= \langle \alpha'; j, m | \, \Big(\frac{\mathcal{J}_x + i\mathcal{J}_y}{\sqrt{2}} \frac{\mathcal{V}_x - i\mathcal{V}_y}{\sqrt{2}} + \frac{\mathcal{J}_x - i\mathcal{J}_y}{\sqrt{2}} \frac{\mathcal{V}_x + i\mathcal{V}_y}{\sqrt{2}} + \mathcal{J}_z \mathcal{V}_z \Big) \, | \alpha; j, m \rangle \\
&= \langle \alpha'; j, m | \, \mathcal{J}_1^0 \mathcal{V}_1^0 - \mathcal{J}_1^+ \mathcal{V}_1^- - \mathcal{J}_1^- \mathcal{V}_1^+ \, | \alpha; j, m \rangle \\
&= \hbar m \, \langle \alpha'; j, m | \, \mathcal{V}_1^0 \, | \alpha; j, m \rangle - \frac{\hbar}{\sqrt{2}} \sqrt{j(j+1) - m(m-1)} \, \langle \alpha'; j, m-1 | \, \mathcal{V}_1^- \, | \alpha; j, m \rangle \\
&\quad - \frac{\hbar}{\sqrt{2}} \sqrt{j(j+1) - m(m+1)} \, \langle \alpha'; j, m+1 | \, \mathcal{V}_1^+ \, | \alpha; j, m \rangle
\end{aligned} \tag{17.59}$$

Les éléments de matrice apparaissant dans cette dernière expression peuvent s'exprimer à l'aide du théorème de Wigner–Eckart (théorème 1) comme étant tous les trois proportionnels à l'élément de matrice réduit $\langle \alpha', j || \mathcal{V}_1 || \alpha, j \rangle$.

Étant donné que $\vec{\mathcal{J}} \cdot \vec{\mathcal{V}}$ est un opérateur scalaire, nous savons d'après l'équation (17.54) que le coefficient de proportionnalité global ne dépend que de j et que $m = m'$ obligatoirement :

$$\exists c_j \in \mathbb{R} \ , \quad \langle \alpha'; j, m | \, \vec{\mathcal{J}} \cdot \vec{\mathcal{V}} \, | \alpha; j, m \rangle = c_j \, \langle \alpha', j || \mathcal{V}_1 || \alpha, j \rangle \tag{17.60}$$

Pour trouver une expression simple de c_j remarquons que cette égalité est vraie pour tout opérateur vectoriel, en particulier pour $\vec{\mathcal{J}}$ lui-même (auquel cas nous pouvons nous restreindre à l'espace hermitien associé à D^j et supprimer l'étiquette α pour les états). Nous avons d'une part

$$\langle j, m | \, \vec{\mathcal{J}} \cdot \vec{\mathcal{J}} \, | j, m \rangle = \langle j, m | \, \mathcal{J}^2 \, | j, m \rangle = \hbar^2 j(j+1) \tag{17.61}$$

et d'autre part

$$\langle j, m | \, \vec{\mathcal{J}} \cdot \vec{\mathcal{J}} \, | j, m \rangle = c_j \, \langle j || \mathcal{J}_1 || j \rangle \quad \Longrightarrow \quad c_j = \frac{\hbar^2 j(j+1)}{\langle j || \mathcal{J}_1 || j \rangle} \tag{17.62}$$

Finalement, en comparant le théorème Wigner–Eckart appliqué à $\mathcal{V}_1$ d'une part et à $\mathcal{J}_1$ d'autre part, nous avons

$$\langle \alpha', j, m' | \, \mathcal{V}_1^r \, | \alpha, j, m \rangle = C^{j\,m'}_{j\,m,\,1r} \, \langle \alpha', j || \mathcal{V}_k || \alpha, j \rangle \tag{17.63a}$$

$$\langle j, m' | \, \mathcal{J}_1^r \, | j, m \rangle = C^{j\,m'}_{j\,m,\,1r} \, \langle j || \mathcal{J}_k || j \rangle \tag{17.63b}$$

nous en déduisons que :

$$C^{j\,m'}_{j\,m,\,1r} = \frac{\langle \alpha', j, m' | \, \mathcal{V}_1^r \, | \alpha, j, m \rangle}{\langle \alpha', j || \mathcal{V}_1 || \alpha, j \rangle} = \frac{\langle j, m' | \, \mathcal{J}_1^r \, | j, m \rangle}{\langle j || \mathcal{J}_1 || j \rangle} \tag{17.64}$$

Nous avons donc, en rassemblant tous les résultats :

$$\begin{aligned} \langle \alpha', j, m' | \, \mathcal{V}_1^r \, | \alpha, j, m \rangle &= \frac{\langle j, m' | \, \mathcal{J}_1^r \, | j, m \rangle}{\langle j || \mathcal{J}_1 || j \rangle} \frac{\langle \alpha'; j, m | \, \vec{\mathcal{J}} \cdot \vec{\mathcal{V}} \, | \alpha; j, m \rangle}{c_j} \\ &= \frac{\langle j, m' | \, \mathcal{J}_1^r \, | j, m \rangle}{\langle j || \mathcal{J}_1 || j \rangle} \, \langle \alpha'; j, m | \, \vec{\mathcal{J}} \cdot \vec{\mathcal{V}} \, | \alpha; j, m \rangle \, \frac{\langle j || \mathcal{J}_1 || j \rangle}{\langle j, m | \, \mathcal{J}^2 \, | j, m \rangle} \\ &= \langle j, m' | \, \mathcal{J}_1^r \, | j, m \rangle \, \frac{\langle \alpha'; j, m | \, \vec{\mathcal{J}} \cdot \vec{\mathcal{V}} \, | \alpha; j, m \rangle}{\hbar^2 j(j+1)} \end{aligned} \tag{17.65}$$

qui constitue le résultat attendu. □

17.3 L'effet Zeeman

L'effet Zeeman, nommé d'après le physicien néerlandais Pieter Zeeman, consiste en la séparation des raies spectrales des atomes en présence d'un champ magnétique extérieur. Ce phénomène, qui joua un rôle majeur dans le découverte du spin, est aussi un outil très précieux en astrophysique car il permet de mesurer le champ magnétique des astres à partir de l'étude de leur spectre ; on peut par exemple obtenir la valeur du champ magnétique à l'intérieur des taches solaires, de l'ordre de $0,3$ Tesla, voir la figure 17.2.

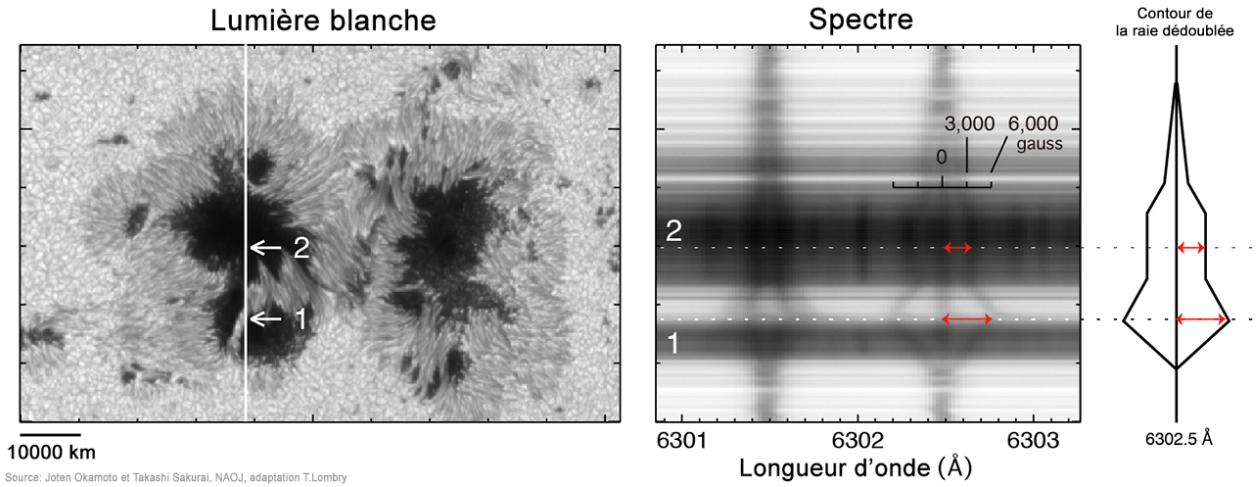

FIGURE 17.2 – Division des raies spectrales du Soleil par effet Zeeman.

On considère un atome d'hydrogène placé dans un champ magnétique extérieur $\vec{B}$ uniforme, choisi conventionnellement suivant l'axe $\vec{e}_z$. L'électron de l'atome d'hydrogène va coupler au champ mangétique en raison des deux type de moments magnétiques que porte la particule, son moment magnétique de spin et son moment magnétique orbital, associés aux opérateurs :

$$\vec{\mu}_s = -\frac{2\mu_B}{\hbar}\vec{S}\,, \tag{17.66a}$$

$$\vec{\mu}_l = -\frac{\mu_B}{\hbar}\vec{\mathcal{L}}\,, \tag{17.66b}$$

en termes du magnéton de Bohr $\mu_B = \frac{|q|\hbar}{2m_e}$. À noter que nous ne prenons pas en compte les moments magnétique de spin et orbital du protons, qui donnent des contributions beaucoup plus faibles car $m_p \gg m_e$.

Le couplage Zeeman entre l'atome d'hydrogène et le champ magnétique est donc de la forme :

$$\mathcal{H}_1^z = -(\vec{\mu}_s + \vec{\mu}_l)\cdot\vec{B} \implies \boxed{\mathcal{H}_1^z = \frac{\mu_B B}{\hbar}(\mathcal{L}_z + 2\mathcal{S}_z)} \tag{17.67}$$

à noter qu'en raison du facteur de Landé pour le spin ($g_s = 2$ en première approximation), le couplage avec le spin est deux fois plus intense qu'avec le moment cinétique orbital.

Nous souhaitons étudier l'effet du champ magnétique sur les niveaux de structure fine. Nous considérons donc le hamiltonien complet

$$\mathcal{H} = \underbrace{\frac{\vec{\mathcal{P}}^2}{2m_e} - \frac{q^2}{4\pi\epsilon_0 \mathcal{R}}}_{\mathcal{H}_0} \underbrace{-\frac{\vec{\mathcal{P}}^4}{8m_e^3 c^2}}_{\mathcal{H}_1^c} \underbrace{-\frac{\hbar^2 q^2}{32\pi\epsilon_0 m_e^2 c^2}\Delta\left(\frac{1}{\mathcal{R}}\right)}_{\mathcal{H}_1^d} + \underbrace{\frac{q^2}{8\pi\epsilon_0 m_e^2 c^2}\frac{\vec{\mathcal{L}}\cdot\vec{\mathcal{S}}}{\mathcal{R}^3}}_{\mathcal{H}_1^{so}} + \underbrace{\frac{\mu_B B}{\hbar}\left(\mathcal{L}_z + 2\mathcal{S}_z\right)}_{\mathcal{H}_1^z} \tag{17.68}$$

Déterminons l'effet du champ magnétique dans la limite de champ faible, c.-à.-d. que le terme $\mathcal{H}_1^z$ de couplage au champ magnétique est considéré comme une perturbation de $\mathcal{H}_0 + \mathcal{H}_1^f$. Nous savons qu'une base d'états propres de $\mathcal{H}_0 + \mathcal{H}_1^f$ est fournie par les vecteurs $|n, l; J, M\rangle$. On écrira alors le terme de perturbation comme

$$\mathcal{H}_1^z = \frac{\mu_B B}{\hbar}\left(\mathcal{J}_z + \mathcal{S}_z\right) \tag{17.69}$$

Les états de même n et même J sont dégénérés, d'après l'équation (17.36). Il faut donc appliquer la théorie des perturbations stationnaires dégénérées developpée dans la section 11.3 du cours 11. Rappelons les principes importants permettant de calculer la correction au spectre en énergie au premier ordre :

- les élements de matrice entre des sous-espaces propres différents du hamiltonien non perturbé ne contribuent pas ; dans le cas présent, les énergies propres du hamiltonien $\mathcal{H}_0 + \mathcal{H}_1^f$ dépendant de n et de J, cela signifie qu'il suffit, au premier ordre, de considérer des éléments de matrice de la forme $\langle n, l'; J, M'|\, \mathcal{H}_1^z\, |n, l; J, M\rangle$;
- Il faut diagonaliser la perturbation dans chaque sous-espace propre, c.-à.-d. dans le cas présent dans l'espace de dimension $2(2J+1)$ dont une base est

$$\left\{ |n, l; J, M\rangle\,, M \in \{-J, \dots J\}\, l \in \{J + 1/2, J - 1/2\}\right\}. \tag{17.70}$$

Dans notre cas, il se trouve que l'opérateur de perturbation est déjà diagonal dans la base $|n, l; J, M\rangle$. Premièrement nous constatons que $\mathcal{J}_z + \mathcal{S}_z$ correspond à la composante selon z d'un opérateur vectoriel. Nous avons en effet :

$$[\,\mathcal{J}_i,\, \mathcal{J}_j + \mathcal{S}_j\,] = [\,\mathcal{J}_i,\, \mathcal{J}_j\,] + [\,\mathcal{L}_i + \mathcal{S}_i,\, \mathcal{S}_j\,] = [\,\mathcal{J}_i,\, \mathcal{J}_j\,] + [\,\mathcal{S}_i,\, \mathcal{S}_j\,] = i\hbar \sum_k \epsilon^{ijk}\left(\mathcal{J}_k + \mathcal{S}_k\right) \tag{17.71}$$

Nous pouvons ensuite appliquer la règle de sélection 1 pour les opérateurs sphériques (en remarquant qu'on peut identifier $\mathcal{J}_z + \mathcal{S}_z$ à la composante $(\mathcal{J} + \mathcal{S})_1^{r=0}$ de l'opérateur vectoriel sphérique correspondant, voir éqn. (17.55)) :

$$\langle n, l'; J, M'|\, \mathcal{H}_1^z\, |n, l; J, M\rangle = \delta_{M,M'}\, \langle n, l'; J, M|\, \mathcal{H}_1^z\, |n, l; J, M\rangle \tag{17.72}$$

Calcul par les coefficients de Clebsch–Gordan

Le calcul du déplacement des niveaux d'énergie se ramène donc au calcul des éléments de matrice diagonaux en M et M' :

$$\begin{aligned}\langle n, l'; J, M|\, \mathcal{H}_1^z \,|n, l; J, M\rangle &= \frac{\mu_B B}{\hbar} \langle n, l'; J, M|\, (\mathcal{L}_z + 2\mathcal{S}_z) \,|n, l; J, M\rangle \\ &= \frac{\mu_B B}{\hbar} \langle n, l'; J, M|\, (\mathcal{J}_z + \mathcal{S}_z) \,|n, l; J, M\rangle \quad (17.73)\end{aligned}$$

On peut calculer directement ces éléments de matrice en utilisant les coefficients de Clebsch–Gordan pour la composition $D^l \otimes D^{1/2} = D^{J=l-1/2} + D^{J=l+1/2}$, voir sous-section 16.1.1 du cours n° 16. Nous avons :

$$\begin{aligned}|l, 1/2; J = l \pm 1/2, M\rangle = \pm\sqrt{\frac{l \pm M + 1/2}{2l+1}}\, |l, m = M - 1/2; 1/2, m_s = 1/2\rangle \\ + \sqrt{\frac{l \mp M + 1/2}{2l+1}}\, |l, m = M + 1/2; 1/2, m_s = -1/2\rangle \quad (17.74)\end{aligned}$$

Les éléments de matrice non-nuls sont donc donnés par

$$\begin{aligned}&\langle n, l; l \pm 1/2, M|\, \mathcal{H}_1^z \,|n, l; l \pm 1/2, M\rangle = \frac{\mu_B B}{\hbar} \langle n, l; J, M|\, (\mathcal{L}_z + 2\mathcal{S}_z) \,|n, l; J, M\rangle \\ &= \frac{\mu_B B}{\hbar}\Big(\frac{l \pm M + 1/2}{2l+1}(M + 1/2) + \frac{l \mp M + 1/2}{2l+1}(M - 1/2)\Big) = \mu_B B \left(1 \pm \frac{1}{2l+1}\right) M\,. \quad (17.75)\end{aligned}$$

Il reste à vérifier que les éléments de matrice avec $l \neq l'$, pour J et M fixés, sont nuls. Il suffit de remarquer que ni l'action de $\mathcal{L}_z$, ni évidemment l'action de $\mathcal{S}_z$, ne peuvent changer la valeur de l.

Pour le voir de manière plus explicite, calculons l'action de la perturbation sur le vecteur avec $l = J - 1/2$:

$$\begin{aligned}\frac{1}{\hbar}\,(\mathcal{L}_z + 2\mathcal{S}_z)\,|n, l = J - 1/2; J, M\rangle = (M + 1/2)\sqrt{\frac{J + M}{2J}}\, |l = J - 1/2, M - 1/2; 1/2, 1/2\rangle \\ + (M - 1/2)\sqrt{\frac{J - M}{2J}}\, |l = J - 1/2, M + 1/2; 1/2, -1/2\rangle \quad (17.76)\end{aligned}$$

qui ne contient que des vecteurs de la forme $|l = J - 1/2, \star; 1/2, \star\rangle$, alors que le vecteur avec $l = J + 1/2$:

$$\begin{aligned}|l = J + 1/2, 1/2; J, M\rangle = -\sqrt{\frac{J + 1 - M}{2J + 2}}\, |l = J + 1/2, m = M - 1/2; 1/2, m_s = 1/2\rangle \\ + \sqrt{\frac{J + 1 + M}{2J + 2}}\, |l = J + 1/2, m = M + 1/2; 1/2, m_s = -1/2\rangle \quad (17.77)\end{aligned}$$

ne contient que des vecteurs de la forme $|l = J + 1/2, \star; 1/2, \star\rangle$. Ces deux vecteurs sont dont bien orthogonaux.

Calcul par le théorème de projection

Ce résultat peut également se trouver plus directement en utilisant le théorème de projection (17.58), ce qui évite de calculer explicitement les coefficients de Clebsch–Gordan.

Étant donné que $\mathcal{S}_z$ correspond à la composante $\mathcal{S}_1^0$ de la base sphérique, nous avons directement : [3]

$$\begin{aligned}\langle l, 1/2; J, M|\, \mathcal{S}_z \,|l, 1/2; J, M\rangle &= \frac{\langle l, 1/2; J, M|\, \vec{\mathcal{J}} \cdot \vec{\mathcal{S}} \,|l, 1/2; J, M\rangle}{\hbar^2 J(J+1)} \langle J, M|\, \mathcal{J}_z \,|J, M\rangle \\ &= \frac{\langle l, 1/2; J, M|\, \left(\mathcal{J}^2 + \mathcal{S}^2 - \mathcal{L}^2\right) |l, 1/2; J, M\rangle}{2\hbar^2 J(J+1)} \langle J, M|\, \mathcal{J}_z \,|J, M\rangle \end{aligned} \quad (17.78)$$

Avec $J = l \pm 1/2$ nous obtenons ensuite

$$\begin{aligned}\langle l, 1/2; l \pm 1/2, M|\, \mathcal{S}_z \,|l, 1/2; l \pm 1/2, M\rangle &= \frac{(l \pm 1/2)(l + 1 \pm 1/2) + 3/4 - l(l+1)}{2(l \pm 1/2)(l + 1 \pm 1/2)} \hbar M \\ &= \left(1 - \frac{(l + 3/2)(l - 1/2)}{(l \pm 1/2)(l + 1 \pm 1/2)}\right) \frac{\hbar M}{2} = \pm \frac{\hbar M}{2l+1} \end{aligned} \quad (17.79)$$

Nous en déduisons que

$$\boxed{\langle \mathcal{H}_1^z \rangle_{n,l;l\pm1/2,M} = \frac{\mu_B B}{\hbar} \langle n, l; l \pm 1/2, M|\, (\mathcal{J}_z + \mathcal{S}_z) \,|n, l; l \pm 1/2, M\rangle = \mu_B B \left(1 \pm \frac{1}{2l+1}\right) M} \quad (17.80)$$

qui correspond bien au résultat attendu. Nous voyons que le champ magnétique extérieur, qui brise explicitement la symétrie de rotation du système, lève complètement la dégénérescence de la structure fine de l'atome d'hydrogène représentée sur la figure 17.1.

Examinons par exemple l'effet du champ magnétique sur les niveaux de structure fine $(2s)_{1/2}$, $(2p)_{1/2}$ tous les deux d'énergie

$$E = -\frac{1}{8} m_e c^2 \alpha^2 - \frac{5}{128} m_e c^2 \alpha^4$$

Nous avons :

— pour le niveau $(2s)_{1/2}$, $l = 0$ donc $J = 1/2 = l + 1/2$. La formule (17.80) donne deux niveaux ($M = 1/2$ et $M = -1/2$) écartés de $\Delta E = 2\mu_B B$;

— pour le niveau $(2p)_{1/2}$, $l = 1$ donc $J = 1/2 = l - 1/2$. La formule (17.80) donne deux niveaux ($M = 1/2$ et $M = -1/2$) écartés de $\Delta E = \frac{2}{3}\mu_B B$.

La double dégénérescence (en J et en M) est donc bien levée par le champ magnétique.

3. Si $l' \neq l$ on voit immédiatement que $\langle l', 1/2; J, M|\, \left(\mathcal{J}^2 + \mathcal{S}^2 - \mathcal{L}^2\right) |l, 1/2; J, M\rangle = 0$.

À retenir

— Les trois contributions à la structure finie du spectre de l'atome d'hydrogène ;
— la dépendance du résultat final en n et J uniquement ;
— la levée partielle de dégénérescence par les corrections de structure fine ;
— la notion d'opérateur tensoriel sphérique ;
— le théorème de Wigner–Eckart ;
— règles de sélection pour les opérateurs scalaires et vectoriels ;
— le théorème de projection ;
— la levée totale de dégénérescence par effet Zeeman.

Appendices

17.A Calcul d'intégrales radiales

L'objectif de cet appendice et de présenter le calcul des valeurs moyennes de certains opérateurs dans des états propres de l'atome d'hydrogène :

$$\left\langle \frac{1}{\mathcal{R}^2} \right\rangle_{nlm} , \quad \left\langle \frac{1}{\mathcal{R}^3} \right\rangle_{nlm}$$

utilisé dans la section 17.1 pour calculer la correction en énergie due au couplage spin-orbite.

Premièrement, écrivons le hamiltonien non perturbé (15.42), dans l'approximation $\mu \simeq m_e$, sous la forme

$$\mathcal{H}_0 = \frac{1}{2m_e}\mathcal{P}_r^2 + \frac{\vec{\mathcal{L}}^2}{2m_e\mathcal{R}^2} - \frac{q^2}{4\pi\epsilon_0}\frac{1}{\mathcal{R}} \tag{17.81}$$

où l'opérateur d'impulsion radiale $\mathcal{P}_r$ est défini par

$$\langle r|\,\mathcal{P}_r\,|f\rangle = -i\hbar\left(\frac{1}{r}\frac{\partial}{\partial r}r\right) f(r) = -i\hbar\big(f''(r) + \tfrac{2}{r}f'(r)\big)\,. \tag{17.82}$$

En effet nous avons bien

$$\langle r|\,[\,\mathcal{R},\,\mathcal{P}_r\,]\,|f\rangle = -i\hbar\left(\frac{\partial}{\partial r}r\,f(r) - \frac{1}{r}\frac{\partial}{\partial r}r^2\,f(r)\right) = i\hbar f(r)\,. \tag{17.83}$$

Comme tout hamiltonien $\mathcal{H}_0$ est un opérateur hermitien, donc nous avons :

$$\begin{aligned} 0 &= \langle n,l,m|\,[\,\mathcal{H}_0,\,\mathcal{P}_r\,]\,|nlm\rangle \\ &= \frac{\hbar^2 l(l+1)}{2m_e}\,\langle n,l,m|\left[\frac{1}{\mathcal{R}^2},\,\mathcal{P}_r\right]|n,l,m\rangle - \frac{q^2}{4\pi\epsilon_0}\,\langle n,l,m|\left[\frac{1}{\mathcal{R}},\,\mathcal{P}_r\right]|n,l,m\rangle \end{aligned} \tag{17.84}$$

Étant donné que, pour toute fonction f dérivable, $[\mathcal{P}_r, f(\mathcal{R})] = -i\hbar f'(\mathcal{R})$, nous obtenons :

$$\frac{\hbar^2 l(l+1)}{m_e} \langle n, l, m| \frac{1}{\mathcal{R}^3} |nlm\rangle = \frac{q^2}{4\pi\epsilon_0} \langle n, l, m| \frac{1}{\mathcal{R}^2} |n, l, m\rangle \ . \tag{17.85}$$

Le membre de droite de cette égalité peut se calculer avec l'astuce suivante. Dans un sous-espace $\mathcal{H}_l$ de moment cinétique orbital donné le hamltionien $\mathcal{H}_0$ s'écrit

$$\mathcal{H}_0\Big|_{\mathcal{H}_l} = \frac{1}{2m_e}\mathcal{P}_r^2 + \frac{\hbar^2 l(l+1)}{2m_e}\frac{1}{\mathcal{R}^2} - \frac{q^2}{4\pi\epsilon_0}\frac{1}{\mathcal{R}} \tag{17.86}$$

Si, formellement, on effectue la substitution $l \mapsto l + \epsilon$ nous obtenons, au premier ordre en ϵ

$$\mathcal{H}_0\Big|_{\mathcal{H}_l} \longmapsto \mathcal{H}_0\Big|_{\mathcal{H}_l} + \epsilon\frac{\hbar^2(l+1/2)}{m_e}\frac{1}{\mathcal{R}^2} + \mathcal{O}(\epsilon^2)\,. \tag{17.87}$$

Si nous considérons le deuxième terme comme une perturbation du hamiltonien, la théorie des perturbations stationnaires au premier ordre donne la correction suivante à l'énergie d'un état $|nlm\rangle$:

$$\delta E_{nlm} = \epsilon\frac{\hbar^2(l+1/2)}{m_e} \langle n, l, m| \frac{1}{\mathcal{R}^2} |n, l, m\rangle + \mathcal{O}(\epsilon^2)\,. \tag{17.88}$$

D'un autre point de vue, la correction en énergie peut se trouver en partant de l'expression exacte (15.71) des niveaux d'énergie, puis en faisant la substitution $l \mapsto l+\epsilon$ et en développant au premier ordre en ϵ. La contrainte $n-l-1 \in \mathbb{N}$, voir éqn. (15.58), indique que les variations de l et n sont reliées par $\delta l = \delta n$ et nous trouvons

$$\delta E_{nlm} = \delta\left(-\frac{E_I}{n^2}\right) = 2\frac{E_I}{n^3}\epsilon \tag{17.89}$$

Soit en égalisant les deux expressions de δE_{nlm} :

$$\langle n, l, m| \frac{1}{\mathcal{R}^2} |n, l, m\rangle = \frac{m_e}{\hbar^2(l+1/2)}\frac{2E_I}{n^3} \tag{17.90}$$

Reportant ce résultat dans l'équation (17.85) nous trouvons :

$$\langle n, l, m| \frac{1}{\mathcal{R}^3} |nlm\rangle = \frac{m_e}{\hbar^2 l(l+1)}\frac{q^2}{4\pi\epsilon_0} \langle n, l, m| \frac{1}{\mathcal{R}^2} |nlm\rangle = \frac{m_e}{\hbar^2 l(l+1)}\frac{q^2}{4\pi\epsilon_0}\frac{m_e}{\hbar^2(l+1/2)}\frac{2E_I}{n^3} \tag{17.91}$$

Soit finalement :

$$\langle n, l, m| \frac{1}{\mathcal{R}^3} |n, l, m\rangle = \frac{1}{n^3 a_0^3 l(l+1/2)(l+1)}\,. \tag{17.92}$$

Bibliographie

[1] J. J. Sakurai, "Modern Quantum mechanics," Cambridge University Press, 1982

Cours n°18

Perturbations dépendantes du temps

Dans le cours n° 17 nous avons étudié l'effet Zeeman, c.-à.-d. l'évolution des niveaux atomiques en prsence d'un champ magnétique extérieur uniforme et constant dans le temps. Plus généralement, l'analyse de la réaction d'un atome à un champ électromagnétique extérieur est un thème central de la mécanique quantique.

Lorsque ce champ extérieur dépend explicitement du temps, il est nécessaire de développer des techniques adaptées de développement perturbatif. De manière générale, soit un système physique décrit par un hamiltonien $\mathcal{H}_0$ indépendant du temps dont nous connaissons une base orthonormée $|n\rangle$ de vecteurs propres :

$$\mathcal{H}_0 |n\rangle = E_n |n\rangle \ , \quad \langle n|m\rangle = \delta_{n,m} \tag{18.1}$$

À l'instant $t = 0$, on prépare le système dans un état $|\psi_0\rangle$ et on « allume » une perturbation dépendant explicitement du temps, par exemple en ajoutant un champ magnétique extérieur. Nous avons donc

$$\mathcal{H}(t) = \begin{cases} \mathcal{H}_0 & t < 0 \\ \mathcal{H}_0 + \mathcal{V}(t) & t \geqslant 0 \end{cases} \tag{18.2}$$

L'objectif du problème est, naturellement, de connaître le vecteur d'état $|\psi(t)\rangle$ décrivant le système à tout instant $t > 0$. On supposera implicitement que, même en présence de $\mathcal{V}(t)$, une base orthonormée de vecteurs propres de $\mathcal{H}_0$ reste une base de l'espace des états $\mathcal{H}$ du système.

18.1 Représentation d'interaction

La dépendance temporelle du hamiltonien (18.2) venant exclusivement du terme $\mathcal{V}(t)$, l'évolution temporelle des états physiques aura deux origines :

— l'évolution temporelle due au hamiltonien non perturbé $\mathcal{H}_0$, qui se fait évoluer chaque état propre $|n\rangle$ comme $|n\rangle \mapsto \exp\left(-\frac{i}{\hbar} E_n t\right) |n\rangle$;

— une évolution temporelle plus complexe, en général non soluble exactement, due au terme $\mathcal{V}(t)$ dépendant explicitement du temps.

Pour prendre en compte d'emblée la dépendance temporelle « triviale » due au hamiltonien $\mathcal{H}_0$ il est utile d'introduire une représentation intermédiaire entre la représentation de Schrödinger et la représentation de Heisenberg.

Définition 1 (représentation d'interaction). *Soit un système quantique associé, pour* $t \geqslant 0$*, à un hamiltonien de la forme* $\mathcal{H}(t) = \mathcal{H}_0 + \mathcal{V}(t)$ *où* $\mathcal{H}_0$ *ne dépend pas explicitement du temps et soit* $|\psi_0; t\rangle_S$ *le vecteur décrivant l'état du système à l'instant* t *dans la représentation de Schrödinger, tel que* $|\psi_0; 0\rangle_S = |\psi_0\rangle$*. L'état du système dans la représentation d'interaction est donné par :*

$$\boxed{|\psi_0; t\rangle_I \overset{\text{déf.}}{=} e^{\frac{it}{\hbar}\mathcal{H}_0} |\psi_0; t\rangle_S \ , \quad \forall t \geqslant 0} \tag{18.3}$$

La représentation d'interaction revient à ne garder que la dépendance dans le temps provenant de $\mathcal{V}(t)$; elle est en quelque sorte intermédiaire entre la représentation de Schrödinger et la représentation de Heisenberg (voir le cours n° 8, section 8.2). Comme pour ces dernières les quantités physiques ne doivent pas dépendre de la représentation choisie. Si on considère les éléments de matrice d'une observable $\mathcal{O}$ nous avons :

$$\langle \mathcal{O} \rangle_{\psi_0} = {}_S\langle \psi_0; t| \, \mathcal{O}|_S \, |\psi_0; t\rangle_S = {}_I\langle \psi_0; t| \, e^{\frac{it}{\hbar}\mathcal{H}_0} \mathcal{O}|_S e^{\frac{-it}{\hbar}\mathcal{H}_0} \, |\psi_0; t\rangle_I \tag{18.4}$$

Nous en déduisons la relation entre les opérateurs dans les représentations de Schrödinger et d'interaction :

$$\boxed{\mathcal{O}|_I \overset{\text{déf.}}{=} e^{\frac{it}{\hbar}\mathcal{H}_0} \mathcal{O}|_S e^{\frac{-it}{\hbar}\mathcal{H}_0}} \tag{18.5}$$

Propriété 1 (évolution temporelle dans la représentation d'interaction). *Le vecteur* $|\psi_0; t\rangle_I$ *représentant l'état d'un système en représentation d'intraction est solution de l'équation différentielle :*

$$\boxed{i\hbar \frac{\partial}{\partial t} |\psi_0; t\rangle_I = e^{\frac{it}{\hbar}\mathcal{H}_0} \mathcal{V}(t) e^{-\frac{it}{\hbar}\mathcal{H}_0} \, |\psi_0; t\rangle_I = \mathcal{V}(t)|_I \, |\psi_0; t\rangle_I} \tag{18.6}$$

Nous avons immédiatement d'après la définition 1 :

$$i\hbar \frac{\partial}{\partial t} |\psi_0; t\rangle_I = i\hbar \frac{\partial}{\partial t} \left(e^{\frac{it}{\hbar}\mathcal{H}_0} \, |\psi_0; t\rangle_S \right) = -\mathcal{H}_0 e^{\frac{it}{\hbar}\mathcal{H}_0} \, |\psi_0; t\rangle_S + e^{\frac{it}{\hbar}\mathcal{H}_0} \big(\mathcal{H}_0 + \mathcal{V}(t) \big) \, |\psi_0; t\rangle_S \tag{18.7}$$

D'où nous tirons l'équation (18.6). □

Pour obtenir, au moins formellement, le vecteur d'état $|\psi_0; t\rangle_I$ représentant le système dans la représentation d'interaction on peut le développer sur la base – fixe au cours du temps – des états propres de $\mathcal{H}_0$. Nous posons :

$$\boxed{|\psi_0; t\rangle_I = \sum_n c_n(t) \, |n\rangle \ ,} \tag{18.8}$$

Avec la condition initiale :

$$\sum_n c_n(0) \, |n\rangle = |\psi_0\rangle \ . \tag{18.9}$$

L'opérateur d'évolution étant unitaire, $\| \, |\psi\rangle_0 \, \| = 1$ implique nécessairement $\sum_n |c_n(t)|^2 = 1$.

Partant de cette décomposition la résolution du problème revient à trouver la solution d'un système d'équations différentielles linéaires du premier ordre. Nous avons en effet d'une part :

$$i\hbar \frac{\partial}{\partial t} |\psi_0; t\rangle_I = i\hbar \sum_m \dot{c}_m(t) \, |m\rangle \tag{18.10}$$

et d'autre part, d'après l'éqn. (18.6) :

$$i\hbar \frac{\partial}{\partial t} |\psi_0; t\rangle_I = \mathcal{V}(t)|_I \, |\psi_0; t\rangle_I = \sum_m c_m(t) e^{\frac{it}{\hbar}\mathcal{H}_0} \mathcal{V}(t) e^{-\frac{it}{\hbar}\mathcal{H}_0} \, |m\rangle = \sum_m e^{\frac{it}{\hbar}\mathcal{H}_0} \mathcal{V}(t) e^{-\frac{iE_m t}{\hbar}} c_m(t) \, |m\rangle \tag{18.11}$$

Projetons ces deux equations sur $|n\rangle$. Nous obtenons :

$$i\hbar \sum_m \dot{c}_m(t) \, \langle n|m\rangle = \sum_m e^{-\frac{iE_m t}{\hbar}} \, \langle n| \, e^{\frac{it}{\hbar}\mathcal{H}_0} \mathcal{V}(t) c_m(t) \, |m\rangle \tag{18.12}$$

Soit, en faisant agir l'opérateur d'évolution à gauche et en utilisant l'orthonormalité de la base :

$$\boxed{i\hbar \, \dot{c}_n(t) = \sum_m e^{-\frac{i(E_m - E_n)t}{\hbar}} c_m(t) \, \langle n| \, \mathcal{V}(t) \, |m\rangle} \tag{18.13}$$

Si nous introduisons les pulsations :

$$\omega_{nm} \overset{\text{déf.}}{=} \frac{E_n - E_m}{\hbar} \tag{18.14}$$

ce système prend la forme différentiel prend la forme :

$$i\hbar \, \dot{c}_n(t) = \sum_m e^{i\omega_{nm} t} c_m(t) \, \langle n| \, \mathcal{V}(t) \, |m\rangle \ . \tag{18.15}$$

Une fois le système différentiel (18.13) résolu pour les fonctions $c_n(t)$ il est aisé de donner le vecteur d'état solution dans la représentation de Schrödinger. D'après léqn. (18.3) nous avons

$$|\psi_0; t\rangle_S = e^{-\frac{it}{\hbar}\mathcal{H}_0} \, |\psi_0; t\rangle_I = \sum_n c_n(t) e^{-\frac{it}{\hbar}\mathcal{H}_0} \, |n\rangle \quad \Longrightarrow \quad \boxed{|\psi_0; t\rangle_S = \sum_n e^{-\frac{iE_n t}{\hbar}} c_n(t) \, |n\rangle} \tag{18.16}$$

Nous voyons bien dans cette expression la factorisation de la dépendance temporelle en une dépendance triviale donnée par l'évolution par $\mathcal{H}_0$ et la dépendance due au terme $\mathcal{V}(t)$ du hamiltonien, encodée dans les fonctions $c_n(t)$ (qui deviennent des constantes si $\mathcal{V} = 0$).

Exemple 1 (résonance magnétique).

Considérons un spin 1/2 (par exemple porté par un électron) dans un champ magnétique composé d'un champ constant selon l'axe Oz et d'un champ tournant dans le plan (Ox, Oy) avec une pulsation ω :

$$\vec{B} = B_0 \vec{e}_z + B_1 \big(\vec{e}_x \cos \omega t + \vec{e}_y \sin \omega t \big) \, . \tag{18.17}$$

Le hamiltonien du système est donné par

$$\mathcal{H} = -\vec{\mu}\cdot\vec{B} = \frac{g_s\mu_B B_0}{\hbar}\mathcal{S}_z + \frac{g_s\mu_B B_1}{\hbar}\big(\mathcal{S}_x\cos\omega t + \mathcal{S}_y\sin\omega t\big) \tag{18.18}$$

Introduisons les pulsations

$$\omega_0 \overset{\text{déf.}}{=} \frac{g_s\mu_B B_0}{\hbar}\,, \quad \omega_1 \overset{\text{déf.}}{=} \frac{g_s\mu_B B_1}{\hbar}\,. \tag{18.19}$$

Le hamiltonien (18.19) prend alors la forme matricielle

$$\mathcal{H} \mathrel{\widehat{=}} \frac{\hbar}{2}\begin{pmatrix}\omega_0 & \omega_1 e^{-i\omega t}\\ \omega_1 e^{i\omega t} & -\omega_0\end{pmatrix} = \underbrace{\frac{\hbar\omega_0}{2}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}}_{\mathcal{H}_0} + \underbrace{\frac{\hbar\omega_1}{2}\begin{pmatrix}0 & e^{-i\omega t}\\ e^{i\omega t} & 0\end{pmatrix}}_{\mathcal{V}(t)}\,. \tag{18.20}$$

Une base de vecteurs propres de $\mathcal{H}_0$ est évidemment fournie par les états propres de $\mathcal{S}_z$, $\mathcal{S}_z|\pm\rangle = \pm\frac{\hbar}{2}|\pm\rangle$ correspondant respectivement à $\binom{1}{0}$ et $\binom{0}{1}$. On suppose le système initialement préparé dans l'état $|+\rangle$. Nous posons donc :

$$|\psi_0; t\rangle_I = c_+(t)\,|+\rangle + c_-(t)\,|-\rangle \;, \quad c_+(0) = 1\,, \quad c_-(0) = 0\,, \tag{18.21}$$

et les coefficients $c_\pm(t)$ sont solutions du système différentiel :

$$\left\{\begin{array}{rcl} i\dot{c}_+(t) &=& \frac{\omega_1}{2}e^{i(\omega_0-\omega)t}c_-(t)\\ i\dot{c}_-(t) &=& \frac{\omega_1}{2}e^{-i(\omega_0-\omega)t}c_+(t)\end{array}\right. \implies \frac{d}{dt}\begin{pmatrix}c_+\\ c_-\end{pmatrix} = -\frac{i\omega_1}{2}\begin{pmatrix}0 & e^{i(\omega_0-\omega)t}\\ e^{-i(\omega_0-\omega)t} & 0\end{pmatrix}\begin{pmatrix}c_+\\ c_-\end{pmatrix} \tag{18.22}$$

La résolution de ce système différentiel, avec les conditions initiales $c_+(0) = 1$ et $c_-(0) = 0$, est laissée en exercice (voir également le sujet de TD n°11). On trouve en particulier :

$$c_-(t) = \frac{-i\omega_1}{\sqrt{\omega_1^2 + (\omega-\omega_0)^2}}e^{\frac{i}{2}(\omega-\omega_0)t}\sin\left(\frac{t}{2}\sqrt{\omega_1^2 + (\omega-\omega_0)^2}\right)\,, \tag{18.23}$$

D'où nous déduisons la probabilité de retournement du spin en fonction du temps (oscillations de Rabi) :

$$p_{+\to-}(t) = |c_-(t)|^2 = \frac{\omega_1^2}{\omega_1^2 + (\omega-\omega_0)^2}\,\sin^2\left(\frac{t}{2}\sqrt{\omega_1^2 + (\omega-\omega_0)^2}\right)\,. \tag{18.24}$$

Cette probabilité est périodique, avec une période $T = 2\pi(\omega_1^2 + (\omega-\omega_0)^2)^{-1/2}$. Dans l'intervalle de temps $0 \leqslant t < T/2$ la probabilité de retournement croît car le spin émet de l'énergie et le système peuple le niveau $|-\rangle$ d'énergie la plus basse. Ensuite dans l'intervalle $T/2 \leqslant t < T$ la probabilité de retournement décroît car le spin absorbe de l'énergie ; la population du niveau $|+\rangle$ d'énergie la plus grande augmente.

Lorsque $\omega \to \omega_0$ il apparaît un phénomène de résonance car la probabilité oscille alors entre 0 et 1. La mesure de ce phénomène permet en particulier des mesures très précises du moment magnétique nucléaire des atomes, comme le moment magnétique de spin est proportionnel à ω_1, voir éqn. (18.19).

Nous retrouverons plus loin ce modèle pour décrire un atome à deux niveaux en présence d'un champ électromagnétique classique oscillant.

18.2 Théorie des perturbations dépendantes du temps

En général la résolution exacte du système différentiel (18.13) est impossible et il faut faire appel à des méthodes d'approximation. Si on peut considérer le terme dépendant du temps $\mathcal{V}(t)$ dans (18.2), comme une perturbation du hamiltonien indépendant du temps $\mathcal{H}_0$, dans un sens qu'il faudra préciser, on peut envisager de rechercher une solution du système différentiel (18.13) comme un développement en série.

Pour organiser le développement perturbatif de la solution, nous réexprimons le hamiltonien comme

$$\mathcal{H} = \mathcal{H}_0 + \epsilon \mathcal{V}(t)\,, \tag{18.25}$$

où ϵ est un paramètre sans dimension, et nous développons les fonctions $c_n(t)$ en puissances de ϵ :

$$c_n(t) = c_n^{(0)}(t) + \epsilon c_n^{(0)}(t) + \epsilon^2 c_n^{(2)}(t) + \cdots \tag{18.26}$$

À la fin du calcul il est possible de prendre la limite $\epsilon \to 1$ pour retrouver le hamiltonien initial. Nous avons à résoudre le système différentiel :

$$\dot{c}_n^{(0)} + \epsilon \dot{c}_n^{(0)} + \epsilon^2 \dot{c}_n^{(0)} + \cdots = \frac{\epsilon}{i\hbar} \sum_m e^{-\frac{i(E_m - E_n)t}{\hbar}} \langle n|\, \mathcal{V}(t)\, |m\rangle \left(c_m^{(0)} + \epsilon c_m^{(1)} + \epsilon^2 c_m^{(2)} + \cdots \right) . \tag{18.27}$$

Les conditions initiales du problème sont prises en compte par le terme d'ordre zéro du développement. Développons l'état initial sur la base des états propres de $\mathcal{H}_0$:

$$|\psi_0\rangle = \sum_n \alpha_n |n\rangle \ . \tag{18.28}$$

Nous imposons alors

$$c_n^{(0)}(t) = \alpha_n \quad , \qquad \forall t \geqslant 0\,, \tag{18.29a}$$

$$c_n^{(k)}(t=0) = 0 \quad , \qquad \forall k \geqslant 1\,. \tag{18.29b}$$

en effet, pour la résolution de (18.27) à l'ordre zéro en ϵ le terme $\mathcal{V}(t)$ du hamiltonien n'intervient pas donc, travaillant dans la représentation d'interaction, nous n'avons pas d'évolution temporelle et les coefficients $c_n^{(0)}$ sont constants dans le temps.

La résolution de ce problème est en un certain sens plus simple que pour la théorie des perturbations stationnaires car, pour un ordre k donné, nous avons une équation différentielle du premier ordre pour chaque fonction $c_n^{(k)}$ avec un terme source dépendant de la solution à l'ordre précédent :

$$\forall k \in \mathbb{N}^\star \ , \quad \boxed{\dot{c}_n^{(k)} = \frac{1}{i\hbar} \sum_m e^{-\frac{i(E_m - E_n)t}{\hbar}} \langle n|\, \mathcal{V}(t)\, |m\rangle\, c_m^{(k-1)}} \tag{18.30}$$

qui s'intègre immédiatement, avec les conditions initiales (18.29), comme :

$$c_n^{(k)}(t) = \frac{1}{i\hbar} \sum_m \int_0^t dt'\, e^{-\frac{i(E_m - E_n)t'}{\hbar}} \langle n|\, \mathcal{V}(t')\, |m\rangle\ c_m^{(k-1)}(t') \tag{18.31}$$

Ainsi, partant des conditions initiales (18.29), on peut premièrement calculer la correction au premier ordre,

$$c_n^{(1)}(t) = \frac{1}{i\hbar} \sum_m \alpha_m \int_0^t dt' \, e^{-\frac{i(E_m - E_n)t'}{\hbar}} \langle n | \mathcal{V}(t') | m \rangle \tag{18.32}$$

Puis, itérativement, les ordres suivants.

Probabilité de transition Si on prépare le système initialement dans un état propre $|m\rangle$ de $\mathcal{H}_0$, il peut effectuer une transition vers un état $|n\rangle$ d'énergie différente, car l'énergie n'est plus conservée en présence de la perturbation dépendante du temps. On peut calculer en fonction du temps la probabilité que le système fasse une transition vers l'état $n \neq m$:

$$p_{m\to n}(t) = |\langle n|m;t\rangle|^2 = \left| \sum_p \left(c_p^{(0)}(t) + c_p^{(1)}(t) + \cdots \right) \langle n|p\rangle \right|^2 = \left| c_n^{(1)}(t) + c_n^{(2)}(t) + \cdots \right|^2 , \tag{18.33}$$

où les fonctions $c_p^{(k)}(t)$ ont été obtenues avec les conditions initiales $c_p^{(0)} = \delta_{m,p}$. Au premier ordre en théorie des perturbations cela donne :

$$p_{m\to n}(t) = \frac{1}{\hbar^2} \left| \int_0^t dt' \, e^{-i\omega_{mn}t'} \langle n | \mathcal{V}(t') | m \rangle \right|^2 . \tag{18.34}$$

On peut en déduire, à cet ordre, la probabilité de quitter l'état $|m\rangle$ (pour un état quelconque) en fonction du temps :

$$p_{m\to\star}(t) = \frac{1}{\hbar^2} \sum_{n\neq m} \left| \int_0^t dt' \, e^{-i\omega_{mn}t'} \langle n | \mathcal{V}(t') | m \rangle \right|^2 . \tag{18.35}$$

On peut estimer que la théorie des perturbations dépendantes du temps est valable tant que $p_{m\to\star}(t) \ll 1$, ce qui pose des contraintes à la fois sur $\mathcal{V}(t)$ et sur l'intervalle de temps $0 \leqslant t \leqslant t_{MAX}$ où nous pouvons utiliser cette méthode.

La contribution au second ordre en théorie des perturbations est également intéressante. Nous avons

$$\begin{aligned} c_n^{(2)}(t) &= \frac{1}{i\hbar} \sum_p \int_0^t dt' \, e^{-i\omega_{pn}t'} \langle n | \mathcal{V}(t') | p \rangle \; c_p^1(t') \\ &= -\frac{1}{\hbar^2} \int_0^t dt' \int_0^{t'} dt'' \sum_p e^{-i\omega_{mp}t'' - i\omega_{pn}t'} \langle n | \mathcal{V}(t') | p \rangle \langle p | \mathcal{V}(t'') | m \rangle \end{aligned} \tag{18.36}$$

On peut interpréter ce résultat de la manière suivante : la transition de l'état $|m\rangle$ vers l'état $|n\rangle$ reçoit au second ordre la contribution de *transitions virtuelles* de l'état $|m\rangle$ vers un état $|p\rangle$ intermédiaire, puis de l'état $|p\rangle$ vers l'état $|n\rangle$.

Exemple 2 (retour sur la résonance magnétique). Reprenons le problème de la résonance magnétique, voir l'exemple 1. Bien que le problème admette une solution exacte, voir éqn. (18.23), nous pouvons l'examiner du point de vue perturbatif et en chercher une solution approchée. Avec les conditions initiales

$$|\psi_0; t=0\rangle = |+\rangle \quad \Longrightarrow \quad c_+^{(0)} = 1 \ , \quad c_-^{(0)} = 0 \tag{18.37}$$

nous obtenons directement au premier ordre en théorie des perturbations dépendant du temps (si $\omega \neq \omega_0$) :

$$c_+^{(1)} = \frac{1}{i\hbar}\int_0^t dt'\, e^{i\omega_0 t'} \langle +|\, \mathcal{V}(t')\, |-\rangle\, c_-^{(0)} = 0 \tag{18.38a}$$

$$\begin{aligned} c_-^{(1)} &= \frac{1}{i\hbar}\int_0^t dt'\, e^{-i\omega_0 t'} \langle -|\, \mathcal{V}(t')\, |+\rangle\, c_+^{(0)} \\ &= \frac{\omega_1}{2i}\int_0^t dt'\, e^{-i\omega_0 t'} e^{i\omega t'} \\ &= \frac{-i\omega_1}{\omega-\omega_0} e^{\frac{i}{2}(\omega-\omega_0)t} \sin\left(\frac{(\omega-\omega_0)t}{2}\right) \end{aligned} \tag{18.38b}$$

Le résultat obtenu pour c_- correspond bien au développement de la solution exacte, voir l'éqn. (18.23), au premier ordre dans le paramètre sans dimension $\omega_1/|\omega-\omega_0|$.

Nous en déduisons une formule perturbative pour la probabilité de retournement au premier ordre :

$$p_{+\to-}(t) = |c_-(t)|^2 = \left(\frac{\omega_1}{\omega-\omega_0}\right)^2 \sin^2\left(\frac{(\omega-\omega_0)t}{2}\right) \tag{18.39}$$

Examinons la validité du développement perturbatif dans le cadre de cet example. Il faut distinguer deux cas de figure :

— lorsque $\omega_1/|\omega-\omega_0| \ll 1$ le système est loin de la résonance et la probabilité de retournement reste faible quelque soit t. L'approche perturbative reste alors valable à tout instant ;

— À la résonance, $\omega = \omega_0$ et on trouve

$$p_{+\to-}(t) = |c_-(t)|^2 = \left(\frac{\omega_1 t}{2}\right)^2 . \tag{18.40}$$

cette fonction n'est pas bornée ce qui invalide l'interprétation comme probabilité si t est arbitraire. En conséquence, l'approche perturbative au problème n'est valable qu'aux temps courts :

$$p_{+\to-}(t) \ll 1 \quad \Longrightarrow \quad t \ll 1/\omega_1 \tag{18.41}$$

En dehors de la résonance, un développement limité de l'éqn. (18.39) au premier ordre en $(\omega-\omega_0)t$ nous indique que l'approche perturbative est valide si $t \ll 1/\omega_1$ et en même temps $t \ll 1/|\omega-\omega_0|$, indépendamment du fait que $\omega_1/|\omega-\omega_0|$ soit petit ou non.

18.3 Règle d'or de Fermi

Le formalisme développé dans la section précédente permet de calculer la probabilité pour un système quantique de faire une transition d'un niveau à un autre de son spectre discret. Pour traiter de phénomènes comme l'ionisation d'un atome, où un électron passe d'un état lié à un état du continuum sous l'effet d'une perturbation causée par un champ électromagnétique, nous obtiendrons dans cette section une formule importante, dite *règle d'or de Fermi*, donnant le taux de transition du système aux temps longs.

18.3.1 Perturbation constante au cours du temps

Considérons un système possédant un spectre discret, comme précédemment, et une perturbation constante pour $t \geqslant 0$:

$$\mathcal{H}(t) = \begin{cases} \mathcal{H}_0 & t < 0 \\ \mathcal{H}_0 + \mathcal{V}_0 & t \geqslant 0 \end{cases} \tag{18.42}$$

En pratique, nous pouvons bien sûr imaginer « allumer » la perturbation, par exemple un champ constant, à $t = 0$, mais nous pouvons aussi appliquer les resultats que nous obtiendrons lorsque le hamiltonien est donné par $\mathcal{H}_0 + \mathcal{V}_0$ à tout instant, si le système est préparé à $t = 0$ dans un état propre de $\mathcal{H}_0$.

Si le système est initialement dans un état $|n\rangle$, la probabilité de transition vers un état $|m\rangle$ au premier ordre est donnée, suivant la formule (18.32), par :

$$p_{n\to m}(t) = \frac{1}{\hbar^2}\left|\int_0^t dt'\, e^{-\frac{i(E_n-E_m)t'}{\hbar}} \langle m|\,\mathcal{V}_0\,|n\rangle\right|^2 = \mathrm{sinc}^2\left(\frac{(E_n-E_m)t}{2\hbar}\right)\,\frac{|\langle m|\,\mathcal{V}_0\,|n\rangle|^2}{\hbar^2}\,t^2\,. \tag{18.43}$$

écrit en terme de la fonction *sinus cardinal*, $\mathrm{sinc}(x) \overset{\text{déf.}}{=} \sin(x)/x$.

Vue comme une fonction de la différence en énergie $E_n - E_m$ entre l'état initial et l'état final, la probabilité (18.43) est de plus en plus piquée au voisinage de $E_n = E_m$, sa largeur étant proportionnelle à $1/t$,. En d'autres termes la violation de la conservation de l'énergie s'attenue avec le temps. Ce résultat n'est guère surprenant, la dépendance temporelle du système étant localisée en $t = 0$.

Nous pouvons également considérer la probabilité pour le système de quitter l'état n vers un état quelconque, qui est donnée par :

$$p_{n\to\star}(t) = \frac{t^2}{\hbar^2}\sum_m |\langle m|\,\mathcal{V}_0\,|n\rangle|^2\,\mathrm{sinc}^2\left(\frac{(E_n-E_m)t}{2\hbar}\right) \tag{18.44}$$

Au temps longs, cette somme est dominée par les contributions des états finaux dont l'énergie E_m proche de l'énergie E_n de l'état initial.

Considérons un système pour lequel il existe un grand nombre d'états d'énergie proche de celle de l'état initial. Nous appelerons dorénavant $|I\rangle$ l'état initial et E_I son énergie. On

peut alors remplacer dans l'expression (18.44) la somme sur un continuum d'états finaux $|F\rangle$, d'énergie E_F :

$$\sum_m \longrightarrow \int dE_F\, \rho(E_F) \tag{18.45}$$

où $\rho(E)$ représente la densité d'états du système ; autrement dit, $\rho(E_F)\, dE_F$ correspond au nombre d'états dans l'intervalle $]E_F - dE_F/2, E_F + dE_F/2[$. La probabilité de quitter l'état initial $|n\rangle$ est donc donnée par : [1]

$$p_{I\to\star}(t) = \frac{t^2}{\hbar^2} \int dE_F\, \rho(E_F)\, |\langle F|\, \mathcal{V}_0\, |I\rangle|^2 \,\mathrm{sinc}^2\left(\frac{(E_F - E_I)t}{2\hbar}\right) \tag{18.46}$$

Aux temps longs cette formule peut se simplifier drastiquement car, comme nous l'avons vu, seul le voisinage de $E_I = E_F$ contribue. Utilisons pour cela la propriété

$$\lim_{\alpha\to\infty} \int_{-\infty}^{+\infty} dx\, \alpha\, \mathrm{sinc}^2(\alpha x) f(x) = \pi f(0)\,, \tag{18.47}$$

ou, exprimé autrement,

$$\lim_{\alpha\to\infty} \alpha\, \mathrm{sinc}^2(\alpha x) = \pi\delta(x)\,. \tag{18.48}$$

Nous obtenons alors la probabilité pour le système d'effectuer une transition de l'état $|I\rangle$ vers les états du continuum est donnée par :

$$p_{I\to\star}(t) \overset{t\to\infty}{\sim} \frac{\pi t}{\hbar^2} \int dE_F\, \rho(E_F)\, |\langle F|\, \mathcal{V}_0\, |I\rangle|^2\, \delta\left(\frac{E_F - E_I}{2\hbar}\right) = \frac{2\pi t}{\hbar} \rho(E_I)\, |\langle F|\, \mathcal{V}_0\, |I\rangle|^2 \Big|_{E_F = E_I}\,. \tag{18.49}$$

Ainsi, aux temps longs, on n'observe de transition que de l'état initial $|I\rangle$ vers des états du continuum de même énergie.

Nous voyons également que la probabilité de transition croît linéairement avec le temps aux temps longs. On considère alors plutôt le taux de transition, qui consiste en la dérivée temporelle de la probabilité :

$$\Gamma_I(t) \overset{\text{déf.}}{=} \frac{d}{dt} p_{I\to\star}(t)\,, \tag{18.50}$$

et le calcul précédent nous indique que, aux temps longs le taux de transition tend vers une constante :

$$\boxed{\Gamma_I \overset{t\to\infty}{\sim} \frac{2\pi}{\hbar} \rho(E_I)\, |\langle F|\, \mathcal{V}_0\, |I\rangle|^2 \Big|_{E_F = E_I}} \tag{18.51}$$

Cette formule importante porte le nom de *règle d'or de Fermi*.

De manière alternative, on peut écrire la règle d'or de Fermi sous la forme suivante :

$$\Gamma_I \overset{t\to\infty}{\sim} \frac{2\pi}{\hbar}\, |\langle F|\, \mathcal{V}_0\, |I\rangle|^2\, \delta(E_F - E_I) \tag{18.52}$$

étant entendu que cette quantité doit s'intégrer avec $\int dE_F\, \rho(E_F)$.

Examinons plus en détail le domaine de validité de la formule perturbative (18.51) en fonction du temps. Nous devons satisfaire deux exigences a priori contradictoires :

1. On suppose implicitement que les états finaux dont l'énergie est dans l'intervalle $]E_F - dE_F/2, E_F + dE_F/2[$ ont tous les mêmes éléments de matrice avec l'état initial $|I\rangle$, du moins à une phase près.

— la probabilité de transition doit rester très inférieure à l'unité, ce qui donne une borne supérieure sur t :

$$t \ll 1/\Gamma_I \tag{18.53}$$

— pour que la formule (18.48) donne une bonne approximation du résultat, il faut que la largeur $\hbar/t$ de la distribution de probabilité vue comme fonction de E_F soit beaucoup plus faible que l'échelle caractéristique de variation de la fonction $E_F \mapsto \rho(E_F)|\langle F|\mathcal{V}_0|I\rangle|^2$ autour du maximum en E_I ; ce second critère fournit une borne inférieure pour t.

18.3.2 Perturbation sinusoïdale

Reprenons le raisonnement de la sous-seciton précédente, en ajoutant cette fois une perturbation sinusoïdale pour pour $t \geqslant 0$, à la pulsation ω :

$$\mathcal{H}(t) = \begin{cases} \mathcal{H}_0 & t < 0 \\ \mathcal{H}_0 + \mathcal{V}_0 e^{i\omega t} + \mathcal{V}_0^\dagger e^{-i\omega t} & t \geqslant 0 \end{cases} \tag{18.54}$$

Ce type de situation se rencontrera fréquemment dans l'étude de l'interaction d'un atome avec un champ électromagnétique classique, voir le cours n° 19.

Premièrement la probabilité de transition d'un niveau $|n\rangle$ à un niveau $|m\rangle$ est donnée au premier ordre par :

$$\begin{aligned} p_{n\to m}(t) &= \frac{1}{\hbar^2}\left|\int_0^t dt'\, e^{-\frac{i(E_n-E_m)t'}{\hbar}}\left(\langle m|\mathcal{V}_0|n\rangle\, e^{i\omega t'} + \langle m|\mathcal{V}_0^\dagger|n\rangle\, e^{-i\omega t'}\right)\right|^2 \\ &= \frac{1}{\hbar^2}\left|\int_0^t dt'\,\left(\langle m|\mathcal{V}_0|n\rangle\, e^{-\frac{i(E_n-E_m-\hbar\omega)t'}{\hbar}} + \langle m|\mathcal{V}_0^\dagger|n\rangle\, e^{-\frac{i(E_n-E_m+\hbar\omega)t'}{\hbar}}\right)\right|^2 \\ &= \frac{1}{\hbar}\left|\langle m|\mathcal{V}_0|n\rangle\,\frac{e^{-\frac{i(E_n-E_m-\hbar\omega)t'}{\hbar}}-1}{E_n-E_m-\hbar\omega} + \langle m|\mathcal{V}_0^\dagger|n\rangle\,\frac{e^{-\frac{i(E_n-E_m+\hbar\omega)t'}{\hbar}}-1}{E_n-E_m+\hbar\omega}\right|^2 \end{aligned} \tag{18.55}$$

Pour donner le comportement du système aux temps longs nous pouvons reprendre avec quelques modifications l'étude faite précédemement dans le cas d'une perturbation constante. Il faut ici distinguer deux cas de figure, suivant que l'un ou l'autre des deux termes dans l'éqn. (18.55) est le terme dominant.

FIGURE 18.1 – Émission stimulée (gauche) et absorption (droite).

Émission stimulée

Considérons que le système, en présence de la perturbation sinusoïdale, évolue d'un état initial d'énergie E_I vers un état d'énergie E_F telle que :

$$E_F = E_I - \hbar\omega \tag{18.56}$$

Si nous voyons la perturbation $\mathcal{V}(t)$ comme un champ électromagnétique extérieur, on interprète cette relation comme l'*émission stimulée* par le système d'un quantum de rayonnement d'énergie $\hbar\omega$, lorsqu'il décroît d'un état d'énergie E_I vers un état d'énergie E_F inférieure (voir fig. 18.1).

On peut alors aisément calculer, en présence d'un continuum d'états finaux, le taux de transition associé par un raisonnement similaire à celui conduisant à l'éqn. (18.51) :

$$\boxed{\Gamma_I \overset{t\to\infty}{\sim} \frac{2\pi}{\hbar}\rho(E_I - \hbar\omega)\left|\langle F|\,\mathcal{V}_0\,|I\rangle\right|^2\Big|_{E_F=E_I-\hbar\omega}} \tag{18.57}$$

Absorption

Considérons que le système, en présence de la perturbation sinusoïdale, évolue d'un état initial d'énergie E_I vers un état d'énergie E_F telle que :

$$E_F = E_I + \hbar\omega \tag{18.58}$$

Si nous voyons la perturbation $\mathcal{V}(t)$ comme un champ électromagnétique extérieur, on interprète cette relation comme l'*absorption* par le système d'un quantum de rayonnement d'énergie $\hbar\omega$, lorsqu'il qui lui permet de passer de l'état d'énergie E_I vers un état d'énergie E_F supérieure (voir fig. 18.1).

On peut alors aisément calculer, en présence d'un continuum d'états finaux, le taux de transition :

$$\boxed{\Gamma_I \overset{t\to\infty}{\sim} \frac{2\pi}{\hbar}\rho(E_I + \hbar\omega)\left|\langle F|\,\mathcal{V}_0^\dagger\,|I\rangle\right|^2\Big|_{E_F=E_I+\hbar\omega}} \tag{18.59}$$

ou bien :

$$\Gamma_I \overset{t\to\infty}{\sim} \frac{2\pi}{\hbar}\left|\langle F|\,\mathcal{V}_0^\dagger\,|I\rangle\right|^2\delta(E_F - E_I - \hbar\omega) \tag{18.60}$$

À retenir

— La représentation d'interaction et le système différentiel associé ;

— la théorie des perturbations dépendantes du temps ;

— la règle d'or de Fermi pour des perturbations constantes ;

— la règle d'or de Fermi pour des perturbations sinusoïdales.

Cours n°19

Interaction matière-rayonnement

Une application majeure de la théorie des perturbations dépendant du temps est l'étude des interactions entre la matière et le rayonnement électromagnétique, en particulier les phénomènes d'absorption et d'émission spontanée. Dans ce cours nous aborderons permièrement l'étude du phénomène en considérant le champ électromagnétique comme un objet classique, puis nous donnerons une présentation succinte de sa quantification canonique.

19.1 Hamiltonien d'une particule chargée

Le hamiltonien d'une particule chargée dans un champ électromagnétique s'obtient plus naturellement en partant de la formulation relativiste du problème, dans laquelle il est simple d'introduire un couplage minimal entre la particule et le champ.

Considérons une particule libre relativiste de coordonnées $x^\mu = (ct, \vec{x}(t))$, plongée dans un champ électromagnétique décrit par le quadri-potentiel $A^\mu = (\phi/c, \vec{A})$. L'action d'une particule libre relativiste de masse m et de charge q est donnée par

$$\mathcal{S} = -mc \int ds - q \int A_\mu dx^\mu , \tag{19.1}$$

avec l'élement d'intégration sur la ligne d'univers :

$$ds = \sqrt{c^2 dt^2 - d\vec{x}^2} = c dt \sqrt{1 - \vec{v}^2/c^2} \tag{19.2}$$

Le second terme représente l'intégrale du quadri-potentiel sur la ligne d'univers et peut s'exprimer comme :

$$A_\mu dx^\mu = \big(\phi - \vec{A} \cdot \vec{v}\big) dt \tag{19.3}$$

Notons que le couplage au champ électromagnétique est complètement fixé, à une constante multiplicative près, par les symétries du problème (groupe de Lorentz). Nous introduirons également comme il est d'usage le facteur de Lorentz

$$\gamma \overset{\text{déf.}}{=} \frac{1}{\sqrt{1 - \vec{v}^2/c^2}} . \tag{19.4}$$

Nous en déduisons donc le lagrangien de la particule :

$$\mathcal{L} = -\frac{mc^2}{\gamma} - q\phi + q\vec{A}\cdot\vec{v}, \tag{19.5}$$

puis l'impulsion généralisée de la particule relativiste :

$$\vec{P} = \frac{\partial\mathcal{L}}{\partial\vec{v}} = \gamma m\vec{v} + q\vec{A} = \vec{p} + q\vec{A}\,, \tag{19.6}$$

où

$$\vec{p} = \gamma m\vec{v} \tag{19.7}$$

désigne la *quantité de mouvement* qui satisfait à la relation

$$\gamma mc^2 = \sqrt{m^2c^4 + p^2c^2} \tag{19.8}$$

Le hamiltonien de la particule est finalement donné par :

$$\mathcal{H}_r = \vec{v}\cdot\vec{P} - \mathcal{L} = \gamma mv^2 + q\vec{A}\cdot\vec{v} + \frac{mc^2}{\gamma} + q\phi - q\vec{A}\cdot\vec{v} = \gamma mc^2 + q\phi$$

$$\Longrightarrow \boxed{\mathcal{H}_r = \sqrt{m^2c^4 + \left(\vec{P} - q\vec{A}\right)^2 c^2} + q\phi} \tag{19.9}$$

Dans la limite non-relativiste $v^2/c^2 \ll 1$, nous obtenons, en soustrayant l'énergie au repos de la particule,

$$\boxed{\mathcal{H} = \frac{\left(\vec{P} - q\vec{A}\right)^2}{2m} + q\phi} \tag{19.10}$$

donnant le couplage minimal entre le champ électromagnétique et la particule.

L'impulsion généralisée $\vec{P}$ étant la variable canoniquement conjuguée à la position $\vec{x}$, le principe de quantification canonique nous donne en mécanique quantique non-relativiste l'opérateur hamiltonien suivant :

$$\mathcal{H} = \frac{\left(\vec{\mathcal{P}} - q\vec{A}(\vec{\mathcal{X}})\right)^2}{2m} + q\phi(\vec{\mathcal{X}})\,, \tag{19.11}$$

avec la relation de commutation canonique [1]

$$[\,\mathcal{X}_i,\,\mathcal{P}_j\,] = i\hbar\delta_{ij}\,. \tag{19.12}$$

Si la particule possède en outre un spin 1/2, nous obtenons l'opérateur hamiltonien complet

$$\boxed{\mathcal{H} = \frac{\left(\vec{\mathcal{P}} - q\vec{A}(\vec{\mathcal{X}})\right)^2}{2m} + q\phi(\vec{\mathcal{X}}) - g_s\,\frac{q}{2m}\,\vec{\mathcal{S}}\cdot\vec{B}(\vec{\mathcal{X}})} \tag{19.13}$$

Rappelons que le couplage entre le spin et le champ magnétique s'obtient à partir de l'équation de Dirac.

1. Même si $\vec{\mathcal{P}}$ n'est pas invariant par transformation de jauge, voir l'éqn. (19.24), le commutateur est invariant car, pour toute transformation de jauge $\Lambda(\vec{x},t)$, $\left[\,\vec{\nabla}\Lambda(\vec{\mathcal{X}},t),\,\mathcal{X}_i\,\right] = 0$.

Approche non-relativiste

On peut naturellement retrouver le hamiltonien (19.10) par une approche non-relativiste. Les équations du mouvement pour une particule chargée,

$$m\ddot{\vec{x}} = q\big(\vec{E} + \dot{\vec{x}} \wedge \vec{B}\big)\,, \tag{19.14}$$

découlent du lagrangien

$$\mathcal{L} = \frac{1}{2} m\dot{\vec{x}}^2 + q\dot{\vec{x}} \cdot \vec{A} - q\phi\,. \tag{19.15}$$

En effet

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}_i} - \frac{\partial \mathcal{L}}{\partial x_i} = m\ddot{x}_i + q\partial_t A_i + q\sum_j \dot{x}_j \partial_j A_i - q\sum_j \dot{x}_j \partial_i A_j - q\partial_i \phi \tag{19.16}$$

qui redonne bien le résultat attendu en reconnaissant le double produit vectoriel $\vec{u} \wedge (\vec{v} \wedge \vec{w}) = (\vec{u}\cdot\vec{w})\vec{v} - (\vec{u}\cdot\vec{v})\vec{w}$. Finalement le lagrangien (19.15) donne l'impulsion canoniquement conjuguée au vecteur position,

$$\vec{P} = \frac{\partial \mathcal{L}}{\partial \dot{\vec{x}}} = m\dot{\vec{x}} + q\vec{A}\,, \tag{19.17}$$

d'où on obtient le hamiltonien

$$\mathcal{H} = \dot{\vec{x}} \cdot \big(m\dot{\vec{x}} + q\vec{A}\big) - \frac{1}{2} m\dot{\vec{x}}^2 - q\dot{\vec{x}} \cdot \vec{A} + q\phi = \frac{1}{2} m\dot{x}^2 - q\phi = \frac{1}{2m}\big(\vec{P} - q\vec{A}\big)^2 + q\phi\,. \tag{19.18}$$

19.2 Invariance de jauge

Le champ électromagnétique est décrit localement par le potentiel vecteur $\vec{A}$ et le potentiel scalaire ϕ, tels que

$$\vec{B} = \vec{\nabla} \wedge \vec{A}\,, \tag{19.19a}$$

$$\vec{E} = -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t}\,. \tag{19.19b}$$

Cette description en termes des potentiels est redondante car les champs $\vec{E}$ et $\vec{B}$, et en conséquence les équations de Maxwell, sont invariants par les *transformations de jauge* :

$$\vec{A} \mapsto \vec{A}_\Lambda \overset{\text{déf.}}{=} \vec{A} + \vec{\nabla}\Lambda\,, \tag{19.20a}$$

$$\phi \mapsto \phi_\Lambda \overset{\text{déf.}}{=} \phi - \frac{\partial \Lambda}{\partial t}\,, \tag{19.20b}$$

où $\Lambda(\vec{x}, t)$ est une fonction différentiable quelconque.

L'opérateur hamiltonien (19.11) d'une particule classique dépendant explicitement des potentiels électromagnétiques, il est nécessaire de comprendre le rôle des transformations

de jauge en mécanique quantique. Ces transformations devant laisser invariantes les quantités physiques mesurables, les valeurs moyennes d'observables doivent correspondre à des grandeurs *invariantes de jauge*.

Les transformations de jauge devant en particuler préserver le produit scalaire entre vecteurs de l'espace des états, nous pouvons associer à l'action d'une transformation de jauge un opérateur unitaire $\mathcal{G}$:

$$|\psi(t)\rangle \longmapsto |\psi(t)\rangle_\Lambda = \mathcal{G}[\Lambda(\vec{\mathcal{X}},t)]\,|\psi(t)\rangle \ , \quad \mathcal{G}^\dagger\mathcal{G} = \mathfrak{I}\,. \tag{19.21}$$

La valeur moyenne de l'opérateur position dans un état quelconque doit naturellement être invariante par transformation de jauge. Nous avons donc

$$\langle\psi(t)|\,\vec{\mathcal{X}}\,|\psi(t)\rangle = {}_\Lambda\langle\psi(t)|\,\vec{\mathcal{X}}\,|\psi(t)\rangle_\Lambda \quad\Longrightarrow\quad \mathcal{G}^\dagger\vec{\mathcal{X}}\mathcal{G} = \vec{\mathcal{X}} \tag{19.22}$$

L'opérateur d'impulsion généralisée $\vec{\mathcal{P}}$ n'est pas invariant de jauge, mais par contre l'opérateur quantité de mouvement

$$\vec{\Pi} \overset{\text{déf.}}{=} \vec{\mathcal{P}} - q\vec{A}(\vec{\mathcal{X}}) \tag{19.23}$$

est nécessairement invariant de jauge, car sa valeur moyenne donne la vitesse de la particule (correspondant à la quantité classique $\vec{p}$ de l'éqn. (19.7)). Le potentiel vecteur se transformant selon (19.20a), la loi de transformation de l'opérateur $\vec{\mathcal{P}}$ s'obtient directement :

$$\vec{\mathcal{P}} \longmapsto \vec{\mathcal{P}} + q\vec{\nabla}\Lambda(\vec{\mathcal{X}})\,. \tag{19.24}$$

En raisonnant avec les valeurs moyennes nous obtenons donc

$$\mathcal{G}^\dagger\vec{\mathcal{P}}\mathcal{G} = \vec{\mathcal{P}} + q\vec{\nabla}\Lambda(\vec{\mathcal{X}}) \tag{19.25}$$

La résolution des équations (19.22,19.25) est aisée. Premièrement, l'équation (19.22) est résolue si $\mathcal{G} = \exp if(\vec{\mathcal{X}})$, où f est une fonction réelle de l'opérateur $\vec{\mathcal{X}}$. Deuxièmement, la propriété $[\mathcal{P}_i,\, F(\vec{\mathcal{X}})] = -i\hbar\partial_i F(\vec{\mathcal{X}})$ implique :

$$\vec{\mathcal{P}}\,\mathcal{G}(\vec{\mathcal{X}}) = \mathcal{G}(\vec{\mathcal{X}})\,\vec{\mathcal{P}} - i\hbar\vec{\nabla}\mathcal{G}(\vec{\mathcal{X}}) \tag{19.26}$$

d'où nous déduisons par identification :

$$\boxed{\mathcal{G}(\vec{\mathcal{X}},t) = e^{\frac{iq}{\hbar}\Lambda(\vec{\mathcal{X}},t)}} \tag{19.27}$$

Nous pouvons aussi examiner la transformation de l'équation de Schrödinger, c.-à.-d. l'équation satisfaite par le vecteur transformé $|\psi(t)\rangle_\Lambda$. Nous avons

$$i\hbar\frac{\partial}{\partial t}e^{\frac{iq}{\hbar}\Lambda(\vec{\mathcal{X}},t)}\,|\psi(t)\rangle = e^{\frac{iq}{\hbar}\Lambda(\vec{\mathcal{X}},t)}\left(i\hbar\frac{\partial}{\partial t}\,|\psi(t)\rangle - q\frac{\partial\Lambda(\vec{\mathcal{X}},t)}{\partial t}\,|\psi(t)\rangle\right) \tag{19.28}$$

et

$$\frac{\left(\vec{\mathcal{P}} - q\vec{A}(\vec{\mathcal{X}}) - q\vec{\nabla}\Lambda(\vec{\mathcal{X}},t)\right)^2}{2m}e^{\frac{iq}{\hbar}\Lambda(\vec{\mathcal{X}},t)}\,|\psi(t)\rangle = e^{\frac{iq}{\hbar}\Lambda(\vec{\mathcal{X}},t)}\frac{\left(\vec{\mathcal{P}} - q\vec{A}(\vec{\mathcal{X}})\right)^2}{2m}\,|\psi(t)\rangle \tag{19.29}$$

où nous avons utilisé

$$\begin{aligned}\big(\vec{\mathcal{P}} - q\vec{A}(\vec{\mathcal{X}}) - q\vec{\nabla}\Lambda(\vec{\mathcal{X}},t)\big)^2 \mathcal{G} &= \big(\vec{\mathcal{P}} - q\vec{A}(\vec{\mathcal{X}}) - q\vec{\nabla}\Lambda(\vec{\mathcal{X}},t)\big)\mathcal{G}\,\mathcal{G}^\dagger\big(\vec{\mathcal{P}} - q\vec{A}(\vec{\mathcal{X}}) - q\vec{\nabla}\Lambda(\vec{\mathcal{X}},t)\big)\mathcal{G} \\ &= \big(\vec{\mathcal{P}} - q\vec{A}(\vec{\mathcal{X}}) - q\vec{\nabla}\Lambda(\vec{\mathcal{X}},t)\big)\mathcal{G}\big(\vec{\mathcal{P}} - q\vec{A}(\vec{\mathcal{X}})\big) = \mathcal{G}\big(\vec{\mathcal{P}} - q\vec{A}(\vec{\mathcal{X}})\big)^2 \quad (19.30)\end{aligned}$$

Nous en déduisons donc que :

$$\begin{aligned}&\mathcal{G}\Big(i\hbar\frac{\partial}{\partial t} - \frac{\big(\vec{\mathcal{P}} - q\vec{A}(\vec{\mathcal{X}})\big)^2}{2m} - q\phi(\vec{\mathcal{X}})\Big)\,|\psi(t)\rangle \\ &\qquad = \Big(i\hbar\frac{\partial}{\partial t} - \frac{\big(\vec{\mathcal{P}} - q\vec{A}(\vec{\mathcal{X}}) - q\vec{\nabla}\Lambda(\vec{\mathcal{X}},t)\big)^2}{2m} - q\phi(\vec{\mathcal{X}}) + q\frac{\partial}{\partial t}\Lambda(\vec{\mathcal{X}},t)\Big)\,|\psi(t)\rangle_\Lambda \quad (19.31)\end{aligned}$$

c.-à.-d. que l'image par une transformation de jauge de l'équation de Schrödinger est l'équation de Schrödinger pour les potentiels transformés appliqués à la fonction d'onde transformée.

L'action d'une transformation de jauge sur les vecteurs de l'espace des états peut faire penser à l'action d'une symétrie :

$$|\psi\rangle \;\longmapsto\; e^{\frac{iq}{\hbar}\Lambda(\vec{\mathcal{X}},t)}\,|\psi\rangle\ , \qquad (19.32)$$

qu'on pourrait voir formellement comme une symétrie $U(1)$ (c.-à.-d. de rotation dans le plan complexe), dont le « paramètre » serait une fonction réelle arbitraire $-q\Lambda(\vec{\mathcal{X}},t)$ de l'opérateur position et du paramètre temps. Pour cette raison, on parle souvent de *symétrie de jauge*, et la transformation (19.32) est souvent interprétée comme l'action d'une symétrie *locale* (c.-à.-d. dont le paramètre dépend de la position).

Il existe cependant une différence fondamentale entre les deux concepts. Une symétrie va transformer un système en un système équivalent (dans le point de vue actif). Par exemple, l'action d'une rotation sur une orbitale atomique va donner la même orbitale atomique mais orientée différemment dans l'espace. En comparaison, une transformation de jauge (19.32) n'aura absolument aucun effet sur la physique du problème car toutes les quantité observables sont, par définition, invariantes de jauge. En d'autre termes aucune expérience ne pourra jamais « mesurer » Λ. Il est préférable de voir l'invariance de jauge comme une redondance dans la description du problème.

19.2.1 Effet Aharonov–Bohm

Le hamiltonien (19.11) indique qu'une particule couple à un champ électromagnétique non pas à travers les champs électromagnétiques $(\vec{E},\vec{B})$ mais à travers les potentiels $(\phi,\vec{A})$. Même si, au final, la physique est invariante de jauge, ce couplage a des conséquences surprenantes.

La plus spectaculaire d'entre elle, l'effet Aharonov–Bohm [1] est vérifiée expérimentalement avec une grande précision. Considérons un expérience d'interférences à deux ondes pour des électrons, représentée sur la figure 19.1. Le dispositif est mis en place de telle sorte que les deux chemins possibles pour les électrons soient de même longueur ; en l'absence de champ magnétique on doit donc observer des interférences constructiives.

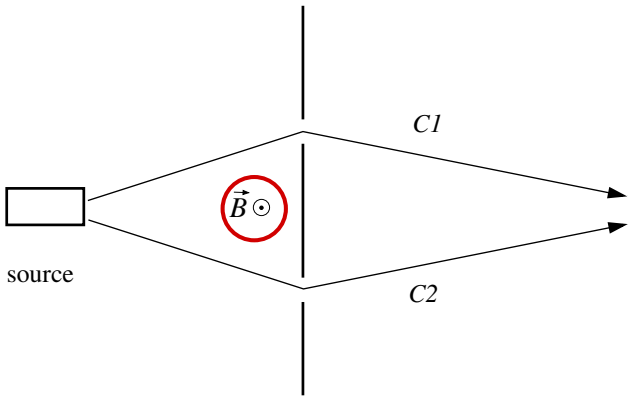

FIGURE 19.1 – Effet Aharonov–Bohm

On place maintenant entre les deux fentes un solénoïde parcouru par un courant électrique, générant un champ magnétique selon l'axe Oz, perpendiculaire au plan de la figure. Crucialement, le solénoïde est confiné dans une enceinte hermétique au champ magnétique, de telle sorte que *les électrons se propagent dans une région ou le champ magnétique est nul.*

Même si le champ magnétique est nul en dehors de l'enceinte, le potentiel vecteur $\vec{A}$ ne peut y être nul partout. Nous avons en effet d'après le théorème de Stockes, pour tout contour C entourant le solénoïde,

$$\oint_C \vec{A} \cdot d\vec{x} = \iint \vec{B} \cdot d\vec{S} = \Phi_B \,, \tag{19.33}$$

où Φ_B est le flux magnétique. Si nous prenons des coordonnées cylindriques (ρ, θ, z) autour de l'axe solénoïde, un potentiel vecteur $\vec{A}$ correspondant ce flux magnétique est donné, à l'extérieur du solénoïde, par :

$$\vec{A} = \frac{\Phi_B}{2\pi\rho}\vec{e}_\theta \,. \tag{19.34}$$

On vérifie que, pour tout $\rho \neq 0$, $\vec{B} = \vec{\nabla} \wedge \vec{A} = \frac{1}{\rho}\partial_\rho(\rho A_\theta)\vec{e}_z = 0$. [2]

Déterminons maintenant les fonctions d'ondes des électrons ayant suivi les chemins 1 et 2 à l'aide de la méthode WKB développée au cours n° 12, avec le remplacement $\vec{p} \to \vec{p} - q\vec{A}$ dû au couplage avec le champ électromagnétique. La fonction d'onde associée à une onde progressive pour une particule d'énergie E ayant suivi un chemin $C_{\vec{x}}$ finissant en $\vec{x}$ est donnée par (voir éqn. (12.30)) :

$$u(\vec{x}) = \frac{c}{\sqrt{2mE}} \exp \frac{i}{\hbar} \int_{C_{\vec{x}}} \left(\vec{p}(\vec{x}') - q\vec{A}(\vec{x'})\right) \cdot d\vec{x}' \tag{19.35}$$

Pour être plus précis, la fonction d'onde WKB étant évaluée le long d'un trajet classique, il faudrait considérer séparément le trajet de la particule de la source à la fente puis de la fente à l'écran, mais la conclusion resterait la même.

2. La situation est analogue au potentiel gravitationnel d'un astre à symétrie sphérique. Le potentiel à l'extérieur de l'astre est identique à celui obtenu pour une source singulière à l'origine.

Considérons la fonction d'onde obtenue par la superposition des fonctions d'onde des électrons ayant suivi les chemins C_1 et C_2 sur la figure 19.1. Nous nous plaçons au centre de l'écran, de telle sorte que leur trajet géométrique soit le même. Nous avons donc :

$$\begin{aligned} u(\vec{x}) &= \frac{c}{\sqrt{2mE}} \left(e^{\frac{i}{\hbar}\int_{C_1}\left(\vec{p}(\vec{x}')-q\vec{A}(\vec{x}')\right)\cdot d\vec{x}'} + e^{\frac{i}{\hbar}\int_{C_2}\left(\vec{p}(\vec{x}')-q\vec{A}(\vec{x}')\right)\cdot d\vec{x}'} \right) \\ &= \frac{c}{\sqrt{2mE}} e^{\frac{i}{\hbar}\int_{C_1}\left(\vec{p}(\vec{x}')-q\vec{A}(\vec{x}')\right)\cdot d\vec{x}'} \left(1 + e^{-\frac{iq}{\hbar}\int_{C_2}\vec{A}(\vec{x}')\cdot d\vec{x}' + \frac{iq}{\hbar}\int_{C_1}\vec{A}(\vec{x}')\cdot d\vec{x}'} \right) \\ &= \frac{c}{\sqrt{2mE}} e^{\frac{i}{\hbar}\int_{C_1}\left(\vec{p}(\vec{x}')-q\vec{A}(\vec{x}')\right)\cdot d\vec{x}'} \left(1 + e^{-\frac{iq}{\hbar}\oint_{C}\vec{A}(\vec{x}')\cdot d\vec{x}'} \right) , \end{aligned} \tag{19.36}$$

où $C = C_2 - C_1$ est un contour fermé tournant une fois autour du solénoïde dans le sens trigonométrique.

D'après l'éqn. (19.33) nous obtenons donc un déphasage entre les deux fonctions d'onde donné par

$$\Delta\varphi \overset{\text{déf.}}{=} \varphi_1 - \varphi_2 = \frac{q}{\hbar}\oint_C \vec{A}(\vec{x}')\cdot d\vec{x}' \implies \boxed{\Delta\varphi = \frac{q\Phi_B}{\hbar}} \tag{19.37}$$

Définissons le « quantum » de flux magnétique par

$$\Phi_0 \overset{\text{déf.}}{=} \frac{2\pi\hbar}{|q|} . \tag{19.38}$$

Lorsque le champ magnétique est ajouté au sytème, l'intensité observée au centre de l'écran diminue jusqu'à obtenir, pour un flux $\Phi_B = \Phi_0/2$, des interférences destructives. En revanche dès que le flux magnétique est un multiple de Φ_0, le champ magnétique n'a plus d'effet sur la figure d'interférences.

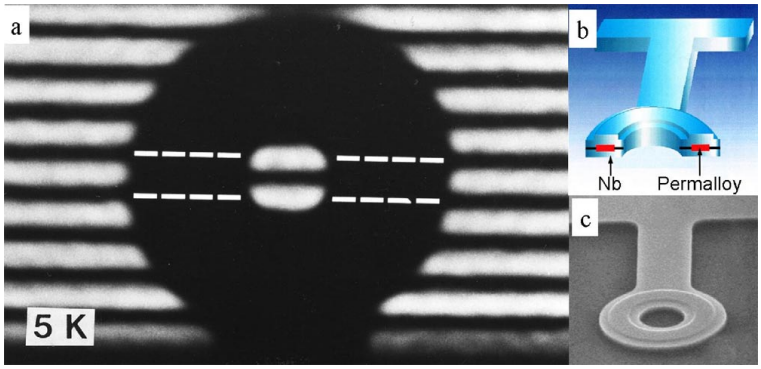

FIGURE 19.2 – Observation expérimentale de l'effet Aharonov–Bohm (crédit [2]).

La première observation expérimentale indiscutable de cet effet quantique spectaculaire est due à l'équipe de A. Tonomura [2]. Le dispositif expérimental, représenté sur la figure 19.2, consiste en un aimant toroïdal, recouvert d'une couche de niobium supraconductrice empêchant tout champ magnétique de s'en échapper.

Un faisceau cohérent d'électrons est envoyé sur l'aimant, perpendiculairement au plan de la figure. Une partie des électrons va passer autour de l'aimant, et une autre partie par le trou en son centre. On fait ensuite interférer ce faisceau avec un faisceau de référence. La figure correspond à l'image des franges d'interférences pour un flux magnétique de la forme $\Phi_B = (\frac{1}{2} + n)\Phi_0$ avec n entier.

Nous conclurons cette discussion par deux remarques :

— l'expresson de la fonction d'onde (19.35) dépend de la jauge choisie pour le potentiel vecteur $\vec{A}$; en revanche le déphasage $\Delta\varphi$, quantité mesurable, n'en dépend pas ;

— du point de vue de l'intégrale de chemin, la méthode WKB ne donne que la contribution du chemin dominant, et on peut se demander quelle est la modification due aux autres chemins. Si C_1' est un chemin quelconque pour un électron passant par la première fente et C_2' un chemin quelconque pour un électron passant par la deuxième fente, le déphasage dû au champ magnétique sera

$$\Delta\varphi = \frac{q}{\hbar}\oint_{C_2'-C_1'} \vec{A}(\vec{x'}) \cdot d\vec{x}' = \frac{q}{\hbar}\Phi_B \,, \tag{19.39}$$

qui est indépendant du chemin fermé suivi autour du solénoïde. Il s'agit d'une quantité purement topologique. Le point crucial est que le chemin ne puisse être contracté en un point, en raison de la singularité créée en $\rho = 0$ par le potentiel vecteur (19.34).

19.3 Absorption et émission spontanée

L'objectif de cette section est d'étudier l'interaction entre un électron et un champ électromagnétique classique, correspondant à une onde plane monochromatique, en particulier concernant les transitions entre niveaux atomiques. Le spin de l'électron ne sera pas ici pris en compte.

Reprenons le hamiltonien d'une particule chargée et développons le terme cinétique. Nous avons :

$$\begin{aligned}\mathcal{H} &= \frac{1}{2m_e}\left(\vec{\mathcal{P}} - q\vec{A}(\mathcal{X})\right)^2 + q\phi(\vec{\mathcal{X}}) \\ &= \frac{1}{2m_e}\vec{\mathcal{P}}^2 - \frac{q}{2m_e}\left(\vec{A}\cdot\vec{\mathcal{P}} + \vec{\mathcal{P}}\cdot\vec{A}\right) + \frac{q^2}{2m_e}\vec{A}^2 + q\phi(\vec{\mathcal{X}})\,.\end{aligned} \tag{19.40}$$

Nous pouvons simplifier le second terme si nous choissons la *jauge de Coulomb*, la plus employée dans ce contexte :

$$\vec{\nabla}\cdot\vec{A} = 0\,. \tag{19.41}$$

Étant donné que

$$\left[\vec{\mathcal{P}},\, \vec{A}(\vec{\mathcal{X}})\right] = -i\hbar\vec{\nabla}\cdot\vec{A} = 0 \tag{19.42}$$

nous obtenons le hamiltonien :

$$\boxed{\mathcal{H} = \frac{1}{2m_e}\vec{\mathcal{P}}^2 - \frac{q}{m_e}\vec{A}\cdot\vec{\mathcal{P}} + +\frac{q^2}{2m_e}\vec{A}^2 + q\phi(\vec{\mathcal{X}})} \tag{19.43}$$

Le terme en A^2 ne sera pas pris en compte dans le traitement perturbatif du couplage au rayonnement, étant sous-dominant. [3]

Considérons une onde plane monochromatique électromagnétique, donnée par le potentiel vecteur :

$$\vec{A}(\vec{x},t) = \frac{1}{2}\vec{A}_0\left(e^{i(\vec{k}\cdot\vec{x}-\omega t)} + e^{-i(\vec{k}\cdot\vec{x}-\omega t)}\right), \tag{19.44}$$

avec la condition de transversalité

$$\vec{k}\cdot\vec{A}_0 = 0 \tag{19.45}$$

imposée par la jauge de Coulomb et la relation de dispersion

$$\vec{k}^2 = \omega^2/c^2\,. \tag{19.46}$$

Nous souhaitons traiter l'effet du terme d'interaction entre l'onde et l'électron à l'aide de la théorie des perturbations dépendant du temps développée dans le cours n° 18, appliquée au hamiltonien

$$\mathcal{H}_0 = \frac{1}{2m_e}\vec{\mathcal{P}}^2 + q\phi(\vec{\mathcal{X}})\,, \tag{19.47}$$

avec, dans le cas de l'atome d'hydrogène, un potentiel scalaire $\phi(\vec{\mathcal{X}})$ coulombien généré par le proton. L'opérateur de perturbation est alors donné par

$$\mathcal{V}(t) = -\frac{q}{2m_e}\left(e^{i(\vec{k}\cdot\vec{\mathcal{X}}-\omega t)} + e^{-i(\vec{k}\cdot\vec{\mathcal{X}}-\omega t)}\right)\vec{A}_0\cdot\vec{\mathcal{P}}\,. \tag{19.48}$$

19.3.1 Section efficace d'absorption

En comparant avec l'étude générale d'une perturbation sinusoïdale, voir l'éqn. (18.55), nous voyons que le premier terme correspond à une situation d'absorption, et le second à une situation d'émission spontanée. En reprenant l'éqn. (18.59) nous en déduisons donc le taux d'absorption du rayonnement par l'électron est donné par :

$$\boxed{\Gamma_I \overset{t\to\infty}{\sim} \frac{2\pi}{\hbar}\frac{q^2|A_0|^2}{4m_e^2}\left|\langle F|\, e^{i\vec{k}\cdot\vec{\mathcal{X}}}\vec{a}\cdot\vec{\mathcal{P}}\,|I\rangle\right|^2 \rho(E_F)\Big|_{E_F=E_I+\hbar\omega}} \tag{19.49}$$

3. Remarquons que, pour l'étude de l'effet Zeeman dans le cours n° 17, nous avions également négligé l'influence du terme quadratique en $\vec{A}$, qui donne, pour un champ magnétique uniforme selon O_z,

$$\vec{A} = -\frac{1}{2}\vec{x}\wedge\vec{B} \implies \frac{q^2}{2m_e}\vec{A}^2 = \frac{q^2B^2}{8m_e}(x^2+y^2)\,.$$

Ce terme étant quadratique en B, est sous-dominant par rapport au terme linéraire (17.67) en théorie des perturbations.

où nous avons posé $\vec{A}_0 = A_0\vec{a}$ avec $\vec{a}^2 = 1$.

Une quantité physique directement mesurable est la *section efficace d'absorption* σ_a, définie comme l'énergie absorbée par l'atome par unité de temps divisée par le flux d'énergie électromagnétique ; comme son nom l'indique, cette quantité a la dimension d'une surface et est d'autant plus grande que la probabilité d'absorption est élevée.

Le flux d'énergie électromagnétique est donnée par valeur moyenne temporelle de la norme du vecteur de Poynting $\vec{\Pi} = c^2\epsilon_0\vec{E}\wedge\vec{B}$. Pour une onde plane progressive nous avons :

$$\vec{A} = \vec{A}_0\cos(\omega t - \vec{k}\cdot\vec{x}) \implies \left\{\begin{array}{rcl} \vec{E} &=& -\partial_t\vec{A} = \omega\vec{A}_0\sin(\omega t - \vec{k}\cdot\vec{x}) \\ \vec{B} &=& \vec{\nabla}\wedge\vec{A} = \left(\vec{k}\wedge\vec{A}_0\right)\sin(\omega t - \vec{k}\cdot\vec{x}) \end{array}\right. \tag{19.50}$$

avec $\vec{k}\cdot\vec{A}_0 = 0$ dans la jauge de Coulomb. On obtient :

$$\Phi_{em} = \left\langle ||\vec{\Pi}|| \right\rangle_t = c^2\epsilon_0||\vec{k}||\omega||\vec{A}_0||^2\left\langle \sin^2(\omega t - \vec{k}\cdot\vec{x})\right\rangle_t = c^2\epsilon_0\frac{\omega^2}{c}A_0^2\times\frac{1}{2} = \frac{\epsilon_0 c}{2}\omega^2A_0^2\,. \tag{19.51}$$

Étant donné que le phénomène considéré correspond à l'absorption d'une énergie $\hbar\omega$, nous trouvons pour la section efficace :

$$\begin{aligned}\sigma_a &= \frac{\hbar\omega\frac{2\pi}{\hbar}\frac{q^2|A_0|^2}{4m_e^2}\left|\langle F|\, e^{i\vec{k}\cdot\vec{x}}\vec{a}\cdot\vec{\mathcal{P}}\,|I\rangle\right|^2}{\frac{\epsilon_0 c}{2}\omega^2|\vec{A}_0|^2}\rho(E_F)\Big|_{E_F=E_I+\hbar\omega} \\ &= \frac{\pi q^2}{m_e^2\epsilon_0 c\omega}\left|\langle F|\, e^{i\vec{k}\cdot\vec{x}}\vec{a}\cdot\vec{\mathcal{P}}\,|I\rangle\right|^2\rho(E_F)\Big|_{E_F=E_I+\hbar\omega}\end{aligned} \tag{19.52}$$

soit, en faisant apparaître la constante de structure fine $\alpha = q^2/(4\pi\epsilon_0\hbar c)$,

$$\boxed{\sigma_a = \frac{4\pi^2\hbar}{m_e^2\omega}\alpha\left|\langle F|\, e^{i\vec{k}\cdot\vec{x}}\vec{a}\cdot\vec{\mathcal{P}}\,|I\rangle\right|^2\rho(E_F)\Big|_{E_F=E_I+\hbar\omega}} \tag{19.53}$$

Il peut paraître étrange d'utiliser la règle d'or de Pauli alors qu'à la fois le niveau de départ et le niveau d'arrivée sont des niveaux discrets. On peut se convaincre que le résultat obtenu est néanmoins correct :

— les niveaux atomiques ne sont jamais parfaitement ponctuels, en raison des collisions entre atomes et des autres types d'effets qui se manifestent expérimentalement par l'élargissement les raies spectrales ;

— on peut voir le taux de transition non pas comme fonction de l'énergie E_F de l'état final mais en tant que fonction de la pulsation ω de l'onde électromagnétique, qui est bien une variable continue. L'onde incidente n'étant jamais parfaitement monochromatique nous avons une distributions de fréquences du rayonnement de largeur finie.

Nous pourrions naturellement traiter de manière exactement similaire le phénomène d'émission spontanée.

19.3.2 Approximation dipolaire

Pour des niveaux atomiques l'élément de matrice de la perturbation apparaissant dans l'éqn. (19.53) n'est pas simple à calculer. Il existe cependant une approximation naturelle, dite *approximation dipolaire électrique*, rendant l'étude du problème beaucoup plus simple. Nous négligerons simplement la dépendance spatiale de l'onde plane à l'échelle atomique :

$$e^{i\vec{k}\cdot\vec{x}} = 1 + \mathcal{O}(\vec{x})\,. \tag{19.54}$$

En effet, pour qu'une transition soit possible il faut que $\hbar\omega$ soit égal à la différence d'énergie entre deux niveaux atomiques, soit en ordre de grandeur pour l'atome d'hydrogène

$$\hbar\omega \sim E_I = \frac{\hbar^2}{2m_e a_0^2} \tag{19.55}$$

Soit une longueur d'onde

$$\lambda \sim 2\pi c \frac{2m_e}{\hbar} a_0^2 = \frac{4\pi m_e c}{\hbar} \frac{4\pi\epsilon_0\hbar^2}{m_e q^2} a_0 = \frac{4\pi}{\alpha} a_0\,. \tag{19.56}$$

Nous en concluons donc que $\lambda \gg a_0$, rayon classique de l'atome, et l'approximation dipolaire est justifiée. [4]

L'élément de matrice que nous devons calculer pour obtenir la section efficace (19.53) est donc simplement, pour une transition entre deux niveaux de l'atome d'hydrogène,

$$\langle n'l'm'|\,\vec{a}\cdot\vec{\mathcal{P}}\,|nlm\rangle = \vec{a}\cdot\langle n'l'm'|\,\vec{\mathcal{P}}\,|nlm\rangle\ . \tag{19.57}$$

Cette écriture indique que, contrairement à ce qu'on pourrait naïvement penser, il ne s'agit pas de la valeur moyenne d'un opérateur scalaire mais bien de la valeur moyenne d'un opérateur vectoriel, car le vecteur constant $\vec{a}$ n'est pas un opérateur vectoriel.

On peut écrire cet élément de matrice sous une forme plus commode en utilisant la propriété suivante :

$$[\,\mathcal{X},\,\mathcal{H}_0\,] = i\hbar\frac{\vec{P}}{m_e} \quad\Longrightarrow\quad \boxed{\langle n'l'm'|\,\vec{\mathcal{P}}\,|nlm\rangle = \frac{m_e(E_n - E'_n)}{i\hbar}\,\langle n'l'm'|\,\vec{\mathcal{X}}\,|nlm\rangle} \tag{19.58}$$

Dans la base sphérique on peut calculer cet élément de matrice par le théorème de Wigner–Eckart, voir l'éqn. (17.56) du cours n° 17 :

$$\langle n', l', m'|\,\mathcal{X}_1^r\,|n, l, m\rangle = C^{l'\,m'}_{l\,m,\,1r}\,\langle n', l'||\mathcal{X}_1||n, l\rangle \tag{19.59}$$

Le théorème de Wigner–Eckart nous apprend en particulier (voir sous l'éqn. (17.56) du cours n° 17) que les transitions d'une orbitale s (avec $l = 0$) vers une autre orbitale s (soit

4. Notons que pour un atome hydrogénoïde de numéro atomique Z, qui a un rayon classique a_0/Z, le critère de validité devient $4\pi Z/\alpha \ll 1$ qui est satisfait pour des noyaux légers.

$l' = 0$) ne sont pas permises dans l'approximation dipolaire. Examinons également l'action de la parité qui fournit des contraintes supplémentaires. Nous avons

$$\langle n', l', m'| \mathcal{X}_1^r |n, l, m\rangle \mapsto \langle n', l', m'| \mathcal{P}^\dagger \mathcal{X}_1^r \mathcal{P} |n, l, m\rangle = (-)^{l+l'} \langle n', l', m'| \mathcal{X}_1^r |n, l, m\rangle \tag{19.60}$$

L'opérateur $\vec{\mathcal{X}}$ étant impair par parité, nous devons obligatoirement avoir $l + l'$ impair.

Les transitions possibles dans l'approximation électrique dipolaire sont donc caractérisées par les règles de sélection (d'après le théorème de Wigner–Eckart plus l'étude de l'action de la parité) :

$$\boxed{l - l' = \pm 1 \ , \quad m' - m \in \{-1, 0, 1\}} \tag{19.61}$$

Pour calculer la section efficace correspondant à une telle transition il suffit selon (19.59) de calculer dans la base sphérique l'élément de matrice associé à $\mathcal{X}_1^0 = \mathcal{Z}$. En termes d'harmoniques sphériques nous pouvons écrire (cf. éqn. (14.91) du cours n° 14) :

$$z = \sqrt{\frac{4\pi}{3}} r Y_1^0(\theta) \tag{19.62}$$

L'élément de matrice à calculer est donc de la forme (avec nécessairement $m' = m$) :

$$\boxed{\begin{aligned}\langle n'l'm| \mathcal{Z} |nlm\rangle = \sqrt{\frac{4\pi}{3}} \int r^2 dr \, \rho_{n',l'}(r) \rho_{n',l'}(r) \, r \\ \times \int \sin\theta d\theta d\phi \ \overline{Y_{l'}^{m'}(\theta,\phi)} Y_1^0(\theta) \, Y_l^m(\theta,\phi) \, .\end{aligned}} \tag{19.63}$$

Prenons comme exemple, pour mener le calcul à son terme, une transition $(1s) \to (2p)$; étant donné que $m = 0$ nous avons également $m' = 0$. Il faut premièrement calculer l'intégrale radiale

$$\sqrt{\frac{4\pi}{3}} \int r^2 dr \, r \, \rho_{2,1}(r) \rho_{1,0}(r) = \sqrt{\frac{4\pi}{3}} \frac{2}{\sqrt{a_0^3}} \frac{1}{\sqrt{24 a_0^3}} \int r^2 dr \, r \frac{r}{a_0} e^{-\frac{r}{2a_0}} e^{-\frac{r}{a_0}} = \frac{\sqrt{4\pi}}{3a_0^4} \int dr \, r^4 e^{-\frac{3r}{2a_0}} \tag{19.64}$$

Le calcul explicite de cette dernière intégrale donne finalement

$$\sqrt{\frac{4\pi}{3}} \int_0^\infty r^2 dr \, r \, \rho_{2,1}(r) \rho_{1,0}(r) = \frac{256}{243} \sqrt{4\pi} a_0 \, . \tag{19.65}$$

Nous avons ensuite à calculer l'intégrale sur les harmoniques sphériques. On peut utiliser la formule (16.76) pour décomposer le produit d'harmoniques sphériques ou bien faire le calcul direct. Nous avons

$$\int \sin\theta d\theta d\phi \ \overline{Y_1^0(\theta)} Y_1^0(\theta) \, Y_0^0(\theta) = \frac{3}{(4\pi)^{3/2}} \int \cos^2\theta \sin\theta d\theta d\phi = \frac{3}{\sqrt{4\pi}} \frac{2}{3} \tag{19.66}$$

En combinant ces résultats nous avons donc :

$$\langle 210| \mathcal{Z} |100\rangle = \frac{512}{243} a_0 \tag{19.67}$$

Si le potentiel vecteur de l'onde incidente est orienté selon Oz, on obtient le résultat final pour la section efficace :

$$\sigma_a = \frac{4\pi\hbar}{m_e^2\omega}\alpha\left(\frac{m_e}{\hbar}\frac{3E_I}{4}\right)^2\left(\frac{512}{243}a_0\right)^2 \rho(E_I/4)\Big|_{\omega=\frac{3E_I}{4\hbar}}$$

$$\Longrightarrow \quad \boxed{\sigma_a = \frac{4096}{6561}4\pi^2\,\frac{\hbar c^2}{\omega}\,\alpha^3\,\rho(E_I/4)\Big|_{\omega=\frac{3E_I}{4\hbar}}} \quad (19.68)$$

Il est aisé de vérifier que ce résultat a bien la dimension d'une surface (comme $\rho(E)$ a la dimension de l'inverse d'une énergie).

À retenir

— Le couplage d'une particule chargée au champ électromagnétique ;

— invariance de jauge en mécanique quantique ;

— l'effet Aharonov–Bohm ;

— notion de section efficace ;

— absorption par un atome dans l'approximation dipolaire et règles de sélection.

Bibliographie

[1] Y. Aharonov et D. Bohm, "Significance of Electromagnetic Potentials in the Quantum Theory", Physical Review 115-3 (1959)

[2] A. Tonomura, et al., "Evidence for Aharonov-Bohm effect with magnetic field completely shielded from electron wave," Phys. Rev. Lett. 56 (1986)

Cours n°20

Quantification du champ électromagnétique

La mécanique quantique non-relativiste de Schrödinger et Heisenberg est adaptée à l'étude d'un système constitué d'une particule, ou d'un nombre de particule fixé. Ce type d'approche n'est pas pertinent dans le cadre relativiste, dans lequel des particules peuvent être constamment créées ou annihilées, ce qui nécessite d'utiliser les méthodes de la théorie quantique des champs.

Alors que, pour des particules massives comme les électrons et les protons, il existe un grand nombre de phénomènes pour lesquels la limite non-relativiste est pertinente, ce n'est évidemment pas le cas pour les photons, particules de masse nulle. Nous construirons, par quantification canonique, une théorie quantique du champ électromagnétique dont l'espace des états possède un nombre indéfini de particules.

Considérons le champ électromagnétique dans le vide, à l'intérieur d'une cavité cubique de côté L avec des conditions aux bords périodiques. Nous considérons les équations de Maxwell dans le vide :

$$\vec{\nabla} \cdot \vec{E} = 0 \tag{20.1a}$$

$$\vec{\nabla} \cdot \vec{B} = 0 \tag{20.1b}$$

$$\vec{\nabla} \wedge \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{20.1c}$$

$$\vec{\nabla} \wedge \vec{B} = \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t} \tag{20.1d}$$

qui peuvent se réexprimer en termes d'un potentiel vecteur $\vec{A}$ tel que, localement,

$$\vec{B} = \vec{\nabla} \wedge \vec{A} \tag{20.2}$$

afin de satisfaire l'éqn. (20.1b). Nous imposons, comme nous l'avions déjà mentionné, la *jauge de Coulomb* :

$$\vec{\nabla} \cdot \vec{A} = 0\,. \tag{20.3}$$

En l'absence de charges on peut prendre un potentiel scalaire nul et le champ électrique est donné en jauge de Coulomb par :

$$\vec{E} = -\frac{\partial \vec{A}}{\partial t}\,. \tag{20.4}$$

Il suffit donc d'analyser le potentiel vecteur $\vec{A}(\vec{x},t)$. L'équation (20.1d) donne :

$$\vec{\nabla} \wedge \left(\vec{\nabla} \wedge \vec{A}\right) = -\frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} \tag{20.5}$$

soit, dans la jauge de Coulomb,

$$\left(\Delta - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\vec{A}(\vec{x},t) = 0\,. \tag{20.6}$$

20.1 Développement en modes

Les solutions de l'équation d'onde (20.6) sont évidemment des ondes planes $e^{i(\vec{k}\cdot\vec{x}\pm\omega t)}$, avec la relation de dispersion pour le vecteur d'onde $\vec{k}$: $\omega_k^2 = c^2\vec{k}^2$. Les vecteurs d'onde sont contraints par les conditions aux limites périodiques d'être de la forme

$$\vec{k} = \frac{2\pi}{L}(n_1,n_2,n_3)\ , \quad n_i \in \mathbb{Z}\,. \tag{20.7}$$

Dans la jauge de Coulomb (20.3) les ondes planes sont transverses, c.-à.-d. qu'une onde plane $\vec{A}(\vec{k})$ de vecteur d'onde $\vec{k}$ satisfait

$$\vec{k}\cdot\vec{A}(\vec{k}) = 0\,. \tag{20.8}$$

Les solutions se développement donc sur une base orthonormée de polarisations transverses à la direction de propagation $\{e^1_{\vec{k}}, e^2_{\vec{k}}\}$.

En prenant en compte la réalité du potentiel vecteur $\vec{A}$, le développement le plus général sur les modes propres de la cavité est donné par

$$\vec{A}(\vec{x},t) = \sum_{n_i,\epsilon}\left(A_\epsilon(\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega_k t)} + \overline{A_\epsilon(\vec{k})}e^{-i(\vec{k}\cdot\vec{x}-\omega_k t)}\right)\vec{e}^{\,\epsilon}_{\vec{k}}\,. \tag{20.9}$$

Nous pouvons en déduire premièrement l'expression du champ électrique, d'après (20.4) :

$$\vec{E}(\vec{x},t) = i\sum_{n_i,\epsilon}\omega_k\left(A_\epsilon(\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega_k t)} - \overline{A_\epsilon(\vec{k})}e^{-i(\vec{k}\cdot\vec{x}-\omega_k t)}\right)\vec{e}^{\,\epsilon}_{\vec{k}} \tag{20.10}$$

et pour le champ magnétique :

$$\begin{aligned}\vec{B}(\vec{x},t) &= \vec{\nabla}\wedge\sum_{n_i,\epsilon}\left(A_\epsilon(\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega_k t)} + \overline{A_\epsilon(\vec{k})}e^{-i(\vec{k}\cdot\vec{x}-\omega_k t)}\right)\vec{e}^{\,\epsilon}_{\vec{k}}\\ &= i\sum_{n_i,\epsilon}\left(A_\epsilon(\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega_k t)} - \overline{A_\epsilon(\vec{k})}e^{-i(\vec{k}\cdot\vec{x}-\omega_k t)}\right)\vec{k}\wedge\vec{e}^{\,\epsilon}_{\vec{k}}\end{aligned} \tag{20.11}$$

L'énergie du champ électromagnétique dans l'enceinte est donnée par :

$$U(t) = \iiint_V d^2x\, u(\vec{x},t) = \frac{\epsilon_0}{2}\iiint_V d^3x \left(\vec{E}^2 + c^2\vec{B}^2\right). \tag{20.12}$$

Pour ce calcul nous pourrons utiliser

$$\iiint_V d^3x\, e^{i(\vec{k}-\vec{k}')\cdot\vec{x}} = V\delta_{\vec{k},\vec{k}'} \tag{20.13}$$

La première contribution provient du terme en $\vec{E}^2$. Nous avons

$$\begin{aligned}\iiint_V d^3x\, \vec{E}^2 = &-\sum_{\vec{k},\vec{k}',\epsilon,\epsilon'} \omega_k\omega_{k'} \iiint_V d^3x \left(A_\epsilon(\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega_k t)} - \overline{A_\epsilon(\vec{k})}e^{-i(\vec{k}\cdot\vec{x}-\omega_k t)}\right)\\ &\times\left(A_{\epsilon'}(\vec{k}')e^{i(\vec{k}'\cdot\vec{x}-\omega_{k'}t)} - \overline{A_{\epsilon'}(\vec{k}')}e^{-i(\vec{k}'\cdot\vec{x}-\omega_{k'}t)}\right)\vec{e}^{\,\epsilon}_{\vec{k}}\cdot\vec{e}^{\,\epsilon'}_{\vec{k}'}\\ = V\sum_{\vec{k},\epsilon,\epsilon'}&\omega_k^2\Big\{\left(A_\epsilon(\vec{k})\overline{A_{\epsilon'}(\vec{k})} + \overline{A_\epsilon(\vec{k})}A_{\epsilon'}(\vec{k})\right)\vec{e}^{\,\epsilon}_{\vec{k}}\cdot\vec{e}^{\,\epsilon'}_{\vec{k}}\\ &- A_\epsilon(\vec{k})A_{\epsilon'}(-\vec{k})\vec{e}^{\,\epsilon}_{\vec{k}}\cdot\vec{e}^{\,\epsilon'}_{-\vec{k}}e^{-2i\omega_k t} - \overline{A_\epsilon(\vec{k})A_{\epsilon'}(-\vec{k})}\vec{e}^{\,\epsilon}_{\vec{k}}\cdot\vec{e}^{\,\epsilon'}_{-\vec{k}}e^{2i\omega_k t}\Big\}\end{aligned} \tag{20.14}$$

Nous avons ensuite la contribution de $\vec{B}^2$ qui donne

$$\begin{aligned}\iiint_V d^3x\, \vec{B}^2 = &-\sum_{\vec{k},\vec{k}',\epsilon,\epsilon'} \iiint_V d^3x \left(A_\epsilon(\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega_k t)} - \overline{A_\epsilon(\vec{k})}e^{-i(\vec{k}\cdot\vec{x}-\omega_k t)}\right)\\ &\times\left(A_{\epsilon'}(\vec{k}')e^{i(\vec{k}'\cdot\vec{x}-\omega_{k'}t)} - \overline{A_{\epsilon'}(\vec{k}')}e^{-i(\vec{k}'\cdot\vec{x}-\omega_{k'}t)}\right)(\vec{k}\wedge\vec{e}^{\,\epsilon}_{\vec{k}})\cdot(\vec{k}\wedge\vec{e}^{\,\epsilon'}_{\vec{k}'})\\ = V\sum_{\vec{k},\epsilon,\epsilon'}&\Big\{\left(A_\epsilon(\vec{k})\overline{A_{\epsilon'}(\vec{k})} + \overline{A_\epsilon(\vec{k})}A_{\epsilon'}(\vec{k})\right)(\vec{k}\wedge\vec{e}^{\,\epsilon}_{\vec{k}})\cdot(\vec{k}\wedge\vec{e}^{\,\epsilon'}_{\vec{k}})\\ &+ A_\epsilon(\vec{k})A_{\epsilon'}(-\vec{k})(\vec{k}\wedge\vec{e}^{\,\epsilon}_{\vec{k}})\cdot(\vec{k}\wedge\vec{e}^{\,\epsilon'}_{-\vec{k}})e^{-2i\omega_k t}\\ &+ \overline{A_\epsilon(\vec{k})A_{\epsilon'}(-\vec{k})}(\vec{k}\wedge\vec{e}^{\,\epsilon}_{\vec{k}})\cdot(\vec{k}\wedge\vec{e}^{\,\epsilon'}_{-\vec{k}})e^{2i\omega_k t}\Big\}\end{aligned} \tag{20.15}$$

En combinant les deux contributions nous constatons l'annulation des deux termes dépendant du temps,

$$A_\epsilon(\vec{k})A_{\epsilon'}(-\vec{k})e^{-2i\omega_k t}\left(\omega_k^2\vec{e}^{\,\epsilon}_{\vec{k}}\cdot\vec{e}^{\,\epsilon'}_{-\vec{k}} - c^2(\vec{k}\wedge\vec{e}^{\,\epsilon}_{\vec{k}})\cdot(\vec{k}\wedge\vec{e}^{\,\epsilon'}_{-\vec{k}})\right) = 0, \tag{20.16}$$

et de même pour son complexe conjugué.

Il nous reste finalement, avec $\vec{e}^{\,\epsilon}_{\vec{k}}\cdot\vec{e}^{\,\epsilon'}_{\vec{k}} = \delta_{\epsilon,\epsilon'}$ la somme des contributions des deux termes constants, soit :

$$U = \frac{\epsilon_0 V}{2}\sum_{\vec{k},\epsilon}\left(A_\epsilon(\vec{k})\overline{A_\epsilon(\vec{k})} + \overline{A_\epsilon(\vec{k})}A_\epsilon(\vec{k})\right)\left(\omega_k^2 + c^2\vec{k}^2\right) \tag{20.17}$$

donnant, à l'aide de la relation de dispersion,

$$\boxed{U = \epsilon_0 V\sum_{\vec{k},\epsilon}\omega_k^2\left(A_\epsilon(\vec{k})\overline{A_\epsilon(\vec{k})} + \overline{A_\epsilon(\vec{k})}A_\epsilon(\vec{k})\right)} \tag{20.18}$$

20.2 Quantification canonique

Nous souhaitons associer à à l'expression (20.18) de l'énergie du champ électromagnétique classique le hamiltonien quantique du système. La similarité entre cette expression et l'oscillateur harmonique nous suggère d'associer au coefficient $A_\epsilon(\vec{k})$ (resp. au coefficient $\overline{A_\epsilon(\vec{k})}$) un opérateur $a^\epsilon_{\vec{k}}$ (resp. $(a^\epsilon_{\vec{k}})^\dagger$) annihilant (resp. créant) un photon de vecteur d'onde $\vec{k}$ et de polarisation ϵ. Ces opérateurs sont conjugués hermitiens l'un de l'autre et obéissent aux relations de commutation :

$$\left[\, a^\epsilon_{\vec{k}},\, (a^{\epsilon'}_{\vec{k}'})^\dagger \,\right] = \delta_{\vec{k},\vec{k}'}\delta_{\epsilon,\epsilon'} \quad , \qquad \left[\, a^\epsilon_{\vec{k}},\, a^{\epsilon'}_{\vec{k}'} \,\right] = 0 \tag{20.19}$$

Un photon de pulsation ω_k étant d'énergie $\hbar\omega_k$, cela fixe la normalisation

$$A_\epsilon(\vec{k}) \quad \longleftrightarrow \quad \sqrt{\tfrac{\hbar}{2\epsilon_0\omega_k V}}\; a^\epsilon_{\vec{k}} \tag{20.20}$$

et nous arrivons au hamiltonien :

$$\boxed{\mathcal{H} = \frac{1}{2}\sum_{\vec{k},\epsilon} \hbar\omega_k \Big(a^\epsilon_{\vec{k}}(a^\epsilon_{\vec{k}})^\dagger + (a^\epsilon_{\vec{k}})^\dagger a^\epsilon_{\vec{k}} \Big)} \tag{20.21}$$

qui apparaît comme une somme sur une infinité d'oscillateurs découplés, chacun correspondant à un état possible du photon (vecteur d'onde et polarisation).

Remarquons qu'il est possible d'obtenir la quantification du champ électromagnétique de manière moins heuristique dans le cadre de la théorie quantique des champs, en partant de l'action classique pour le champ électromagnétique.

Le nombre de photons à l'intérieur de l'enceinte dans une certain état est valeur propre de l'opérataeur :

$$\mathcal{N}_{\vec{k},\epsilon} \overset{\text{déf.}}{=} (a^\epsilon_{\vec{k}})^\dagger a^\epsilon_{\vec{k}}\,, \tag{20.22}$$

et les opérateurs $\mathcal{N}_{\vec{k},\epsilon}$ associés à des états $(\vec{k},\epsilon)$ différents commutent entre eux. L'état du champ électromagnétique est donc spécifié par un vecteur propre commun à tous ces opérateurs spécifiant le nombre de photons dans chacun des états possibles :

$$\big|\{n_{\vec{k},\epsilon}\}\big\rangle \;\;,\quad \forall\vec{k}',\ \forall\epsilon',\ \mathcal{N}_{\vec{k}',\epsilon'}\big|\{n_{\vec{k},\epsilon}\}\big\rangle = n_{\vec{k}',\epsilon'}\big|\{n_{\vec{k},\epsilon}\}\big\rangle \tag{20.23}$$

Les entiers positifs $n_{\vec{k}',\epsilon'}$ sont les *nombres d'occupation* associés aux différents états possibles des photons. Le nombre total de photons N, qui n'est pas fixé, est donné dans un état donné par la valeur moyenne de l'opérateur

$$\mathcal{N} \overset{\text{déf.}}{=} \sum_{\vec{k},\epsilon}\mathcal{N}_{\vec{k},\epsilon} \quad\Longrightarrow\; N = \sum_{\vec{k},\epsilon}\big\langle\{n_{\vec{k}',\epsilon'}\}\big|\,\mathcal{N}_{\vec{k},\epsilon}\,\big|\{n_{\vec{k}',\epsilon'}\}\big\rangle = \sum_{\vec{k},\epsilon} n_{\vec{k},\epsilon} \tag{20.24}$$

Les relations de commutation (20.19) indiquent que, à l'instar d'un oscillateur harmonique unique, on peut définir un état fondamental $|\Omega\rangle$ par :

$$\forall\vec{k},\ \forall\epsilon\ ,\quad a^\epsilon_{\vec{k}}|\Omega\rangle = 0\,. \tag{20.25}$$

cet état est caractérisé par des nombres d'occupation tous nuls

$$\forall \vec{k},\ \forall \epsilon\ ,\quad \mathcal{N}_{\vec{k}}^{\epsilon}\,|\Omega\rangle = 0 \implies n_{\vec{k}',\epsilon'} = 0 \tag{20.26}$$

À partir de cet état fondamental, on peut construire un état pour un nombre de particules N quelconque comme

$$\left|\{n_{\vec{k},\epsilon}\}\right\rangle \propto \prod_{\vec{k},\epsilon}\Big(\left(a_{\vec{k}}^{\epsilon}\right)^{\dagger}\Big)^{r_{\vec{k},\epsilon}}|\Omega\rangle\ ,\quad \sum_{\vec{k},\epsilon} r_{\vec{k},\epsilon} = N\,. \tag{20.27}$$

Par exemple, l'état correspondant à un photon unique de vecteur d'onde $\vec{k}$ et de polarisation ϵ est donné par le vecteur $\left|0,\ldots,0,n_{\vec{k},\epsilon}=1,0,\ldots,0\right\rangle = \left(a_{\vec{k}}^{\epsilon}\right)^{\dagger}|\Omega\rangle$.

Le potentiel vecteur classique (20.9) est associé par quantification canonique un *opérateur de champ* $\vec{\mathcal{A}}$:

$$\vec{\mathcal{A}}(\vec{x},t) = \sum_{\vec{k},\epsilon}\sqrt{\tfrac{\hbar}{2\epsilon_0\omega_k V}}\left(a_{\vec{k}}^{\epsilon}\,e^{i(\vec{k}\cdot\vec{x}-\omega_k t)} + \left(a_{\vec{k}}^{\epsilon}\right)^{\dagger}e^{-i(\vec{k}\cdot\vec{x}-\omega_k t)}\right)\vec{e}_{\vec{k}}^{\,\epsilon}\,, \tag{20.28}$$

dont l'action sur l'état fondamental va créer un état à un photon de la forme :

$$\vec{\mathcal{A}}(\vec{x},t)\,|\Omega\rangle = \sum_{\vec{k},\epsilon}\sqrt{\tfrac{\hbar}{2\epsilon_0\omega_k V}}\,e^{-i(\vec{k}\cdot\vec{x}-\omega_k t)}\,\vec{e}_{\vec{k}}^{\,\epsilon}\,\underbrace{\left(a_{\vec{k}}^{\epsilon}\right)^{\dagger}|\Omega\rangle}_{\text{état }\left|0,\ldots,0,n_{\vec{k},\epsilon}=1,0,\ldots,0\right\rangle} \tag{20.29}$$

Le formalisme développé ici correspond à la représentation de Heisenberg, comme l'opérateur de champ dépend du temps et les vecteurs d'état de l'espace de Fock $\mathcal{H}_f$ n'en dépendent pas. Cependant, le champ electromagnétique n'interagissant pas avec lui-même, l'évolution d'un état de la particule dans le temps est un simple facteur de phase.

Remarquons que nous avons implicitement supposé que les photons suivaient la statistique de Bose–Einstein car il est possible de mettre un nombre arbitraire de photons dans chaque état $(\vec{k},\epsilon)$, et que les permutations entre les photons n'ont aucun effet sur les vecteurs d'état (20.27). [1] L'espace des états ainsi construit est un exemple d'*espace de Fock*,

$$\mathcal{H}_f \overset{\text{déf.}}{=} \bigoplus_{N=0}^{\infty}\mathfrak{S}_N\mathcal{H}^{\otimes N}\,, \tag{20.30}$$

pour lequel le nombre de particules n'est pas fixé, voir le chapitre n° 6.

1. Si nous avions supposé que les photons étaient des fermions, il aurait fallu remplacer la relation de commutation (20.19) par des relations d'anticommutation

$$\left\{\,a_{\vec{k}}^{\epsilon},\,(a_{\vec{k}'}^{\epsilon'})^{\dagger}\,\right\} = \delta_{\vec{k},\vec{k}'}\delta_{\epsilon,\epsilon'}\quad,\qquad \left\{\,a_{\vec{k}}^{\epsilon},\,a_{\vec{k}'}^{\epsilon'}\,\right\} = 0$$

qui auraient impliqué que le hamiltonien (20.21) ne possède qu'une énergie propre $\hbar\omega_k/2$ pour un pulsation ω_k donnée :

$$\mathcal{H} = \frac{1}{2}\sum_{\vec{k},\epsilon}\hbar\omega_k\delta_{\vec{k},\vec{k}'}\delta_{\epsilon,\epsilon'}$$

ce qui n'est pas raisonnable physiquement (l'énergie d'une onde monochromatique n'est pas fixée !).

20.3 Hélicité du photon

Le théorème spin-statistique (voir sous-section 6.3 du cours n° 6) liant le spin d'une particule à sa statistique (Bose–Einstein ou Fermi–Dirac) nous pouvons nous demander quel est le spin des photons. Cependant, la notion de spin proprement dite n'existe pas pour les photons. Pour définir le spin d'une particule, il faut se placer dans le réferentiel dans lequel la particule est au repos et considérer la transformation de l'état de la particule sous des rotations de l'espace, donc des transformations de $SO(3)$ autour de sa position dans ce référentiel.

Pour les photons, ou plus généralement les particules de masse nulle, il n'existe aucun référentiel dans lequel la particule est au repos et il existera toujours une direction privilégiée dans l'espace, associée au vecteur d'onde $\vec{k}$ et donnant la direction de propagation. Le groupe de rotation pertinent est donc le groupe $SO(2)$ des rotations dans le plan orthogonal à $\vec{k}$.[2] Si $\vec{e}^{\,1}_{\vec{k}}$ et $\vec{e}^{\,2}_{\vec{k}}$ est une base orthonormée du plan orthogonal au vecteur $\vec{k}$, nous pouvons définir une base de vecteurs de polarisation circulaire $\vec{e}^{\,\epsilon}_{\vec{k}}$ avec $\epsilon = \pm$,

$$\vec{e}^{\,\pm}_{\vec{k}} = \mp \frac{\vec{e}^{\,1}_{\vec{k}} \pm i\vec{e}^{\,2}_{\vec{k}}}{\sqrt{2}} \tag{20.31}$$

tels que

$$\overline{\vec{e}^{\,\epsilon}_{\vec{k}}} \cdot \vec{e}^{\,\epsilon'}_{\pm\vec{k}} = \pm\delta_{\epsilon\epsilon'} \tag{20.32a}$$

$$\overline{\vec{e}^{\,\prime\epsilon}_{\vec{k}}} \wedge \vec{e}^{\,\epsilon'}_{\pm\vec{k}} = \pm i\epsilon\delta_{\epsilon\epsilon'}\vec{n}_{\vec{k}} \tag{20.32b}$$

ou nous avons $\vec{k} = k\vec{n}_{\vec{k}}$ avec $\vec{n}^{\,2}_{\vec{k}} = 1$. Les matrices de rotation dans le plan orthogonal au vecteur $\vec{k}$ sont donc de la forme :

$$R(\vec{n}_{\vec{k}}, \phi) = \exp \phi \vec{n}_{\vec{k}} \cdot \vec{j}, \tag{20.33}$$

en fonction des matrices 3×3 j_i correspondant à la représentation vectorielle de l'algèbre de Lie $\mathfrak{so}_3$, voir l'éqn. (4.7) du cours n° 4.

Les états de polarisation circulaire (20.32) sont états propres des rotations d'axe $\vec{n}_{\vec{k}}$ de valeurs propres $m = \pm 1$. En effet nous avons

$$\begin{aligned} \exp\left(\vec{n}_{\vec{k}} \cdot \vec{j}\,\phi\right) \vec{e}^{\,\pm}_{\vec{k}} &= \mp \frac{\left(\vec{e}^{\,1}_{\vec{k}} \cos\phi - \vec{e}^{\,2}_{\vec{k}} \sin\phi\right) \pm i\left(\vec{e}^{\,2}_{\vec{k}} \cos\phi + \vec{e}^{\,1}_{\vec{k}} \sin\phi\right)}{\sqrt{2}} \\ &= e^{\pm i\phi} \vec{e}^{\,\pm}_{\vec{k}}\,. \end{aligned} \tag{20.34}$$

2. Pour être plus précis, un photon d'énergie E est caractérisée par sa 4-impulsion qu'on peut choisir telle que $\vec{k}$ est selon l'axe Oz, c.-à.-d. $p^\mu_\circ = \frac{E}{c}(1,0,0,1)$. On définit alors le *petit groupe* associé comme l'ensemble des éléments du groupe de Lorentz laissant ce 4-vecteur invariant. En termes des générateurs infinitésimaux $\omega^\mu{}_\nu$ (avec $\omega^{\mu\nu} = -\omega^{\nu\mu}$) nous obtenons les conditions $\omega^\mu{}_\nu p^\nu_\circ = 0$, $\mu = 0, \ldots, 3$, qui réduisent le nombre de générateurs indépendants de 6 à 3. On montre qu'ils sont associés aux rotations et translations dans le plan (Ox, Oy), et donc aux générateurs infinitésimaux du groupe euclidien à deux dimensions $E(2) = O(2) \ltimes \mathbb{R}^2$.

Alors qu'une particule massive de spin $s = 1$ possède trois états de spin possibles, $m \in \{-1, 0, 1\}$, une particule de masse nulle et d'hélicité $h = 1$ ne possède que deux états d'hélicité, $m \in \{-1, 1\}$.

Il existe une relation profonde entre l'absence de masse du photon et l'invariance de jauge (19.19) de l'électromagnétisme. L'invariance de jauge de l'électromagnétisme implique qu'il n'existe que deux polarisations physiques du photons ; nous l'avons montré dans la jauge de Coulomb mais ceci reste vrai pour tout choix de jauge. Ce comptage est en accord avec l'existence de deux états d'hélicité seulement du photon. Ajouter un terme de masse au champ électromagnétique brise l'invariance de jauge (voir cours de théorie classique des champs), ce qui enlève la possibilité d'éliminer le troisième état de polarisation, et dans le même temps pour une particule vectorielle massive les deux états d'hélicité sont remplacé par les trois états de spin.

20.4 Énergie du vide

Le hamiltonien (20.21) peut, au moins formellement, se mettre sous la forme canonique de l'oscillateur harmonique pour chaque terme de la somme :

$$\mathcal{H} = \sum_{\vec{k},\epsilon} \hbar\omega_k \left((a_k^\epsilon)^\dagger a_k^\epsilon + 1/2 \right) = \sum_{\vec{k},\epsilon} \hbar\omega_k \left(\mathcal{N}_{\vec{k},\epsilon} + 1/2 \right). \tag{20.35}$$

Le terme constant est naturellement identifié comme l'énergie de l'état fondamental $|\Omega\rangle$:

$$\mathcal{H}\,|\Omega\rangle = \left(\sum_{\vec{k},\epsilon} \frac{\hbar\omega_k}{2} \right) |\Omega\rangle = \left(\hbar c \sum_{\vec{k}} \left\| \vec{k} \right\| \right) |\Omega\rangle \tag{20.36}$$

Ce terme est associé, comme pour l'oscillateur harmonique, aux fluctuations quantiques du vide, mais la somme sur tous les vecteurs d'ondes possibles produit un résultat infini ! Cette divergence est associée aux fluctuations du vide pour des très petites longueurs d'onde, et porte pour cette raison le nom de *divergence ultraviolette*.

On s'attend à ce que la physique soit modifier à très haute énergie, donc cette divergence n'est pas forcément en soi problématique – est-ce raisonnable de considérer des photons possédant une énergie supérieure à l'énergie de Planck ? Néamnoins il faut développer des méthodes pour extraire de ces expressions des quantités physiques.

Effet Casimir ⋆

L'énergie infinie du vide pour le champ électromagnétique ne possède pas en soi de signification physique (en tout cas lorsque le champ électromagnétique n'est pas couplé à la relativité générale) car nous pouvons redéfinir l'origine des énergies.

Cette énergie du vide est néanmoins responsable d'effets observables, comme l'*effet Casimir* [1], voir la figure 20.1, responsable de l'attraction entre deux plaques conductrices à faible distances.

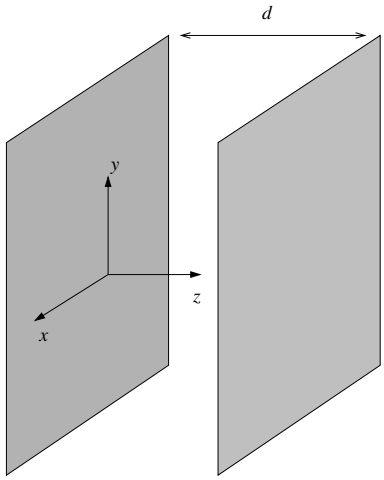

FIGURE 20.1 – Effet Casimir

On considère deux plaques parallèles supraconductrices, donc parfaitement conductrices, séparées selon l'axe Oz d'une distance d très petite devant la taille L de la cavité. Négligeant les effets de bord, le champ électromagnétique entre les plaques est soumis aux conditions aux limites suivantes :

— le champ électrique tangent aux plaques doit s'annuler, soit

$$E_{x,y}(x, y, z = 0, t) = E_{x,y}(x, y, z = L, t) = 0$$

— le champ magnétique normal aux plaques doit s'annuler, soit

$$B_z(x, y, z = 0, t) = B_z(x, y, z = L, t) = 0$$

Pour satisfaire ces conditions nous considérons, pour $k_z \neq 0$, un potentiel vecteur de la forme :

$$\vec{A} = \vec{e}^{\,\epsilon}_{\vec{k}} e^{i(k_x x + k_y y)} \sin(k_z z) e^{-i\omega_k t}$$

$$k_x = \frac{2\pi n_1}{L}\,,\ k_y = \frac{2\pi n_2}{L}\,,\ k_z = \frac{\pi n_3}{D}\,,\quad n_{1,2} \in \mathbb{Z}\,,\quad n_3 \in \mathbb{N}^\star \quad (20.37)$$

Alors que l'annulation du champ électrique tangent est automatique, pour annuler le champ magnétique normal il faut en outre imposer :

$$\vec{e}_z \cdot (\vec{k} \wedge \vec{e}^{\,\epsilon}_{\vec{k}}) = 0 \ \Leftrightarrow k_x e^{\epsilon}_{\vec{k},y} - k_y e^{\epsilon}_{\vec{k},x} = 0\,, \quad (20.38)$$

donnant au final deux polarisations possibles par valeur de $\vec{k}$.

Lorsque $k_z = 0$ on considère plutôt un potentiel vecteur de la forme $\vec{A} = \vec{e}^{\,\epsilon}_{\vec{k}} e^{i(k_x x + k_y y)} e^{-i\omega_k t}$ et l'annulation du champ électrique tangent aux deux plaques supraconductrices impose alors $\vec{e}^{\,\epsilon}_{\vec{k}} = \vec{e}_z$; il n'existe qu'une polarisation possible.

L'énergie électromagnétique du vide contenue dans le volume délimité selon Oz par les deux plaques, et selon les autres axes par la cavité, est donc donné par :

$$E_0(d) = \hbar c \sum_{n_1,n_2=\infty}^{\infty} \left(\frac{1}{2}\sqrt{\left(\frac{2\pi n_1}{L}\right)^2 + \left(\frac{2\pi n_2}{L}\right)^2} + \sum_{n_3=1}^{\infty} \sqrt{\left(\frac{2\pi n_1}{L}\right)^2 + \left(\frac{2\pi n_2}{L}\right)^2 + \left(\frac{\pi n_3}{d}\right)^2} \right) \tag{20.39}$$

Dans la limite $L \gg d$ on peut remplacer la somme sur n_1 et n_2 par une intégrale sur k_x et k_y :

$$E_0(d) = \hbar c L^2 \int \frac{dk_x}{2\pi} \int \frac{dk_y}{2\pi} \left(\frac{1}{2}\sqrt{k_x^2 + k_y^2} + \sum_{n_3=1}^{\infty} \sqrt{k_x^2 + k_y^2 + \left(\frac{\pi n_3}{d}\right)^2} \right) \tag{20.40}$$

On peut ensuite faire l'intégrale sur (k_x, k_y) en coordonnées polaires pour calculer directement l'intégrale angulaire et, en introduisant la variable radiale sans dimension $\nu = (k_x^2 + k_y^2)d^2/\pi^2$ on obtient l'expression suivante pour l'énergie par unité de surface :

$$\frac{1}{L^2}E_0(d) = \frac{\pi^2 \hbar c}{4d^3} \left\{ \frac{1}{2}\int_0^{\infty} d\nu\, \sqrt{\nu} + \sum_{n=1}^{\infty} \int_0^{\infty} d\nu\, \sqrt{\nu + n^2} \right\} . \tag{20.41}$$

Toutes les manipulations que nous avons faites dans le calcul sont complètement formelles, car la quantité que nous manipulons est divergente. Néanmoins on peut comparer le résultat (20.41) au résultat obtenu en l'absence des plaques et espérer que la différence entre les deux soit finie, car les modes responsables de la divergence, qui ont une longueur d'onde infiniment petite, existent dans les deux configurations.

En l'absence de plaques, l'énergie du vide dans le même volume est directement donnée par l'éqn. (20.36), soit, dans la limite continue :

$$E_0(\infty) = \hbar c \left(\frac{L}{2\pi}\right)^2 \left(\frac{d}{2\pi}\right) \int d^3k\, \sqrt{k_x^2 + k_y^2 + k_z^2} = \hbar c \frac{L^2 d}{8\pi^3} 2\pi \frac{\pi^2}{2d^2} \int d\nu \int dk_z\, \sqrt{\frac{\pi^2}{d^2}\nu + k_z^2} \tag{20.42}$$

En introduisant la variable sans dimension $x = k_z d/\pi$ on trouve finalement :

$$\frac{1}{L^2}E_0(\infty) = \frac{\pi^2 \hbar c}{4d^3} \int_0^{\infty} dx \int_0^{\infty} du\, \sqrt{u + x^2}\,. \tag{20.43}$$

La différence d'énergie par unité de surface est donc :

$$\frac{1}{L^2}E_0(d) - \frac{1}{L^2}E_0(\infty) = \frac{\pi^2 \hbar c}{4d^3} \left\{ \frac{1}{2}\int_0^{\infty} d\nu\, \sqrt{\nu} + \sum_{n=1}^{\infty} \int_0^{\infty} d\nu\, \sqrt{\nu + n^2} - \int_0^{\infty} dx \int_0^{\infty} du\, \sqrt{u + x^2} \right\} \tag{20.44}$$

Nous remarquons que l'expression entre accolades correspond à la différence entre la somme des valeurs entières positives de la fonction $f(x) = \int_0^{\infty} du\sqrt{u + x^2}$ et son intégrale sur l'intervalle $[0, +\infty[$. Les deux termes ont la même divergence ultraviolette que l'on peut régulariser en introduisant par exemple la fonction

$$f_\lambda(x) \overset{\text{déf.}}{=} \int_0^{\infty} du \sqrt{u + x^2} e^{-\lambda(u+x^2)} = \int_{x^2}^{\infty} \sqrt{t} e^{-\lambda t} \tag{20.45}$$

et en prenant la limite $\lambda \to 0^+$ à la fin du calcul ; naturellement le résultat final ne doit pas dépendre de la méthode de régularisation employée.

Pour évaluer la différence entre la somme et l'intégrale on peut évoquer le *théorème de Euler–Maclaurin* qui donne le développement asymptotique :

$$\frac{f_\lambda(0)}{2} + \sum_{n=1}^{\infty} f_\lambda(n) - \int_0^\infty f_\lambda(x)dx = -\frac{f'_\lambda(0)}{12} + \frac{f'''_\lambda(0)}{720} + \cdots \tag{20.46}$$

la somme se poursuivant sur les dérivées impaires et les coefficients étant reliés aux nombres de Bernoulli. Nous avons :

$$f'_\lambda(x) = -2x^2 e^{-\lambda x^2} \ , \quad f'''_\lambda(x) = -4e^{-\lambda x^2}\left(2x^4\lambda^2 - 5x^2\lambda + 1\right) \tag{20.47}$$

et toutes les dérivées d'ordre supérieur sont d'ordre $\mathcal{O}(\lambda)$ et s'annulent donc lorsque le régulateur est enlevé à la fin du calcul. On remarque également que le terme $f_\lambda(0)/2$ dans (20.46) correspond au premier terme entre accolades dans (20.44), provenant des modes avec $k_z = 0$.

Nous obtenons donc finalement dans la limite $\lambda \to 0^+$ la différence entre l'énergie du vide par unité de surface en présence et en l'absence de plaques :

$$\boxed{\frac{1}{L^2}\big(E_0(d) - E_0(\infty)\big) = -\frac{\pi^2}{720}\frac{\hbar c}{d^3}} \tag{20.48}$$

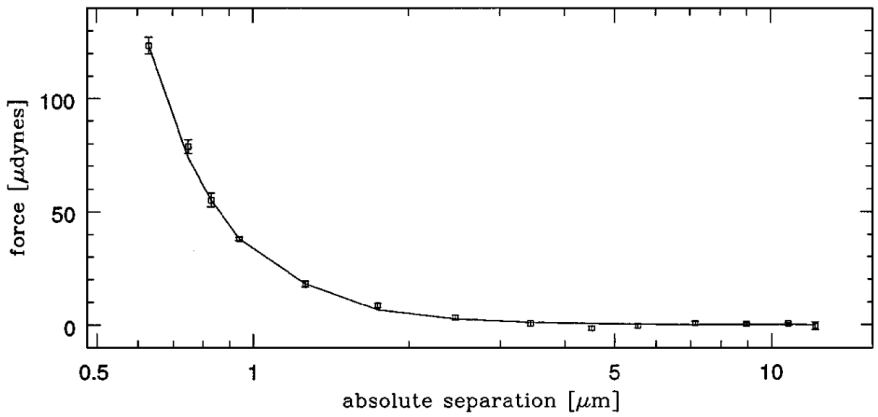

FIGURE 20.2 – Mesure de l'effet Casimir : comparaison des points expérimentaux avec la courbe théorique (source : S. K. Lamoreaux [2])

Ce résultat remarquable permet de prédire l'existence d'une force macroscopique entre les deux plaques d'origine quantique. Cette force par unité de surface est donnée simplement par la dérivée par rapport à d de l'énergie surfacique obtenue, soit :

$$\frac{F_z}{L^2} = -\frac{\partial}{\partial z}\frac{E_0(d) - E_0(\infty)}{L^2} \quad \Longrightarrow \quad \boxed{\frac{F_z}{L^2} = -\frac{\pi^2}{240}\frac{\hbar c}{d^4}} \tag{20.49}$$

Cette force est extrêmement faible tant que la distance entre les plaques reste macroscopique. Par exemple, pour une distance d'un micromètre on trouve que $F/L^2 \simeq 4 \times 10^{-4}$ N/m^2. En conséquence sa mesure expérimentale indiscutable est assez récente [2], voir fig. 20.2.

20.5 Le modèle de Jaynes–Cummings

La quantification du champ électromagnétique modifie la physique de son interaction avec les atomes que nous avions étudiée dans le cours n° 19. Nous prendrons ici un modèle très simplifié, le *modèle de Jaynes–Cummings* [3], pour analyser certains des nouveaux phénomènes dus à la quantification du champ. Nous suivons ici de près l'exposition de David Tong [4]

Premièrement, on modélise l'atome par un simple système à deux niveaux $|\downarrow\rangle$ et $|\uparrow\rangle$, étant supposé que les autres niveaux ne participent pas aux transitions étudiées dans le modèle. Le hamiltonien de l'atome est donc simplement : [3]

$$\mathcal{H}_a = \frac{\hbar\omega_0}{2}\big(\,|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|\,\big)\,. \tag{20.50}$$

Deuxièmement, l'atome est placé dans une cavité supraconductrice de taille L, de telle sorte que seuls les photons de longueur d'onde $\lambda = 2L/n$ puissent s'y trouver. On choisit les dimensions de la cavité de telle sorte que $\omega = 4\pi L/c$ soit du même ordre de grandeur que ω_0, et dans ces conditions seul le mode fondamental ($n = 1$) jouera un rôle significatif dans l'interaction avec l'atome. Si nous ignorons les effets de la polarisation (et utilisons l'approximation dipolaire pour s'affranchir de la dépendance dans la direction du vecteur d'onde $\vec{k}$ du photon) le hamiltonien simplifié du champ électromagnétique est donc un simple oscillateur harmonique :

$$\mathcal{H}_e = \hbar\omega(a^\dagger a + 1/2)\,, \tag{20.51}$$

avec $\omega = 4\pi L/c \sim \omega_0$. Une base de l'espace des états du champ est donc $\{|n\rangle\,,\, n \in \mathbb{N}\}$.

Enfin, l'interaction entre l'atome et le champ, analogue du couplage dipolaire (19.57) dans ce modèle simplifié, doit pouvoir rendre compte de la transition $|\downarrow\rangle \to |\uparrow\rangle$ avec absorption d'un photon, et de la transition $|\uparrow\rangle \to |\downarrow\rangle$ avec émission d'un photon. Nous écrirons donc :

$$\mathcal{H}_i = \frac{\hbar\omega_1}{2}\big(a^\dagger\,|\downarrow\rangle\langle\uparrow| + a\,|\uparrow\rangle\langle\downarrow|\,\big)\,. \tag{20.52}$$

ou la pulsation ω_1 représente l'intensité du couplage entre l'atome et le champ.

Le hamiltonien total du système, appelé hamiltonien de Jaynes–Cummings, est donné par la somme de ces trois termes :

$$\boxed{\mathcal{H}_{JC} = \frac{\hbar\omega_0}{2}\big(\,|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|\,\big) + \hbar\omega(a^\dagger a + 1/2) + \frac{\hbar\omega_1}{2}\big(a^\dagger\,|\downarrow\rangle\langle\uparrow| + a\,|\uparrow\rangle\langle\downarrow|\,\big)} \tag{20.53}$$

Ce modèle simple a été beaucoup étudié et utilisé car il est exactement soluble. Il permet également d'analyser la physique des exprériences d'électrodynamique en cavité, comme celles menées par l'équipe de Serge Haroche [5]. Dans ces expériences, le champ électromagnétique dans la cavité est préparé dans un certain état quantique, avant d'être sondé par le passage d'un atome.

Nous introduirons pour la suite le paramètre de *désaccord* entre la cavité et l'atome à deux niveaux, défini par :

$$\delta \overset{\text{déf.}}{=} \omega - \omega_0\,. \tag{20.54}$$

3. Nous utilisons par commodité la notation $|\uparrow\rangle$, $|\downarrow\rangle$ dans ce problème même si le spin n'y joue pas de rôle.

L'espace des états du système est naturellement formé du produit tensoriel des états de l'atome et de ceux du champ électromagnétique, et nous choisirons la base orthonormée :

$$\{|\uparrow\rangle \otimes |n\rangle\,,\, |\downarrow\rangle \otimes |n\rangle\,;\, n \in \mathbb{N}\} \tag{20.55}$$

et nous noterons de manière condensée $|\uparrow, n\rangle \overset{\text{déf.}}{=} |\uparrow\rangle \otimes |n\rangle$ et $|\downarrow, n\rangle \overset{\text{déf.}}{=} |\downarrow\rangle \otimes |n\rangle$.

Contrairement au hamiltonien de couplage avec le champ classique, le hamiltonien (20.53) est indépendant du temps (dans la représentation choisie). L'état fondamental du système est naturellement $|\downarrow; 0\rangle$ où l'atome est dans son état fondamental et où le nombre de photons est nul :

$$\mathcal{H}_{\text{JC}}\,|\downarrow; 0\rangle = \frac{\hbar\delta}{2}\,|\downarrow; 0\rangle\ . \tag{20.56}$$

Les autres états de la base (20.55) ne sont pas états propres du hamiltonien mais le mélange concerne simplement les états par paires. En effet un état $|\downarrow, n\rangle$ avec $n \geqslant 1$ sera apparié avec un état $|\uparrow, n-1\rangle$, le processus correspondant étant l'absorption ou l'émission d'un photon.

Le hamiltonien est donc diagonal par blocs 2×2, et il suffit de rechercher les vecteurs propres pour chacun des blocs. Nous avons :

$$\mathcal{H}_{\text{JC}}\,|\downarrow, n\rangle = \left(\tfrac{\hbar\delta}{2} + \hbar\omega n\right)|\downarrow, n\rangle + \tfrac{\hbar\omega_1}{2}\sqrt{n}\,|\uparrow, n-1\rangle\ , \tag{20.57a}$$

$$\mathcal{H}_{\text{JC}}\,|\uparrow, n-1\rangle = \left(-\tfrac{\hbar\delta}{2} + \hbar\omega n\right)|\uparrow, n-1\rangle + \tfrac{\hbar\omega_1}{2}\sqrt{n}\,|\downarrow, n\rangle\ . \tag{20.57b}$$

La diagonalisation du hamiltonien revient donc dans chaque bloc à la diagonalisation de la matrice

$$M = \begin{pmatrix} \delta & \omega_1\sqrt{n} \\ \omega_1\sqrt{n} & -\delta \end{pmatrix} = \delta\sigma_3 + \omega_1\sqrt{n}\sigma_1\,. \tag{20.58}$$

La résolution de ce problème est bien connue. On introduit θ, « angle de mélange » entre les deux états, tel que :

$$\tan 2\theta = \frac{\omega_1\sqrt{n}}{\delta}\,, \tag{20.59}$$

et les deux vecteurs propres, d'énergies propres

$$E_\pm = \hbar\omega n \pm \frac{\hbar}{2}\sqrt{\delta^2 + n\omega_1^2}\,, \tag{20.60}$$

sont donnés par

$$|n_+\rangle = \cos\theta\,|\downarrow, n\rangle + \sin\theta\,|\uparrow, n-1\rangle\ , \tag{20.61a}$$

$$|n_-\rangle = -\sin\theta\,|\downarrow, n\rangle + \cos\theta\,|\uparrow, n-1\rangle + \ . \tag{20.61b}$$

Les vecteurs propres du hamiltonien (20.53) étant trouvés, nous pouvons analyser, dans le cadre du modèle simplifié de Jaynes–Cummings, l'interaction entre l'atome et le champ électromagnétique quantique présent dans la cavité.

20.5.1 Oscillations de Rabi

L'analyse du modèle de Jaynes–Cummings permet de retrouver le phénomène d'oscillations de Rabi déjà étudié en présence d'un champ classique. Préparons le système dans l'état initial $|\downarrow, n\rangle$, c.-à.-d. l'atome dans l'état fondamental en présence d'un nombre fixe n de photons. Le système est décrit à un instant $t \geqslant 0$ quelconque par :

$$\begin{aligned}|\psi(t)\rangle = e^{-\frac{it}{\hbar}\mathcal{H}_{JC}}|\downarrow, n\rangle = e^{-\frac{it}{\hbar}\mathcal{H}_{JC}}(\cos\theta\,|n_+\rangle - \sin\theta\,|n_-\rangle)\\ = \cos\theta e^{-\frac{itE_+}{\hbar}}|n_+\rangle - \sin\theta e^{-\frac{itE_-}{\hbar}}|n_-\rangle\ , \quad (20.62)\end{aligned}$$

et la probabilité de transition vers l'état $|\uparrow, n-1\rangle$ où l'atome est passé à l'état excité en absorbant un photon est donnée par :

$$\begin{aligned}p(t) = |\langle\uparrow, n-1|\psi(t)\rangle|^2 = \left|\cos\theta\sin\theta e^{-\frac{itE_+}{\hbar}} - \cos\theta\sin\theta e^{-\frac{itE_-}{\hbar}}\right|^2\\ = \sin^2(2\theta)\sin^2\frac{(E_+ - E_-)t}{2\hbar} = \frac{\omega_1^2 n}{\delta^2+\omega_1^2 n}\sin^2\left(\frac{t}{2}\sqrt{\delta^2+n\omega_1^2}\right)\ . \quad (20.63)\end{aligned}$$

cela correspond bien au résultat (18.24) avec le remplacement $\omega_1^2 \to n\omega_1^2$ car l'énergie du champ électromagnétique est proportionnelle au nombre de photons.

20.5.2 Résurgences

Mis à part les oscillations de Rabi le modèle de Jaynes–Cummings met en évidence des effets liés au caractère quantique du champ électromagnétique. Considérons que le champ électromagnétique est dans un état cohérent avec un grand nombre moyen de photons :

$$|\alpha\rangle = e^{-|\alpha|^2/2}e^{\alpha a^\dagger}|0\rangle\ , \quad \langle\alpha|\mathcal{N}|\alpha\rangle = |\alpha|^2 \gg 1\,. \quad (20.64)$$

Rappelons qu'un état cohérent est, pour un oscillateur harmonique, un état minimisant à tout instant la relation d'incertitude et dont l'évolution au cours du temps des valeurs moyennes de position et d'impulsion correspond à celle d'un oscillateur classique. Dans le cadre de la quantification du champ électromagnétique, les états cohérents correspondent au rayonnement émis par les lasers.

Si nous préparons le système (atome+champ) dans un état $|\downarrow, \alpha\rangle$ son évolution au cours du temps donne :

$$\begin{aligned}|\psi(t)\rangle = e^{-\frac{it}{\hbar}\mathcal{H}_{JC}}e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}|\downarrow, n\rangle\\ = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\alpha^n}{\sqrt{n!}}\left(\cos\theta e^{-\frac{itE_+}{\hbar}}|n_+\rangle - \sin\theta e^{-\frac{itE_-}{\hbar}}|n_-\rangle\right)\\ = e^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{\left(\alpha e^{-i\omega t}\right)^n}{\sqrt{n!}}\left(\cos\theta e^{-\frac{it}{2}\sqrt{\delta^2+n\omega_1^2}}|n_+\rangle - \sin\theta e^{\frac{it}{2}\sqrt{\delta^2+n\omega_1^2}}|n_-\rangle\right)\ , \quad (20.65)\end{aligned}$$

la probabilité de passer à l'état excité de l'atome est la somme sur tous les états correspondants de la base produit tensoriel, c.-à.-d.

$$\begin{aligned} p(t) &= \sum_{n=1}^{\infty} |\langle n-1, \uparrow | \psi(t) \rangle|^2 \\ &= \sum_{n=1}^{\infty} e^{-|\alpha|^2} \left| \frac{\left(\alpha e^{-i\omega t}\right)^n}{\sqrt{n!}} \right|^2 \left| \cos\theta \sin\theta e^{-\frac{it}{2}\sqrt{\delta^2+n\omega_1^2}} - \cos\theta \sin\theta e^{\frac{it}{2}\sqrt{\delta^2+n\omega_1^2}} \right|^2, \end{aligned} \quad (20.66)$$

soit finalement :

$$\boxed{p(t) = e^{-|\alpha|^2} \sum_{n=1}^{\infty} \frac{\omega_1^2 n}{\delta^2 + \omega_1^2 n} \frac{|\alpha|^{2n}}{n!} \sin^2\left(\frac{t}{2}\sqrt{\delta^2 + n\omega_1^2}\right)} \quad (20.67)$$

Examinons pour simplifier la discussion le système lorsque la cavité est résonante avec l'atome ($\omega = \omega_0$). Nous avons alors plus simplement :

$$p(t) = e^{-|\alpha|^2} \sum_{n=1}^{\infty} \frac{|\alpha|^{2n}}{n!} \sin^2\left(\frac{\omega_1 t}{2}\sqrt{n}\right) \quad (20.68)$$

Cette probabilité est une somme de sinusoïdes dont les fréquences ne sont pas dans des rapports rationnels ; en conséquence, nous n'aurons pas de superposition cohérente de ces différents termes aux temps longs mais plutôt des déphasages aléatoires.

Pour étudier plus quantitativement les phénomènes apparaissant aux temps longs réécrivons l'éqn. (20.68) comme

$$p(t) = e^{-|\alpha|^2} \sum_{n=1}^{\infty} \frac{|\alpha|^{2n}}{n!} \left(\frac{1}{2} - \frac{1}{2}\cos\left(\sqrt{n}\omega_1 t\right)\right) = \frac{1}{2} - \frac{1}{2} e^{-|\alpha|^2} \sum_{n=1}^{\infty} \frac{|\alpha|^{2n}}{n!} \cos\left(\sqrt{n}\omega_1 t\right) . \quad (20.69)$$

La somme dans le second terme est fortement dominée par les termes avec $n \sim |\alpha|^2$. En effet nous avons, à l'aide de la formule de Stirling,

$$\ln \frac{|\alpha|^{2n}}{n!} = n \ln |\alpha|^2 + n - n \ln n + \mathcal{O}(\ln n) , \quad (20.70)$$

qui possède un maximum en $n = |\alpha|^2$. [4] Un développement de Taylor autour du maximum donne :

$$e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!} \simeq \frac{1}{\sqrt{2\pi|\alpha|^2}} e^{-\frac{(n-|\alpha|^2)^2}{2|\alpha|^2}} . \quad (20.71)$$

La probabilité de transition est donc finalement donnée dans cette approximation, valide pour $|\alpha| \gg 1$, par :

$$\boxed{p(t) \simeq \frac{1}{2} - \frac{1}{\sqrt{8\pi|\alpha|^2}} \sum_{n=-\infty}^{\infty} e^{-\frac{(n-|\alpha|^2)^2}{2|\alpha|^2}} \cos\left(\sqrt{|\alpha|^2 + (n-|\alpha|^2)}\omega_1 t\right)} \quad (20.72)$$

4. Mathématiquement, nous avons donné une approximation de la loi de Poisson par une loi normale, ce qui est valable lorsque son espérance est grande.

où nous avons étendu sans inconvénient, la distribution étant très piquée et n'étant pas centrée en zéro, la borne inférieure de la somme à $-\infty$.

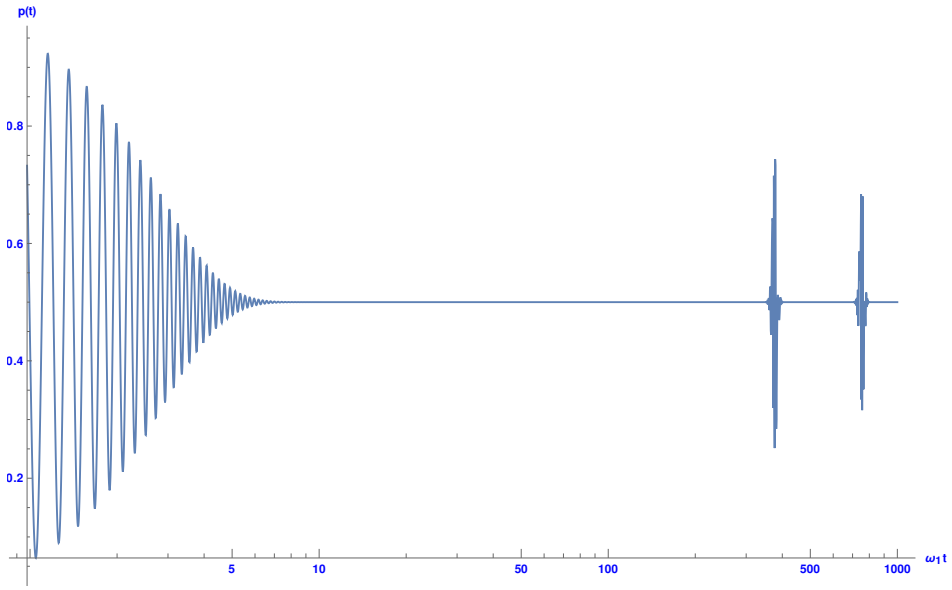

FIGURE 20.3 – Résurgences dans le modèle de Jaynes–Cummings, pour $|\alpha| = 30$ (échelle logarithmique)

Comme nous l'avons vu, les contributions significatives à cette somme sont dans l'intervalle $\left|n - |\alpha|^2\right| < |\alpha|$. Au bout d'un temps t le déphasage maximal entre les différents termes composant le signal est donc :

$$\Delta\varphi = \sqrt{|\alpha|^2 + |\alpha|}\,\omega_1 t - \sqrt{|\alpha|^2 - |\alpha|}\,\omega_1 t = \omega_1 t + \mathcal{O}(1/|\alpha|)\,. \tag{20.73}$$

Nous nous attendons alors au régimes suivants, voir la figure 20.3 :

— pour $\omega_1 t \ll 2\pi$ le déphasage entre les différents termes de la somme (20.72) sont faibles et nous trouvons des oscillations de Rabi amorties de période $T \sim \frac{2\pi}{\omega_1|\alpha|}$;

— lorsque $\omega_1 t \gtrsim 2\pi$ on observe un effondrement des oscillations dû aux interférences destructrices et la probabilité se stabilise à $p = 1/2$, signifiant que les deux états de l'atome, $\uparrow$ et $\downarrow$, sont équiprobables ;

— à des temps beaucoup plus longs, les phases des termes consécutifs dans la somme (20.72) autour du maximum en $n = |\alpha|$ vont approximativement « s'aligner ». Cela est possible pour la première fois si

$$\omega_1 t_{\text{REG}} \sqrt{|\alpha|^2 + 1} - \omega_1 t_{\text{REG}} \sqrt{|\alpha|^2} = 2\pi \quad \Longrightarrow \quad \boxed{t_{\text{REG}} \simeq \frac{4\pi|\alpha|}{\omega_1}} \tag{20.74}$$

et la probabilité va brusquement osciller avec une amplitude significative (mais inférieure à 1/2). Ce phénomène spectaculaire est appelé *résurgence quantique*, et est une conséquence du caractère discret du spectre du champ électromagnétique quantique.

À retenir

— Décomposition en modes du champ électromagnétique ;

— quantification canonique du champ électromagnétique ;

— helicité des photons ;

— énergie du vide du champ et effet Casimir ;

— modèle de Jaynes–Cummings, propriétés et résurgences quantiques.

Bibliographie

[1] H. B. Casimir et D. Polder, “The Influence of Retardation on the London-van der Waals Forces,” Phys. Rev. 73, 1948

[2] S. K. Lamoreaux, “Demonstration of the Casimir Force in the 0.6 to 6 μm Range,” Phys. Rev. Lett. 78 (1997) year = ”1997”

[3] E. T. Jaynes and F. W. Cummings, “Comparison of quantum and semiclassical radiation theories with application to the beam maser,” IEEE Proc. **51** (1963) 89. doi :10.1109/PROC.1963.1664

[4] D. Tong, “Applications of Quantum Mechanics,” University of Cambridge, voir ce lien.

[5] M. Brune, F. Schmidt-Kaler, A. Maali, J. Dreyer, E. Hagley, J. M. Raimond, et S. Haroche, “Quantum Rabi Oscillation : A Direct Test of Field Quantization in a Cavity,” Phys. Rev. Lett. 76 (1996)