

Alexandre Feller, Clément Cabart et Benjamin Roussel

vous présentent

Need to Know

Version compilée le 25 décembre 2015

La version la plus récente est toujours disponible à l'adresse suivante http://cours.atdot.eu/Need_to_Know.pdf



Richard Phillips Feynman (May 11, 1918 - February 15, 1988)

How it came to be

SPLEEN

*Quand le ciel bas et lourd pèse comme un couvercle
Sur l'esprit gémissant en proie aux longs ennuis,
Et que l'horizon embrassant tout le cercle
Il nous verse un jour noir plus triste que les nuits ;*

*Quand la terre est changée en un cachot humide,
Où l'Espérance, comme une chauve-souris,
S'en va battant les murs de son aile timide
Et se cognant la tête à des plafonds pourris ;*

*Quand la pluie étalant ses immenses traînées
D'une vaste prison imite les barreaux,
Et qu'un peuple muet d'infâmes araignées
Vient tendre ses filets au fond de nos cerveaux,*

*Des cloches tout à coup sautent avec furie
Et lancent vers le ciel un affreux hurlement,
Ainsi que des esprits errant et sans patrie
Qui se mettent à geindre opiniâtrement.*

*- Et de longs corbillards, sans tambours ni musique,
Défilent lentement dans mon âme ; l'Espoir,
Vaincu, pleure, et l'Angoisse atroce, despotique,
Sur mon crâne incliné plante son drapeau noir.*

BAUDELAIRE, *Les Fleurs du Mal*

Ce n'était à la base qu'une simple idée. Une petite envie de faire quelque chose d'utile, pour une fois, un simple résumé. Mais comme beaucoup de choses que je fais, ce n'est jamais si simple. C'est devenu une obsession, une envie furieuse de faire quelque chose de bien, de beau.

Pourquoi? Cette envie m'est venue sur la route, en rentrant chez moi. Comme la plupart du temps, je me sens seul. Et je l'étais. Cette douce mélancolie est toujours présente, parfois cachée, parfois non, comme un fardeau, un démon. À ce moment, j'étais étrange, triste. Avec le temps, cette idée devenue obsession, s'est trouvée être mon moyen d'oublier. Il y a certaines choses que j'ai dites ou faites que je regrette. Que je regrette vraiment énormément. Pourtant, je sais maintenant que cela ne pourra pas changer. Pour l'instant en tout cas. Je pense avoir perdu beaucoup par ma bêtise et quoique je fasse, c'est fini. Je dois vivre avec mes regrets, mes remords et mes souvenirs. Mais le plus dur est de dire au revoir à ce que l'on a perdu. Dans ma faiblesse, j'ai choisi d'essayer d'oublier, mais au fond de moi je sais que j'en suis incapable. Il n'est pas si facile de dire adieu à quelque chose qui nous a beaucoup plus apporté que ce que l'on peut imaginer. Mais cela il faut le comprendre avant de l'avoir détruit.

L'écriture de ce texte m'a, un temps, permis de me concentrer sur autre chose, d'oublier. De croire qu'il était possible qu'en quelques semaines les choses puissent s'améliorer. D'une certaine manière, ce fut le cas. Dans mon égoïsme, je ne voulais rien partager, n'en voyant pas l'intérêt pour quiconque, hormis me permettre de penser que je puisse faire une chose utile, de bien. Je ne pense pas en être capable et je pense toujours que cela ne servira pas à grand monde mais Clément m'a en quelque sorte convaincu que cela pouvait parfois aider. Son investissement dans l'écriture et celui de Benjamin n'a pu qu'améliorer un texte qui sans cela aurait été froid. Je ne sais pas si j'ai déjà réellement aidé quelqu'un. Je pense, mais ce n'est peut-être qu'une illusion.

Pour le moment c'est fini mais je sais qu'elle, ma mélancolie, n'est pas partie...

Table des matières

1	Relativité restreinte	9
1.1	Principe de relativité	9
1.1.1	Énoncé du principe de relativité restreinte	9
1.1.2	Temps propre et dilatation des durées	10
1.1.3	Contraction des longueurs	10
1.1.4	Transformation de Lorentz	10
1.1.5	Composition des vitesses	11
1.1.6	Intervalle	11
1.2	Dynamique relativiste	12
1.2.1	Quadri-vecteurs vitesse et impulsion	12
1.2.2	Particule libre massive	13
1.2.3	Énergie et impulsion	15
1.2.4	Lois de conservation	16
1.3	Espace-temps de Minkowski et relativité restreinte	16
1.3.1	L'espace-temps	16
1.3.2	Ligne d'univers, temps propre and co.	18
1.3.3	Observateurs	20
1.3.4	Cinématique	21
1.3.5	Pour aller plus loin	21
2	Électrodynamique	22
2.1	Les équations de Maxwell et les changements de référentiel	22
2.1.1	Les équations de Maxwell	22
2.1.2	Équations de conservation	23
2.1.3	Potentiels et jauges	25
2.1.4	Résumé	27
2.1.5	Transformation des champs sous une transformation de Galilée	28
2.2	Solutions des équations de Maxwell, Potentiels retardés	29
2.2.1	Potentiels retardés	29
2.2.2	Potentiels de Liénard-Wiechert	30
2.2.3	Charge en mouvement rectiligne uniforme	30
2.3	Rayonnement électromagnétique	31
2.3.1	Les champs électrique et magnétique rayonnés	31
2.3.2	Puissance rayonnée	32
2.4	Formulation relativiste de l'électrodynamique	33
2.4.1	Quadri-vecteur courant	33
2.4.2	Quadri-vecteur potentiel	34
2.4.3	Équations de Maxwell-Lorentz	34
2.4.4	Transformation des champs	35
2.4.5	Résumé	35
2.4.6	Invariant de Lorentz	36
2.4.7	Tenseur énergie-impulsion	37
2.5	Formulation lagrangienne de l'électrodynamique	38
2.5.1	Particules libres dans un champ imposé	39
2.5.2	Champs en fonction des sources	40
2.5.3	Résumé	42
2.5.4	Lois de conservation	42

3	Mécanique quantique	44
3.1	Approche de Feynman	44
3.1.1	Amplitudes de probabilités	44
3.1.2	Intégrale de chemin	45
3.1.3	Fonction d'onde	46
3.2	Postulats de la mécanique quantique	47
3.2.1	États quantiques et règle de Born	47
3.2.2	Équation de Schrödinger	47
3.2.3	La mesure	48
3.2.4	Schémas de Schrödinger et de Heisenberg	50
3.2.5	Représentation, opérateurs position et impulsion, commutation	50
3.2.6	Principe d'indétermination	51
3.3	Propriétés élémentaires	52
3.3.1	Particule libre	52
3.3.2	Équation de Schrödinger indépendante du temps	53
3.3.3	Résolution de cette équation	54
3.3.4	États liés du puits carré	56
3.3.5	Espace des phases et densité de niveaux	56
3.3.6	Diffusion par un potentiel : effet tunnel	57
3.4	Théorie élémentaire de la diffusion	60
3.4.1	Notions générales	61
3.4.2	Méthode par résolvante	61
3.4.3	Approximation de Born	65
3.5	Perturbation dépendante du temps	65
3.5.1	Schéma en interaction	65
3.5.2	Méthode d'approximation	66
3.5.3	Règle d'or de Fermi	69
3.6	Opérateur densité	69
3.6.1	États intriqués	69
3.6.2	Opérateur densité : définitions	70
3.6.3	États purs, cohérences et mélange	70
3.6.4	Opérateur densité réduit	71
3.6.5	Évolution temporelle	72
3.6.6	Mesures et matrice densité	72
3.7	Opérateurs de champs	73
3.7.1	Oscillateur harmonique	73
3.7.2	Système de N particules identiques	75
3.7.3	Particules identiques en nombre indéterminé	76
3.7.4	Opérateurs de champs	79
3.8	Généralités sur les symétries	83
3.8.1	Groupe de transformations	83
3.8.2	Groupe de Lie	83
3.8.3	Représentation, théorème de Wigner	84
3.8.4	Groupe et algèbre de Lie	85
3.8.5	Symétries	87
3.9	Rotations et moments angulaires	87
3.9.1	Groupe des rotations	87
3.9.2	Moment cinétique	89
4	Physique nucléaire	92
4.1	Introduction	92
4.1.1	L'expérience de Rutherford	92
4.1.2	Le noyau atomique	93
4.1.3	Sur la force nucléaire forte	93
4.2	Premières observations expérimentales	94
4.2.1	Masse et énergie de liaison	94
4.2.2	Les lois de la radioactivité	94
4.3	Modèle de la goutte liquide	96
4.3.1	Formule de Weizsäcker	97

4.3.2	Applications	98
4.4	Étude détaillée de la radioactivité	99
4.4.1	Désintégration α , β	99
4.4.2	Fission et fusion nucléaire	102
4.5	Structure nucléaire : modèle en couches	104
4.5.1	Modèle du gaz de Fermi	104
4.5.2	Potentiel nucléaire moyen	105
4.5.3	Modèle en couches	106
4.6	Taille et forme des noyaux	108
4.6.1	Section efficace	109
4.6.2	Section efficace différentielle	109
4.6.3	Taille et densité des noyaux	109
4.6.4	Moments nucléaires	111
5	Physique des particules	113
5.1	Observation du monde des particules élémentaires	113
5.2	Particules et interactions	113
5.2.1	Introduction aux monde des particules	113
5.2.2	Nombre leptonique, nombre de saveur	114
5.2.3	L'isospin	114
5.2.4	Sur les interactions fondamentales	115
5.3	Introduction aux théories de jauge	118
5.3.1	La quantification canonique	118
5.3.2	Théorie de Maxwell et invariance de jauge	119
5.3.3	Invariance de jauge locale	120
5.3.4	Modèle standard	121
5.4	Équation de Klein-Gordon	121
5.4.1	Dérivation de l'équation de Klein-Gordon	121
5.4.2	Interprétation probabiliste	122
5.4.3	Formulation lagrangienne et invariance de jauge	122
5.4.4	Résumé	123
5.5	Équation de Dirac	123
5.5.1	Moment cinétique, spin et matrices de Pauli	123
5.5.2	L'équation de Dirac	127
5.5.3	Interprétation probabiliste	128
5.5.4	États stationnaires	129
5.5.5	Interprétation de Feynman	130
5.5.6	Spineurs et énergie négative	130
5.5.7	Le spin	131
5.5.8	Formulation lagrangienne et invariance de jauge	131
5.6	Interaction électromagnétique et diagramme de Feynman	132
5.6.1	Sur la section efficace	132
5.6.2	Interaction pour un boson de spin 0	133
5.6.3	Interaction pour un fermion de spin 1/2	134
5.6.4	Diagramme de Feynman	136
5.6.5	Diffusion électron-positron	137
6	Physique statistique	141
6.1	Thermodynamique	141
6.1.1	Introduction et définitions	141
6.1.2	Le postulat fondamental et l'entropie	142
6.1.3	Variables conjuguées, état d'équilibre	143
6.1.4	Coefficients calorimétriques	146
6.1.5	Les principes	146
6.1.6	Fonctions et potentiels thermodynamiques	148
6.1.7	Le gaz parfait	151
6.2	Fondements de la physique statistique	151
6.2.1	Description statistique	151
6.2.2	Entropie statistique	152

6.2.3	Ensemble microcanonique	153
6.2.4	Ensemble canonique	154
6.2.5	Ensemble grand-canonique	156
6.3	Physique statistique classique	158
6.3.1	Description statistique	158
6.3.2	Du quantique au classique : formulation de Feynman-Kac	158
6.4	Statistiques quantiques	160
6.4.1	Statistique de Bose-Einstein	161
6.4.2	Statistique de Fermi-Dirac	161
6.5	Gaz réels et liquides	162
6.5.1	Introduction	162
6.5.2	Calculs préliminaires	162
6.5.3	Développement du Viriel	163
6.5.4	Corrélations	165
6.6	Théorie cinétique des gaz	166
7	Matière condensée	169
7.1	Sur la conduction dans les métaux	169
7.1.1	Modèle de Drude	169
7.1.2	Effet Hall	170
7.2	Structure cristalline	171
7.2.1	Cristal et réseau	171
7.2.2	Réseau réciproque et diffraction	172
7.2.3	Détermination d'une structure cristalline par diffusion	173
7.3	Dynamique d'un réseau cristallin	175
7.3.1	Modèle classique du cristal harmonique	175
7.3.2	Modèle quantique du cristal harmonique	179
7.4	Structure électronique	181
7.4.1	Modèle des électrons libres	181
7.4.2	Théorème de Bloch	183
7.4.3	Électrons quasi-libres	184
7.4.4	Modèle des liaisons fortes	186
7.5	Sur la conduction dans les métaux : approche semi-classique	189
7.5.1	Modèle semi-classique	189
7.5.2	Bandes remplies	190
7.5.3	Les électrons et les trous	191
7.6	Énergie de cohésion	192
7.6.1	Cristaux moléculaires	192
7.6.2	Cristaux ioniques	192
7.6.3	Cristaux covalents et métaux	193
8	Transition de phases	194
8.1	Généralités sur les transitions de phases	195
8.2	Le modèle d'Ising	196
8.2.1	Transition ferromagnétique : approche phénoménologique	196
8.2.2	Description du modèle d'Ising	196
8.2.3	Modèle à une dimension	197
8.2.4	La limite thermodynamique	198
8.2.5	Notion de paramètre d'ordre et classification	199
8.2.6	Fonction de corrélation	199
8.2.7	Approximation du champ moyen	200
8.2.8	Exposants critiques	201
8.3	Théorie de Landau	201
8.3.1	Énergie libre de Landau	202
8.3.2	Hamiltonien effectif de Ginzburg-Landau. Approximation de Landau	202
8.3.3	Corrélation et critère de Ginzburg	205
8.4	Groupe de renormalisation	206
8.4.1	Modèle d'Ising unidimensionnel	206
8.4.2	Modèle d'Ising bidimensionnel	207

A	Dérivées et intégrales fonctionnelles	210
A.1	Dérivée fonctionnelle	210
A.1.1	Définition	210
A.1.2	Quelques exemples	211
A.1.3	Résumé	212
A.2	Intégrale fonctionnelle	213
B	Théorie des groupes	214
B.1	Groupes et exemples	214
B.1.1	Définitions	214
B.1.2	Un exemple de groupe	215
B.2	Action de groupes	215
B.3	Représentation de groupe	216
B.4	Groupes de Lie, algèbres de Lie et représentations	216
B.4.1	Groupe de Lie : définition	216
B.4.2	Algèbre et groupe de Lie	216
B.4.3	L'application exponentielle	217
B.4.4	Groupes matriciels	218
B.4.5	Action des groupes de Lie et des algèbres de Lie	219

Préface

Mon introduction n'aura pas le caractère littéraire de celle d'Alexandre, et je m'en excuse. Elle cherchera plutôt à expliquer le but de ce texte. À l'inverse d'Alexandre, je pense que les pages qui vont suivre peuvent être d'une grande utilité. Elles n'ont pas vocation à remplacer les cours que nous avons pu suivre, et encore moins à en être de fidèles transcriptions. La plupart de ce qui est écrit ici a été développé ou supposé acquis dans nos différents cours, mais a été retravaillé à l'aide de différentes références bibliographiques, qui seront données au début de chaque chapitre. Le but d'Alexandre, puis le mien lorsque je l'ai rejoint dans ce travail un peu fou, a été de créer quelque chose de consistant et de cohérent, qui puisse être lu essentiellement dans l'ordre, résumant ou développant selon les cas nos cours et rappelant des données de Licence nécessaires à la compréhension. Les parties les moins utiles à la compréhension générale et les plus éloignées des cours seront signalées tout au long du texte. N'attendez pas néanmoins de ce texte un résumé digeste, prêt à avaler, et permettant de réussir instantanément tout examen, mais plutôt une relecture des cours, plus personnelle et, nous l'espérons, parfois plus claire.

Je remercie Alexandre pour m'avoir embarqué dans son projet en cours de route, et m'avoir poussé à sortir un peu de ma fainéantise habituelle pour travailler. L'écriture d'une (petite) partie de ce que vous avez sous les yeux m'a apporté beaucoup, bien que j'aie eu à faire certaines des parties les plus rébarbatives de ce texte, pour cause de flemme d'Alexandre. Jamais en tout cas je n'aurais pensé taper des cours en L^AT_EX à 1h du matin sans obligations, et pourtant... Je remercie également énormément Benjamin, pour le temps qu'il a bien voulu passer à faire les différentes commandes de schémas que nous lui avons passées, le résultat étant bien au-dessus de mes espoirs initiaux.

Mise à jour à la deuxième édition :

Mais pourquoi est-ce que j'ai accepté de recommencer ce truc ? Ça n'en finira donc jamais ? Profitez bien de cette nouvelle mouture, avec plein de choses en plus (qui a dit que 4 pages de Physique Statistique, c'était ridicule ? 30 ça te va mieux ?) et toujours autant de santé mentale perdue de notre côté !

Chapitre 1

Relativité restreinte

Sommaire

1.1 Principe de relativité	9
1.2 Dynamique relativiste	12
1.3 Espace-temps de Minkowski et relativité restreinte	16

La physique newtonienne est basée sur deux notions. La première est celle d'espace, où l'on repère un point par trois coordonnées ; en ce sens, l'espace physique est assimilé à l'espace euclidien \mathbb{R}^3 . La deuxième est celle de temps qui, pour Newton, est une notion absolue. Dans ce cadre, une particule est décrite par trois coordonnées évoluant au cours du temps. À cette description de l'espace-temps newtonien s'ajoute un principe de symétrie, celui de relativité de Galilée, qui stipule que deux observateurs en mouvement rectiligne uniforme l'un par rapport à l'autre sont équivalents.

L'avènement de la théorie de Maxwell et la preuve de la constance de la vitesse de la lumière quelque soit l'observateur amenèrent Einstein à repenser la structure de l'espace et du temps en 1905. Einstein abandonna la notion de temps absolu et reconstruit toute la mécanique. La nouvelle mécanique, dite relativiste, est compatible avec l'électromagnétisme, et est maintenant basée sur un espace-temps à quatre dimensions particulier. Cette vision géométrique de la relativité restreinte est principalement due à Minkowski et Poincaré.

Pour ce chapitre, nous avons utilisé les références [6], [13] et [11]. Cette dernière part directement de l'approche géométrique et est utile pour approfondir la relativité restreinte. Les écrits d'Einstein sont limpides de clarté, à lire.

1.1 Principe de relativité

Nous allons d'abord dans cette section aborder la relativité restreinte de la même manière qu'Einstein. Nous donnerons les postulats de base puis en déduisons quelques conséquences physiques comme la dilatation des temps ou la contraction des longueurs. Le point de vue géométrique de Minkowski sera abordé dans la suite de manière plus précise. Ici nous allons donc à l'essentiel de la relativité restreinte et de ses conséquences les plus marquantes.

1.1.1 Énoncé du principe de relativité restreinte

Postulat de relativité : il existe des référentiels, appelés référentiels inertiels (galiléens), en mouvement rectiligne uniforme les uns par rapport aux autres, dans lesquels les phénomènes et lois physiques sont identiques.

Constance de c : la vitesse de la lumière c est une constante fondamentale indépendante de tout observateur inertiel

L'idée de la constance de la vitesse de la lumière provient en fait des équations de Maxwell qui sont intrinsèquement des équations relativistes, même si cela ne se voit pas un premier abord. Einstein prit le parti de l'électromagnétisme et non celui de Newton (qui était incompatible avec cette conséquence, même en rajoutant l'idée d'éther). La mécanique obtenue à partir de ces postulats est bien égale, dans une certaine approximation, à savoir $c \rightarrow \infty$, à la mécanique newtonienne.

La première vérification expérimentale de l'invariance de la vitesse de la lumière fût faite par Michelson et Morley en s'appuyant sur le phénomène d'interférence.

1.1.2 Temps propre et dilatation des durées

Analysons une première situation qui montre que la notion de temps absolu doit être abandonnée au profit d'une notion de temps dépendante de l'observateur.

Considérons dans un premier temps deux miroirs parallèles fixes l'un par rapport à l'autre et séparés d'une distance L . Une source lumineuse envoie de la lumière, un photon, sur l'autre miroir. Le temps τ que met ce photon pour atteindre l'autre miroir est donné par $L = c\tau$. Ce temps serait celui mesuré par un observateur inertiel au repos par rapport aux miroirs et est appelé *temps propre*.

Supposons maintenant que ces deux miroirs sont en mouvement rectiligne uniforme à la vitesse u par rapport à un observateur inertiel. D'après le postulat de la constance de la vitesse de la lumière, le temps t mis par le photon est cette fois donné par $L^2 + t^2u^2 = c^2t^2$. En remplaçant L par l'expression en fonction de τ , nous arrivons à la relation

$$t = \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} \tau = \gamma \tau \quad (1.1)$$

où γ est appelé *facteur de Lorentz*. Plusieurs remarques sont à faire. Tout d'abord, les temps mesurés ne sont pas égaux, le temps dépend de l'observateur! Nous avons perdu la notion de temps absolu de la mécanique newtonienne. La positivité du terme sous la racine impose que la vitesse d'un objet ne peut dépasser celle de la lumière et en retour, nous permet d'affirmer que $\gamma \geq 1$. Le temps t mesuré par l'observateur en mouvement est plus long que le temps τ mesuré par l'observateur au repos : c'est le phénomène de *dilatation des durées*.

1.1.3 Contraction des longueurs

On considère une règle de longueur L' au repos dans le référentiel galiléen \mathcal{R}' en mouvement rectiligne uniforme à la vitesse u par rapport au référentiel \mathcal{R} . On cherche la longueur L de cette règle dans le référentiel \mathcal{R} . La manière de mesurer des distances en relativité restreinte peut être réduite à des mesures temporelles *via* l'émission de photons. En effet, l'observateur en \mathcal{R} mesure la longueur de la règle en envoyant des rayons lumineux sur les deux extrémités, par exemple un photon sur chaque extrémité, réfléchis ensuite vers la source. Le décalage temporel entre l'arrivée des deux photons définit la longueur de la règle! La figure FIG schématise une telle expérience. L'égalité $L^2 = c^2\tau^2 - u^2\tau^2$ donne

$$L = \frac{L'}{\gamma} \quad (1.2)$$

où γ représente le facteur de Lorentz. Ce facteur étant toujours plus grand que un, la longueur de la règle en mouvement L apparaît plus courte que sa longueur au repos L' : c'est le phénomène de *contraction des longueurs*.

1.1.4 Transformation de Lorentz

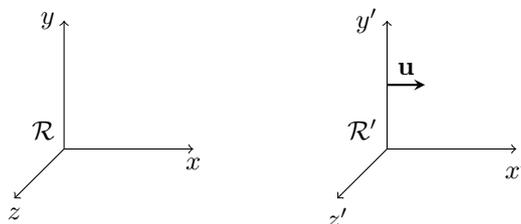


FIGURE 1.1 – La situation considérée pour les transformations de Lorentz : \mathcal{R}' est un référentiel en mouvement à la vitesse \mathbf{u} par rapport au référentiel \mathcal{R}

Soit \mathcal{R}' un référentiel en mouvement à la vitesse \mathbf{u} par rapport au référentiel \mathcal{R} , comme présenté sur la figure 1.1. Le cadre adopté ici de deux référentiels colinéaires en mouvement rectiligne uniforme l'un par rapport à l'autre est un cas particulier (on pourrait avoir une rotation par exemple) mais montre toute la particularité de la relativité restreinte.



Les relations entre les coordonnées d'un événement dans \mathcal{R}' et dans \mathcal{R} sont données par *une transformation de Lorentz*, notée $\mathcal{L}(u)$, particulière (appelée précisément transformation de Lorentz spéciale). On note $\beta = \frac{u}{c}$ et γ le facteur de Lorentz.

Les coordonnées d'un événement dans un référentiel inertiel \mathcal{R}' sont reliées à celles dans le référentiel \mathcal{R} par les relations

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \quad (1.3)$$

où β est la vitesse de \mathcal{R}' relativement à \mathcal{R} normalisée par c et γ est le facteur de Lorentz.

Nous voyons que dans ce cas, le facteur de Lorentz n'affecte que la coordonnée temporelle et celle dans la direction du mouvement. Ces deux coordonnées sont d'ailleurs dépendantes l'une de l'autre, montrant une fois de plus le caractère relatif du temps. Si on regarde la limite non relativiste $c \rightarrow \infty$, $\gamma \rightarrow 1$ et $\beta \rightarrow 0$, on retrouve le cas de la mécanique newtonienne.

Une telle transformation est une symétrie particulière de la relativité restreinte : deux observateurs reliés par une transformation de Lorentz sont équivalents et voient la même physique (les équations du mouvement sont les mêmes). L'ensemble de toutes les symétries de la relativité restreinte forment le *groupe de Lorentz*.

1.1.5 Composition des vitesses

On cherche à trouver la relation de composition des vitesses en relativité restreinte. En d'autres termes, le but est de trouver comment la vitesse d'une particule se transforme par changement de référentiel dans le cadre de la relativité restreinte. On considère une particule en mouvement quelconque à la vitesse \mathbf{v}' dans le référentiel \mathcal{R}' en mouvement rectiligne uniforme à la vitesse \mathbf{u} par rapport au référentiel \mathcal{R} . On suppose la vitesse dirigée selon l'axe (Ox)¹. En écrivant la transformation de Lorentz de manière infinitésimale

$$\begin{pmatrix} cdt \\ dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} \gamma & \gamma\beta & 0 & 0 \\ \gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} cdt' \\ dx' \\ dy' \\ dz' \end{pmatrix} \quad (1.4)$$

on peut déterminer la vitesse de la particule \mathbf{v} dans le référentiel \mathcal{R} , composante par composante. Nous avons

$$\begin{aligned} v_x &= \frac{dx}{dt} = \frac{v'_x + u}{1 + \frac{uv'_x}{c^2}} \\ v_y &= \frac{dy}{dt} = \frac{v'_y}{\gamma \left(1 + \frac{uv'_x}{c^2}\right)} \\ v_z &= \frac{dz}{dt} = \frac{v'_z}{\gamma \left(1 + \frac{uv'_x}{c^2}\right)} \end{aligned} \quad (1.5)$$

Il est donc clair que les lois de composition des vitesses galiléennes ne sont plus valables pour des vitesses relativistes. Pour de faibles vitesses, on retrouve bien sûr la loi galiléenne de sommation des vitesses. Une autre différence par rapport au cas non relativiste se voit sur les composantes de la vitesse orthogonales à la direction du mouvement des référentiels v_y et v_z . Dans le cas non relativiste, ces vitesses sont identiques dans les deux référentiels. Or ceci n'est plus vrai pour des vitesses relativistes et l'origine de cette modification provient du fait que le temps est relatif à l'observateur $dt \neq dt'$.

1.1.6 Intervalle

Comme on le verra, il est facile de voir la relativité restreinte au travers d'un cadre géométrique, celui de l'espace-temps de Minkowski. Disons seulement ici qu'un événement est caractérisé par quatre coordonnées

1. Ceci ne restreint pas la généralité du problème, puisqu'on peut toujours s'y ramener en opérant des rotations.



d'espace-temps (ct, x_1, x_2, x_3) et que les postulats précédents nous permettent de définir une quantité appelée *intervalle* qui est indépendante de tout observateur : c'est un invariant relativiste.

Dans un référentiel donné \mathcal{R} et pour deux événements caractérisés par (ct_1, x_1, y_1, z_1) et (ct_2, x_2, y_2, z_2) , on appelle *intervalle* entre ces deux événements $s_{1,2}$, la quantité définie par

$$s_{1,2}^2 = -c^2(t_1 - t_2)^2 + (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2$$

Notons la présence d'un signe moins qui, on le verra, est au cœur des phénomènes relativistes. La convention de signe choisie est celle la plus répandue pour la relativité générale.

On définit *l'intervalle* entre deux événements de manière infinitésimale, qui est un invariant relativiste, comme

$$ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 \tag{1.6}$$

Le signe de l'intervalle est lié à la causalité. Pour $s_{1,2}^2 < 0$, l'intervalle est dit de *genre temps* ; les événements peuvent être reliés par un signal lumineux. Pour $s_{1,2}^2 > 0$, l'intervalle est dit de *genre espace* ; les événements sont totalement indépendants et ne peuvent s'influencer. Enfin, pour $s_{1,2}^2 = 0$, l'intervalle est dit de *genre lumière*, et définit le *cône de lumière*. Une représentation simple est donnée sur la figure 1.3 dans le cas d'une dimension de temps et une d'espace. Enfin, on appelle *ligne d'univers* la trajectoire d'un point dans l'espace-temps.

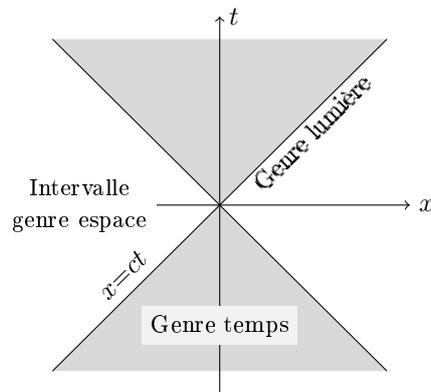


FIGURE 1.2 – Cône de lumière. On représente ici en deux dimensions le cas à quatre dimensions. L'intervalle entre un événement (ct, x) et l'origine $O(0, 0)$ est $s^2 = c^2 t^2 - x^2$. Si $s^2 < 0$, on peut relier l'origine et l'événement considéré par un signal lumineux, et l'intervalle est de genre *temps*. Pour $s^2 > 0$, on a un intervalle de genre *espace*, qui signifie que l'événement et l'origine ne peuvent avoir d'influence l'un sur l'autre, car ils ne peuvent pas être reliés par un signal. Le cas $s^2 = 0$ est le cas limite et définit le cône de lumière.

1.2 Dynamique relativiste

Jusqu'à alors nous avons introduit de façon simple les postulats et leurs conséquences immédiates. En particulier la constance de la vitesse de la lumière entre deux référentiels inertiels modifie la vision newtonienne de l'espace et du temps absolu et doivent être traitées sur le même plan. Ceci est d'autant plus clair lorsqu'on s'attarde sur les nouvelles lois de changement de référentiels inertiels.

Pour tenir compte de façon explicite de ces subtilités, il est commode d'écrire les lois de la dynamique et la cinématique d'un modèle relativiste sous une forme quadridimensionnelle. Les objets ne vivent alors plus dans un espace évoluant dans le temps mais dans un espace-temps quadridimensionnel. Ce point de vue géométrique, dans lequel toute la théorie peut être exprimé, sera développé dans la section suivante. Ici on ira directement à l'essentiel des définitions cinématiques et dynamiques en passant par une formulation variationnelle.

1.2.1 Quadri-vecteurs vitesse et impulsion

De manière quadridimensionnelle, une particule relativiste est décrite par sa position dans l'espace-temps par ses coordonnées x^μ . En choisissant un observateur inertiel \mathcal{O} , on peut écrire $x^\mu = (ct, \mathbf{r})$ où t et \mathbf{r} sont respectivement le temps et la position de l'événement relativement à l'observateur.



La vitesse est en mécanique définie comme étant la dérivée temporelle de la position. Une définition similaire de la quadrivitesse peut être donnée en faisant toutefois attention à donner une définition invariante de Lorentz. La façon la plus naturelle est de définir la quadrivitesse comme étant l'évolution de la quadriposition en fonction du temps propre de la particule. Ceci n'a de sens que pour *une particule massive*. En terme plus géométrique, on paramétrise la ligne d'univers par le temps propre. On définit donc le quadri-vecteur vitesse U^μ de la particule massive comme

$$U^\mu = \frac{dx^\mu}{d\tau} = \gamma \frac{dx^\mu}{dt} = (\gamma c, \gamma \mathbf{v}) \quad (1.7)$$

où la seconde égalité, obtenue en utilisant $t = \gamma\tau$, exprime cette vitesse relativement à l'observateur \mathcal{O} , \mathbf{v} étant le vitesse de la particule relativement à \mathcal{O} . Une propriété cruciale du quadri-vecteur vitesse est la constance de sa norme

$$U^\mu U_\mu = -\gamma^2 c^2 + \gamma^2 v^2 = -c^2 \quad (1.8)$$

Pour *une particule libre massive*, on définit la quadri-impulsion P^μ en multipliant le quadri-vecteur vitesse par la masse m de la particule

$$P^\mu = mU^\mu = (m\gamma c, m\gamma \mathbf{v}) \quad (1.9)$$

Tout comme pour la vitesse, la norme de l'impulsion est constante et vaut

$$P^\mu P_\mu = m^2 U^\mu U_\mu = -m^2 c^2 \quad (1.10)$$

Nous verrons dans la suite le lien entre l'énergie et l'impulsion au sens classique avec le quadri-vecteur impulsion.

Quadri-vecteur position

$$x^\mu = (ct, \mathbf{r})$$

Quadri-vecteur vitesse

$$U^\mu = \frac{dx^\mu}{d\tau} = (\gamma c, \gamma \mathbf{v})$$

$$U^\mu U_\mu = -c^2$$

Quadri-vecteur impulsion

$$P^\mu = mU^\mu = (m\gamma c, m\gamma \mathbf{v})$$

$$P^\mu P_\mu = -m^2 c^2$$

1.2.2 Particule libre massive

Action libre

On cherche l'action ou le lagrangien d'une particule libre relativiste, sachant qu'à la limite non-relativiste, on doit retrouver la mécanique newtonienne. Tout d'abord, les grandeurs que l'on définit doivent être des invariants de Lorentz. Le lagrangien doit être un invariant de Lorentz, tout comme l'action. Pour cela, nous définirons l'action comme l'intégrale du lagrangien sur le temps propre de la particule entre deux événements a et b de sa ligne d'univers.

$$S_{\text{Libre}} = \int_a^b \mathcal{L}(q_i, \dot{q}_i, t) dt = \int_a^b \Lambda \left(x^\mu, \frac{dx^\mu}{d\tau} \right) d\tau \quad (1.11)$$



N particules relativistes et considérer aussi des interactions. L'interaction relativiste la plus simple que l'on connaisse est l'interaction électromagnétique entre particules chargées que l'on verra dans le chapitre suivant.

Lagrangien et action d'une particule libre massive

$$\Lambda_{\text{Libre}} \left(x^\mu, \frac{dx^\mu}{d\tau} \right) = -mc^2 \quad S_{\text{Libre}} = -mc^2 \int_a^b d\tau = -mc \int_a^b ds \quad (1.18)$$

Équation du mouvement

$$\frac{dU^\mu}{d\tau} = 0 \quad (1.19)$$

1.2.3 Énergie et impulsion

En mécanique analytique, on définit l'énergie et l'impulsion d'une particule à partir des dérivées temporelle et spatiale de l'action par rapport au point d'arrivée. En relativité, on définit le quadri-vecteur impulsion à partir du quadri-gradient de l'action par rapport à l'événement final,

$$P_\mu = -\partial_\mu S = \left(-\frac{1}{c} \frac{\partial S}{\partial t}, \nabla S \right) \quad (1.20)$$

où la seconde égalité est relative à un observateur inertiel. Si l'on se réfère au calcul qui précède, la variation de l'action par rapport à l'événement final est donnée par le terme entre crochet, qui n'est dans ce cas plus nul, $P^\mu = mU^\mu$. En continuant l'analogie avec la mécanique classique, nous avons $-\frac{1}{c} \frac{\partial S}{\partial t} = \frac{E}{c}$ et $\nabla S = \mathbf{p}$, où E et \mathbf{p} sont respectivement l'énergie et l'impulsion de la particule relativement à un observateur inertiel. Nous avons donc

$$P^\mu = \left(\frac{E}{c}, \mathbf{p} \right) \quad (1.21)$$

En rapprochant cette expression avec 1.9, on exprime l'énergie et l'impulsion de la particule en fonction de la vitesse, le tout relativement à un observateur inertiel \mathcal{O}

$$E = \gamma mc^2 \quad (1.22a)$$

$$\mathbf{p} = \gamma m \mathbf{v} \quad (1.22b)$$

Finalement, on arrive à la relation la plus connue (ou presque) de la relativité restreinte en calculant la norme de l'impulsion

$$P^\mu P_\mu = -m^2 c^2 = -\frac{E^2}{c^2} + p^2 \quad i.e. \quad E^2 = m^2 c^4 + p^2 c^2 \quad (1.23)$$

Pour une particule au repos par rapport à un observateur inertiel, l'énergie a une valeur non nulle E_0 appelée *énergie de repos ou de masse* et vaut

$$E_0 = mc^2 \quad (1.24)$$

Enfin, pour des vitesses non relativiste $v/c \ll 1$, le développement limité de l'énergie totale relativiste fait apparaître, en plus du terme de repos, l'expression de l'énergie cinétique newtonienne $\frac{1}{2}mv^2$.



L'énergie et l'impulsion d'une particule forment un quadri-vecteur, le quadri-vecteur impulsion P^μ , relié au quadri-vecteur vitesse par

$$P^\mu = \left(\frac{E}{c}, \mathbf{p} \right) = mU^\mu \quad (1.25)$$

D'où les expressions fondamentales de l'énergie et de l'impulsion d'une particule, qui dépendent du facteur de Lorentz γ

$$E = \gamma mc^2 \quad (1.26a)$$

$$\mathbf{p} = \gamma m \mathbf{v} \quad (1.26b)$$

Enfin, ces deux quantités ne sont pas indépendantes l'une de l'autre et sont reliées par

$$E^2 = m^2 c^4 + p^2 c^2 \quad (1.27)$$

1.2.4 Lois de conservation

Utiliser la quadri-impulsion est un atout majeur pour extraire la physique d'un système. Cela provient du fait qu'elle vérifie une loi de conservation que l'on peut voir comme une conséquence du théorème de Noether pour des systèmes invariants par translation spatio-temporelle. On peut aussi considérer cette loi comme un postulat, que l'on doit compléter par le postulat de la conservation du moment cinétique que l'on ne détaillera pas ici.

Considérons un système isolé et une surface de l'espace-temps S . L'impulsion totale contenue dans ce volume est la somme de toutes les impulsions qui traversent cette région (avec la bon signe suivant l'orientation de la région que l'on choisit). Alors cette impulsion totale est nulle.

L'utilisation de cette loi permet de comprendre beaucoup de phénomènes tels que l'effet Doppler, les collisions élastiques et inélastiques (diffusion Compton par exemple)...

1.3 Espace-temps de Minkowski et relativité restreinte

Dans cette section, nous allons tenter de donner un avant-goût d'une formulation complètement géométrique de la relativité restreinte à partir de la notion d'espace-temps. Tout ce que nous avons rappelé plus haut peut être défini de façon géométrique, c'est-à-dire de façon indépendante de tout observateur. L'intérêt principal est de faire ressortir simplement toutes les nouveautés conceptuelles introduites par Einstein et est au cœur de la théorie relativiste de la gravitation, la Relativité Générale.

1.3.1 L'espace-temps

On développe ici la notion d'espace-temps et on donne sa définition formelle. Assez mathématique dans sa formulation, il n'est pas nécessaire d'en comprendre tous les tenants et les aboutissants mais il est nécessaire un jour de les assimiler (surtout si on veut aborder proprement la Relativité Générale).

Cadre newtonien

Dans le cadre de la mécanique newtonienne, les notions d'espace et de temps sont absolues. Les quantités cinématiques telles que la position, la vitesse ou l'accélération sont définies elles aussi de manière absolues : un événement est repéré par des coordonnées d'espace, qui sont des fonctions du temps absolu. Tout comme en relativité, nous pouvons introduire une notion d'espace-temps \mathcal{E} dans le cadre newtonien de la façon suivante

$$\mathcal{E} = \bigcup_{t \in \mathbb{R}} \Sigma_t \quad (1.28)$$

où quelque soit t , Σ_t est un espace euclidien de dimension trois. L'espace-temps newtonien est un feuilletage de ces espaces paramétrés par le temps absolu t . Bien que cela ne soit pas très utile, cela nous permet de voir comment se traduisent, d'un point de vue géométrique, les idées newtoniennes et de les comparer à celles de la



relativité restreinte, toujours du point de vue géométrique. En fait, l'espace-temps relativiste oublie le temps absolu (on n'a pas de structure en feuilletage) ainsi que la métrique euclidienne.

Cadre relativiste

L'approche géométrique de la relativité restreinte, développée par Minkowski, se focalise sur la notion d'espace-temps. Dans le cadre de la relativité restreinte, l'espace-temps \mathcal{E} est un espace affine de dimension quatre muni d'un (pseudo) produit scalaire lorentzien, aussi appelé *métrique lorentzienne* et noté \mathbf{g} . En clair, $\mathcal{E} \simeq (\mathbb{R}^4, \mathbf{g})$. Une métrique lorentzienne \mathbf{g} est une forme bilinéaire symétrique non dégénérée de signature $(-, +, +, +)$ sur l'espace vectoriel E associé à \mathcal{E} . La signature de la métrique est une propriété indépendante de la base de l'espace (théorème d'inertie de Sylvester). Concrètement, cela signifie qu'il existe une base de E telle que $\mathbf{g}(x, y) = -x_0y_0 + x_1y_1 + x_2y_2 + x_3y_3$. La seule différence avec un produit scalaire au sens classique, qualifié dans notre contexte d'euclidien, provient de la signature. Dans le cas euclidien, elle serait composée uniquement de signes positifs. En faisant attention au signe moins, la définition d'orthogonalité est la même que dans le cas traditionnel. Dans une représentation matricielle, une telle métrique se réduit dans une base orthonormée à la *matrice de Minkowski* $\eta_{\mu\nu}$,

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1.29)$$

L'introduction de cette structure est la principale différence avec le cadre newtonien qui, lui, repose sur des notions euclidiennes. Elle permet entre autres de classer les vecteurs \mathbf{x} de l'espace suivant le signe de $\mathbf{g}(\mathbf{x}, \mathbf{x})$, autrement appelé *intervalle*, qui n'a plus de raison d'être toujours positif à cause de la signature! On retrouve la classification faite plus haut : s'il est négatif, on a un vecteur de *genre temps*, s'il est nul de *genre lumière* et s'il est positif de *genre espace*.

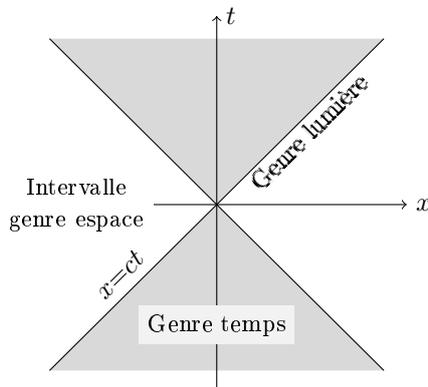


FIGURE 1.3 – Cône de lumière. On représente ici en deux dimensions le cas à quatre dimensions. L'intervalle entre un événement (ct, x) et l'origine $O(0, 0)$ est $s^2 = c^2t^2 - x^2$. Si $s^2 < 0$, on peut relier l'origine et l'événement considéré par un signal lumineux, et l'intervalle est de genre *temps*. Pour $s^2 > 0$, on a un intervalle de genre *espace*, qui signifie que l'événement et l'origine ne peuvent avoir d'influence l'un sur l'autre, car il ne peuvent pas être reliés par un signal. Le cas $s^2 = 0$ est le cas limite et définit le cône de lumière.

Cela étant, nous pouvons à partir d'un point de l'espace-temps, autrement appelé *événement*, déterminer le cône de lumière issu de ce point, distinguant le genre des vecteurs autour de cet événement. Puisque aucune notion de temps absolu n'est présente, il ne nous est pas possible de définir intrinsèquement une notion de passé et de futur, en clair une flèche du temps. Cependant, il est possible de choisir de manière purement arbitraire une telle direction. Le cône de lumière étant composé de deux nappes l'une sera appelée nappe du futur \mathcal{I}^+ et l'autre nappe du passé, déterminant ainsi une flèche du temps.

Pour être complet, il reste à définir une notion d'orientation sur l'espace-temps (en fait sur E). Dans le cadre euclidien, l'orientation d'une base orthonormée est donnée par le signe du déterminant (forme trilinéaire antisymétrique) de cette base : elle est dite directe si le déterminant vaut $+1$ et indirecte sinon. Dans le cadre plus général de la relativité restreinte, une orientation de l'espace nous est fournie lorsqu'on se donne une forme



multilinéaire antisymétrique (quadri-linéaire dans notre cas). En fait, cela revient³ à choisir *le tenseur de Levi-Civita* ϵ , complètement déterminé par $\epsilon_{0123} = 1$. Une base quelconque sera dite directe si ϵ évalué sur cette base est positif et indirecte sinon.

L'espace-temps de Minkowski est donc la donnée d'un espace affine \mathcal{E} de dimension quatre muni d'une métrique \mathbf{g} de signature $(-, +, +, +)$, d'une flèche du temps \mathcal{I}^+ , et d'une orientation ϵ

$$(\mathcal{E}, \mathbf{g}, \mathcal{I}^+, \epsilon) \tag{1.30}$$

C'est cette structure qui est à la base de la relativité restreinte, bien qu'il reste encore à voir comment on fait de la physique avec. Notons bien que l'espace affine ainsi que la métrique sont des notions absolues au même titre que l'espace et le temps dans le cadre newtonien! Par contre, dans le cadre de la relativité générale, l'espace \mathcal{E} (qui n'a plus la structure d'un espace affine mais celle plus générale d'espace courbe, et même plus exactement d'une variété différentielle) et la métrique ne sont pas déterminés *a priori* et doivent être déterminés grâce aux équations d'Einstein.

Dualité et notation quadridimensionnelle

Considérons un élément x de notre espace \mathcal{E} . On note $(e_\mu)_{\mu \in I}$ une base de notre espace. On appelle *coordonnées contravariantes* les coordonnées de x dans cette base

$$x = \sum_{\mu} x^\mu e_\mu = x^\mu e_\mu \tag{1.31}$$

avec la convention d'Einstein de somme sur les indices répétés. Il est possible de représenter l'événement x d'une autre manière, *via* la métrique } définie sur E : c'est la notion de dualité. L'espace dual E^* d'un espace vectoriel E est l'ensemble des formes linéaires sur E (des applications linéaires de E dans \mathbb{R}). Le fait que nous soyons en dimension finie permet d'établir un isomorphisme entre ces deux espaces. La métrique \mathbf{g} permet de définir une dualité particulière appelée *dualité métrique* par

$$\begin{matrix} E & \longrightarrow & E^* \\ x & \longmapsto & \omega : E \longrightarrow \mathbb{R} \\ & & y \longmapsto \mathbf{g}(x, y) \end{matrix} \tag{1.32}$$

On appelle *coordonnées covariantes*, notée x_μ , les coordonnées de la forme linéaire associée à un vecteur x par dualité métrique, c'est-à-dire

$$x_\mu = \mathbf{g}(x, e_\mu) = \underbrace{\mathbf{g}(e_\nu, e_\mu)}_{g_{\nu\mu}} x^\nu \tag{1.33}$$

La figure 1.4 représente ces différentes façons de voir l'événement x , dans le cas plus simple d'un espace de dimension deux avec une métrique euclidienne. Nous voyons qu'à l'aide du tenseur métrique, il est possible de faire le lien entre les coordonnées contravariantes et covariantes. De manière informelle, il permet de faire monter ou descendre les indices. Plusieurs écritures sont alors possibles, par exemple pour le produit scalaire

$$\mathbf{g}(x, y) = g_{\mu\nu} x^\mu y^\nu = \omega(y) = x_\mu y^\mu = x^\mu y_\mu \tag{1.34}$$

1.3.2 Ligne d'univers, temps propre and co.

Ligne d'univers

Une particule de masse non nulle m est décrite par une courbe \mathcal{L} régulière (C^2 par morceaux) de \mathcal{E} de genre temps appelée *ligne d'univers*.

3. L'espace des formes quadri-linéaires antisymétrique est de dimension un. Ceci est un cas particulier valable quelque soit la dimension n de l'espace : l'espace des formes n -linéaires antisymétrique est de dimension un.

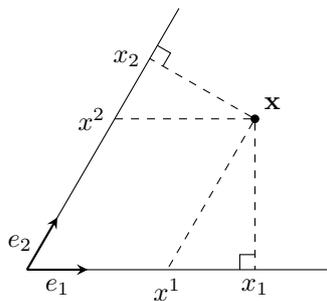


FIGURE 1.4 – Coordonnées contravariantes et covariantes en géométrie euclidienne. On a $\mathbf{x} = x^1\mathbf{e}_1 + x^2\mathbf{e}_2$, tandis que $x_1 = \mathbf{x} \cdot \mathbf{e}_1$ et $x_2 = \mathbf{x} \cdot \mathbf{e}_2$, où \cdot est le produit scalaire associé à l’espace, ici le produit scalaire classique de \mathbb{R}^2 : $\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}||\mathbf{y}| \cos(\mathbf{x}, \mathbf{y})$.

Le fait que la ligne d’univers soit de genre temps signifie que tous les vecteurs tangents $d\mathbf{x}$ à cette courbe sont de genre temps, soit $\mathbf{g}(d\mathbf{x}, d\mathbf{x}) < 0$. La ligne d’univers a un caractère absolu, elle ne dépend pas de l’observateur. Une telle courbe peut être paramétrée de manière purement mathématique par différents paramètres λ sans aucune signification physique. Une courbe est mathématiquement l’image d’une application ϕ , appelée paramétrage, définie comme

$$\phi : \begin{array}{l} \mathbb{R} \longrightarrow \mathcal{E} \\ \lambda \longmapsto \phi(\lambda) \end{array} \tag{1.35}$$

Une particule de masse nulle est décrite par une droite \mathcal{E} ayant pour vecteur directeur un vecteur de genre lumière. C’est une droite du cône de lumière.

En particulier, la ligne d’univers d’un photon rentre dans cette catégorie (d’où les noms donnés). On verra que ce postulat est la traduction géométrique du postulat d’Einstein concernant l’indépendance de la vitesse de la lumière vis-à-vis des observateurs inertiels.

Temps propre

La métrique de Minkowski ne permet pas de définir une norme au sens mathématique du terme puisqu’il existe des vecteurs de "norme" nulle différents du vecteur nul, autrement appelés vecteurs de genre lumière. En revanche, la métrique induit une sorte de mesure des longueurs sur l’ensemble des vecteurs de genre temps (elle n’induit toujours pas une norme puisque l’inégalité triangulaire n’est pas vérifiée) puisqu’elle ne s’annule alors que pour le vecteur nul. On définit alors le temps propre associé à la ligne d’univers \mathcal{L} comme la mesure de cette longueur. Il sert comme paramétrage physique de la ligne d’univers.

Le *temps propre* $d\tau$ en un point d’une ligne d’univers \mathcal{L} est défini à partir de la longueur du vecteur tangent $d\mathbf{x}$ induite par la métrique \mathbf{g} ,

$$\begin{aligned} cd\tau &= \sqrt{\mathbf{g}(d\mathbf{x}, d\mathbf{x})} && \text{Orientation vers le futur} \\ cd\tau &= -\sqrt{\mathbf{g}(d\mathbf{x}, d\mathbf{x})} && \text{Orientation vers le passé} \end{aligned} \tag{1.36}$$

La constante c intervient comme un facteur de conversion permettant au temps propre de bien avoir la dimension d’un temps dans nos unités usuelles et vaut par convention $c = 2.99792458 \cdot 10^8 \text{ m s}^{-1}$. On verra que c’est en fait la vitesse associée à une particule de masse nulle, comme le photon. Nous voyons que le temps propre est relié à l’intervalle $ds^2 = \mathbf{g}(d\mathbf{x}, d\mathbf{x})$. En termes de composantes dans une base donnée,

$$cd\tau = \sqrt{-ds^2} = \sqrt{-g_{\mu\nu}x^\mu x^\nu} \tag{1.37}$$

On définit une *horloge idéale* comme une suite d’événements subdivisant la ligne d’univers en intervalles de temps propre réguliers.

Quadri-vitesse et quadri-accélération

Soit une particule massive de ligne d'univers \mathcal{L} paramétrée par son temps propre τ . La *quadrivitesse* \mathbf{u} de cette particule est définie par

$$\mathbf{u} = \frac{1}{c} \frac{d\mathbf{x}}{d\tau} \quad (1.38)$$

La quadrivitesse est en fait l'unique champ de vecteurs tangents unitaires à \mathcal{L} ; nous pouvons montrer que $\mathbf{u} \cdot \mathbf{u} = -1$. La quadrivitesse est une propriété intrinsèque à la ligne d'univers. Elle n'est pas définie relativement à un observateur particulier et est en ce sens une notion absolue et par là non mesurable⁴! Puisque nous ne parlons que d'une particule massive, la quadrivitesse en un point de la ligne d'univers est un vecteur de genre temps, par définition de la ligne d'univers.

Soit une particule massive de ligne d'univers \mathcal{L} paramétrée par son temps propre τ . La *quadriaccélération* \mathbf{a} de cette particule est définie par

$$\mathbf{a} = \frac{1}{c} \frac{d\mathbf{u}}{d\tau} \quad (1.39)$$

A priori, ce champ de vecteurs n'est pas normalisé. Par contre, en dérivant la relation de normalisation de la quadrivitesse, on trouve que \mathbf{u} et \mathbf{a} sont orthogonaux, $\mathbf{u} \cdot \mathbf{a} = 0$. En d'autres termes, la quadriaccélération est un champ de vecteurs de genre espace. En effet, grâce à l'orthogonalité, nous pouvons choisir \mathbf{u} et \mathbf{a} comme deux des vecteurs d'une base orthonormée de l'espace-temps. Or, comme \mathbf{u} est un vecteur de genre temps et que la signature de la métrique est indépendante de la base, nous avons nécessairement $\mathbf{a} \cdot \mathbf{a} > 0$, d'où le genre espace de la quadriaccélération.

1.3.3 Observateurs

Simultanéité

Avant de pouvoir définir la notion d'observateur en relativité restreinte, il est nécessaire de définir la notion de simultanéité. Cette notion, intuitive et absolue dans le cadre newtonien, devient une notion relative du fait de l'absence d'un temps absolu. Soit une particule massive de ligne d'univers \mathcal{L} paramétrée par son temps propre t . À l'instant t_1 , la particule émet un signal lumineux. On suppose que ce signal est réfléchi à l'événement M et que la particule réceptionne le signal à l'instant t_2 .

Le critère d'Einstein-Poincaré énonce que l'événement $O(t)$ de \mathcal{L} est dit *simultané* à M si

$$t = \frac{t_1 + t_2}{2}$$

Ceci définit pour chaque événement $O(t)$ de la ligne d'univers une hypersurface de simultanéité $\mathcal{S}_{\mathcal{L}}(O(t))$: c'est l'ensemble de tous les événements simultanés à $O(t)$.

Si on raisonne de manière locale en assimilant la ligne d'univers à une droite (on néglige sa courbure, c'est-à-dire physiquement l'accélération), l'hypersurface de simultanéité appelée dans ce cas *espace local de repos* et noté $\mathcal{E}_{\mathcal{L}}(O(t))$ correspond à l'ensemble de événement M tels que $\mathbf{O}(t)\mathbf{M} \perp \mathbf{u}(t)$

$$\mathcal{E}_{\mathcal{L}}(O(t)) = \{M \in \mathcal{E}, \mathbf{O}(t)\mathbf{M} \perp \mathbf{u}(t)\} \quad (1.40)$$

Géométriquement, c'est l'espace orthogonal pour la métrique \mathbf{g} à la ligne d'univers \mathcal{L} au point $O(t)$. On peut aussi le voir comme l'espace tangent au point $O(t)$ à l'hypersurface de simultanéité $\mathcal{S}_{\mathcal{L}}(O(t))$.

Avec cette définition de la simultanéité, il est possible de définir une mesure des distances à partir de celle du temps. La distance $\|\mathbf{O}(t)\mathbf{M}\|$ est mesuré de la façon suivante : un signal lumineux est émis en $O(t_1)$ réfléchi en M et reçu en $O(t_2)$. La distance est alors simplement obtenue en faisant le produit de la vitesse de la lumière c multipliée par le temps de l'aller $(t_2 - t_1)/2$. À partir du critère de simultanéité, on obtient la formule

$$\|\mathbf{O}(t)\mathbf{M}\| = c\sqrt{(t - t_1)(t_2 - t)} \quad (1.41)$$

⁴. Le fait de choisir une normalisation en $1/c$, outre le fait de rendre les vecteurs unitaires, permet d'adimensionner cette "vitesse" et de se rappeler que c'est une notion absolue et non mesurable.



Qu'est ce qu'un observateur ?

On appelle *référentiel local* en $O(t)$ de la ligne d'univers \mathcal{L} , où t est par exemple le temps propre, une base orthonormée directe $(\mathbf{e}_1(t), \mathbf{e}_2(t), \mathbf{e}_3(t), \mathbf{e}_4(t))$ de l'espace temps \mathcal{C} avec la contrainte $\mathbf{e}_1(t) = \mathbf{u}(t)$.

Le premier vecteur de la base, le seul de genre temps, est toujours choisi comme étant la quadrivitesse. Les trois autres, tous de genre espace, forment une base de l'espace local de repos au point considéré. Du point de vue mathématique, on supposera toujours que le champ de vecteurs ainsi défini est de classe C^1 . Pour rappel, nous avons donc

$$\forall t, \mathbf{g}(\mathbf{e}_\alpha, \mathbf{e}_\beta) = \eta_{\alpha\beta} \quad \text{et} \quad \epsilon(\mathbf{e}_1(t), \mathbf{e}_2(t), \mathbf{e}_3(t), \mathbf{e}_4(t)) = 1 \quad (1.42)$$

La différence avec un système de coordonnées mathématique est que chaque événement est repéré par le biais d'opérations physiques, comme le temps propre ou l'émission de signaux lumineux. Les coordonnées attribuées sont donc purement relatives, dépendantes du système physique. Du coup, nous pouvons maintenant définir ce qu'est un observateur.

Un *observateur* est la donnée d'une ligne d'univers \mathcal{L} ainsi que d'un référentiel local en chaque point.

1.3.4 Cinématique

Dorénavant, toutes les définitions de base ont été données pour reconstruire la cinématique et la dynamique de la relativité restreinte.⁵ On ne va pas tout reconstruire mais juste donner le point de vue, la méthode pour faire cela. Par exemple, comment définit-on le facteur de Lorentz ?

On considère donc une particule \mathcal{P} ayant une certaine ligne d'univers $\mathcal{L}_{\mathcal{P}}$, de temps propre t' et un observateur \mathcal{O} de ligne d'univers $\mathcal{L}_{\mathcal{O}}$ de temps propre t . On considère alors deux événements $O(t)$ et $O(t+dt)$ pour l'observateur. A ces deux événements sont associés deux événements simultanés pour la particule $P(t)$ et $P(t+dt)$. Pour la particule, ces deux événements sont séparés par un temps propre dt' qui, contrairement au cas newtonien, n'est *a priori* pas égal au temps propre dt écoulé pour l'observateur. On définit alors le facteur de Lorentz γ de \mathcal{P} par rapport à \mathcal{O} comme

$$dt = \gamma dt' \quad (1.43)$$

Avec cette définition, on peut exprimer le facteur de Lorentz de différentes manières en fonction des données à disposition, par exemple retrouver la formuler habituelle $1/\sqrt{1-v^2/c^2}$.

Comment définit-on la vitesse de la particule relativement à l'observateur ? Simplement, on considère un référentiel local associé à l'observateur, un événement $O(t)$ et son événement simultané $P(t)$ et on définit la vitesse de la particule relativement à l'observateur à l'instant t comme la dérivée des coordonnées spatiales dans le référentiel local par rapport au temps t . Ceci permet de retrouver par exemple la décomposition de la quadri-vitesse de la particule relativement à l'observateur $(\gamma c, \gamma \mathbf{v})$. On pourra faire de même avec l'accélération.

1.3.5 Pour aller plus loin

On vient juste ici d'aborder l'approche géométrique de la relativité restreinte. Outre de finir de définir complètement les notions géométriquement comme exposé ci-dessus, beaucoup de phénomènes propres à la relativité restreinte peuvent être abordés : l'analyse du groupe de Lorentz, la quadri-accélération et sa décomposition relativement à un observateur, les collisions, les observateurs non-inertiels (accélérés ou en rotations), la précession de Thomas... Bref, encore plein de choses que l'on peut aborder principalement dans [11].

⁵. Il faut se rappeler que hormis la définition même de l'espace-temps (et les propriétés qui en découlent) toutes les autres notions comme les lignes d'univers, les vitesses, etc, ne sont intéressantes que pour une théorie classique mais pas pour une théorie quantique. Construire une théorie quantique et relativiste c'est construire une théorie quantique dans l'espace-temps de Minkowski et compatible avec sa structure.



Chapitre 2

Électrodynamique

Sommaire

2.1	Les équations de Maxwell et les changements de référentiel	22
2.2	Solutions des équations de Maxwell, Potentiels retardés	29
2.3	Rayonnement électromagnétique	31
2.4	Formulation relativiste de l'électrodynamique	33
2.5	Formulation lagrangienne de l'électrodynamique	38

Maintenant que nous avons rappelé le formalisme de la relativité restreinte, nous allons nous intéresser à sa principale application : l'étude de l'électromagnétisme, et en particulier la reformulation relativiste des champs. Tout d'abord, nous ferons un rapide retour sur les lois de l'électromagnétisme classique, et nous montrerons que les transformations de Galilée ne peuvent pas être les bonnes équations de changement de référentiels au vu de ces lois. Dès lors, nous suivrons l'idée de formuler les lois de l'électrodynamique de façon relativiste, puis sous une formulation lagrangienne, afin de dériver l'électromagnétisme à partir d'un principe variationnel. Enfin, nous étudierons différentes applications de cette théorie.

Pour ce chapitre, nous avons essentiellement utilisé [13] et [11] qui aborde l'électromagnétisme de façon géométrique dans la continuité de la relativité restreinte. Histoire d'embrouiller les choses, on utilisera non pas la métrique de Minkowski du chapitre précédent mais celle avec les signes opposés $\eta_{\mu\nu} = (1, -1, -1, -1)$:

2.1 Les équations de Maxwell et les changements de référentiel

2.1.1 Les équations de Maxwell

Densité de charge et densité de courant

On considère un ensemble de charges $\{q_a\}$ situées à l'instant t aux positions $\{\mathbf{r}_a(t)\}$, et possédant les vitesses $\{\mathbf{v}_a(t)\}$. On définit alors le champ scalaire *densité de charge* $\rho(\mathbf{r}, t)$, et le champ vectoriel *densité de courant* $\mathbf{j}(\mathbf{r}, t)$ par :

$$\rho(\mathbf{r}, t) = \sum_a q_a \delta[\mathbf{r} - \mathbf{r}_a(t)] \tag{2.1a}$$

$$\mathbf{j}(\mathbf{r}, t) = \sum_a q_a \mathbf{v}_a(t) \delta[\mathbf{r} - \mathbf{r}_a(t)] \tag{2.1b}$$

Équations de Maxwell-Lorentz

A l'aide de ces distributions de charge et de courant, on postule les équations de Maxwell qui forment un ensemble d'équations différentielles. Elles s'écrivent de la manière suivante

$$\begin{cases} \nabla \cdot \mathbf{E}(\mathbf{r}, t) = \frac{\rho(\mathbf{r}, t)}{\epsilon_0} & \text{Équation de Maxwell-Gauss} \\ \nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0 & \text{Absence de monopole magnétique} \\ \nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} & \text{Équation de Maxwell-Faraday} \\ \nabla \times \mathbf{B}(\mathbf{r}, t) = \mu_0 \mathbf{j}(\mathbf{r}, t) + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} & \text{Équation de Maxwell-Ampère} \end{cases} \tag{2.2}$$



Deux d'entre elles relient champs et sources, tandis que les deux autres donnent des relations entre les champs. Il nous manque toujours un moyen de déterminer comment ces champs affectent le mouvement de particules chargées. On postule alors l'équation de Lorentz, régissant le mouvement des charges sous l'action du couple (\mathbf{E}, \mathbf{B})

$$m_a \frac{d}{dt} \left(\frac{\mathbf{v}_a}{\sqrt{1 - \mathbf{v}_a^2/c^2}} \right) = q_a (\mathbf{E}(\mathbf{r}_a, t) + \mathbf{v}_a \times \mathbf{B}(\mathbf{r}_a, t)) \quad (2.3)$$

Ces équations sont normalement données avec un ensemble de conditions initiales et de conditions aux limites sur les champs comme sur les charges pour pouvoir résoudre totalement un problème d'électromagnétisme.

Formulation intégrale de ces équations

Les équations de Maxwell, formulées précédemment sous une forme différentielle, peuvent aussi être écrites sous une forme intégrale grâce à deux théorèmes d'analyse vectorielle que l'on rappelle ci-dessous.

Théorème de Green-Ostrogradski Pour tout champ vectoriel "régulier" \mathbf{A} défini dans un volume V délimité par la surface fermée orientée vers l'extérieur S , on a

$$\int_V \nabla \cdot \mathbf{A} d^3\mathbf{r} = \oint_S \mathbf{A} \cdot \mathbf{nd}S \quad (2.4)$$

Théorème de Stokes

Si S est une surface ouverte arbitraire et \mathcal{C} le contour sur lequel s'appuie cette surface, orienté de façon directe relativement à l'orientation \mathbf{n} de la surface S , alors, pour tout champ vectoriel "régulier" \mathbf{A} , on a :

$$\oint_{\mathcal{C}} \mathbf{A} \cdot d\mathbf{l} = \int_S (\nabla \times \mathbf{A}) \cdot \mathbf{nd}S \quad (2.5)$$

Si on applique les deux théorèmes précédents aux équations de Maxwell, on obtient quatre nouvelles équations :

$$\left\{ \begin{array}{ll} \oint_S \mathbf{E} \cdot \mathbf{nd}S = \frac{Q_{\text{tot}}^V}{\epsilon_0} & \text{Applications en électrostatique} \\ \oint_S \mathbf{B} \cdot \mathbf{nd}S = 0 & \text{Conservation du flux de } \mathbf{B} \text{ le long d'un tube de courant} \\ \oint_{\mathcal{C}} \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \int_S \mathbf{B} \cdot \mathbf{nd}S & \text{Application : loi de l'induction} \\ \oint_{\mathcal{C}} \mathbf{B} \cdot d\mathbf{l} = \mu_0 \int_S (\mathbf{j} + \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}) \cdot \mathbf{nd}S & \text{Applications en magnétostatique} \end{array} \right. \quad (2.6)$$

Une remarque importante à faire concerne l'équivalence des deux formulations. Cette équivalence n'est vraie que dans le cadre de la théorie des distributions. Si l'on se restreint à un cadre moins général (les fonctions par exemple) alors la formulation intégrale contient plus d'information que la formulation différentielle. Par exemple, il peut arriver que pour résoudre un problème *via* la forme différentielle nous soyons obligés de faire des hypothèses sur les champs qui ne sont pas nécessaires en utilisant la formulation intégrale.

2.1.2 Équations de conservation

En appliquant des formules d'analyse vectorielle aux équations de Maxwell, on peut trouver plusieurs équations de conservation. Ces équations seront faciles à vérifier plus loin à l'aide de la formulation lagrangienne de l'électrodynamique, puisqu'elles correspondent à des invariances du Lagrangien sous certaines transformations.

Conservation de la charge

On prend la divergence de l'équation de Maxwell-Ampère, en se souvenant que $\forall \mathbf{A}(\mathbf{r}, t), \nabla \cdot (\nabla \times \mathbf{A}) = 0$:

$$\nabla \cdot (\nabla \times \mathbf{B}) = 0 = \mu_0 (\nabla \cdot \mathbf{j} + \epsilon_0 \partial_t (\nabla \cdot \mathbf{E})) \quad (2.7)$$

et puisque $\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$, on trouve finalement :

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0 \quad (2.8)$$



Cette relation correspond à la conservation de la charge électrique. En effet, sa forme intégrale (obtenue à l'aide du théorème de Stokes) est :

$$\underbrace{\frac{d}{dt} \int_V \rho(\mathbf{r}, t) d^3\mathbf{r}}_{\text{variation de la charge totale contenue dans } V} + \underbrace{\oint_S \mathbf{j}(\mathbf{r}, t) \cdot \mathbf{n} dS}_{\text{flux de charges à travers la surface frontière de } V} = 0 \quad (2.9)$$

Conservation de l'énergie

A partir de l'équation de Maxwell-Ampère, on peut écrire

$$\mathbf{j} \cdot \mathbf{E} = \mathbf{E} \cdot (\nabla \times \mathbf{B}) - \frac{1}{2} \epsilon_0 \frac{\partial \mathbf{E}^2}{\partial t} \quad (2.10)$$

Ce qui peut être réécrit, à l'aide de la formule $\nabla \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot (\nabla \times \mathbf{a}) - \mathbf{a} \cdot (\nabla \times \mathbf{b})$ sous la forme :

$$\mathbf{j} \cdot \mathbf{E} = \epsilon_0 c^2 \nabla \cdot (\mathbf{B} \times \mathbf{E}) - \frac{1}{2} \epsilon_0 \left(\frac{\partial \mathbf{E}^2}{\partial t} + c^2 \frac{\partial \mathbf{B}^2}{\partial t} \right) \quad (2.11)$$

On pose alors

$$u(\mathbf{r}, t) = \frac{1}{2} \epsilon_0 (\mathbf{E}^2 + c^2 \mathbf{B}^2) = \frac{\epsilon_0 \mathbf{E}^2}{2} + \frac{\mathbf{B}^2}{2\mu_0} \quad (2.12a)$$

$$\mathbf{\Pi}(\mathbf{r}, t) = \epsilon_0 c^2 (\mathbf{E} \times \mathbf{B}) = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0} \quad (2.12b)$$

u est la *densité locale d'énergie électromagnétique*, tandis que $\mathbf{\Pi}$ est appelé *vecteur de Poynting* et représente le flux de cette énergie par unité de surface. Ces quantités permettent de réécrire l'équation sous la forme

$$\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{\Pi} + \mathbf{j} \cdot \mathbf{E} = 0 \quad (2.13)$$

ou encore, sous forme intégrale :

$$\underbrace{\frac{d}{dt} \int_V u(\mathbf{r}, t) d^3\mathbf{r}}_{\text{variation de l'énergie totale contenue dans } V} + \underbrace{\oint_S \mathbf{\Pi}(\mathbf{r}, t) \cdot \mathbf{n} dS}_{\text{flux de charges à travers la surface frontière de } V} = \underbrace{- \int_V \mathbf{j} \cdot \mathbf{E} d^3\mathbf{r}}_{\text{puissance EM fournie aux charges par le champ électrique}} \quad (2.14)$$

Conservation de l'impulsion

Si l'on part de l'équation de Lorentz, on peut chercher à dériver une formule identique pour l'impulsion. En appelant $\mathbf{P}_{\text{part}}^V$ l'impulsion totale des particules contenues dans le volume V , on a :

$$\frac{d}{dt} \mathbf{P}_{\text{part}}^V = \sum_{a \in V} q_a (\mathbf{E}(\mathbf{r}_a(t)) + \mathbf{v}_a \mathbf{B}(\mathbf{r}_a(t))) = \int_V (\rho \mathbf{E} + \mathbf{j} \times \mathbf{B}) d^3\mathbf{r} \quad (2.15)$$

On remplace ensuite ρ et \mathbf{j} dans cette équation à l'aide des équations de Maxwell avec sources. Après quelques développements, on arrive à :

$$\rho \mathbf{E} + \mathbf{j} \times \mathbf{B} = \epsilon_0 [\mathbf{E} \cdot (\nabla \cdot \mathbf{E}) + c^2 \mathbf{B} \cdot (\nabla \cdot \mathbf{B}) - \mathbf{E} \times (\nabla \times \mathbf{E}) - c^2 \mathbf{B} \times (\nabla \times \mathbf{B})] - \epsilon_0 \frac{\partial}{\partial t} (\mathbf{E} \times \mathbf{B}) \quad (2.16)$$

Le terme $\nabla \cdot \mathbf{B}$, nul, est rajouté artificiellement pour permettre la suite du calcul. En effet, on peut montrer que

$$[\mathbf{E} \cdot (\nabla \cdot \mathbf{E}) - \mathbf{E} \times (\nabla \times \mathbf{E})]_i = \frac{\partial}{\partial x_j} \left(E_i E_j - \frac{1}{2} (\mathbf{E} \cdot \mathbf{E}) \delta_{ij} \right) \quad (2.17)$$



Jauge de Lorentz

Cette jauge consiste à imposer

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2} \frac{\partial \phi}{\partial t} = 0 \quad (2.26)$$

Avec cette condition, les équations sur les potentiels deviennent

$$\Delta \phi - \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} = -\frac{\rho}{\epsilon_0} \quad (2.27a)$$

$$\Delta \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{j} \quad (2.27b)$$

Ces équations sont découplées, et peuvent être facilement résolues en utilisant la méthode des fonctions de Green. De plus, cette jauge est *covariante* : si elle est satisfaite dans un référentiel \mathcal{R} , alors elle est également satisfaite dans tout référentiel \mathcal{R}' en translation rectiligne uniforme par rapport à \mathcal{R} . Elle est donc adaptée à une formulation relativiste de l'électromagnétisme.



2.1.4 Résumé

Formulation classique de l'électromagnétisme

On définit les densités de charge et de courant par :

$$\rho(\mathbf{r}, t) = \sum_a q_a \delta[\mathbf{r} - \mathbf{r}_a(t)]$$

$$\mathbf{j}(\mathbf{r}, t) = \sum_a q_a \mathbf{v}_a(t) \delta[\mathbf{r} - \mathbf{r}_a(t)]$$

Les équations de Maxwell sont alors :

$$\left\{ \begin{array}{ll} \nabla \cdot \mathbf{E}(\mathbf{r}, t) = \frac{\rho(\mathbf{r}, t)}{\epsilon_0} & \text{Équation de Maxwell-Gauss} \\ \nabla \cdot \mathbf{B}(\mathbf{r}, t) = 0 & \text{Absence de monopole magnétique} \\ \nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} & \text{Équation de Maxwell-Faraday} \\ \nabla \times \mathbf{B}(\mathbf{r}, t) = \mu_0 \mathbf{j}(\mathbf{r}, t) + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} & \text{Équation de Maxwell-Ampère} \end{array} \right.$$

Équation de Lorentz

$$m_a \frac{d}{dt} \left(\frac{\mathbf{v}_a}{\sqrt{1 - \mathbf{v}_a^2/c^2}} \right) = q_a (\mathbf{E}(\mathbf{r}_a, t) + \mathbf{v}_a \times \mathbf{B}(\mathbf{r}_a, t))$$

Conservation de la charge :

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

Conservation de l'énergie :

$$\frac{\partial u}{\partial t} + \nabla \cdot \mathbf{\Pi} + \mathbf{j} \cdot \mathbf{E} = 0$$

avec u la densité locale d'énergie électromagnétique, et $\mathbf{\Pi}$ le vecteur de Poynting

$$u(\mathbf{r}, t) = \frac{1}{2} \epsilon_0 (\mathbf{E}^2 + c^2 \mathbf{B}^2) = \frac{\epsilon_0 \mathbf{E}^2}{2} + \frac{\mathbf{B}^2}{2\mu_0}$$

$$\mathbf{\Pi}(\mathbf{r}, t) = \epsilon_0 c^2 (\mathbf{E} \times \mathbf{B}) = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0}$$

Conservation de l'impulsion :

$$\frac{d}{dt} (\mathbf{P}_{\text{part}}^V)_i = \int_V (\partial_j \tau_{ij} - \epsilon_0 \partial_t (\mathbf{E} \times \mathbf{B})_i) d^3\mathbf{r}$$

avec τ le tenseur de Maxwell

$$\tau_{ij} = \epsilon_0 \left[E_i E_j + c^2 B_i B_j - \frac{1}{2} (\mathbf{E} \cdot \mathbf{E} + c^2 \mathbf{B} \cdot \mathbf{B}) \delta_{ij} \right]$$

On définit de plus les potentiels ϕ et \mathbf{A} par :

$$\mathbf{B} = \nabla \times \mathbf{A}$$

$$\mathbf{E} = -\nabla \phi - \partial_t \mathbf{A}$$

On peut rajouter des conditions sur ces potentiels, appelées jauges, sans perdre d'informations sur les champs
Jauge de Lorentz :

$$\nabla \cdot \mathbf{A} + \frac{1}{c^2} \frac{\partial \phi}{\partial t} = 0$$

2.1.5 Transformation des champs sous une transformation de Galilée

On va désormais montrer pourquoi une formulation relativiste de l'électromagnétisme est nécessaire, en remarquant que les équations de Maxwell ne sont pas invariantes sous l'effet d'une transformation de Galilée. Si l'on considère deux référentiels \mathcal{R} et \mathcal{R}' en translation rectiligne uniforme l'un par rapport à l'autre, et tels que :

$$\mathbf{r}' = \mathbf{r} - \mathbf{v}t \quad (2.28a)$$

$$t' = t \quad (2.28b)$$

on a alors :

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial t'} - \mathbf{v} \cdot \nabla \quad (2.29a)$$

$$\nabla' = \nabla \quad (2.29b)$$

Ceci permet d'écrire par exemple

$$\nabla \cdot \mathbf{E}(\mathbf{r}, t) = \nabla' \cdot \mathbf{E}[\mathbf{r}(\mathbf{r}', t'), t(\mathbf{r}', t')]. \quad (2.30)$$

Dans la suite, on notera $\mathbf{E}[\mathbf{r}(\mathbf{r}', t'), t(\mathbf{r}', t')] = \mathbf{E}[\mathbf{r}, t]$. Par définition, cette quantité est égale à $\mathbf{E}'(\mathbf{r}', t')$. Les équations de Maxwell ainsi que la conservation de la charge sont alors transformées de la façon suivante

$$\begin{aligned} \nabla' \cdot \mathbf{E}[\mathbf{r}, t] &= \frac{\rho(\mathbf{r}, t)}{\epsilon_0} & \nabla' \cdot \mathbf{E}'(\mathbf{r}', t') &= \frac{\rho'(\mathbf{r}', t')}{\epsilon_0} \\ \nabla' \cdot \mathbf{B}[\mathbf{r}, t] &= 0 & \nabla' \cdot \mathbf{B}'(\mathbf{r}', t') &= 0 \\ \nabla' \times (\mathbf{E}[\mathbf{r}, t] + \mathbf{v} \times \mathbf{B}[\mathbf{r}, t]) &= -\frac{\partial \mathbf{B}[\mathbf{r}, t]}{\partial t'} & \nabla' \times \mathbf{E}'(\mathbf{r}', t') &= -\frac{\partial \mathbf{B}'(\mathbf{r}', t')}{\partial t'} \\ \nabla' \times \left(\mathbf{B}[\mathbf{r}, t] - \frac{\mathbf{v} \times \mathbf{E}[\mathbf{r}, t]}{c^2} \right) &= \mu_0(\mathbf{j} - \rho\mathbf{v}) + \mu_0\epsilon_0 \frac{\partial \mathbf{E}[\mathbf{r}, t]}{\partial t'} & \nabla' \times \mathbf{B}'(\mathbf{r}', t') &= \mu_0\mathbf{j}'(\mathbf{r}', t') + \mu_0\epsilon_0 \frac{\partial \mathbf{E}'(\mathbf{r}', t')}{\partial t'} \\ \frac{\partial \rho}{\partial t'} + \nabla' \cdot (\mathbf{j} - \rho\mathbf{v}) &= 0 & \frac{\partial \rho'}{\partial t'} + \nabla' \cdot \mathbf{j}' &= 0 \end{aligned} \quad (2.31)$$

La première colonne donne le résultat du changement de variables en partant des équations de Maxwell dans le référentiel \mathcal{R} . Les équations de Maxwell sont ici exprimées dans les nouvelles variables². La seconde série d'équation donne les équations de Maxwell telles qu'elles devraient être dans le référentiel \mathcal{R}' si nous avions invariance par transformation de Galilée. Regardons par exemple comment se transforme l'équation Maxwell-Faraday. Nous avons

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{B}(\mathbf{r}, t)}{\partial t} \quad (2.32a)$$

$$\nabla' \times \mathbf{E}[\mathbf{r}, t] = -\frac{\partial \mathbf{B}[\mathbf{r}, t]}{\partial t'} + (\mathbf{v} \cdot \nabla')\mathbf{B}[\mathbf{r}, t] \quad (2.32b)$$

En utilisant $\nabla \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a}(\nabla \cdot \mathbf{b}) - \mathbf{b}(\nabla \cdot \mathbf{a}) + (\mathbf{b} \cdot \nabla)\mathbf{a} - (\mathbf{a} \cdot \nabla)\mathbf{b}$, on trouve alors

$$\nabla' \times (\mathbf{E}[\mathbf{r}, t] + \mathbf{v} \times \mathbf{B}[\mathbf{r}, t]) = -\frac{\partial \mathbf{B}[\mathbf{r}, t]}{\partial t'} \quad (2.33)$$

Cela étant, si l'on veut que toutes ces équations soit invariantes, il faut que les charges et les champs vérifient les relations

$$\rho' = \rho \quad (2.34a)$$

$$\mathbf{j}' = \mathbf{j} - \rho\mathbf{v} \quad (2.34b)$$

$$\mathbf{E}' = \mathbf{E} \quad (2.34c)$$

$$\mathbf{B}' = \mathbf{B} \quad (2.34d)$$

$$\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B} \quad (2.34e)$$

$$\mathbf{B}' = \mathbf{B} - \frac{\mathbf{v} \times \mathbf{E}}{c^2} \quad (2.34f)$$

². Dans la première colonne on a choisi de ne pas écrire les équations en fonction des champs E' et B' pour rappeler que nous sommes partis des équations de Maxwell valables dans le référentiel \mathcal{R} pour les champs E et B . Mais il faut se rappeler que $\mathbf{E}[\mathbf{r}, t] = \mathbf{E}'(\mathbf{r}', t')$!



Or, ces relations ne peuvent pas être vérifiées en même temps : les équations de Maxwell ne sont donc pas invariantes sous une transformation de Galilée! En revanche, on peut montrer qu'elles sont invariantes sous une transformation de Lorentz, ce qui amène à chercher une formulation relativiste de l'électromagnétisme. Historiquement, c'est l'étude de l'électromagnétisme et de ses invariances qui a amené à introduire les transformations de Lorentz. La question était alors de savoir laquelle des transformations de Galilée ou de Lorentz était réellement fondamentale. Einstein admit que c'était celle de Lorentz, modifia les équations de la dynamique de Newton et développa ainsi la mécanique relativiste, mécanique invariante sous une transformation de Lorentz.

2.2 Solutions des équations de Maxwell, Potentiels retardés

Nous allons maintenant nous attarder à analyser plus précisément les champs électriques et magnétiques exactes solutions des équations de Maxwell. En particulier, nous verrons une fois de plus au travers du champ produit par une charge en mouvement rectiligne uniforme que l'électromagnétisme ne respecte pas les postulats de la mécanique newtonienne (pas d'invariance sous transformations de Galilée) mais est en fait une théorie relativiste où l'information se propage à une vitesse finie. Les aspects dynamiques seront analysés plus en détails et principalement les effets de rayonnement du champ électromagnétique par une distribution de charges et de courants arbitraire.

2.2.1 Potentiels retardés

On considère une distribution de charges et de courants quelconque. Le but est de déterminer les champs exacts créés par une telle distribution. On se place en jauge de Lorentz et on considère les équations vérifiées par les potentiels

$$\begin{aligned}\square\left(\frac{\Phi(\mathbf{r}, t)}{c}\right) &= -\mu_0\rho(\mathbf{r}, t)c \\ \square\mathbf{A}(\mathbf{r}, t) &= -\mu_0\mathbf{j}(\mathbf{r}, t)\end{aligned}\quad (2.35)$$

on cherche donc les solutions de l'équation d'onde

$$\square\Psi(\mathbf{r}, t) = -f(\mathbf{r}, t) \quad (2.36)$$

qui se résout à l'aide des fonctions de Green. La fonction de Green $\mathcal{G}(\mathbf{r}, t, \mathbf{r}', t')$ vérifie l'équation

$$\square\mathcal{G}(\mathbf{r}, t, \mathbf{r}, t') = \left(\Delta - \frac{1}{c^2}\partial_{tt'}^2\right)\mathcal{G}(\mathbf{r}, t, \mathbf{r}', t') = \delta(\mathbf{r} - \mathbf{r}')\delta(t - t') \quad (2.37)$$

Trouver les solutions d'une telle équation se fait en utilisant la transformée de Fourier. On ne fera pas ces étapes ici et donnons directement les solutions. On obtient deux fonctions de Green nommées \mathcal{G}^+ (retardée) et \mathcal{G}^- (avancée) qui s'écrivent

$$\mathcal{G}^\pm(\mathbf{r}, t, \mathbf{r}', t') = \frac{1}{4\pi\|\mathbf{r} - \mathbf{r}'\|}\delta\left[t' - \left(t \mp \frac{\|\mathbf{r} - \mathbf{r}'\|}{c}\right)\right] \quad (2.38)$$

Le contenu physique de la fonction delta est simple à interpréter. La fonction de Green retardée est causale au sens où un événement de coordonnées (t', \mathbf{r}') ne pourra être influencé par un autre en (t, \mathbf{r}) que si une information a pu se propager à la vitesse c (parcourir la distance entre les deux événements) en un laps de temps $t' - t$. Au contraire la fonction de Green avancée est anti-causale. En choisissant le potentiel retardé pour avoir une solution causale (même si mathématiquement rien n'empêche de choisir l'autre solution), nous avons

$$\Psi(\mathbf{r}, t) = \frac{1}{4\pi}\iint\frac{1}{\|\mathbf{r} - \mathbf{r}'\|}\delta\left[t' - \left(t - \frac{\|\mathbf{r} - \mathbf{r}'\|}{c}\right)\right]f(\mathbf{r}', t')d\mathbf{r}'dt' \quad (2.39)$$

En intégrant la composante temporelle nous avons, en définissant $[\cdot]_{\text{ret}}$ l'estimée en $t' = t - \frac{\|\mathbf{r} - \mathbf{r}'\|}{c}$

$$\Psi(\mathbf{r}, t) = \frac{1}{4\pi}\int\frac{[f(\mathbf{r}', t')]_{\text{ret}}}{\|\mathbf{r} - \mathbf{r}'\|}d\mathbf{r}' \quad (2.40)$$



La vraie solution complète des équations du mouvement devrait contenir un terme supplémentaire $\Psi_0(\mathbf{r}, t)$ solution de l'équation libre. Il s'avère que l'on peut la choisir nulle la plupart du temps si on traite un système isolé sans ondes entrantes.

Dans le cas de l'électromagnétisme, nous obtenons les potentiels retardés, avec lesquels nous pourrions en déduire les champs électriques et magnétiques exactes,

$$\Phi(\mathbf{r}, t) = \frac{1}{4\pi\epsilon_0} \int \frac{[\rho(\mathbf{r}', t')]_{\text{ret}}}{\|\mathbf{r} - \mathbf{r}'\|} d\mathbf{r}' \quad (2.41a)$$

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi} \int \frac{[\mathbf{j}(\mathbf{r}', t')]_{\text{ret}}}{\|\mathbf{r} - \mathbf{r}'\|} d\mathbf{r}' \quad (2.41b)$$

2.2.2 Potentiels de Liénard-Wiechert

On s'intéresse au cas particulier d'une charge ponctuelle en mouvement, dont la position au cours du temps est $\mathbf{r}_q(t')$

$$\rho(\mathbf{r}', t') = q\delta(\mathbf{r}' - \mathbf{r}_q(t')) \quad (2.42)$$

On utilise la forme générale des équations, c'est-à-dire celle où nous avons explicitement toutes les intégrations. Avec cela, nous avons en intégrant sur les variables spatiales

$$\Phi(\mathbf{r}, t) = \frac{q}{4\pi\epsilon_0} \int \frac{\delta\left(t' - \left(t - \frac{\|\mathbf{r} - \mathbf{r}_q(t')\|}{c}\right)\right)}{\|\mathbf{r} - \mathbf{r}_q(t')\|} dt' \quad (2.43)$$

En opérant le changement de variable $s = t' + \frac{\|\mathbf{r} - \mathbf{r}_q(t')\|}{c}$, nous avons $ds = dt' + \frac{1}{c} \frac{dr}{dt'} dt'$ où nous avons défini $r = \|\mathbf{r} - \mathbf{r}_q(t')\|$. Le potentiel s'écrit donc

$$\begin{aligned} \Phi(\mathbf{r}, t) &= \frac{q}{4\pi\epsilon_0} \int \frac{\delta(s - t)}{\left(1 + \frac{1}{c} \frac{dr}{dt'}\right) r} ds \\ &= \frac{q}{4\pi\epsilon_0} \left[\frac{1}{\left(1 + \frac{1}{c} \frac{dr}{dt'}\right) r} \right]_{t' = t - \frac{\|\mathbf{r} - \mathbf{r}_q(t')\|}{c}} \end{aligned} \quad (2.44)$$

Mais

$$\frac{dr}{dt'} = -\frac{d\mathbf{r}_q(t)}{dt'} \cdot \frac{(\mathbf{r} - \mathbf{r}_q(t'))}{\|\mathbf{r} - \mathbf{r}_q(t)\|} = -\mathbf{v}(t') \cdot \mathbf{n} \quad (2.45)$$

d'où la forme finale, appelée potentiels de Liénard-Wiechert

$$\Phi(\mathbf{r}, t) = \frac{q}{4\pi\epsilon_0} \left[\frac{1}{(1 - \boldsymbol{\beta} \cdot \mathbf{n}) r} \right]_{\text{ret}} \quad (2.46a)$$

$$\mathbf{A}(\mathbf{r}, t) = \frac{q\mu_0}{4\pi} \left[\frac{\mathbf{v}}{(1 - \boldsymbol{\beta} \cdot \mathbf{n}) r} \right]_{\text{ret}} \quad (2.46b)$$

2.2.3 Charge en mouvement rectiligne uniforme

Pour illustrer le contenu de ces formules, considérons une charge en mouvement rectiligne uniforme. Si l'électromagnétisme et la relativité Galiléenne étaient compatibles, ce champ pourrait directement être trouvé à partir de la loi de Coulomb valable dans le référentiel inertiel au repos de la charge et des lois de transformation de Galilée. Cet exemple a pour intérêt principal d'illustrer que ces lois de transformation ne s'appliquent pas au champ électromagnétique et que les bonnes lois de transformation entre différents référentiels inertiels sont données par les transformations de Lorentz. De plus il permet de voir aussi que champ électrique et champ magnétique sont en fait intrinsèquement liés.



On considère donc un observateur et une particule chargée en mouvement rectiligne uniforme l'un par rapport à l'autre dans la direction x à la vitesse v . Pour simplifier les calculs, on ne regarde le champ que dans le plan Oxy . On cherche à déterminer le champ au point M de coordonnées (x, y) à l'instant t . La position de la particule sera notée $P(t)$.

Pour obtenir les champs électrique et magnétique, on doit évaluer le facteur $(1 - \boldsymbol{\beta} \cdot \mathbf{n})r$ dans 2.46 à l'instant retardé. Un raisonnement géométrique permet d'obtenir l'égalité $[(1 - \boldsymbol{\beta} \cdot \mathbf{n})r]_{\text{ret}} = r\sqrt{1 - \frac{v^2}{c^2} \sin^2(\theta)}$ (il faut comparer la situation à l'instant retardé et à l'instant t). Ici θ est l'angle entre la direction de la vitesse et la direction $\mathbf{P}(t)\mathbf{M}$, r en est la norme et \mathbf{n} le vecteur unitaire. Donc, en écrivant cela en coordonnées cartésiennes

$$\Phi(\mathbf{r}, t) = \frac{q\gamma}{4\pi\epsilon_0\sqrt{\gamma^2(x-vt)^2 + y^2}} \quad (2.47a)$$

$$\mathbf{A}(\mathbf{r}, t) = \frac{q\mu_0\gamma}{4\pi\sqrt{\gamma^2(x-vt)^2 + y^2}} \mathbf{v} \quad (2.47b)$$

Ayant cela, on trouve les champs électrique et magnétique *via* les formules reliant les champs aux potentiels $\mathbf{E} = -\nabla\Phi - \partial_t\mathbf{A}$ et $\mathbf{B} = \nabla \times \mathbf{A}$.

$$\mathbf{E}(\mathbf{r}, t) = \frac{q\gamma}{4\pi\epsilon_0(\gamma^2(x-vt)^2 + y^2)^{3/2}} [(x-vt)\mathbf{e}_x + y\mathbf{e}_y] \quad (2.48a)$$

$$= \frac{q}{4\pi\epsilon_0\gamma^2 r^2 \left[1 - \frac{v^2}{c^2} \sin^2(\theta)\right]^{3/2}} \mathbf{n} \quad (2.48b)$$

$$\mathbf{B}(\mathbf{r}, t) = \frac{\mu_0}{4\pi} \frac{qv}{\gamma^2 r^2 \left[1 - \frac{v^2}{c^2} \sin^2(\theta)\right]^{3/2}} \mathbf{e}_x \times \mathbf{n} \quad (2.48c)$$

Les conclusions à tirer de ce résultat sont les suivantes. Dans le référentiel au repos le champ électrique suit la loi de Coulomb tandis qu'aucun champ magnétique n'est présent. Dans le référentiel inertiel de l'observateur voyant la particule en mouvement rectiligne uniforme à la vitesse v , les champs solutions des équations de Maxwell ont la forme ci-dessus. On voit clairement qu'un raisonnement à partir de la loi de Coulomb et des lois de transformation de Galilée de la mécanique ne donnent pas du tout le bon champ électrique. Pire que cela elles ne permettent pas de prévoir l'existence d'un champ magnétique mesuré par l'observateur. Les lois de transformation $x' = x - vt$, $t' = t$ et $y' = y$ de Galilée ne marchent pas pour le champ électromagnétique.

2.3 Rayonnement électromagnétique

L'étude de l'électromagnétisme se fait traditionnellement en étudiant d'abord les situations stationnaires (distribution des charges et des courants indépendants du temps) où l'on met en évidence des lois comme celle de Coulomb. Les situations dynamiques sont ensuite abordées en regardant les solutions des équations de Maxwell dans le vide puis dans des milieux. Ceci met en évidence l'existence des ondes électromagnétiques ainsi que leur structure et la manière dont elles se propagent. Reste donc à comprendre la génération de telles ondes électromagnétiques. Pour cela, il est en fait nécessaire d'utiliser les solutions complètes des équations de Maxwell déterminées précédemment.

2.3.1 Les champs électrique et magnétique rayonnés

À partir de l'expression des potentiels précédente, on peut trouver l'expression des champs

$$\mathbf{E}(\mathbf{r}, t) = \frac{q}{4\pi\epsilon_0} \left(\left[\frac{\mathbf{n} - \boldsymbol{\beta}}{\gamma^2 r^2 (1 - \mathbf{n} \cdot \boldsymbol{\beta})^3} \right]_{\text{ret}} + \left[\frac{1}{rc} \frac{\mathbf{n} \times \left((\mathbf{n} - \boldsymbol{\beta}) \times \frac{\partial \boldsymbol{\beta}}{\partial t'} \right)}{(1 - \mathbf{n} \cdot \boldsymbol{\beta})^3} \right]_{\text{ret}} \right) \quad (2.49a)$$

$$\mathbf{B}(\mathbf{r}, t) = \left[\frac{\mathbf{n} \times \mathbf{E}}{c} \right]_{\text{ret}} \quad (2.49b)$$

Le point important ici est l'apparition du terme en $\frac{\partial \boldsymbol{\beta}}{\partial t'}$ (on notera dans la suite cette quantité $\dot{\boldsymbol{\beta}}$). En effet, ce terme, dit *de rayonnement*, est en $1/r$, alors que le premier terme est en $1/r^2$. Ce dernier est, on le voit, égal au champ de Coulomb pour une particule de vitesse nulle et en est la généralisation. On voit que le terme de rayonnement est quant à lui proportionnel à l'accélération du système : des charges accélérées rayonnent du champ électromagnétique et perdent de l'énergie *via* ce dernier.



2.3.2 Puissance rayonnée

Le terme de rayonnement domine le champ à grande distance, et c'est le seul terme que nous étudierons par la suite. On va distinguer la limite non relativiste du calcul complet et analyser la puissance rayonnée. Les résultats trouvés ont des applications physiques très diverses, allant des phénomènes de diffusion permettant de comprendre la couleur du ciel au rayonnement synchrotron dans les accélérateurs de particules.

Cas non relativiste

Dans ce cas, $|\beta| \ll 1$. L'expression approchée pour le champ électrique est alors

$$\mathbf{E}(\mathbf{r}, t) = \frac{q}{4\pi\epsilon_0} \left[\frac{\mathbf{n} \times (\mathbf{n} \times \dot{\beta})}{rc} \right]_{\text{ret}} \quad (2.50)$$

Ce qui s'écrit encore :

$$\mathbf{E}(\mathbf{r}, t) = \frac{q}{4\pi\epsilon_0} \left[\frac{(\mathbf{n} \cdot \dot{\beta}) \mathbf{n} - \dot{\beta}}{rc} \right]_{\text{ret}} \quad (2.51)$$

On s'intéresse ici à la puissance rayonnée. Pour cela, calculons le vecteur de Poynting, en se rappelant que \mathbf{E} et \mathbf{n} sont orthogonaux

$$\begin{aligned} \mathbf{\Pi} &= \frac{\mathbf{E} \times \mathbf{B}}{\mu_0} \\ &= \frac{\mathbf{E} \times (\mathbf{n} \times \mathbf{E})}{\mu_0 c} \\ &= \frac{|\mathbf{E}|^2}{\mu_0 c} \mathbf{n} \end{aligned}$$

On sait que l'énergie traversant une surface par unité de temps est égale au flux du vecteur de Poynting à travers cette surface. Donc, en notant P la puissance rayonnée par la particule, on a

$$\frac{dP}{dS} = \frac{|\mathbf{E}|^2}{\mu_0 c} \quad (2.52a)$$

$$= \frac{q^2 \dot{\beta}^2}{16\pi^2 \epsilon_0 c r^2} (1 - \cos^2(\theta)) \quad (2.52b)$$

où l'on a noté θ l'angle entre $\dot{\beta}$ et \mathbf{n} . Attention, toutes les expressions écrites sont à évaluer au temps retardé!

On trouve alors la formule de Larmor pour une particule non relativiste :

$$\frac{dP}{d\Omega} = \frac{q^2 \sin^2(\theta) \dot{\mathbf{v}}^2}{16\pi^2 \epsilon_0 c^3} \quad (2.53)$$

L'intégration de cette formule sur tout l'angle solide nous donne la puissance totale rayonnée par la particule

$$P_{\text{tot}} = \frac{q^2 \dot{\mathbf{v}}^2}{6\pi\epsilon_0 c^3} \quad (2.54)$$

Les points physiques importants qui sortent de cette formule sont la dépendance quadratique en l'accélération de la puissance rayonnée et celle en sinus de l'angle entre la direction de l'accélération et la direction de visée au temps retardé. On voit en particulier que la puissance rayonnée est maximale dans une direction orthogonale à l'accélération et nulle dans une direction parallèle. Ce résultat est le cas général de ce qui se passe pour le champ rayonné par un dipôle (comme une antenne par exemple...).

Enfin, on définit l'*indicatrice de rayonnement* comme étant la surface vérifiant $\frac{dP}{dS} = \text{constante}$. Ici, elle est de la forme $R = k \sin(\theta)$



Cas relativiste

Passons maintenant au cas relativiste. On évalue toujours le vecteur de Poynting pour la partie radiative seulement. Les champs \mathbf{E} et \mathbf{n} sont toujours orthogonaux. On obtient alors facilement

$$\frac{dP}{dS} = \frac{q^2}{16\pi^2\epsilon_0 c} \frac{\left| \mathbf{n} \times \left((\mathbf{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}} \right) \right|^2}{r^2 (1 - \mathbf{n} \cdot \boldsymbol{\beta})^5} \quad (2.55)$$

On ne va pas dans la suite détailler toutes les expressions utiles pour extraire la physique de cette expression (l'idée est en fait simplement de développer le carré scalaire et d'exprimer cela en terme de l'accélération et la vitesse de la particule relativement à l'observateur au temps retardé; ce n'est pas très dur mais les formules sont un peu grosses).

La physique à retenir est la suivante.

La puissance rayonnée dépend de la direction relative de l'accélération de la particule et de sa vitesse par rapport à l'observateur. Si la vitesse est nulle, on retrouve la formule de Larmor avec une distribution dipolaire du rayonnement. Dans le cas où la vitesse et l'accélération sont colinéaires, les lobes du rayonnement se focalisent dans la direction du mouvement et l'intensité devient très grande. Ce phénomène s'appelle focalisation Doppler. Dans le cas où vitesse et accélération sont orthogonales, le même phénomène se passe, seuls les angles et l'intensité changent. Ce dernier cas est typiquement le rayonnement émis par une particule chargée dans un champ magnétique (mouvement hélicoïdal) appelé alors *rayonnement synchrotron* dont les propriétés sont utilisées dans de nombreux domaines de recherche.

Pour être complet, on donne la puissance totale rayonnée, qui doit bien sûr redonner à la limite non-relativiste la formule de Larmor

$$P_{\text{tot}} = -\frac{q^2}{6\pi\epsilon_0 m^2 c^3} \frac{dP_\mu}{d\tau} \frac{dP^\mu}{d\tau} \quad (2.56)$$

Ceci nous donne, en explicitant P_μ , la formule de Liénard

$$P_{\text{tot}} = \frac{q^2 \gamma^6}{6\pi\epsilon_0 c} \left(\dot{\boldsymbol{\beta}}^2 - (\boldsymbol{\beta} \times \dot{\boldsymbol{\beta}})^2 \right) \quad (2.57a)$$

$$= \frac{q^2}{6\pi\epsilon_0 m^2 c^3} \left(\dot{p}_\parallel^2 + \gamma^2 \dot{p}_\perp^2 \right) \quad (2.57b)$$

2.4 Formulation relativiste de l'électrodynamique

L'électromagnétisme est une théorie relativiste. Cependant, ceci n'est pas évident au premier abord. On souhaite donc reformuler la théorie du champ électromagnétique de manière à faire ressortir cet aspect.

Les postulats fondamentaux sont ceux de la relativité restreinte ainsi que le postulat de l'invariance de la charge. L'invariance des lois de la physique est automatiquement satisfaite si nous les écrivons sous forme tensorielle. Les transformations considérées sont celles du groupe de Lorentz.

L'intérêt d'une telle reformulation provient du fait qu'elle permet de faire clairement ressortir certains aspects physiques de l'électromagnétisme (lois de transformation entre observateurs, propagation à une vitesse finie, lois de conservation) mais également du fait qu'elle est au cœur de la formulation lagrangienne et des généralisations aux interactions faible et forte.

2.4.1 Quadri-vecteur courant

Soit une charge q dans un référentiel \mathcal{R} que l'on modélise par une distribution de charge ρ . Dans \mathcal{R}' en translation rectiligne uniforme la charge q occupe le volume dV' et correspond à ρ' . Par invariance de la charge, $\rho dV = \rho' dV'$. Pendant un intervalle de temps dt , la charge se déplace de dx^μ ³ = $(cdt, d\mathbf{r})$. Alors, $\rho dV dx^\mu$ est un quadri-vecteur. De plus, $cdtdV$ est un invariant de Lorentz. En effet, il s'agit de l'élément d'intégration dans l'espace de Minkowski. Le jacobien du changement de base valant 1 (déterminant de la matrice de Lorentz \mathcal{L}), on a $cdtdV = cd't'dV'$. Finalement, comme $cdtdV$ est un invariant et que $\rho dV dx^\mu$ est un quadri-vecteur, $\rho \frac{dx^\mu}{dt}$ est un quadri-vecteur. On pose :

³. Comme pour les tenseurs précédemment, on utilise la même notation pour un quadri-vecteur et pour sa composante μ , et ce afin de bien indiquer si l'on donne les coordonnées covariantes ou contravariantes du quadri-vecteur



$$J^\mu = \rho \frac{dx^\mu}{dt} = (\rho c, \mathbf{j}) = (\rho c, \rho \mathbf{v}) = \rho_0 (\gamma c, \gamma \mathbf{v}) \quad (2.58)$$

où ρ_0 est la distribution de charge dans le référentiel propre de la particule.

2.4.2 Quadri-vecteur potentiel

En jauge de Lorentz, les équations aux potentiels sont :

$$\square \left(\frac{\Phi(\mathbf{r}, t)}{c} \right) = -\mu_0 \rho(\mathbf{r}, t) c \quad (2.59a)$$

$$\square \mathbf{A}(\mathbf{r}, t) = -\mu_0 \mathbf{j}(\mathbf{r}, t) \quad (2.59b)$$

Ceci amène à considérer $\left(\frac{\Phi}{c}, \mathbf{A} \right)$ comme les composantes contravariantes d'un quadri-vecteur. On appelle quadri-potentiel le quadri-vecteur $\Phi^\mu = \left(\frac{\Phi}{c}, \mathbf{A} \right)$. On le note parfois A^μ . Les équations aux potentiels s'écrivent alors :

$$\square \Phi^\mu = -\mu_0 J^\mu \quad (2.60)$$

La jauge de Lorentz s'écrit quant à elle :

$$\partial_\mu \Phi^\mu = 0 \quad (2.61)$$

2.4.3 Équations de Maxwell-Lorentz

En écrivant les champs \mathbf{E} et \mathbf{B} en fonction des potentiels, on peut faire apparaître les composantes complètement contravariantes d'un tenseur antisymétrique d'ordre 2, noté $F^{\mu\nu}$. On appelle *tenseur électromagnétique* le tenseur antisymétrique $F^{\mu\nu}$ défini par ⁴

$$F^{\mu\nu} = \partial^\mu \Phi^\nu - \partial^\nu \Phi^\mu = \begin{pmatrix} 0 & -\frac{E_x}{c} & -\frac{E_y}{c} & -\frac{E_z}{c} \\ \frac{E_x}{c} & 0 & -B_z & B_y \\ \frac{E_y}{c} & B_z & 0 & -B_x \\ \frac{E_z}{c} & -B_y & B_x & 0 \end{pmatrix} \quad (2.62)$$

Les composantes covariantes de ce tenseur sont :

$$F_{\mu\nu} = g_{\mu\alpha} g_{\nu\beta} F^{\alpha\beta} = \begin{pmatrix} 0 & \frac{E_x}{c} & \frac{E_y}{c} & \frac{E_z}{c} \\ -\frac{E_x}{c} & 0 & -B_z & B_y \\ -\frac{E_y}{c} & B_z & 0 & -B_x \\ -\frac{E_z}{c} & -B_y & B_x & 0 \end{pmatrix} \quad (2.63)$$

Les équations de Maxwell se formulent alors simplement *via* ce tenseur de la manière qui suit. Les équations de Maxwell avec les sources (Maxwell-Gauss et Maxwell-Ampère) sont données par :

$$\partial_\mu F^{\mu\nu} = \mu_0 J^\nu \quad (2.64)$$

Les équations de Maxwell propres aux champs électromagnétiques (Maxwell-Faraday et Maxwell-Flux) sont données par :

$$\partial_\alpha F_{\beta\gamma} + \partial_\gamma F_{\alpha\beta} + \partial_\beta F_{\gamma\alpha} = 0 \quad (2.65)$$

Reste maintenant à reformuler l'équation de la dynamique d'une particule chargée dans un champ électromagnétique. Avec l'équation du mouvement $m \frac{d(\gamma \mathbf{v})}{dt} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ et le bilan en puissance $\frac{d(\gamma m c^2)}{dt} = q(\mathbf{E} \cdot \mathbf{v})$, on peut obtenir l'équation de la dynamique dans le cadre relativiste. La dynamique d'une particule chargée dans un champ est donnée par l'équation :

$$\frac{dP^\mu}{d\tau} = q F^{\mu\nu} U_\nu \quad (2.66)$$

où $P^\mu = (m\gamma c, m\gamma \mathbf{v}) = mU^\mu$ est le vecteur quadri-impulsion, $U^\mu = (\gamma c, \gamma \mathbf{v})$ est le quadri-vecteur vitesse ⁵.

⁴. Pour se rappeler de ce tenseur, on peut se souvenir que la ligne 0 contient uniquement des signes moins, et que les signes moins pour le champ \mathbf{B} sont obtenus pour tous les ensemble de deux éléments dans $\{1, 2, 3\}$ pris dans l'ordre naturel, à savoir (1, 2), (2, 3) et (3, 1).

⁵. Ici, ce vecteur est utilisé en coordonnées covariantes, c'est-à-dire $U_\mu = (\gamma c, -\gamma \mathbf{v})$



$$\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} \quad (2.80)$$

En terme des champs,

$$\epsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} = -\frac{8}{c} \mathbf{E} \cdot \mathbf{B} \quad (2.81)$$

Le produit scalaire des champs est donc un invariant relativiste.

2.4.7 Tenseur énergie-impulsion

Définition

Jusqu'à maintenant, nous n'avons donné une description relativiste que pour la dynamique d'une particule (description microscopique), possiblement en interaction avec le champ électromagnétique. On peut généraliser une telle description relativiste pour des systèmes continus : les principales applications de cette description sont l'hydrodynamique relativiste et la relativité générale, théorie relativiste de la gravitation. L'objet permettant de faire cette généralisation est *le tenseur énergie-impulsion* que l'on va définir.

Considérons un certain volume d'espace-temps V de normale n^μ (dont le genre dépend bien sûr du volume que l'on choisit). Le tenseur énergie impulsion $T_{\mu\nu}$ est défini comme étant "une densité de quadri-impulsion". Plus précisément, l'impulsion étant un vecteur, le passage à la limite continue nous dit que le flux à travers le volume V correspond à la quadri-impulsion totale du volume V . Explicitement

$$p_\mu|_V = \pm \frac{1}{c} \int_V T_{\mu\nu} n^\nu dV \quad (2.82)$$

Regardons maintenant la signification physique de chacune des composantes. Soient un observateur de quadri-vitesse u^μ et un volume dans son espace de repos (de normale de genre temps égale à u^μ donc). Alors $dp_\mu u^\mu = T_{\mu\nu} u^\nu u^\mu / cdV$. Cela nous donne par exemple $dE = T_{00} dV$. Dit autrement $T_{00} = \epsilon$ où ϵ est la densité d'énergie mesurée par l'observateur. En raisonnant de la même façon, on peut montrer que $T_{i0} = c \frac{dp_i}{dv} \equiv c\pi_i$ (proportionnel à la densité d'impulsion π mesurée par l'observateur dans la direction \mathbf{e}_i), $T_{0i} = \frac{1}{c} \frac{dE_i}{dt} \equiv \frac{1}{c} \phi_i$ (proportionnel au flux d'énergie ϕ_i traversant une surface de normale \mathbf{e}_i) et enfin $T_{ij} \equiv -S_{ij}$ se dénomme tenseur des contraintes et donne la force dans la direction \mathbf{e}_i sur une surface de normale \mathbf{e}_j . Pour résumer

$$T_{\mu\nu} = \begin{pmatrix} \epsilon & \phi_x/c & \phi_y/c & \phi_z/c \\ c\pi_x & -S_{xx} & -S_{xy} & -S_{xz} \\ c\pi_y & -S_{yx} & -S_{yy} & -S_{yz} \\ c\pi_z & -S_{zx} & -S_{zy} & -S_{zz} \end{pmatrix} \quad (2.83)$$

Un résultat important (que l'on ne montrera pas et qui provient de la conservation du moment cinétique) est que le tenseur énergie-impulsion est *symétrique*. Ceci a pour conséquence que pour un observateur donné $\phi = \pi c^2$, qui peut être vu comme la généralisation de $E = mc^2$.

Conservation de l'énergie-impulsion

Nous avons vu qu'à partir des équations de Maxwell, on obtient des lois de conservation de l'énergie et de l'impulsion du champ électromagnétique. La conservation de l'énergie s'écrit

$$\int_V \left(\frac{\partial u}{\partial t} + \nabla \cdot \Pi + \mathbf{j} \cdot \mathbf{E} \right) d^3\mathbf{r} = 0 \quad (2.84)$$

avec $u(\mathbf{r}, t) = \frac{1}{2} \epsilon_0 \mathbf{E}^2 + \frac{1}{2\mu_0} \mathbf{B}^2$ la densité d'énergie, $\Pi = \frac{\mathbf{E} \times \mathbf{B}}{\mu_0}$ le vecteur de Poynting⁷. La conservation de l'impulsion s'écrit

⁷. Le flux du vecteur de Poynting à travers une surface S donne la puissance ayant traversé cette surface. Il caractérise le transport de l'énergie électromagnétique.



$$\int_V (\partial_j \tau_{ij} - \partial_t \epsilon_0 (\mathbf{E} \times \mathbf{B})_i) d^3\mathbf{r} = \frac{d\mathbf{P}_i}{dt} \quad (2.85)$$

où \mathbf{P} est l'impulsion totale des particules contenues dans V et τ est le tenseur de Maxwell dont les composantes s'écrivent

$$\tau_{ij} = -\epsilon_0 \left[\frac{1}{2} \mathbf{E}^2 \delta_{ij} - E_i E_j \right] - \frac{1}{\mu_0} \left[\frac{1}{2} \mathbf{B}^2 \delta_{ij} - B_i B_j \right] \quad (2.86)$$

Les conservations de l'énergie et de l'impulsion peuvent être ré-exprimées en terme du seul tenseur énergie-impulsion en une loi de conservation covariante locale. Énoncée de manière simple, celle-ci nous dit qu'un système physique isolé conserve sa quadri-impulsion. S'il est décrit par un tenseur énergie-impulsion cela aboutit à la loi $\partial_\mu T^{\mu\nu} = 0$. Si le système n'est plus isolé mais qu'un champ extérieur est présent, la densité de quadri-force \mathcal{F}^μ associée doit être prise en compte dans le membre de droite.

La conservation de l'énergie-impulsion pour un système isolé se traduit par l'équation $\partial_\mu T_{\text{mat}}^{\mu\nu} = 0$. En présence d'un champ extérieur \mathcal{F}^μ , on a la généralisation

$$\partial_\mu T_{\text{mat}}^{\mu\nu} = \mathcal{F}^\nu \quad (2.87)$$

Considérons un système de particules chargées décrit par un tenseur énergie impulsion $T_{\text{mat}}^{\mu\nu}$. Étant chargées, les particules interagissent entre elles par le force de Lorentz. En densité, celle-ci s'écrit $\mathcal{F}^\mu = F^{\mu\alpha} J_\alpha$. En utilisant les équations de Maxwell, on peut montrer que cette densité peut s'écrire comme la divergence d'un tenseur symétrique $T_{\text{em}}^{\mu\nu}$:

$$\partial_\mu T_{\text{em}}^{\mu\nu} = -F^{\nu\alpha} J_\alpha \quad (2.88)$$

où $T_{\text{em}}^{\mu\nu}$, appelé *tenseur énergie impulsion du champ électromagnétique*, s'écrit

$$T_{\mu\nu}^{\text{em}} = \frac{1}{\mu_0} \left[F_{\mu\beta} F_\nu^\beta - \frac{1}{4} g_{\mu\nu} F_{\alpha\beta} F^{\alpha\beta} \right]. \quad (2.89)$$

Avec cette définition, à partir de l'égalité 2.87, on obtient l'équation $\partial_\mu (T_{\text{mat}}^{\mu\nu} + T_{\text{em}}^{\mu\nu}) = 0$. Plus explicitement, le tenseur énergie-impulsion de la matière chargée n'est pas conservé sauf si on tient compte de la force électromagnétique qui dérive de l'objet $T_{\mu\nu}^{\text{em}}$. C'est la somme des deux tenseurs qui obéit à une loi de conservation et qui permet bien de considérer que $T_{\mu\nu}^{\text{em}}$ est l'énergie-impulsion associé au champ électromagnétique.

De manière explicite, le tenseur énergie-impulsion s'écrit

$$T^{\mu\nu} = \begin{pmatrix} \frac{1}{2} \left(\epsilon_0 E^2 + \frac{1}{\mu_0} B^2 \right) & \Pi_x/c & \Pi_y/c & \Pi_z/c \\ \Pi_x/c & -\tau_{xx} & -\tau_{xy} & -\tau_{xz} \\ \Pi_y/c & -\tau_{yx} & -\tau_{yy} & -\tau_{yz} \\ \Pi_z/c & -\tau_{zx} & -\tau_{zy} & -\tau_{zz} \end{pmatrix} \quad (2.90)$$

2.5 Formulation lagrangienne de l'électrodynamique

L'électromagnétisme est la plus simple des théories des champs dans le cadre de la théorie de la relativité. Le système que l'on considère est un ensemble de particules en interaction avec un champ qui sera représenté par un champ de quadri-vecteurs. *A priori*, l'action peut se décomposer en trois parties : l'une correspond à l'action d'une particule libre, une autre au champ libre et une troisième qui décrit l'interaction entre les particules et le champ.

On va procéder en deux étapes. La première consiste à considérer des particules libres dans un champ imposé. On doit retrouver les équations du mouvement et les équations de Maxwell homogènes. La deuxième sera de considérer la dynamique des particules donnée et les champs comme des inconnues. Ceci nous permettra de retrouver les équations de Maxwell avec les sources.



2.5.1 Particules libres dans un champ imposé

Le but de ce paragraphe est d'exhiber un lagrangien tel que les équations d'Euler-Lagrange, ou de manière équivalente le Principe de Moindre Action, nous redonnent l'équation de la dynamique mais aussi les équations de Maxwell homogènes. On peut, sans rien supposer de connu, postuler le lagrangien de la théorie (ou le trouver par des arguments heuristiques...) et retrouver toute la théorie de l'électromagnétisme.

Remarques préliminaires

En relativité restreinte, on écrit l'action S de la façon suivante

$$S = \int_a^b L(q_i, \dot{q}_i, t) dt = \int_a^b \Lambda \left(x^\mu, \frac{dx^\mu}{d\tau} \right) d\tau \quad (2.91)$$

En mécanique classique, l'action est l'intégrale du lagrangien sur le temps. Or, ceci n'a aucune raison d'être dans le cadre de la théorie de la relativité restreinte, où le temps n'a plus de rôle particulier. De plus, le quadri-vecteur vitesse a la particularité d'être de norme constante égale à c^2 . En relativité restreinte, une courbe n'est donc plus décrite par le paramètre temporel t mais par un invariant de Lorentz, τ ⁸. On impose de plus que l'action soit un invariant de Lorentz pour que les équations gardent un sens lors d'un changement de référentiel. Ainsi, $\Lambda \left(x^\mu, \frac{dx^\mu}{d\tau} \right)$ est aussi un invariant de Lorentz.

En tenant compte de toutes ces contraintes, le Principe de Moindre Action nous donne les équations de la dynamique du système, appelées équations d'Euler-Lagrange,

$$\frac{d}{d\tau} \left(\frac{\partial \Lambda}{\partial \dot{x}^\mu} \right) - \frac{\partial \Lambda}{\partial x^\mu} = 0 \quad (2.92)$$

Lagrangien et équations d'Euler-Lagrange

L'action d'une particule libre s'écrit simplement

$$S_{\text{Particules Libres}} = -mc \int_a^b ds = -mc^2 \int_a^b d\tau \quad (2.93)$$

où a et b sont deux événements. Le lagrangien associé est donc

$$L_{\text{Particules Libres}} = -\frac{mc^2}{\gamma}$$

$$\Lambda_{\text{Particules Libres}} = -mc \sqrt{\frac{dx^\alpha}{d\tau} \frac{dx_\alpha}{d\tau}}. \quad (2.94)$$

On postule maintenant l'action associée à l'interaction entre les particules et le champ $S_{\text{Interaction}}$. On suppose le champ décrit par le quadri-potentiel $A^\mu = \left(\frac{\Phi}{c}, \mathbf{A} \right)$. L'invariant de Lorentz scalaire le plus simple que l'on puisse trouver est

$$S_{\text{Interaction}} = -q \int_a^b A^\mu dx_\mu \quad (2.95)$$

dont le lagrangien associé est

$$\Lambda_{\text{Interaction}} = -qA^\mu \frac{dx_\mu}{d\tau} = -qA^\mu U_\mu \quad (2.96)$$

Finalement, le lagrangien total que l'on considère ici, c'est-à-dire pour un système de particules libres dans un champ imposé, est donné par

⁸. τ est le plus souvent pris égal au temps propre, mais pas toujours puisque celui-ci ne permet pas de décrire une particule sans masse.



Densité lagrangienne et équations du champ

Voyons maintenant comment s'écrivent les équations de Lagrange dans le cas où nous avons une densité de lagrangien. On considère donc un champ $u(x, y, z, t)$. On peut définir une densité de lagrangien $\mathcal{L}(u, \partial_x u, \partial_y u, \partial_z u, \partial_t u)$ et un lagrangien $L = \int \mathcal{L}(u, \partial_x u, \partial_y u, \partial_z u, \partial_t u) dx dy dz$. Pour déterminer les équations du mouvement, on applique le principe de moindre action.

$$\delta S = 0 = \int \left[\frac{\partial \mathcal{L}}{\partial u} \delta u + \frac{\partial \mathcal{L}}{\partial (\partial_t u)} \delta (\partial_t u) + \frac{\partial \mathcal{L}}{\partial \partial_x u} \delta (\partial_x u) + \dots \right] dx dy dz dt \quad (2.103)$$

Nous devons effectuer une intégration par partie pour isoler les variations $\delta_i u$. Par exemple, nous avons

$$\int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial (\partial_t u)} \delta (\partial_t u) dt = \underbrace{\left[\frac{\partial \mathcal{L}}{\partial (\partial_t u)} \delta u \right]_{t_1}^{t_2}}_{=0} - \int_{t_1}^{t_2} \frac{\partial}{\partial t} \left(\frac{\partial \mathcal{L}}{\partial (\partial_t u)} \right) \delta u dt \quad (2.104)$$

La variation de l'action s'écrit finalement comme

$$\delta S = \int \left[\frac{\partial \mathcal{L}}{\partial u} - \frac{\partial}{\partial t} \left(\frac{\partial \mathcal{L}}{\partial (\partial_t u)} \right) - \frac{\partial}{\partial x} \left(\frac{\partial \mathcal{L}}{\partial (\partial_x u)} \right) - \dots \right] \delta u dx dy dz dt = 0 \quad (2.105)$$

Cette équation doit être vérifiée quelle que soit la variation δu , ce qui implique que l'intégrande doit être nul. Nous obtenons au final l'équation d'Euler-Lagrange

$$\frac{\partial \mathcal{L}}{\partial u} - \frac{\partial}{\partial t} \left(\frac{\partial \mathcal{L}}{\partial (\partial_t u)} \right) - \frac{\partial}{\partial x_i} \left(\frac{\partial \mathcal{L}}{\partial (\partial_{x_i} u)} \right) = 0 \quad (2.106)$$

En utilisant les notations quadri-dimensionnelles, nous pouvons réécrire simplement cette équation de la manière suivante

$$\frac{\partial}{\partial x^\mu} \left(\frac{\partial \mathcal{L}}{\partial (\frac{\partial u}{\partial x^\mu})} \right) - \frac{\partial \mathcal{L}}{\partial u} = 0 \quad (2.107)$$

Ces équations peuvent être généralisées au cas où le champ est vectoriel. Pour un système à N particules, on avait un système de N équations différentielles. Pour un champ, on a une unique équation aux dérivées partielles et non pas un système d'équations différentielles ordinaires.

Lagrangien du champ électromagnétique

On cherche à déterminer le lagrangien du champ électromagnétique. Pour cela, on rappelle que puisque l'action est un invariant de Lorentz, tout comme l'élément d'intégration, la densité de Lagrangien doit être un invariant de Lorentz. Elle doit être de plus invariante de jauge; seuls les champs doivent apparaître. Si l'on cherche à obtenir des équations linéaires, nous devons avoir une dépendance quadratique en champ. Or, nous disposons de deux invariants satisfaisant ces exigences. L'un d'eux, qui correspond au produit scalaire des champs ne peut pas convenir puisque, entre autres, on ne peut pas décrire un problème d'électrostatique. Finalement, on postule la densité de lagrangien du champ électromagnétique

$$\mathcal{L}_{\text{Champ Libre}} = -\frac{1}{4\mu_0} F^{\mu\nu} F_{\mu\nu} \quad (2.108)$$

En utilisant les équations d'Euler-Lagrange avec cette densité, nous retrouvons les équations de Maxwell sans sources.

Enfin, si l'on fait la somme de cette densité lagrangienne du champ libre et de celle correspondant à l'interaction, on obtient la nouvelle densité lagrangienne

$$\mathcal{L}_{\text{Champ} + \text{Interaction}} = -\frac{1}{4\mu_0} F^{\mu\nu} F_{\mu\nu} - A_\mu J^\mu \quad (2.109)$$

Cette nouvelle densité redonne alors les équations de Maxwell avec sources lorsque l'on utilise les équations d'Euler-Lagrange, à savoir

$$\partial_\mu F^{\mu\nu} = \mu_0 J^\mu \quad (2.110)$$



2.5.3 Résumé

Formulation lagrangienne de l'électrodynamique

Lagrangien d'une particule libre

$$\Lambda_L \left(x^\mu, \frac{dx^\mu}{d\tau} \right) = -mc \sqrt{g_{\alpha\beta} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau}} \quad (2.111a)$$

$$= -mc \sqrt{\frac{dx^\alpha}{d\tau} \frac{dx_\alpha}{d\tau}} \quad (2.111b)$$

Lagrangien et densité lagrangienne d'interaction

$$L_I \left(x^\mu, \frac{dx^\mu}{d\tau} \right) = -q A^\mu \frac{dx_\mu}{d\tau} \quad (2.112a)$$

$$\mathcal{L}_I \left(x^\mu, \frac{dx^\mu}{d\tau} \right) = -\mathcal{A}_\mu J^\mu \quad (2.112b)$$

Densité lagrangienne du champ électromagnétique

$$\mathcal{L}_{EM} \left(x^\mu, \frac{dx^\mu}{d\tau} \right) = -\frac{1}{4\mu_0} F^{\mu\nu} F_{\mu\nu} \quad (2.113)$$

2.5.4 Lois de conservation

Théorème de Noether

Le pendant en théorie des champs du théorème de Noether en mécanique classique est que l'on effectue des transformations infinitésimales sur les champs, que le rôle de dt est joué par d^4x et que si l'on est capable de trouver $\phi^\alpha(\psi, x)$ tel que

$$\left(\frac{d\mathcal{L}}{d\epsilon} \right)_{\epsilon=0} = \partial_\alpha \phi^\alpha \quad (2.114)$$

alors il existe un courant G^α ¹⁰ dont la quadri-divergence est nulle, $\partial_\alpha G^\alpha = 0$, avec

$$G^\alpha = \frac{\partial \mathcal{L}}{\partial (\partial_\alpha \psi_l)} \delta \psi_l - \phi^\alpha \quad (2.115)$$

Invariance de jauge et conservation de la charge

On rappelle que la théorie de l'électromagnétisme est une théorie invariante de jauge. Voyons quelle en est la conséquence. Pour cela, on va appliquer le théorème de Noether. Tout d'abord, on rappelle l'action de la théorie ainsi que la dépendance en fonction de la jauge de chacun des termes.

$$S = \underbrace{S_{\text{Part.}}}_{\text{Indépendant de la jauge}} + \int \left(\underbrace{-\frac{1}{4\mu_0} F^{\mu\nu} F_{\mu\nu}}_{\text{Indépendant de la jauge}} \underbrace{-\mathcal{A}_\mu J^\mu}_{\text{Dépendant de la jauge}} \right) d^4x \quad (2.116)$$

Considérons alors la transformation de jauge suivante

$$\frac{\Phi}{c} \rightarrow \frac{\Phi}{c} - \partial_t \Lambda \quad (2.117)$$

$$\mathbf{A} \rightarrow \mathbf{A} + \nabla \Lambda \quad (2.118)$$

¹⁰. *A priori*, ce courant n'a rien à voir avec le quadri-vecteur courant.



Avec cette transformation, le terme d'interaction se réécrit comme

$$-\mathcal{A}_\mu J^\mu \rightarrow -\mathcal{A}_\mu J^\mu - \Lambda \partial_\mu J^\mu + \underbrace{\partial_\mu (\Lambda J^\mu)}_{\substack{\text{Divergence :} \\ \text{Ne contribue pas}}} \quad (2.119)$$

On montre que le courant conservé de Noether, noté G^α , vaut

$$G^\alpha = J^\alpha \quad (2.120)$$

Ainsi, d'après le théorème de Noether, la conservation de la charge est une conséquence de l'invariance de jauge de la théorie.



Chapitre 3

Mécanique quantique

Sommaire

3.1	Approche de Feynman	44
3.2	Postulats de la mécanique quantique	47
3.3	Propriétés élémentaires	52
3.4	Théorie élémentaire de la diffusion	60
3.5	Perturbation dépendante du temps	65
3.6	Opérateur densité	69
3.7	Opérateurs de champs	73
3.8	Généralités sur les symétries	83
3.9	Rotations et moments angulaires	87

Le but de ce chapitre sera, après avoir redonné les postulats et le formalisme de base de la mécanique quantique, d'étudier différents phénomènes de mécanique quantique, en lien avec divers autres chapitres, notamment en physique nucléaire et en physique des particules. Le premier phénomène étudié sera la réflexion ou la transmission d'une onde sur un potentiel, qui se terminera par une courte étude de l'effet tunnel. Ensuite, l'étude de la diffusion, c'est-à-dire de l'envoi d'un flux de particules incidentes sur des centres diffuseurs, sera effectuée. Puis nous étudierons la théorie des perturbations dépendantes du temps, qui est la base du formalisme des diagrammes de Feynman. Après ces parties en lien avec la physique des particules, nous étendrons un peu plus le formalisme et l'étude de la mécanique quantique en introduisant la matrice densité, qui est la description réelle d'un système quantique, puisqu'elle permet notamment d'expliquer le phénomène d'intrication. Puis nous introduirons les opérateurs de champs, qui sont nécessaires pour la description d'un système à nombre de particules indéterminé. Enfin, nous nous pencherons sur les symétries, qui sont la base d'une description des interactions fondamentales.

Les références bibliographiques que nous avons utilisées pour cette partie sont [14], [5], [7] et [16].

Comme précédemment, la section de révisions peut être passée si l'on est déjà familier avec tout le formalisme de la mécanique quantique.

3.1 Approche de Feynman

Une très grande partie de l'étrangeté et de la richesse des phénomènes quantiques proviennent d'effets d'interférences analogues aux interférences lumineuses¹. C'est à partir de ce constat que l'approche de Feynman de la mécanique quantique s'est développée, comme nous allons le voir ici. Cela nous permettra de définir les amplitudes de probabilité de transition, qui sont au cœur de la mécanique quantique, ainsi que la notion d'intégrale de chemin².

3.1.1 Amplitudes de probabilités

Intéressons-nous à une expérience particulièrement simple, les fentes de Young. Une source S envoie des particules à travers un mur percé de deux trous 1 et 2 ; ce mur est considéré comme classique, c'est-à-dire que

1. La deuxième grande cause de l'étrangeté du quantique est liée au phénomène d'intrication.
2. Qui resservira notamment dans la plupart des cours de M2... et pas uniquement !



les seules trajectoires autorisées pour les particules sont celles passant par l'un des trous. Les particules sont ensuite détectées par un dispositif \mathcal{D} placé à une position donnée.

L'information expérimentale que l'on mesure est de nature probabiliste ; on cherche la probabilité que la particule arrive sur le détecteur. Il s'avère que cette probabilité n'a pas l'allure de l'intuition classique que l'on pourrait avoir : on observe une figure d'interférence.

La description de ce type de figure en optique se fait à l'aide des champs électrique et magnétique. Pour trouver l'intensité de la figure d'interférence, on somme tous les champs pertinents puis on évalue le module au carré. La mécanique quantique se base sur une description similaire. À chaque chemin γ que la particule est en mesure de suivre, on associe un nombre complexe appelé *amplitude de probabilité du chemin* γ ³, $\mathcal{A}[\gamma] \in \mathbb{C}$. Il faut bien comprendre que ce chemin est un chemin dans l'espace-temps. Le module au carré de cette amplitude nous donne *la probabilité que la particule suive effectivement ce chemin*.

Lorsqu'on s'intéresse à l'évolution à partir d'un état initial i à un état final f , nous devons sommer sur tous les chemins reliant ces deux états. La probabilité de transition entre ces états est toujours donnée par le module au carré de l'amplitude totale, qui est donc le module au carré de la somme sur les amplitudes de tous les chemins.

En guise d'exemple d'application, revenons à l'expérience des fentes d'Young. En notant \mathcal{A}_1 l'amplitude de probabilité associée au chemin qui va de la source vers le détecteur en passant par le trou 1 et \mathcal{A}_2 celle en passant par le trou 2, la probabilité I_D que la particule soit détectée par \mathcal{D} est

$$I_D = |\mathcal{A}_1 + \mathcal{A}_2|^2 = |\mathcal{A}_1|^2 + |\mathcal{A}_2|^2 + \mathcal{A}_1^* \mathcal{A}_2 + \mathcal{A}_2^* \mathcal{A}_1 \tag{3.1}$$

Les deux premiers termes du membre de droites sont les probabilités que la particule passe respectivement à travers le trou 1 et le trou 2. Les deux derniers sont d'origine purement quantique et sont la signature des interférences entre les différentes trajectoires.

Une idée naturelle est de chercher à savoir par quel trou la particule passe. Si une telle expérience est faite, par exemple en plaçant un détecteur à un des trous, le résultat est tout autre : aucune interférences n'apparaît ! La détection du chemin d'une manière ou d'une autre détruit les interférences quantiques. Ainsi, dans l'expérience des fentes d'Young exposée plus haut, il nous est impossible de dire ni même de savoir par un quelconque procédé si la particule passe par l'un *ou* par l'autre des trous sans changer la physique de l'expérience. Formellement, si l'on sait détecter par quel chemin passe la particule, on ne doit plus sommer l'amplitude de probabilité des alternatives mais les probabilités elles-mêmes. Nous aurons alors

$$I_D = |\mathcal{A}_1|^2 + |\mathcal{A}_2|^2 \tag{3.2}$$

Cette forte influence de la mesure en mécanique quantique contraste avec notre expérience quotidienne où il est possible d'effectuer des mesures sans pour autant affecter le système.

3.1.2 Intégrale de chemin

Les amplitudes \mathcal{A}_1 et \mathcal{A}_2 sont en fait obtenues en faisant la somme sur tous les chemins γ_1 et γ_2 passant respectivement à travers le trou 1 et 2. Elles s'expriment sous forme d'intégrales fonctionnelle appelées *intégrales de chemins*. Cette formulation de la mécanique quantique fut développée par Feynman au cours de sa thèse et un exposé indirect est fait dans *la Bible*. Nous pouvons donc écrire

$$\mathcal{A}_1 = \sum_{\gamma_1} \mathcal{A}[\gamma_1] \quad \mathcal{A}_2 = \sum_{\gamma_2} \mathcal{A}[\gamma_2] \tag{3.3}$$

La notation traditionnelle n'est en revanche pas celle-ci. L'amplitude \mathcal{A} de transition entre un état initial i et un état final f s'écrit comme l'intégrale sur tous les chemins reliant ces deux états,

$$\mathcal{A} = \int_{\gamma_{if}} \mathcal{A}[\gamma_{if}] \mathcal{D}\gamma \tag{3.4}$$

Tout le problème est maintenant de savoir que vaut $\mathcal{A}[\gamma]$ pour un chemin donné. Après cela, nous aurons à disposition tout ce qu'il faut pour pouvoir décrire le monde quantique. Mais les résultats obtenus jusqu'ici ne nous permettent pas d'inférer la forme de ces amplitudes de probabilité.

3. On parlera également d'amplitude de transition



Une des propriétés importantes des interactions en physique est leur localité (TROUVER UNE BONNE FAÇON DE L'EXPLIQUER). En termes des amplitudes de probabilité, si on se donne un chemin γ_{12} reliant les état 1 et 2 et un chemin γ_{23} , alors l'amplitude de probabilité du chemin concaténé γ_{13} est donnée par

$$\mathcal{A}[\gamma_{13}] = \mathcal{A}[\gamma_{23}] \mathcal{A}[\gamma_{12}] \tag{3.5}$$

Cette propriété permet de déterminer l'amplitude associé à un chemin à partir de la connaissance de l'amplitude au niveau infinitésimal. Enfin, Feynman postule que l'amplitude de probabilité associée au chemin γ est reliée à l'action S calculée le long de cette trajectoire⁴

$$\mathcal{A}[\gamma] \propto e^{\frac{i}{\hbar} S[\gamma]} \tag{3.6}$$

où \hbar est la *constante de Planck* normalisée et vaut $1,05457173 \cdot 10^{-34} \text{ m}^2\text{kgs}^{-1}$. Le facteur de proportionalité est appelé *phase topologique* puisqu'il ne dépend pas vraiment du chemin suivi ; il est identique pour deux chemins déformable l'un en l'autre de manière continue⁵. Au final, l'amplitude de probabilité \mathcal{A} vaut, en introduisant \mathcal{N} comme facteur de normalisation,

$$\mathcal{A} = \mathcal{N} \int_{if} e^{\frac{i}{\hbar} S[\gamma_{if}]} \mathcal{D}\gamma \tag{3.7}$$

Le point de vue de Feynman de la mécanique quantique est réminiscent du lien entre l'optique géométrique et l'optique ondulatoire. En optique géométrique, la lumière est décrite par un rayon obéissant au principe de Fermat. Ce dernier est un principe variationnel stipulant que la trajectoire des rayons optimise le chemin optique. Notons que nous raisonnons ici dans l'espace, le chemin optique étant *grosso modo* égal à la distance parcourue. Cette description n'est valable que lorsque la longueur d'onde est petite devant les dimensions caractéristiques du système. Si cela n'est pas le cas, elle est supplantée par une description ondulatoire. Celle-ci repose sur le principe de Huygens qui énonce que l'amplitude de l'onde reçue est une somme de contributions issues de la source. Ce principe contient celui de Fermat lorsque la longueur d'onde tend vers zéro.

La description de Feynman repose sur les mêmes idées. La mécanique classique est basée sur un principe variationnel (de Hamilton) : la trajectoire suivie par une particule est un extremum de l'action, intégrale du Lagrangien. Rappelons que cette fois on raisonne dans l'espace-temps. Une description quantique repose en revanche sur une somme de contributions, une pour chaque chemin reliant les états initiaux et finals. À la limite où \hbar tend vers 0, on peut retrouver la description classique en terme d'extremum d'action. En un sens, une expérience de mécanique quantique est une expérience de diffraction dans l'espace-temps !

3.1.3 Fonction d'onde

Jusqu'à présent, l'objet de base que nous avons considéré (l'amplitude de probabilité) était associé à une histoire particulière de la particule, c'est-à-dire une trajectoire donnée de l'espace-temps. Le plus souvent, cependant, nous n'avons pas besoin d'une information aussi particulière ; on s'intéresse surtout à l'amplitude de probabilité que la particule soit à l'instant t à la position x sans se soucier de son histoire, c'est-à-dire sans se préoccuper de la manière dont elle y parvient. Une telle amplitude est appelée *fonction d'onde* et est notée $\psi(x, t)$. Son interprétation est toujours la même, à savoir que son module au carré nous donne la probabilité que la particule soit en x à l'instant t . Pour être plus précis, la fonction d'onde représente une densité de probabilité de présence : l'intégrale sur un volume donné de son module au carré nous donne la probabilité que la particule se trouve dans ce volume à un instant donné.

De part sa définition, la fonction d'onde obéit à une équation simple. Pour plus de clarté, on notera l'amplitude de transition $\langle x', t' | x, t \rangle$, auparavant notée \mathcal{A} . La fonction d'onde obéit alors à l'équation intégrale

$$\psi(\mathbf{r}, t) = \int_{\mathbb{R}^3} \langle \mathbf{r}, t | \mathbf{r}', t' \rangle \psi(\mathbf{r}', t') d\mathbf{r}' \tag{3.8}$$

L'intégrande représente l'amplitude de probabilité que la particule suive un chemin reliant (\mathbf{r}', t') à (\mathbf{r}, t) : $\psi(\mathbf{r}', t')$ représente l'amplitude que la particule se trouve à l'instant t' en \mathbf{r}' tandis que $\langle \mathbf{r}, t | \mathbf{r}', t' \rangle$ est associée à la transition vers l'état final. Par définition de la fonction d'onde, l'histoire de la particule ne nous intéresse pas, d'où la présence de l'intégrale qui nous permet de tenir compte de tous les chemins pertinents.

4. Cette idée provient de l'analyse approfondie de la mécanique quantique formulée par Schrödinger, Heisenberg et Dirac.

5. En terme mathématiques, la phase topologique ne dépend que de la classe d'homotopie des chemins.



On voit au final que la connaissance de la fonction d'onde à un instant donné nous permet de déterminer l'évolution future de la particule grâce aux amplitudes de transition calculables *via* les postulats de Feynman. Ces derniers nous permettent d'avoir un nouveau point de vue en permettant de dériver une équation différentielle du premier ordre pour la fonction d'onde, *l'équation de Schrödinger*. En fait, on glisse de point de vue en passant d'une représentation lagrangienne de la mécanique quantique au point de vue traditionnel hamiltonien que nous allons explorer plus en détails dans la suite. Un dernier point que l'on peut enfin montrer, fort utile pour les calculs, est que l'amplitude de transition $\langle \mathbf{r}, t | \mathbf{r}', t' \rangle$ est en fait une fonction de Green de l'équation de Schrödinger ! Ceci se reflète d'ailleurs dans l'équation intégrale 3.8 où l'amplitude apparaît comme un noyau intégral.

3.2 Postulats de la mécanique quantique

Après cette première partie un peu éloignée de la présentation "standard" de la mécanique quantique, nous allons ici redonner les postulats de la mécanique quantique dans leur énoncé traditionnel. Ils nous donnent accès à la formulation hamiltonienne de la théorie, qui est plus formelle que le point de vue lagrangien de Feynman.

3.2.1 États quantiques et règle de Born

Un système quantique est décrit à l'instant t par un élément $|\psi(t)\rangle$ appelé *ket* d'un espace de Hilbert complexe \mathcal{H} , aussi appelé *espace des états*.

Le principe de superposition est contenu dans la structure de l'espace des états, qui est un espace vectoriel. Pour être plus précis, un état du système est donné par *un rayon* de l'espace des états. En fait, deux vecteurs de l'espace de Hilbert proportionnels décrivent *le même* état physique; $|\psi\rangle$ et $\lambda|\psi\rangle$ avec $\lambda \neq 0$ décrivent le même état physique (si les vecteurs sont normés, comme c'est souvent le cas, ils diffèrent juste d'une phase). De manière intuitive, tous ces vecteurs sont dits "égaux" et définissent un même objet que l'on nomme rayon ⁶

Interprétation : règle de Born

Le lien avec la physique du système est fait grâce à l'interprétation probabiliste de la mécanique quantique faite par Born, à savoir

La norme au carré de l'état du système $\| |\psi\rangle \|^2$ est la probabilité que le système se trouve dans l'état $|\psi\rangle$.

Une chose est essentielle à comprendre : l'interprétation statistique que l'on fait n'est pas due à un manque d'information sur le système comme en mécanique statistique, mais est au contraire inhérente au caractère quantique du système.

3.2.2 Équation de Schrödinger

Postulat de la dynamique

La dynamique de l'état $|\psi(t)\rangle$ est donnée par l'équation de Schrödinger

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H(t) |\psi(t)\rangle \tag{3.10}$$

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

L'équation de Schrödinger est une équation différentielle du premier ordre (un système différentiel du premier ordre pour être plus correct). L'évolution de l'état d'un système est donc **parfaitement déterministe**.

⁶. Mathématiquement, un rayon est une classe d'équivalence de la relation

$$|\psi\rangle \sim |\phi\rangle \Leftrightarrow \exists \lambda \in \mathbb{C}^*, |\phi\rangle = \lambda |\psi\rangle \tag{3.9}$$

Un état est décrit par un élément de l'espace projectif de Hilbert $P(\mathcal{H})$. C'est cette nuance qui nous permet, par exemple, de représenter l'état d'un qubit, dont l'espace des états est le plus simple possible, à savoir un espace de Hilbert de dimension deux, par un vecteur de la sphère de Bloch (aussi appelé sphère de Riemann).



Conservation des probabilités

L'équation de Schrödinger implique une équation de conservation similaire à l'équation de conservation de la charge en électromagnétisme. Cette équation représente la conservation des probabilités. En posant $\rho(\mathbf{r}, t) = \psi^*(\mathbf{r}, t)\psi(\mathbf{r}, t)$ la densité de probabilité et $\mathbf{j}(\mathbf{r}, t) = \frac{\hbar}{2m}\text{Im}(\psi^*(\mathbf{r}, t)\nabla\psi(\mathbf{r}, t))$ le courant de probabilité, on montre à partir de l'équation de Schrödinger que nous avons l'équation de conservation suivante

$$\frac{\partial\rho(\mathbf{r}, t)}{\partial t} + \nabla \cdot \mathbf{j}(\mathbf{r}, t) = 0 \tag{3.11}$$

C'est en fait l'existence d'une telle équation de conservation qui permet de donner du sens à l'interprétation probabiliste de $\rho(\mathbf{r}, t)$ et $\mathbf{j}(\mathbf{r}, t)$.

Opérateur d'évolution

L'équation de Schrödinger est une équation différentielle. Cependant, il est possible de donner une forme intégrale du postulat de la dynamique en utilisant l'opérateur d'évolution.

La dynamique de l'état $|\psi(t)\rangle$ au temps t se déduit de l'état $|\psi(t_0)\rangle$ au temps t_0 par l'action de l'opérateur d'évolution $U(t, t_0)$, opérateur unitaire, comme $|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle$. En outre, l'opérateur d'évolution vérifie l'équation

$$i\hbar \frac{d}{dt}U(t, t_0) = H(t)U(t, t_0) \tag{3.12}$$

$$U(t, t_0) = \mathbb{1} \tag{3.13}$$

L'opérateur évolution est l'objet qui se rapproche le plus de la formulation par intégrale de chemin. En effet, la représentation position de cet opérateur n'est rien d'autre que l'amplitude de probabilité de Feynman associée à l'évolution du système entre les état \mathbf{r}_0, t_0 et \mathbf{r}, t . L'opérateur d'évolution possède la propriété de semi-groupe :

$$U(t_3, t_2)U(t_2, t_1) = U(t_3, t_1) \tag{3.14a}$$

$$U(t_0, t) = U^{-1}(t, t_0) \tag{3.14b}$$

La première égalité caractérise le fait que les interactions sont locales! Notons que ceci n'est pas une spécificité quantique mais du type d'équation différentielle qu'est l'équation de Schrödinger⁷. C'est cette propriété qui nous permet de retomber, à partir de la formulation opératorielle de la mécanique quantique, sur l'intégrale de chemin et la formulation de Feynman (voir annexe sur l'intégrale fonctionnelle pour un exemple en mécanique statistique A.2).

La solution de l'équation 3.12 n'est pas simple dans le cas général. Cependant, dans le cas où le hamiltonien est indépendant du temps, nous avons la solution

$$U(t, t_0) = e^{-\frac{i}{\hbar}(t-t_0)H} \tag{3.15}$$

3.2.3 La mesure

Jusqu'alors, nous n'avons pas discuté de la manière dont s'effectuent les mesures en mécanique quantique. Une caractéristique singulière de la mesure quantique est qu'elle influence et perturbe le système, contrairement au cas classique où il est possible de mesurer sans perturber. Les postulats sur la mesure nous permettent de dire comment on décrit une grandeur physique, quels sont les résultats possibles, avec quelle probabilité on peut les obtenir et enfin comment l'état du système est modifié par le processus de mesure.

⁷. C'est en fait une conséquence du théorème de Cauchy puisqu'il suffit de montrer que les deux côtés de l'égalité vérifient le même problème de Cauchy.

Une grandeur physique est décrite par un opérateur hermitien M de l'espace des états \mathcal{H} , dont les valeurs propres sont les résultats possibles de mesure. Si le système physique est dans l'état normé $|\psi\rangle$, la probabilité (ou la densité de probabilité pour un spectre continu) d'obtenir le résultat $m \in \text{sp}(M)$ est donnée par la norme de la projection de $|\psi\rangle$ sur le sous-espace propre de m . Juste après une telle mesure, en notant P le projecteur sur le sous-espace propre de m , l'état du système se réduit à

$$|\psi(t^+)\rangle = \frac{P|\psi(t^-)\rangle}{\sqrt{\langle\psi(t^-)|P|\psi(t^-)\rangle}} \tag{3.16}$$

Il convient ici de faire quelques remarques. Le caractère hermitien des opérateurs assure que les résultats d'une mesure (les valeurs propres) sont bien réels. Dans le cas particulier d'un spectre discret, les résultats sont même quantifiés!

La dernière partie du postulat est souvent appelé *postulat de réduction du paquet d'onde*. Il rend l'évolution d'un système quantique soumis à une mesure non unitaire! Cette non-unitarité contraste avec l'équation de Schrödinger, où l'évolution est unitaire.

Spectre discret

Supposons le spectre de A discret. On note $\text{sp}(A) = \{a_n, n \in \mathbb{Z}\}$. Pour $n \in \mathbb{Z}$, la valeur propre a_n peut être dégénérée. On note $(e_{n_i})_{i \in [1, p_n]}$ une base du sous-espace propre associé à la valeur propre a_n , avec p_n la dimension de cet espace. La probabilité de trouver la valeur propre a_n est, d'après le postulat :

$$\mathcal{P}(a_n) = \|P_n|\psi\rangle\|^2 \tag{3.17}$$

où P_n est le projecteur orthogonal sur le sous-espace propre associé à la valeur propre a_n . Explicitons un peu mieux $\|P_n|\psi\rangle\|^2$ en utilisant les propriétés de P_n ($P_n^\dagger = P_n, P_n^2 = P_n$).

$$\|P_n|\psi\rangle\|^2 = \langle\psi|P_n^\dagger P_n|\psi\rangle = \langle\psi|P_n|\psi\rangle \tag{3.18}$$

Enfin, si on écrit $P_n = \sum_{k=1}^{p_n} \langle e_{n_k} | \cdot \rangle e_{n_k}$ ou en notation de Dirac $P_n = \sum_{k=1}^{p_n} |e_{n_k}\rangle\langle e_{n_k}|$ on obtient :

$$\mathcal{P}(a_n) = \sum_{k=1}^{p_n} |\langle e_{n_k} | \psi \rangle|^2 \tag{3.19}$$

Valeur moyenne d'une observable

Soit une observable A dont le spectre est supposé discret $\text{sp}(A) = \{a_n, n \in \mathbb{Z}\}$ (le résultat final sera aussi valable pour un spectre continu). Le système physique étant dans l'état $|\psi\rangle$, on définit la valeur moyenne de A dans cet état par :

$$\langle A \rangle_\psi = \sum_{n \in \mathbb{Z}} a_n \mathcal{P}(a_n) \tag{3.20}$$

où $\mathcal{P}(a_n)$ désigne la probabilité d'obtenir la valeur a_n lorsque l'on mesure A dans l'état $|\psi\rangle$. En utilisant la formule 3.18, et le fait que l'on dispose d'une base de l'espace, on obtient :

$$\langle A \rangle_\psi = \langle \psi | A | \psi \rangle \tag{3.21}$$

Interprétation

Le point de vue sur la mesure adopté ici, celui de Copenhague, fait une distinction nette entre le système quantique et l'observateur (l'appareil de mesure). L'observateur dans ce point de vue n'est pas soumis aux lois quantiques mais aux lois classiques. C'est une manière de justifier le fait que l'appareil de mesure n'est pas soumis à une évolution unitaire.

Si on veut garder le point de vue que la mécanique quantique décrit le monde microscopique, l'observateur doit être aussi décrit par ses lois. Une approche purement quantique pose plusieurs questions, en partie résolues par



la décohérence où l’environnement du système et de l’appareil de mesure, tous décrits par les lois quantiques, joue un rôle majeur. L’un des problèmes d’une description quantique du système est que rien n’interdit en principe que l’observateur et le système soient dans une superposition d’état. C’est le coeur du paradoxe du chat de Schrödinger. Or ceci n’est clairement pas observé pour des objets macroscopiques. L’idée de la décohérence et de l’approche de Zurek est de ne pas oublier que dans la vraie vie, le système observateur-système n’est pas isolé de son environnement. Sa présence permet de détruire les états “chats de Schrödinger” et de retrouver une description classique. Intuitivement, une partie de l’information sur l’état du système et de l’appareil de mesure fuit dans les degrés de libertés environnementaux qui ne sont pas observés : c’est cette information perdue, non observée, qui permet de retrouver (moyennant quelques subtilités) la réalité classique⁸. Notons que cette l’information n’est pas détruite, ceci n’est pas possible en mécanique quantique, mais juste non observée. En principe, il est possible de retrouver toute l’information sur l’état dans l’environnement et voir les “chats de Schrödinger”.

3.2.4 Schémas de Schrödinger et de Heisenberg

Ce sont deux représentations complètement équivalentes des états et des observables, mais qui diffèrent dans leur manière de décrire l’évolution dans le temps. Un troisième schéma, le schéma en interaction, sera discuté dans la section 3.5. On a déjà vu, dans le cas de l’équation de Schrödinger, que cette représentation suppose des états évoluant dans le temps, tandis que les observables n’évoluent pas dans temps. Le schéma de Heisenberg, lui, fait la supposition inverse. Dans le cas de Schrödinger, l’évolution des états dans le temps est décrite par un opérateur d’évolution $U(t, t_0)$, tel que

$$|\psi^S(t)\rangle = U(t, t_0)|\psi^S(t_0)\rangle \tag{3.22}$$

où, ici, la notation $|\psi^S(t)\rangle$ dénote l’état dans le schéma de Schrödinger. Pour Heisenberg, on utilisera $|\psi^H\rangle$. Pour passer du schéma de Schrödinger au schéma de Heisenberg, les relations à utiliser sont :

$$|\psi^H\rangle = U^\dagger(t, t_0)|\psi^S(t)\rangle = |\psi^S(t_0)\rangle \tag{3.23a}$$

$$A^H(t) = U^\dagger(t, t_0)A^S(t)U(t, t_0) \tag{3.23b}$$

L’équation d’évolution est alors

$$i\hbar \frac{dA^H(t)}{dt} = i\hbar \frac{\partial A^H(t)}{\partial t} + [A^H(t), H^H] \tag{3.24}$$

Une propriété à noter est que, si H^S ne dépend pas explicitement du temps, on a $H^S = H^H$.

3.2.5 Représentation, opérateurs position et impulsion, commutation

En mécanique classique, une particule est décrite par la donnée de sa position et de son impulsion. En mécanique quantique, tel n’est plus le cas. La position et l’impulsion sont maintenant des opérateurs qui agissent dans un espace de Hilbert. Dans la suite, nous allons chercher à déterminer comment ces opérateurs agissent. Pour cela, on se place dans le cas particulier de la représentation position où l’opérateur position est facilement identifié. Dans la représentation position que nous allons étudier, l’état $|\psi(t)\rangle$ est équivalent à la donnée d’une fonction $\psi(t, \mathbf{r}) \in \mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$ ⁹. L’état quantique n’est donc plus un objet abstrait, mais bien une fonction au sens classique. En fait l’espace $\mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$ est trop gros pour les besoins de la physique. Des propriétés de régularité, notamment à l’infini, sont en générale exigées. L’espace des états est donc un sous-espace de $\mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$.

⁸. Si on se rappelle l’expérience des fentes d’Young, c’est le simple fait de placer un détecteur sur l’un des bras qui détruit la figure d’interférence, indépendamment de si le résultat de détection est connu de l’expérimentateur ou non. Cette prise d’information sur l’état du système, bien que non observée, modifie tout de même son comportement physique.

⁹. La notation peut être ambiguë. Étant donné un instant t , l’état est la donnée d’une fonction de carré intégrable sur les coordonnées spatiales, à valeurs complexes. Mathématiquement, on écrirait plutôt :

$$\forall t, \psi_t \in \mathcal{L}^2(\mathbb{R}^3, \mathbb{C}) \tag{3.25}$$



Opérateur position

Soit $\psi(t, \mathbf{r}) \in \mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$ une fonction d'onde. L'action de l'opérateur position \mathbf{R} est la suivante

$$\mathbf{R}\psi(t, \mathbf{r}) = \mathbf{r} \cdot \psi(t, \mathbf{r}) \quad (3.26)$$

En représentation "classique", les vecteurs propres de l'opérateur position sont notés $|\mathbf{r}\rangle$, et on écrit la fonction d'onde comme

$$\psi(t, \mathbf{r}) = \langle \mathbf{r} | \psi(t) \rangle \quad (3.27)$$

Opérateur impulsion

Reste maintenant à déterminer l'action de l'opérateur impulsion en représentation position (on pourrait se placer en représentation impulsion, auquel cas l'action de l'opérateur serait de la même forme que ci-dessus). Considérons pour cela un système physique **isolé** caractérisé par la fonction d'onde $\psi(t, \mathbf{r})$. Ce système est donc invariant par translation spatiale. On considère dans la suite une translation infinitésimale $\delta\mathbf{r}$. Alors, au premier ordre, nous avons

$$\psi(t, \mathbf{r} + \delta\mathbf{r}) = (1 + \delta\mathbf{r} \cdot \nabla)\psi(t, \mathbf{r}) \quad (3.28)$$

L'invariance se traduit par le fait que l'opérateur ∇ commute avec l'opérateur hamiltonien du système. En d'autres termes nous avons une quantité conservée. Par analogie avec la mécanique classique où la quantité conservée associée à l'invariance par translation spatiale est l'impulsion totale du système, l'opérateur ∇ est associé, à une constante multiplicative près, à l'opérateur impulsion du système.

En représentation position, l'opérateur impulsion s'écrit

$$\mathbf{p} = -i\hbar\nabla \quad (3.29)$$

Relation de commutation

Un résultat majeur que l'on peut tirer de ces résultats est la relation de commutation des opérateurs position et impulsion. Un calcul simple avec une fonction d'onde dans le bon espace donne

$$[p_i, x_k] = -i\hbar\delta_{ik} \quad (3.30)$$

où p_i et x_k désignent les composantes des opérateurs impulsion et position.

3.2.6 Principe d'indétermination

Dans la section 3.2.3, nous avons défini la valeur moyenne d'une observable. Une information qui serait intéressante à obtenir est l'écart à la moyenne. On cherche à avoir une idée de la dispersion des résultats que l'on peut s'attendre à obtenir. L'écart quadratique moyen ΔA dans l'état $|\psi\rangle$ permet d'accéder à cette donnée. Par définition

$$\Delta A = \sqrt{\langle (A - \langle A \rangle)^2 \rangle} = \sqrt{\langle A^2 \rangle - \langle A \rangle^2} \quad (3.31)$$

Considérons deux observables \mathcal{A} et \mathcal{B} dont les opérateurs associés sont A et B . Alors on peut montrer que :

$$\Delta A \Delta B \geq \left| \frac{\langle [A, B] \rangle}{2} \right| \quad (3.32)$$

Ceci est l'énoncé exact du principe d'incertitude de Heisenberg, directement conséquence de la définition exacte de la variance (ou écart quadratique moyen) donnant l'information sur la dispersion des résultats que l'on peut mesurer. Par exemple, pour les opérateurs \mathbf{x} et \mathbf{p} , on a



$$\Delta \mathbf{x} \Delta \mathbf{p} \geq \frac{\hbar}{2} \tag{3.33}$$

En fait, ce résultat permet de définir des variables conjuguées : A et B sont dites conjuguées si $[A, B] = i\hbar$, ce qui donne alors $\Delta A \Delta B \geq \frac{\hbar}{2}$. Ce résultat est par exemple valable pour E et t .

3.3 Propriétés élémentaires

Deux points de vue ont été développés pour la mécanique quantique, celui de Feynman où la fonction d'onde est donnée de façon intégrale, le noyau étant calculé à partir des postulats de l'intégrale de chemin, et le point de vue de l'équation de Schrödinger. Ces deux points de vue sont en fait complémentaires : celui de Feynman est plus physique et adapté aux généralisations tandis que celui de Schrödinger est plus adapté pour les calculs. Ainsi, une manière intéressante de voir les choses est de décrire l'évolution d'un système du point de vue intégral tandis que les calculs sont fait *via* l'équation différentielle. Les amplitudes de probabilités de Feynman sont, comme nous l'avons dit sans le démontrer, des fonctions de Green de l'équation de Schrödinger.

Ici, nous allons nous intéresser au calcul de certaines propriétés qui sont des conséquences directe de la formulation quantique, tels que l'effet tunnel ou la forme des états liés dans un potentiel carré. Notons que le calcul des fonctions de Green est loin de toujours être facile et que, par conséquent, une résolution plus traditionnelle des équations est souvent plus directe (l'atome d'hydrogène par exemple). C'est pourquoi dans la suite nous ne regarderons que la fonction de Green de la particule libre pour ensuite revenir à des choses plus classiques.

3.3.1 Particule libre

Le problème à résoudre est le suivant. On cherche à déterminer la fonction d'onde d'une particule de masse m . Pour plus de simplicité, on travaillera à une dimension, la généralisation étant facile à obtenir. Nous devons donc trouver la fonction de Green que l'on notera $\mathcal{G}(x_2, t_2; x_1, t_1)$ solution de

$$i\hbar \frac{\partial}{\partial t_2} \mathcal{G}(x_2, t_2; x_1, t_1) + \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x_2^2} \mathcal{G}(x_2, t_2; x_1, t_1) = i\hbar \delta(x_2 - x_1) \delta(t_2 - t_1) \tag{3.34}$$

$$\mathcal{G}(x_2, t_2; x_1, t_1) = 0 \text{ si } t_2 < t_1 \tag{3.35}$$

On rappelle que la détermination d'une fonction de Green d'une équation ne peut être complète que lorsque des conditions aux limites sont imposées, ce que fait la seconde partie du système. Le plus simple pour résoudre est de passer dans le domaine de Fourier avec $\mathcal{G}(k, \omega)$, puisque la fonction de Green réelle ne dépend que de la différence des temps et des positions du fait de l'invariance du problème.

$$\left(\hbar\omega - \frac{\hbar^2 k^2}{2m} \right) \mathcal{G}(k, \omega) = 1 \tag{3.36}$$

Cette équation est à prendre *au sens des distributions*. Les solutions de l'équation de Fourier font intervenir une solution particulière, donnée par la distribution *valeur principale de Cauchy* $i\hbar \mathcal{P} \left(\frac{1}{\hbar\omega - \frac{\hbar^2 k^2}{2m}} \right)$ ainsi que les solutions de l'équation homogène donnée par la distribution de Dirac $i\hbar C \delta \left(\hbar\omega - \frac{\hbar^2 k^2}{2m} \right)$, où C est une constante. Une telle distribution (qui est tempérée) admet une transformée de Fourier. La condition aux limites permet de déterminer la constante C . Ici nous pouvons montrer que la constante $-i\pi$ convient. Une écriture plus commode de la distribution solution est donnée grâce à l'égalité $\lim_{\epsilon \rightarrow 0} \frac{1}{x+i\epsilon} = \mathcal{P} \left(\frac{1}{x} \right) - i\pi \delta(x)$. La fonction de Green solution dans le domaine de Fourier vaut donc

$$\mathcal{G}(k, \omega) = i\hbar \lim_{\epsilon \rightarrow 0} \frac{1}{\hbar\omega - \frac{\hbar^2 k^2}{2m} + i\epsilon} \tag{3.37}$$

La fonction de Green dite *retardée* dans l'espace réel est

$$\mathcal{G}(x, t) = \frac{i\hbar}{(2\pi)^2} \lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}^2} \frac{e^{i(kx - \omega t)}}{\hbar\omega - \frac{\hbar^2 k^2}{2m} + i\epsilon} d\omega dk \tag{3.38}$$



L'intégrale peut se calculer grâce au théorème des résidus. L'introduction de ϵ permet de décaler le pôle hors de l'axe réel dans le plan complexe ce qui aurait sinon posé quelques problèmes. Ou encore nous pouvons dire que l'intégration se fait sur un chemin contournant le pôle par un cercle de rayon ϵ centré sur le pôle qui n'est plus déplacé. Du coup, l'intégrale sur ω vaut

$$\lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}} \frac{e^{-i\omega t}}{\omega - \frac{\hbar k^2}{2m} + i\epsilon} d\omega = \begin{cases} 0 & \text{si } t < 0 \\ -\frac{2i\pi}{\hbar} e^{-i\frac{k^2 \hbar t}{2m}} & \text{si } t > 0 \end{cases} \quad (3.39)$$

Finalement,

$$\mathcal{G}(x, t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{i(kx - \frac{\hbar k^2 t}{2m})} dk \quad (3.40)$$

Nous devons calculer une intégrale proche d'une intégrale gaussienne. En fait, il est possible de se ramener à ce cas en mettant le polynôme du second degré sous forme canonique. Notons que l'intégrale à calculer est une exponentielle complexe gaussienne. Le résultat est le même que dans le cas réel et se montre encore à l'aide du théorème des résidus en intégrant la fonction e^{-z^2} sur un contour adéquat. Pour conclure, la fonction de Green retardée vaut

$$\mathcal{G}(x, t) = \begin{cases} 0 & \text{si } t < 0 \\ \sqrt{\frac{m}{2\pi i \hbar t}} e^{im\frac{x^2}{2\hbar t}} & \text{si } t > 0 \end{cases} \quad (3.41)$$

Connaissant la fonction d'onde à un instant t_0 , il est possible de la déterminer à tous les instants ultérieurs grâce à la formule intréale. Du point de vue différentiel, la prise en compte des conditions initiales se fait en rajoutant un terme $\psi(x, t_0) \delta(t - t_0)$ au membre de droite. Cela étant, nous avons

$$\psi(x, t) = \int_{\mathbb{R}} \mathcal{G}(x - x_0, t - t_0) \psi(x_0, t_0) dx_0 = \sqrt{\frac{m}{2\pi i \hbar t}} \int_{\mathbb{R}} e^{im\frac{(x-x_0)^2}{2\hbar(t-t_0)}} \psi(x_0, t_0) dx_0 \quad (3.42)$$

Paquet d'onde gaussien

Pour avoir une vision plus concrète de l'évolution, on considère qu'à l'instant initial t_0 la fonction d'onde est une gaussienne centrée en 0 de largeur l . Une telle fonction d'onde décrit une particule libre au repos.

$$\psi(x_0, t_0) = \sqrt{\pi} l e^{-\frac{x_0^2}{2l^2}} \quad (3.43)$$

L'évolution est alors donnée par 3.42. L'intérêt de choisir un paquet d'onde gaussien est qu'il reste gaussien lors de son évolution. Le calcul des intégrales est similaire à celui qui a conduit à la formule 3.42. La largeur du paquet d'onde est alors à l'instant t donnée par

$$l(t) = \sqrt{l + \frac{\hbar^2 l^2}{m^2}} \quad (3.44)$$

La largeur du paquet s'étale donc au cours du temps. Ce phénomène, qui se produit lors de n'importe quelle évolution libre, est appelé *étalement du paquet d'onde*. On peut le comprendre comme issu du principe d'indétermination de Heisenberg.

Si on considère que la particule possède une impulsion p_0 , la fonction d'onde initiale est multipliée par la phase $e^{-i\frac{p_0 x_0}{\hbar}}$. Ceci a pour effet de décentrer la gaussienne. À l'instant t , le centre du paquet d'onde est en $p_0 t / m$. Ceci est en accord avec les lois classiques du mouvement !

3.3.2 Équation de Schrödinger indépendante du temps

Dans toute cette partie, on se placera en dimension un. Le hamiltonien indépendant du temps le plus général possible est

$$H = \frac{P^2}{2m} + V(x) \quad (3.45)$$



où $K = \frac{P^2}{2m}$ est l'opérateur *énergie cinétique* et $V(x)$ l'opérateur *potentiel*. L'équation d'évolution est

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H(t) |\psi(t)\rangle \tag{3.46}$$

ou encore, en représentation position (obtenue en multipliant l'équation précédente à gauche par le bra $\langle x|$)

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} + V(x) \psi(x, t) \tag{3.47}$$

La partie droite de cette équation étant indépendante du temps, on sait qu'on peut trouver des solutions sous la forme

$$|\psi(t, x)\rangle = e^{-i\frac{Et}{\hbar}} |\psi(0, x)\rangle \tag{3.48}$$

avec $H|\psi(0, x)\rangle = E|\psi(0, x)\rangle$ comme nouvelle équation à résoudre, c'est-à-dire ici

$$\left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right] \psi(x) = E\psi(x) \tag{3.49}$$

3.3.3 Résolution de cette équation

On voit très facilement que la résolution de l'équation 3.49 va dépendre du signe de $E - V(x)$. Dans le cas le plus simple, $V = 0$, et l'équation devient

$$\left[\frac{\partial^2}{\partial x^2} + \frac{2mE}{\hbar^2} \right] \psi(x) = 0 \tag{3.50}$$

dont la solution est une combinaison d'onde planes

$$\psi(x) = Ae^{ipx/\hbar} + Be^{-ipx/\hbar} \tag{3.51}$$

avec $p = \sqrt{2mE} > 0$. Maintenant, passons au cas où $V(x)$ n'est pas nul. En mécanique classique, on aurait deux types d'états selon si $E > V(x)$ ou $E < V(x)$. Cette distinction va à nouveau apparaître ici. Prenons pour commencer le cas d'une marche de potentiel

$$V(x) = 0 \text{ pour } x < 0 \tag{3.52a}$$

$$V(x) = V_0 \text{ pour } x > 0 \tag{3.52b}$$

On va résoudre l'équation

$$\left[\frac{\partial^2}{\partial x^2} + \frac{2m(E - V(x))}{\hbar^2} \right] \psi(x) = 0 \tag{3.53}$$

dans les deux régions. Supposons tout d'abord $0 < E < V_0$. On définit k et κ par

$$k = \sqrt{\frac{2mE}{\hbar^2}} \quad \kappa = \sqrt{\frac{2m(V_0 - E)}{\hbar^2}} \tag{3.54}$$

Les solutions de l'équation de Schrödinger sont alors

$$\psi(x) = Ae^{ikx} + Be^{-ikx} \text{ pour } x < 0 \tag{3.55a}$$

$$\psi(x) = Ce^{-\kappa x} + De^{\kappa x} \text{ pour } x > 0 \tag{3.55b}$$



On a de plus des conditions de continuité pour ψ et pour sa dérivée en $x = 0$, ainsi qu'une condition de non divergence à l'infini (le module de ψ ne peut pas tendre vers l'infini). Ces conditions imposent :

$$D = 0 \tag{3.56a}$$

$$C = A + B \tag{3.56b}$$

$$-\kappa = \frac{ik(A - B)}{A + B} \tag{3.56c}$$

Réflexion et transmission d'une onde incidente

On peut interpréter l'onde associée au coefficient A (que nous prendrons dans la suite égal à 1) comme correspondant à une onde incidente de particules. Dans ce cas, l'onde "B" est une onde réfléchie et l'onde "C" est une onde transmise. Dans le cas que nous étudions jusqu'ici ($E < V_0$), l'étude classique du problème donnerait une réflexion totale de l'onde. Ici, les fonctions d'onde obtenues sont :

$$\psi(x) = e^{ikx} + Be^{-ikx} \text{ pour } x < 0 \tag{3.57a}$$

$$\psi(x) = Ce^{-\kappa x} \text{ pour } x > 0 \tag{3.57b}$$

avec

$$B = -\frac{\kappa + ik}{\kappa - ik} \quad C = -\frac{2ik}{\kappa - ik} \tag{3.58}$$

On peut démontrer facilement que $|B| = 1$, ce qui montre que l'onde réfléchie a une intensité égale à l'onde incidente, et qu'il y a bien réflexion totale. Néanmoins, on voit également apparaître une onde évanescente dans la zone $x > 0$, ce qui signifie que la probabilité de présence n'est pas nulle dans cette région, contrairement au cas classique. On a une profondeur de pénétration $l = 1/\kappa$ (voir figure 3.1).

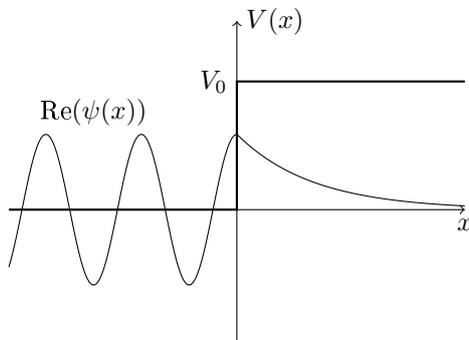


FIGURE 3.1 – Onde incidente sur une marche de potentiel, dans le cas d'une énergie $E < V_0$. On a apparition d'une onde évanescente dans la zone $x > 0$, ce qui n'est pas le cas dans le cas classique.

Nous allons maintenant nous intéresser au cas où $E > V_0$. Dans le cas classique, on aurait une transmission complète, et on va montrer que ce n'est pas le cas ici. On pose

$$k' = \sqrt{\frac{2m(E - V_0)}{\hbar^2}} \tag{3.59}$$

Les nouvelles solutions sont (l'absence de solutions "D" provient du fait que l'on suppose ne pas avoir de sources de particules ou de réflexions dans la zone $0 < x < \infty$)

$$\psi(x) = e^{ikx} + Be^{-ikx} \text{ pour } x < 0 \tag{3.60a}$$

$$\psi(x) = Ce^{ik'x} \text{ pour } x > 0 \tag{3.60b}$$

avec les conditions

$$B = \frac{k - k'}{k + k'} \quad C = \frac{2k}{k + k'} \tag{3.61}$$

On voit que le coefficient de réflexion dans le cas quantique n'est pas nul. On définit alors les coefficients de réflexion et de transmission

$$R = |B|^2 = \left(\frac{k - k'}{k + k'} \right)^2 \tag{3.62a}$$

$$T = 1 - R = \frac{4kk'}{(k + k')^2} = \frac{k'}{k} |C|^2 \tag{3.62b}$$

Le fait que $T \neq |C|^2$ provient du fait que c'est le courant de particules qui est conservé lors du passage de la barrière et non la probabilité de présence.

3.3.4 États liés du puits carré

On se place dans le cas d'un puits carré infini, c'est-à-dire

$$V(x) = 0 \text{ pour } 0 \leq x \leq a \tag{3.63a}$$

$$V(x) = +\infty \text{ pour } x < 0 \text{ ou } x > a \tag{3.63b}$$

D'après l'étude précédente, la fonction d'onde s'annule hors du puits de potentiel, et est de la forme $\psi(x) = Ae^{ikx} + Be^{-ikx}$ à l'intérieur du puits. Pour des raisons de continuité en 0 et en a , il faut que ψ s'annule. On a alors

$$\psi(x) = A \sin(kx) \text{ où } k = \sqrt{\frac{2mE}{\hbar^2}} \tag{3.64}$$

Et, puisque $\psi(a) = 0$, k ne peut prendre que des valeurs *quantifiées*, k_n . On a donc, $\forall n \in \mathbb{N}$

$$k_n = \frac{\pi(n + 1)}{a} \tag{3.65a}$$

$$E_n = \frac{\hbar^2 k_n^2}{2m} \tag{3.65b}$$

Les niveaux d'énergie disponibles dans un puits infini sont quantifiés. Il faut de plus normaliser les fonctions d'ondes, c'est-à-dire imposer que $\int_{\mathbb{R}} \psi^*(x)\psi(x)dx = 1$. Enfin, les calculs menés ici sont adaptables au cas d'un puits fini, c'est-à-dire de potentiel fixé $V = V_0$ en dehors du puits. Dans ce cas, qui ne sera pas traité ici, on a apparition d'ondes évanescentes des deux côtés du puits, et une distinction entre solutions paires et impaires de l'équation de Schrödinger qui n'a pas lieu pour le puits infini. La forme des quatre premières solutions est donnée sur la figure 3.2.

3.3.5 Espace des phases et densité de niveaux

On cherche à déterminer la densité des niveaux d'énergie. On considère k dans l'intervalle $[k_n; k_n + \Delta k]$. En utilisant la relation donnant les niveaux d'énergie dans un puits de potentiel de longueur l_x , en le supposant cette fois périodique, nous avons, en généralisant à trois dimensions, un nombre de solutions Δn dans cet intervalle

$$\Delta n = \frac{l_x l_y l_z}{(2\pi)^3} \Delta k_x \Delta k_y \Delta k_z \tag{3.66a}$$

On en déduit la *densité d'états*,

$$\mathcal{D}(\mathbf{k}) = \frac{V}{(2\pi)^3} \tag{3.67}$$



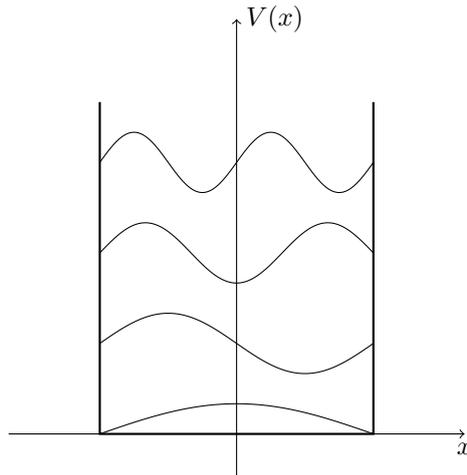


FIGURE 3.2 – Forme des quatre premiers états liés du puits carré infini.

où V est le volume. Dans l'espace des impulsions

$$\mathcal{D}(\mathbf{p}) = \frac{V}{(2\pi\hbar)^3} = \frac{V}{h^3}. \tag{3.68}$$

On peut, à partir de là, exprimer la densité d'énergie en utilisant la relation (valable uniquement en trois dimensions)

$$\mathcal{D}(E)dE = \mathcal{D}(p)4\pi p^2 dp. \tag{3.69}$$

La relation liant l'énergie et l'impulsion d'une particule libre $E = \frac{p^2}{2m}$ permet de conclure,

$$\mathcal{D}(E) = \frac{Vm}{2\pi^2\hbar^3} \sqrt{2mE} \tag{3.70}$$

On peut déterminer enfin le nombre de niveaux dans un élément de volume de l'espace des phases. Nous avons

$$dN = \frac{d^3\mathbf{r}d^3\mathbf{p}}{h^3} \tag{3.71}$$

Ceci peut s'interpréter grâce au principe d'incertitude de Heisenberg. En effet, ce principe nous dit que connaissant la position à Δx près, on connaît l'impulsion au mieux à $\Delta p \sim \frac{\hbar}{\Delta x}$. Une particule quantique, contrairement à une particule classique, doit occuper au minimum un volume de l'espace des phases de h^3 .

3.3.6 Diffusion par un potentiel : effet tunnel

Dans cette section, on va étudier le cas où l'on définit trois zones de potentiels

$$V(x) = 0 \text{ pour } x \leq -\frac{a}{2} \tag{3.72a}$$

$$V(x) = V_0 \text{ pour } -\frac{a}{2} < x < \frac{a}{2} \tag{3.72b}$$

$$V(x) = 0 \text{ pour } \frac{a}{2} \leq x \tag{3.72c}$$



Matrice de passage

Tout d'abord, on va étudier un cas général où $E > V_0$. En définissant $k = \sqrt{\frac{2mE}{\hbar^2}}$ et $k' = \sqrt{\frac{2m(V_0+E)}{\hbar^2}}$, on obtient les fonctions d'onde dans les trois régions :

$$x < -\frac{a}{2} \quad \psi(x) = Ae^{ikx} + Be^{-ikx} \quad (3.73a)$$

$$-\frac{a}{2} \leq x \leq \frac{a}{2} \quad \psi(x) = Ce^{ik'x} + De^{-ik'x} \quad (3.73b)$$

$$x > \frac{a}{2} \quad \psi(x) = Fe^{ikx} + Ge^{-ikx} \quad (3.73c)$$

On commence par étudier le passage de la première région à la deuxième, c'est-à-dire le lien entre les coefficients A, B et C, D . L'équation de Schrödinger étant linéaire, on peut trouver une relation matricielle entre ces coefficients.

$$\begin{pmatrix} A \\ B \end{pmatrix} = R \begin{pmatrix} C \\ D \end{pmatrix} \quad (3.74)$$

R est une matrice 2×2 . En remarquant que si $\psi(x)$ est solution de l'équation de Schrödinger, alors $\psi^*(x)$ l'est également, on obtient

$$\begin{pmatrix} B^* \\ A^* \end{pmatrix} = R \begin{pmatrix} D^* \\ C^* \end{pmatrix} \quad (3.75)$$

Ceci nous donne $R_{11}^* = R_{22}$ et $R_{12}^* = R_{21}$. On peut donc écrire la matrice sous la forme

$$R = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} \quad (3.76)$$

À l'aide des relations de continuité sur ψ et sa dérivée en $-a/2$, on va pouvoir trouver les coefficients α et β . On se place dans le cas $C = 1, D = 0$, qui est le plus pratique pour le calcul. On trouve alors

$$\alpha = \frac{k + k'}{2k} e^{\frac{ia}{2}(k-k')} \quad (3.77a)$$

$$\beta = \frac{k - k'}{2k} e^{\frac{ia}{2}(k+k')} \quad (3.77b)$$

Avec ces coefficients, on trouve que le déterminant de R est $\frac{k'}{k}$, ce qui est une condition nécessaire à la conservation des courants à l'interface.

On passe alors à la seconde interface, où l'on définit de la même façon R' par :

$$\begin{pmatrix} C \\ D \end{pmatrix} = R' \begin{pmatrix} F \\ G \end{pmatrix} \quad (3.78)$$

On montre facilement (en remarquant que les conditions de continuité s'obtiennent en échangeant a et $-a$ ainsi que k et k') que

$$R' = \begin{pmatrix} \alpha & -\beta^* \\ -\beta & \alpha^* \end{pmatrix} \quad (3.79)$$

La matrice totale reliant le couple (A, B) à (F, G) est alors $M = RR'$. Avec les mêmes arguments que pour R , on trouve sa forme générale :

$$M = \begin{pmatrix} \gamma & \delta \\ \delta^* & \gamma^* \end{pmatrix} \quad (3.80)$$



On peut montrer facilement que son déterminant est égal à 1. On a de plus :

$$M_{11} = \gamma = \frac{k'}{k}(\alpha^2 - \beta^2) \tag{3.81a}$$

$$M_{12} = \delta = \frac{k'}{k}(-\alpha\beta^* + \alpha^*\beta) \tag{3.81b}$$

Ce sont ces coefficients qui nous serviront dans la suite.

Effet tunnel

On s'intéresse ici au cas $E < V_0$. La quantité k' précédente est alors imaginaire pure :

$$k' = i\kappa \quad \kappa = \sqrt{\frac{2m(V_0 - E)}{\hbar^2}} \tag{3.82}$$

Avec le formalisme des matrices de passage évoqué précédemment, on va pouvoir trouver très facilement le coefficient de transmission à travers le potentiel considéré. On se place dans le cas où la source de particules est située en $x = -\infty$, et on prend $A = 1$ par convention. On suppose de plus qu'il n'y a pas de sources de particules en $x = +\infty$, ce qui impose $G = 0$. On a alors

$$\begin{pmatrix} 1 \\ B \end{pmatrix} = M \begin{pmatrix} F \\ 0 \end{pmatrix} = \begin{pmatrix} M_{11}F \\ M_{21}F \end{pmatrix} \tag{3.83}$$

Et donc, $F = 1/M_{11}$. Le coefficient de transmission que nous cherchons est $T = |F|^2$. On le calcule alors explicitement :

$$\begin{aligned} T &= |1/M_{11}|^2 && (3.84) \\ &= \frac{1}{|\frac{k'}{k}(\alpha^2 - \beta^2)|^2} \\ &= \frac{1}{\left| \frac{(k+k')^2}{4kk'} e^{ia(k-k')} - \frac{(k-k')^2}{4kk'} e^{ia(k+k')} \right|^2} \end{aligned}$$

On se rappelle alors que $k' = i\kappa$, pour écrire :

$$\begin{aligned} \frac{1}{T} &= \left| \frac{(k + i\kappa)^2 e^{\kappa a} - (k - i\kappa)^2 e^{-\kappa a}}{4ik\kappa} \right|^2 \\ &= \left| \frac{(k^2 - \kappa^2) \sinh(\kappa a) + 2ik\kappa \cosh(\kappa a)}{2k\kappa} \right|^2 \\ &= \frac{1}{4k^2\kappa^2} ((k^2 - \kappa^2) \sinh^2(\kappa a) + 4k^2\kappa^2 \cosh^2(\kappa a)) \\ &= \frac{1}{4k^2\kappa^2} ((k^2 + \kappa^2) \sinh^2(\kappa a) + 4k^2\kappa^2) \end{aligned}$$

Ce qui nous donne enfin, en posant $q^2 = k^2 + \kappa^2 = 2mV_0/\hbar^2$:

$$T = \frac{1}{1 + \frac{q^4}{4k^2\kappa^2} \sinh^2(\kappa a)} \tag{3.85}$$

On voit alors que la probabilité de traverser une barrière de potentiel, qui était absolument nulle dans le cas classique, ne l'est plus ici. Cet effet est appelé *effet tunnel* et est représenté graphiquement sur la figure 3.3.

Un cas extrêmement fréquent va être celui d'une barrière "épaisse", c'est-à-dire telle que $\kappa a \gg 1$. Dans ce cas, on peut obtenir l'expression approchée :



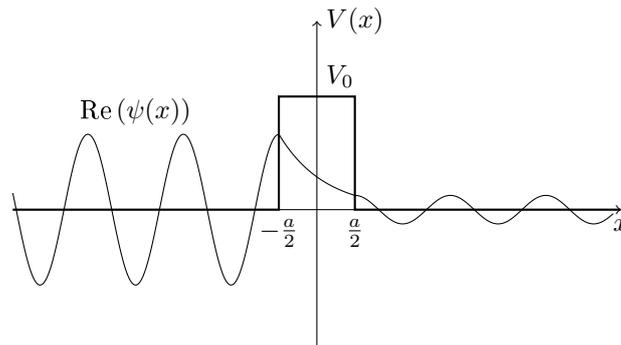


FIGURE 3.3 – Exemple d’effet tunnel. L’onde incidente a une énergie $E < V_0$. On voit que, contrairement au cas classique, une particule peut traverser la barrière de potentiel. Cette probabilité de passage reste néanmoins faible dans le cas général.

$$T \simeq \frac{16k^2\kappa^2}{q^4} e^{-2\kappa a} \tag{3.86}$$

Le terme dominant étant l’exponentielle, une formule approchée est souvent utilisée dans le cas où $E < \text{Max}(V(x))$. On décompose alors le potentiel en crêneaux de longueur Δx , et on calcule le facteur de transmission pour chaque intervalle $[x_i, x_i + \Delta x]$

$$T(x_i) \simeq e^{-2\kappa(x_i)\Delta x} \quad \kappa(x_i) = \sqrt{\frac{2m(V(x_i) - E)}{\hbar^2}} \tag{3.87}$$

Le facteur de transmission total est alors, à la limite $\Delta x \rightarrow 0$,

$$T \simeq \exp\left(-2 \int_{x_1}^{x_2} \sqrt{\frac{2m(V(x) - E)}{\hbar^2}} dx\right) \tag{3.88}$$

C’est cette formule qui sera utilisée en physique nucléaire pour le cas de la désintégration α ou de la fission.

3.4 Théorie élémentaire de la diffusion

On va s’intéresser ici à un phénomène qui est au cœur de nombreuses applications de la mécanique quantique, et dont nous nous réservons dans plusieurs autres chapitres de ce Need to Know : la diffusion. Il s’agit tout simplement du bombardement d’un ensemble de particules cibles par un faisceau de particules incidentes, et l’on veut regarder la façon dont l’émission se déroule après ce contact. Cela est utilisé notamment pour des méthodes spectroscopiques permettant de sonder la structure de la matière (que nous reverrons en partie lors du chapitre 7 de Matière Condensée), ou pour calculer les collisions entre particules (voir le chapitre 5 de Physique des Particules). On étudiera ici la diffusion stationnaire¹⁰.

La diffusion permet de sonder la structure de la matière ; elle est utile notamment en physique des particules. Elle consiste en

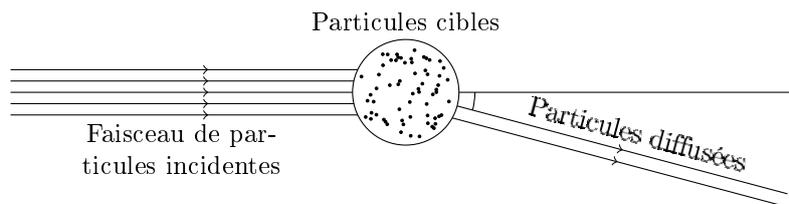


FIGURE 3.4 – Expérience de diffusion.

¹⁰. De manière plus correcte, on a une dépendance en temps et la prise en compte du potentiel se fait adiabatiquement.

3.4.1 Notions générales

Difficultés et hypothèses simplificatrices

On considère une diffusion élastique : l'énergie cinétique totale est conservée, cependant les directions de propagation sont modifiées. Il existe des interactions à trois corps ; mais on néglige ces interactions, la densité des particules incidentes étant faible. Il existe aussi des diffusions multiples. Ici, on suppose la densité des particules cibles suffisamment faible, pour pouvoir les négliger (si λ est la longueur d'onde des particules incidentes, alors λ est très petite devant la distance typique entre les particules cibles). Enfin, on suppose la diffusion incohérente. En d'autres termes, on néglige les phénomènes d'interférences qui peuvent exister entre les différentes ondes diffusées. Les particules loin de la cible peuvent être considérées comme libres. On se ramène donc à un problème à deux corps donné par un potentiel $V(\mathbf{r})$.

Section efficace

Soit ρ la densité de particules incidentes. La densité de courant \mathbf{j} de particules incidentes s'écrit $\mathbf{j} = \rho \mathbf{v} = \rho \frac{\mathbf{p}}{m}$. La norme de \mathbf{j} donne le flux de particules incidentes. Le nombre de particules cibles par unité de surface perpendiculaire au faisceau et de temps est noté N_T . Si on note $dN_s(\Omega)$ le nombre de particules diffusées par unité de temps dans un angle solide $d\Omega$, nous avons

$$dN_s(\Omega) = N_T \|\mathbf{j}\| \frac{d\sigma}{d\Omega} d\Omega \tag{3.89}$$

où $\frac{d\sigma}{d\Omega}$ est la section efficace différentielle.

Par définition, on appelle *section efficace différentielle* $\frac{d\sigma}{d\Omega}$ définie par

$$\frac{d\sigma}{d\Omega} = \frac{dN_s(\Omega)}{N_T \|\mathbf{j}\|} \tag{3.90}$$

On appelle section efficace totale σ l'intégrale sur tous les angles solides de la section efficace différentielle

$$\sigma = \int \frac{d\sigma}{d\Omega} d\Omega \tag{3.91}$$

On démontre (les prochaines pages sont consacrées à cette démonstration) que la section efficace de collision est :

$$\frac{d\sigma}{d\Omega} = |f_{\mathbf{k}}(\Omega)|^2 \tag{3.92}$$

avec

$$f_{\mathbf{k}}(\Omega) = -\frac{m}{2\pi\hbar^2} \int d\mathbf{r}' e^{-i\hat{\mathbf{k}} \cdot \mathbf{r}'} V(\mathbf{r}') \psi_{\mathbf{k}}(\mathbf{r}') \tag{3.93}$$

et \mathbf{k} le vecteur d'onde de l'onde incident, Ω l'angle solide correspondant à la "direction" de diffusion considérée.

3.4.2 Méthode par résolvante

On va chercher à résoudre un problème à deux corps en présence d'un potentiel $V(\mathbf{r})$, c'est-à-dire trouver les états propres du hamiltonien suivant.

$$H = -\frac{\hbar^2}{2m} \nabla_{\mathbf{r}} + V(\mathbf{r}) \tag{3.94}$$

En l'absence de V , ou quand \mathbf{r} tend vers l'infini dans le cas général, on a le hamiltonien non perturbé H_0 , dont les solutions sont $\psi_{\mathbf{k}}^{(0)}$

$$H_0 = -\frac{\hbar^2}{2m} \nabla_{\mathbf{r}} \quad (3.95a)$$

$$\psi_{\mathbf{k}}^{(0)} = e^{i\mathbf{k}\cdot\mathbf{r}} \quad (3.95b)$$

$$E_{\mathbf{k}}^{(0)} = \frac{\hbar^2 k^2}{2m} \quad (3.95c)$$

On va maintenant chercher les solutions du hamiltonien total, $\psi_{\mathbf{k}}$ telles que $H\psi_{\mathbf{k}} = E_{\mathbf{k}}\psi_{\mathbf{k}}$. Puisque la diffusion est élastique, l'énergie est inchangée ($E_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m}$). Pour cela, on cherche $\psi_{\mathbf{k}}$ sous la forme :

$$\psi_{\mathbf{k}} = \psi_{\mathbf{k}}^{(0)} + \delta\psi_{\mathbf{k}} \quad (3.96)$$

On a alors :

$$(H_0 + V)(\psi_{\mathbf{k}}^{(0)} + \delta\psi_{\mathbf{k}}) = E_{\mathbf{k}}(\psi_{\mathbf{k}}^{(0)} + \delta\psi_{\mathbf{k}}) \quad (3.97)$$

$$(H_0 - E_{\mathbf{k}})\delta\psi_{\mathbf{k}} = -V\psi_{\mathbf{k}} \quad (3.98)$$

On aimerait alors pouvoir inverser $(H_0 - E_{\mathbf{k}})$, ce qui permettrait d'obtenir un développement perturbatif en puissances de V pour $\delta\psi_{\mathbf{k}}$. Malheureusement, ce n'est pas le cas. On pose alors $\lambda = -E_{\mathbf{k}} + i\epsilon$, et l'opérateur $(H_0 + \lambda)$ est alors inversible. En faisant tendre ϵ vers 0, on aura le résultat cherché.

Calcul de l'inverse de $H_0 + \lambda$

On appelle cet inverse $G_{\lambda}^{(0)}$, et on note $G_{\lambda}^{(0)}(\mathbf{r}, \mathbf{r}')$ l'élément de matrice de $G_{\lambda}^{(0)}$.

$$G_{\lambda}^{(0)}(\mathbf{r}, \mathbf{r}') = \langle \mathbf{r} | G_{\lambda}^{(0)} | \mathbf{r}' \rangle \quad (3.99)$$

On a

$$(H_0 + \lambda \mathbb{1})G_{\lambda}^{(0)} = \mathbb{1} \quad (3.100)$$

$$\langle \mathbf{r} | (H_0 + \lambda \mathbb{1})G_{\lambda}^{(0)} | \mathbf{r}' \rangle = \langle \mathbf{r} | \mathbf{r}' \rangle = \delta(\mathbf{r} - \mathbf{r}') \quad (3.101)$$

On utilise la relation de fermeture de la base $|\mathbf{r}\rangle$ pour réécrire le terme de gauche

$$\int d\mathbf{r}'' \langle \mathbf{r} | (H_0 + \lambda \mathbb{1}) | \mathbf{r}'' \rangle \langle \mathbf{r}'' | G_{\lambda}^{(0)} | \mathbf{r}' \rangle = \delta(\mathbf{r} - \mathbf{r}') \quad (3.102)$$

Or, $\langle \mathbf{r} | \frac{\mathbf{p}^2}{2m} | \mathbf{r}'' \rangle = -\frac{\hbar^2}{2m} \Delta_{\mathbf{r}} \delta(\mathbf{r} - \mathbf{r}'')$ et de façon évidente, $\langle \mathbf{r} | \lambda \mathbb{1} | \mathbf{r}'' \rangle = \lambda \delta(\mathbf{r} - \mathbf{r}'')$. L'équation à résoudre devient donc, après intégration sur \mathbf{r}'' ,

$$\left(-\frac{\hbar^2}{2m} \Delta_{\mathbf{r}} + \lambda \right) G_{\lambda}^{(0)}(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') \quad (3.103)$$

Cette équation montre que $G_{\lambda}^{(0)}(\mathbf{r}, \mathbf{r}')$ est en fait la fonction de Green de l'équation de Schrödinger considérée.

Pour résoudre cette équation, on va chercher $G_{\lambda}^{(0)}(\mathbf{r}, \mathbf{r}')$ sous la forme $G_{\lambda}^{(0)}(\mathbf{r} - \mathbf{r}')$, et on va passer par les transformées de Fourier.

$$G_{\lambda}^{(0)}(\mathbf{r} - \mathbf{r}') = \frac{1}{(2\pi)^3} \int e^{i\mathbf{k}' \cdot (\mathbf{r} - \mathbf{r}')} \mathcal{G}_{\lambda}^{(0)}(\mathbf{k}') d\mathbf{k}' \quad (3.104)$$

En passant dans le domaine de Fourier l'équation 3.103, on trouve que :

$$\mathcal{G}_\lambda^{(0)}(\mathbf{k}') = \frac{1}{\frac{\hbar^2 \mathbf{k}'^2}{2m} + \lambda} \tag{3.105}$$

On trouve alors, en choisissant comme axe principal de notre base sphérique pour l'intégration la direction de $\mathbf{r} - \mathbf{r}'$

$$\begin{aligned} G_\lambda^{(0)}(\mathbf{r} - \mathbf{r}') &= \frac{1}{(2\pi)^3} \int_0^\infty k'^2 dk' \int_0^\pi d\theta \sin(\theta) \int_0^{2\pi} d\phi \frac{e^{ik'|\mathbf{r}-\mathbf{r}'|\cos(\theta)}}{\frac{\hbar^2 \mathbf{k}'^2}{2m} + \lambda} \\ &= \frac{1}{(2\pi)^2} \int_0^\infty k'^2 dk' \frac{1}{\frac{\hbar^2 \mathbf{k}'^2}{2m} + \lambda} \left[-\frac{e^{ik'|\mathbf{r}-\mathbf{r}'|\cos(\theta)}}{ik'|\mathbf{r}-\mathbf{r}'|} \right]_0^\pi \\ &= -\frac{i}{(2\pi)^2} \frac{1}{|\mathbf{r} - \mathbf{r}'|} \int_{\mathbb{R}} \frac{k' e^{ik'|\mathbf{r}-\mathbf{r}'|}}{\frac{\hbar^2 \mathbf{k}'^2}{2m} + \lambda} \end{aligned} \tag{3.106}$$

Le passage à la dernière ligne se fait en remarquant que, en notant A le terme apparaissant pour $\theta = \pi$ dans l'intégrande et B celui pour $\theta = 0$, on a $\int_0^{+\infty} A = \int_{-\infty}^0 B$, et donc que $\int_0^{+\infty} A + B = \int_{-\infty}^0 B + \int_0^{+\infty} B = \int_{\mathbb{R}} B$. Pour résoudre l'intégrale finale, on va devoir passer par l'analyse complexe et par le théorème des résidus. Les pôles de l'intégrande sont solutions de $k'^2 = -\lambda \frac{2m}{\hbar^2}$, ou encore $k'^2 = k^2 - i\epsilon'$ (où $\epsilon' = \epsilon \frac{2m}{\hbar^2}$). Les solutions sont

$$k_\pm = \pm k \left(1 - i \frac{\epsilon'}{k^2} \right)^{\frac{1}{2}} \tag{3.107a}$$

$$\simeq \pm \left(k - i \frac{\epsilon'}{2k} \right) \tag{3.107b}$$

On choisit alors $\epsilon' < 0$, ce qui permet que la solution k_+ soit à l'intérieur du contour d'intégration défini sur la figure 3.5.

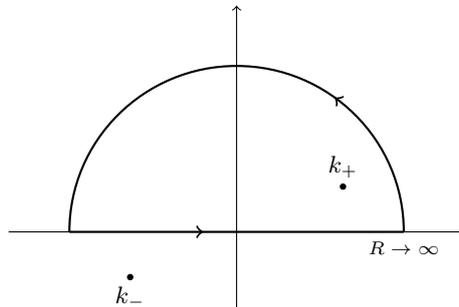


FIGURE 3.5 – Contour d'intégration pour le théorème des résidus.

Le lemme de Jordan nous permet d'affirmer que la partie circulaire de ce contour ne contribue pas à l'intégrale, tandis que la partie correspondant à la droite réelle est l'intégrale recherchée. Le théorème des résidus nous donne alors

$$G_\lambda^{(0)}(\mathbf{r} - \mathbf{r}') = -\frac{i}{(2\pi)^2} \frac{1}{|\mathbf{r} - \mathbf{r}'|} 2i\pi \text{Res}_{k_+}(f) \tag{3.108}$$

$$= \frac{1}{2\pi} \frac{1}{|\mathbf{r} - \mathbf{r}'|} \frac{k_+ e^{ik_+|\mathbf{r}-\mathbf{r}'|}}{\frac{\hbar^2}{2m} (k_+ - k_-)} \tag{3.109}$$

Ce qui donne finalement, en faisant tendre ϵ vers 0,

$$G_\lambda^{(0)}(\mathbf{r} - \mathbf{r}') = \frac{m}{2\pi \hbar^2} \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \tag{3.110}$$

Ce qui correspond à une onde sphérique sortante.

Retour à la fonction d'onde

$G_\lambda^{(0)}(\mathbf{r} - \mathbf{r}')$ étant la fonction de Green de l'équation 3.103, on a le résultat :

$$\delta\psi_{\mathbf{k}}(\mathbf{r}) = -\frac{m}{2\pi\hbar^2} \int_{\mathbb{R}^3} d\mathbf{r}' \frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} V(\mathbf{r}')\psi_{\mathbf{k}}(\mathbf{r}') \tag{3.111}$$

Cette équation est une équation intégrale pour $\delta\psi_{\mathbf{k}}$. Dans le cas de la diffusion, ce qui nous intéresse est le cas $r \rightarrow \infty$. Dans ce cas, on a

$$\begin{aligned} k|\mathbf{r} - \mathbf{r}'| &\simeq kr - k\frac{\mathbf{r} \cdot \mathbf{r}'}{r} \\ &\simeq kr - \hat{\mathbf{k}} \cdot \mathbf{r}' \end{aligned} \tag{3.112}$$

Ce qui donne

$$\delta\psi_{\mathbf{k}}(\mathbf{r}) \simeq -\frac{m}{2\pi\hbar^2} \frac{e^{ikr}}{r} \int d\mathbf{r}' e^{-i\hat{\mathbf{k}} \cdot \mathbf{r}'} V(\mathbf{r}')\psi_{\mathbf{k}}(\mathbf{r}') \tag{3.113}$$

On peut montrer que l'intégrale ne dépend que de l'angle solide Ω , et on l'appelle $f_{\mathbf{k}}(\Omega)$. Alors, on a :

$$\psi_{\mathbf{k}}(\mathbf{r}) \simeq \underbrace{e^{i\mathbf{k} \cdot \mathbf{r}}}_{\substack{\text{Onde} \\ \text{plane} \\ \text{entrante}}} + f_{\mathbf{k}}(\Omega) \underbrace{\frac{e^{ikr}}{r}}_{\substack{\text{Onde} \\ \text{sphérique} \\ \text{sortante}}} \tag{3.114a}$$

$$f_{\mathbf{k}}(\Omega) = -\frac{m}{2\pi\hbar^2} \int d\mathbf{r}' e^{-i\hat{\mathbf{k}} \cdot \mathbf{r}'} V(\mathbf{r}')\psi_{\mathbf{k}}(\mathbf{r}') \tag{3.114b}$$

Calcul de la section efficace différentielle

Pour ce calcul, on va avoir besoin du flux de particules incidentes $\mathbf{j}^{(\text{in})}$, ainsi que du nombre de particules diffusées dans l'angle solide Ω $dN_s(\Omega)$, qui dépend du flux de particules sortantes $\mathbf{j}^{(\text{out})}$. Pour cela, on va traiter d'un côté l'onde plane entrante, et d'un autre l'onde sphérique sortante. On rappelle que, pour une fonction d'onde ψ , on a $\mathbf{j} = \frac{\hbar}{m} \text{Im}(\psi^* \nabla \psi)$. On a alors, avec l'axe \mathbf{z} choisi dans la direction de \mathbf{k} :

$$\psi^{(\text{in})} = Ae^{ikz} \tag{3.115a}$$

$$\mathbf{j}^{(\text{in})} = |A|^2 \frac{\hbar\mathbf{k}}{m} \mathbf{e}_z \tag{3.115b}$$

Pour l'onde sortante, on a, en se limitant au premier ordre en $\frac{1}{r}$:

$$\psi^{(\text{out})} = Af_{\mathbf{k}}(\Omega) \frac{e^{ikr}}{r} \tag{3.116a}$$

$$\nabla\psi^{(\text{out})} \simeq ik\psi^{(\text{out})} \mathbf{e}_r \tag{3.116b}$$

On a alors :

$$\mathbf{j}^{(\text{out})} = \frac{\hbar}{m} k |\psi^{(\text{out})}|^2 \mathbf{e}_r \tag{3.117a}$$

$$\mathbf{j}^{(\text{out})} = \frac{\hbar}{m} k |A|^2 \frac{|f_{\mathbf{k}}(\Omega)|^2}{r^2} \mathbf{e}_r \tag{3.117b}$$

On peut alors écrire $dN_s(\Omega)$ de deux façons :

$$dN_s(\Omega) = |\mathbf{j}^{(\text{out})}| r^2 d\Omega \tag{3.118a}$$

$$dN_s(\Omega) = |\mathbf{j}^{(\text{in})}| \frac{d\sigma}{d\Omega} d\Omega \tag{3.118b}$$



Ce qui nous donne alors le résultat général suivant pour la diffusion :

$$\frac{d\sigma}{d\Omega} = |f_{\mathbf{k}}(\Omega)|^2 \tag{3.119}$$

Lors de la collision d'un faisceau de particules de vecteur d'onde \mathbf{k} sur des cibles, on obtient pour la fonction d'onde des particules :

$$\psi_{\mathbf{k}}(\mathbf{r}) \simeq \underbrace{e^{i\mathbf{k}\cdot\mathbf{r}}}_{\substack{\text{Onde} \\ \text{plane} \\ \text{entrante}}} + f_{\mathbf{k}}(\Omega) \underbrace{\frac{e^{ikr}}{r}}_{\substack{\text{Onde} \\ \text{sphérique} \\ \text{sortante}}} \tag{3.120}$$

où

$$f_{\mathbf{k}}(\Omega) = -\frac{m}{2\pi\hbar^2} \int d\mathbf{r}' e^{-i\mathbf{k}\cdot\mathbf{r}'} V(\mathbf{r}') \psi_{\mathbf{k}}(\mathbf{r}') \tag{3.121}$$

On démontre alors que la section efficace de collision est :

$$\frac{d\sigma}{d\Omega} = |f_{\mathbf{k}}(\Omega)|^2 \tag{3.122}$$

3.4.3 Approximation de Born

L'approximation de Born consiste à ne considérer que le terme du premier ordre du développement en puissance du potentiel $V(\mathbf{r})$. Ceci donne

$$\begin{aligned} f_{\mathbf{k}}(\Omega) &\approx -\frac{m}{2\pi\hbar^2} \int e^{-i\mathbf{k}\cdot\mathbf{r}'} V(\mathbf{r}') \psi_k^0(\mathbf{r}') d^3\mathbf{r}' \\ &\approx -\frac{m}{2\pi\hbar^2} \int e^{-i\mathbf{q}\cdot\mathbf{r}'} V(\mathbf{r}') d^3\mathbf{r}', \quad \mathbf{q} = \mathbf{k}' - \mathbf{k} \\ &\approx -\frac{m}{2\pi\hbar^2} \langle \psi_{k'}^0 | V | \psi_k^0 \rangle \end{aligned}$$

Dans le cadre de l'approximation de Born, la section efficace différentielle vaut

$$\frac{d\sigma}{d\Omega} = \frac{m^2}{4\pi^2\hbar^4} \left| \int e^{-i\mathbf{q}\cdot\mathbf{r}'} V(\mathbf{r}') d^3\mathbf{r}' \right|^2, \quad \mathbf{q} = \mathbf{k}' - \mathbf{k} \tag{3.123}$$

3.5 Perturbation dépendante du temps

Dans cette section, on va s'intéresser à l'évolution d'un système dont l'évolution libre est connue sous l'effet d'un potentiel dépendant du temps. Ce potentiel est supposé faible, ce qui permet de le traiter comme perturbant assez faiblement le système.

Cette étude est extrêmement importante, puisqu'elle va nous permettre de comprendre quelle est la probabilité de transition entre deux états sous l'effet de ce potentiel perturbatif, ce qui culminera avec le résultat dit de la "règle d'or de Fermi", que nous utiliserons de manière extensive dans de nombreux autres chapitres.

3.5.1 Schéma en interaction

On considère un système décrit par un hamiltonien $H_s(t)$ constitué de deux termes, l'un indépendant du temps H_s^0 dont le spectre est connu et l'autre qui est une perturbation dépendante du temps $H'_s(t)$. On notera $|n\rangle$ le vecteur propre de H_s^0 associé à la valeur propre E_n^0 .



$$H_s(t) = H_s^0 + H'_s(t) \tag{3.124}$$

On suppose qu'à t_0 le système est dans l'état propre $|i\rangle$ de H_s^0 . On cherche à savoir quelle est la probabilité qu'à l'instant t le système soit dans l'état propre $|f\rangle$.

L'opérateur d'évolution pour le hamiltonien non perturbé s'écrit $U^0(t, t_0)$ tandis que l'on notera celui du hamiltonien total $U(t, t_0)$. On définit alors

$$W_I(t, t_0) = U^0(t_0, t)U(t, t_0) = (U^0(t, t_0))^{-1}U(t, t_0) \tag{3.125}$$

L'idée est d'isoler la contribution perturbative dans l'évolution et d'éliminer la contribution non perturbée liée à H_s^0 . L'état en schéma d'interaction est défini par

$$|\psi_I(t)\rangle = W_I(t, t_0)|\psi_s(t_0)\rangle. \tag{3.126}$$

La figure 3.6 présente ces différentes quantités.

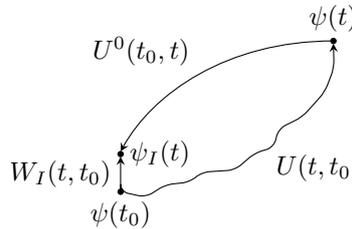


FIGURE 3.6 – Représentation de l'action de l'opérateur $W_I(t, t_0) = U^0(t_0, t)U(t, t_0)$

On peut déterminer l'équation d'évolution de l'opérateur W_I à partir de l'équation de Schrödinger. On obtient

$$i\hbar \frac{dW_I(t, t_0)}{dt} = H'_I(t)W_I(t, t_0) \tag{3.127}$$

où nous avons défini le hamiltonien en schéma d'interaction ; plus généralement, un opérateur $A_s(t)$ en schéma de Schrödinger s'écrira par définition en schéma d'interaction comme

$$A_I(t) = U^0(t_0, t)A_s(t)U^0(t, t_0) \tag{3.128}$$

et vérifiera l'équation d'évolution suivante

$$i\hbar \frac{dA_I(t)}{dt} = i\hbar \frac{\partial A_I(t)}{\partial t} + [A_I, H_s^0] \tag{3.129}$$

Isoler la contribution perturbative est particulièrement efficace lorsque l'on souhaite appliquer des méthodes perturbatives pour résoudre un problème où la perturbation peut être considérée comme "petite".

3.5.2 Méthode d'approximation

Probabilité de transition et graphes de Feynman

Le problème était de trouver la probabilité qu'à l'instant t le système soit dans l'état propre $|f\rangle$ sachant qu'il était dans l'état propre $|i\rangle$ de H_s^0 à t_0 . On sait que cette probabilité s'écrit

$$p_{i \rightarrow f} = |\langle f | \psi_s(t) \rangle|^2, \psi_s(t_0) = |i\rangle. \tag{3.130}$$

Il faut donc déterminer l'état à l'instant t . Faisons le lien avec le schéma en interaction. On rappelle que par construction du schéma en interaction et par définition de l'opérateur d'évolution, nous avons

$$|\psi_I(t)\rangle = W_I(t, t_0)|\psi_s(t_0)\rangle = U^0(t_0, t)U(t, t_0)|\psi_s(t_0)\rangle = W_I(t, t_0)|i\rangle \quad (3.131a)$$

$$|\psi_s(t)\rangle = U(t, t_0)|\psi_s(t_0)\rangle \quad (3.131b)$$

Ces deux équations donnent donc

$$|\psi_s(t)\rangle = U^0(t, t_0)|\psi_I(t)\rangle = U^0(t, t_0)W_I(t, t_0)|i\rangle \quad (3.132)$$

Faisons maintenant un développement perturbatif. En intégrant l'équation d'évolution de l'opérateur W_I , nous obtenons l'équation intégrale suivante

$$W_I(t, t_0) = \mathbb{1} + \frac{1}{i\hbar} \int_{t_0}^t H'_I(t')W_I(t', t_0)dt' \quad (3.133)$$

Mais nous pouvons réinjecter cette expression pour réécrire $W_I(t', t_0)$, et à nouveau recommencer. Cela donne

$$W_I(t, t_0) = \mathbb{1} + \frac{1}{i\hbar} \int_{t_0}^t H'_I(t')dt' + \left(\frac{1}{i\hbar}\right)^2 \int_{t_0}^t \int_{t_0}^{t'} H'_I(t')H'_I(t'')W_I(t'', t_0)dt''dt' \quad (3.134)$$

$$= \mathbb{1} + \frac{1}{i\hbar} \int_{t_0}^t H'_I(t')dt' + \left(\frac{1}{i\hbar}\right)^2 \int_{t_0}^t \int_{t_0}^{t'} H'_I(t')H'_I(t'')dt''dt' + \dots \quad (3.135)$$

Au premier ordre, nous avons une expression de $W_I(t, t_0)$ simplement en fonction de la perturbation

$$W_I(t, t_0) = \mathbb{1} + \frac{1}{i\hbar} \int_{t_0}^t H'_I(t')dt' \quad (3.136)$$

Cette expression va nous permettre de donner au premier ordre la probabilité de transition que l'on recherche. En effet,

$$\begin{aligned} |\psi_s(t)\rangle &= U^0(t, t_0)W_I(t, t_0)|i\rangle = U^0(t, t_0) \left(\mathbb{1} + \frac{1}{i\hbar} \int_{t_0}^t H'_I(t')dt' \right) |i\rangle \\ &= U^0(t, t_0) \left(|i\rangle + \frac{1}{i\hbar} \int_{t_0}^t H'_I(t')|i\rangle dt' \right) \\ &= e^{\frac{E_i^0}{i\hbar}(t-t_0)}|i\rangle + U^0(t, t_0) \frac{1}{i\hbar} \int_{t_0}^t H'_I(t')|i\rangle dt' \end{aligned} \quad (3.137)$$

Dans le cas où $|i\rangle \neq |f\rangle$, nous avons donc

$$\begin{aligned} \langle f|\psi_s(t)\rangle &= \frac{1}{i\hbar} \int_{t_0}^t \langle f|U^0(t, t_0)H'_I(t')|i\rangle dt' \\ &= \frac{1}{i\hbar} \int_{t_0}^t \langle f|U^0(t, t_0)U^0(t_0, t')H'_I(t')U^0(t', t_0)|i\rangle dt' \\ &= \frac{1}{i\hbar} \int_{t_0}^t \langle f|U^0(t, t')H'_I(t')U^0(t', t_0)|i\rangle dt' \end{aligned} \quad (3.138)$$

$$\begin{aligned} &= \frac{1}{i\hbar} \int_{t_0}^t \langle f|U^0(t, t')H'_I(t')e^{\frac{E_i^0}{i\hbar}(t'-t_0)}|i\rangle dt' \\ &= \frac{1}{i\hbar} \int_{t_0}^t e^{\frac{E_i^0}{i\hbar}(t'-t_0)} \langle U^0(t', t)f|H'_I(t')|i\rangle dt' \\ &= \frac{1}{i\hbar} \int_{t_0}^t e^{\frac{E_i^0}{i\hbar}(t'-t_0)} e^{\frac{E_f^0}{i\hbar}(t-t')} \langle f|H'_I(t')|i\rangle dt' \\ &= e^{-\frac{E_i^0}{i\hbar}t_0} e^{-\frac{E_f^0}{i\hbar}t} \frac{1}{i\hbar} \int_{t_0}^t e^{\frac{E_i^0 - E_f^0}{i\hbar}t'} \langle f|H'_I(t')|i\rangle dt' \end{aligned} \quad (3.139)$$



Une représentation des termes perturbatifs, notamment de l'équation 3.138 peut être faite. Ces diagrammes sont des diagrammes de Feynman, visibles sur la figure 3.7.

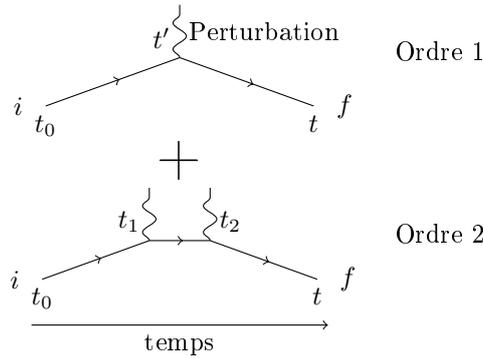


FIGURE 3.7 – Graphes de Feynman donnant les deux premiers termes perturbatifs. La branche i est associée à l'état initial $|i\rangle$. Cet état évolue librement entre t_0 et t' , via l'opérateur d'évolution $U^0(t', t_0)$. À l'instant t' , il y a une perturbation représentée par $H'_I(t')$. Puis on a à nouveau propagation libre entre t' et t , vers l'état final $|f\rangle$. La perturbation est possible pour tout t' . Ceci permet d'écrire de façon graphique l'équation 3.138. À l'ordre deux, on a deux perturbations à prendre en compte.

Finalement, le module au carré de cette expression nous donne la probabilité de transition au premier ordre que l'on recherche

$$p_{i \rightarrow f} = |\langle f | \psi_s(t) \rangle|^2 = \frac{1}{\hbar^2} \left| \int_{t_0}^t e^{\frac{E_i^0 - E_f^0}{i\hbar} \tau} \langle f | H'_s(\tau) | i \rangle d\tau \right|^2 \tag{3.140}$$

Cas où H'_s ne dépend pas explicitement du temps

Pour simplifier, prenons $t_0 = 0$ et supposons que la perturbation ne dépend pas explicitement du temps. La probabilité de transition vaut alors

$$p_{i \rightarrow f} = |\langle f | \psi_s(t) \rangle|^2 = \frac{1}{\hbar^2} |\langle f | H'_s | i \rangle|^2 f(t, \omega_{if}), \text{ avec } f(t, \omega) = \left[\frac{2}{\omega} \sin\left(\frac{\omega t}{2}\right) \right]^2 \tag{3.141}$$

où par définition, $\hbar\omega_{if} = E_i^0 - E_f^0$. On peut montrer la propriété suivante

$$\lim_{t \rightarrow \infty} f(t, \omega) = 2\pi t \delta(\omega) \tag{3.142}$$

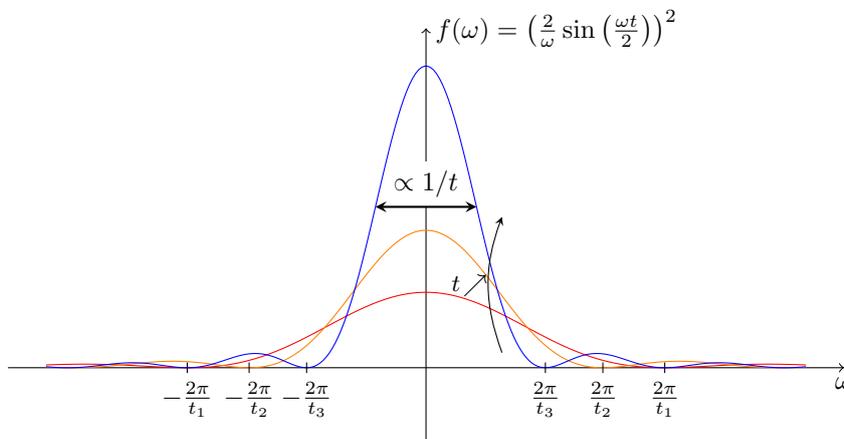


FIGURE 3.8 – Représentation de la fonction $f(t, \omega)$.



De plus, la forme de $f(t, \omega)$, visible sur la figure 3.8 permet d'affirmer qu'il n'y a essentiellement que des transitions pour $\omega_{if} \in [-\frac{2\pi}{t}, \frac{2\pi}{t}]$ ou de manière équivalente pour un écart d'énergie entre les états $|i\rangle$ et $|f\rangle$, $\Delta E \sim \hbar \Delta \omega \leq \frac{\hbar}{t}$. À la limite où t tend vers l'infini, les transitions ne se font donc que pour des états de même énergie. Expérimentalement, ceci permet de mesurer l'énergie de l'état final.

3.5.3 Règle d'or de Fermi

Dans ce qui précède, nous avons examiné la probabilité de transition d'un état propre à un autre. Cependant, si l'état final est un état de diffusion, il est impossible de l'isoler parfaitement. Nous allons donc nous intéresser à la probabilité de transition vers un ensemble \mathcal{F} d'états finals. Il faut donc sommer sur toutes les probabilités ou de manière équivalente intégrer en tenant compte de la densité d'énergie.

$$p_{i \rightarrow \mathcal{F}} = \int_{\mathcal{F}} p_{i \rightarrow f} \mathcal{D}(E_f^0, \beta) dE_f^0 d\beta \tag{3.143}$$

$$= \frac{1}{\hbar^2} \int_{\mathcal{F}} \mathcal{D}(E_f^0, \beta) |\langle f, \beta | H'_s | i \rangle|^2 f(t, \omega_{if}) dE_f^0 d\beta \tag{3.144}$$

où β est un paramètre nécessaire pour décrire complètement la base des états propres. Pour t assez grand, nous pouvons approximer la fonction f par la distribution de Dirac

$$f(t, \omega) = 2\pi t \hbar \delta(E_i^0 - E_f^0). \tag{3.145}$$

Ceci nous donne enfin

$$p_{i \rightarrow \mathcal{F}} = \begin{cases} 0 & \text{s'il n'y a pas d'état d'énergie } E_i^0 \text{ dans } \mathcal{F} \\ \frac{2\pi t}{\hbar} \int_{\mathcal{F}} \mathcal{D}(E_i^0, \beta) |\langle f, \beta | H'_s | i \rangle|^2 d\beta & \text{sinon} \end{cases} \tag{3.146}$$

Généralement, on discute de la probabilité par unité de temps. On énonce donc la *règle d'or de Fermi* en donnant

$$\Gamma_{i \rightarrow \mathcal{F}} = \frac{1}{\Delta \beta} \frac{dp_{i \rightarrow \mathcal{F}}}{dt} = \frac{2\pi}{\hbar} \mathcal{D}(E_i^0, \beta) |\langle f, \beta | H'_s | i \rangle|^2 \tag{3.147}$$

Dans cette formule, on ne tient compte que des états $|f, \beta\rangle$ d'énergie E_i^0 . Cette formule est valable pour t suffisamment grand pour pouvoir approximer la fonction f par la distribution de Dirac mais aussi pour t suffisamment petit pour que le développement perturbatif soit valable.

3.6 Opérateur densité

Jusqu'ici, nous n'avons pu nous intéresser qu'à un ensemble limité de situations décrites par la mécanique quantique : états à une seule particule, systèmes isolés, états "purs"... Or, il serait intéressant de se détacher de ces limites et d'avoir un formalisme permettant de décrire de manière complète tout type de système quantique. C'est ce que va nous permettre l'opérateur densité. Celui-ci va nous permettre notamment de décrire de manière très simple à la fois les états purs mais aussi les mélanges statistiques d'états. Ces états "mélanges" sont très souvent rencontrés expérimentalement, que ce soit parce que la préparation du système peut créer plusieurs états avec une certaine probabilité¹¹, ou suite à l'effet d'une interaction avec un environnement (ceci correspond au phénomène de décohérence, qui ne sera pas traité ici car allant bien au delà du programme du cours...). Le fait de pouvoir décrire une situation dans laquelle plusieurs états sont possibles avec une probabilité associée à chaque état permet de se convaincre que l'opérateur densité est également la formulation adaptée pour la mécanique statistique.

3.6.1 États intriqués

L'intrication est au cœur de la mécanique quantique¹² : c'est une conséquence du principe de superposition. Elle apparaît en fait dès qu'on considère deux systèmes en interaction : dans ce cas, il n'est plus possible de parler des états des systèmes de manière séparée mais seulement du système total.

¹¹. Attention, cette situation est totalement différente de celle d'une superposition quantique d'état, comme on le verra dans cette section !

¹². Comme on l'a dit en introduction de ce chapitre, il s'agit même de l'une des deux grandes raisons de l'étrangeté quantique.



On considère deux systèmes physiques \mathcal{S}_1 et \mathcal{S}_2 décrits respectivement par les espaces des états \mathcal{H}_1 et \mathcal{H}_2 . On s'intéresse maintenant au système complet $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2\}$, et on fait le postulat suivant concernant l'espace des états de ce système, motivé par le principe de superposition :

L'espace des états \mathcal{H} du système $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2\}$ est le produit tensoriel des espaces \mathcal{H}_1 et \mathcal{H}_2 .

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \tag{3.148}$$

Un cas particulier important est celui où un état $|\psi\rangle$ de \mathcal{H} est factorisable, c'est-à-dire qu'il s'écrit comme le produit tensoriel d'un élément de \mathcal{H}_1 et d'un de \mathcal{H}_2 , $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$. Dans ce cas, il est possible de parler de l'état d'un des sous-systèmes. On dira que le système \mathcal{S}_1 est dans l'état $|\psi_1\rangle$ et que le système \mathcal{S}_2 est dans l'état $|\psi_2\rangle$. Les deux sous-systèmes sont *indépendants*. Cependant, dans le cas général, dès que l'on a une interaction entre les deux systèmes, rien ne permet de dire qu'un état de \mathcal{H} puisse être factorisé de cette manière. Si l'état n'est pas factorisable, on dit que nous avons *un état intriqué*.

Considérons une base pour chacun des espaces de Hilbert, $(|\phi_i^{(1)}\rangle)_i$ et $(|\phi_j^{(2)}\rangle)_j$, l'état du système \mathcal{S} s'écrit sous la forme

$$|\phi\rangle = \sum_{ij} c_{ij} (|\phi_i^{(1)}\rangle)_i \otimes (|\phi_j^{(2)}\rangle)_j \tag{3.149}$$

Ce n'est que lorsque les coefficients c_{ij} peuvent se mettre sous la forme $c_{ij} = \alpha_i \beta_j$ que l'état du système est factorisé.

En guise d'exemple, considérons deux spins demi-entier dans l'état $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ où $(|0\rangle, |1\rangle)$ désigne une base de l'espace des états d'un spin. Un tel état est appelé *état de Bell* et est au coeur par exemple du paradoxe d'Einstein, Podolsky, et Rosen (EPR). Tout d'abord, cet état n'est pas factorisé et il nous est donc impossible de parler de l'état d'un seul des spins. De plus, cette intrication à la particularité de montrer une corrélation parfaite entre les résultats de mesure sur les spins¹³.

3.6.2 Opérateur densité : définitions

La notion d'état intriqué montre qu'il n'est pas possible dans le cas général de parler de l'état d'un des sous-systèmes pour un système composé. On aimerait pourtant en avoir une description.

Dans un autre registre, lorsque l'on dispose d'une information complète sur un système, son état est décrit par un état de l'espace de Hilbert. Mais que se passe-t-il lorsque ne nous disposons que d'une information partielle sur le système? Ce manque d'information provient de la mesure quantique qui est de nature probabiliste, contrairement à une mesure classique où aucune limitation sur la précision n'existe. L'analogie de la densité de probabilité classique est la matrice densité.

Un système quantique est décrit par un opérateur, appelé *opérateur densité*, $\rho(t)$ vérifiant les propriétés suivantes

$$\rho^\dagger(t) = \rho(t) \quad \text{Tr}(\rho(t)) = 1 \quad \langle \psi | \rho(t) | \psi \rangle \geq 0 \quad \text{Tr}(A\rho(t)) = \langle A \rangle_{\rho(t)} \tag{3.150}$$

$\rho(t)$ étant un opérateur hermitien, il est diagonalisable en base orthonormée et ses valeurs propres sont réelles. Dans une telle base $\{|\psi_i\rangle\}_i$, nous avons

$$\rho(t) = \sum_i p_i |\psi_i\rangle \langle \psi_i| \tag{3.151}$$

D'après les propriétés de l'opérateur densité, les p_i sont positifs et de somme unité. On peut donc les interpréter comme des probabilités. L'opérateur densité représente donc un mélange statistique des états ψ_i avec des probabilités p_i . Cette interprétation n'est valable que dans une base orthonormée.

3.6.3 États purs, cohérences et mélange

En mécanique quantique, l'état d'un système est décrit par un élément d'un espace de Hilbert. Dans le contexte de l'opérateur densité, un élément de l'espace des états est appelé *état pur*. Un système dans l'état pur normé $|\psi\rangle$ est alors décrit par l'opérateur densité $\rho = |\psi\rangle \langle \psi|$ qui est un projecteur. De manière générale

¹³. Voir le cours de *Physique, Information et Calcul* pour plus d'informations sur l'intrication.



L'opérateur densité décrit un *état pur* s'il vérifie $\rho(t)^2 = \rho(t)$, en d'autres termes, si $\rho(t)$ est un projecteur.

D'après cette définition, nous avons donc : $\forall i, p_i^2 = p_i$ donc $p_i = 0, 1$. Or comme la somme des probabilités vaut 1, un seul des p_i n'est pas nul. Finalement, pour un état pur, $\rho(t) = |\psi\rangle\langle\psi|$. Une autre manière de caractériser un état pur est de montrer que $Tr(\rho^2(t)) = 1$.

Dans le cas où l'opérateur densité ne décrit pas un état pur, on dit que le système est dans un *mélange statistique*. Les propriétés de l'opérateur densité nous permettent, à tout instant, de le diagonaliser, ses valeurs propres étant toutes positives sommées à l'unité. En notant $|\phi_n\rangle$ les vecteurs propres associés aux valeurs propres p_n à l'instant t , nous avons $\rho(t) = \sum_n p_n |\phi_n\rangle\langle\phi_n|$. Sous cette forme, il est du coup possible de dire que l'opérateur densité est un mélange statistique des états $|\phi_n\rangle$ avec les poids p_n .

Analysons un exemple. Considérons un système, par exemple un spin demi-entier, dans la superposition d'état $|\psi\rangle = \sqrt{\frac{3}{2}} (|0\rangle + \frac{1}{2}|1\rangle)$. C'est un état pur représenté par l'opérateur densité

$$\rho = |\psi\rangle\langle\psi| \tag{3.152}$$

$$= \frac{3}{2} \left(|0\rangle\langle 0| + \frac{1}{4}|1\rangle\langle 1| + \frac{1}{2}|0\rangle\langle 1| + \frac{1}{2}|1\rangle\langle 0| \right) \tag{3.153}$$

Matriciellement, dans la base $(|0\rangle, |1\rangle)$:

$$\rho = \frac{3}{2} \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} \end{pmatrix} \tag{3.154}$$

Chaque terme de la matrice densité possède une interprétation physique claire. Les *populations* de chaque état sont données par $\text{Prob}(|0\rangle) = \text{tr}(|0\rangle\langle 0|\rho) = \frac{3}{2}$ et $\text{Prob}(|1\rangle) = \text{tr}(|1\rangle\langle 1|\rho) = \frac{3}{2} \cdot \frac{1}{4}$. Ils correspondent en fait aux termes diagonaux de l'opérateur densité; la partie diagonale correspond aux projecteurs sur chacun des états. Ils est alors clair que les termes hors diagonaux sont caractéristiques des interférences entre deux états et sont par conséquent appelé *cohérences*.

3.6.4 Opérateur densité réduit

L'opérateur densité total d'un système composé nous permet d'en définir un nouveau pour chacun des sous-systèmes, ce que ne nous permettait pas de faire la description en terme de kets.

Considérons deux systèmes agissant sur l'espace de Hilbert $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. On cherche à définir un opérateur densité associé à l'un des sous-systèmes, par exemple le premier. Pour cela, considérons une grandeur physique agissant uniquement sur ce système 1. L'opérateur associé à cette grandeur s'écrit donc $A_1 \otimes \mathbb{1}_2$. On cherche un opérateur densité $\rho^{(1)}$ agissant sur \mathcal{H}_1 tel que

$$\langle A_1 \rangle = \text{Tr}(\rho^{(1)} A_1). \tag{3.155}$$

Une piste est donnée grâce aux définitions des moyennes

$$\begin{aligned} \langle A_1 \otimes \mathbb{1}_2 \rangle &= \text{Tr}(\rho A_1 \otimes \mathbb{1}_2) = \sum_{n_1 m_1 n_2 m_2} A_{n_1 m_1} \delta_{n_2 m_2} \rho_{n_1 m_1 n_2 m_2} \\ &= \sum_{n_1 m_1} A_{n_1 m_1} \sum_{n_2 m_2} \delta_{n_2 m_2} \rho_{n_1 m_1 n_2 m_2} \end{aligned} \tag{3.156a}$$

$$\begin{aligned} \langle A_1 \rangle &= \langle A_1 \otimes \mathbb{1}_2 \rangle = \text{Tr}(A \rho^{(1)}) \\ &= \sum_{n_1 m_1} A_{n_1 m_1} \rho_{n_1 m_1}^{(1)} \end{aligned} \tag{3.156b}$$

L'opérateur densité associé à la particule 1 est donc la *trace partielle* de l'opérateur densité ρ sur l'espace \mathcal{H}_2 :

$$\rho_{n_1 m_1}^{(1)} = \sum_{n_2 m_2} \delta_{n_2 m_2} \rho_{n_1 m_1 n_2 m_2} \tag{3.157}$$



La trace partielle nous permet d'avoir une description d'un des sous-systèmes seul en tenant compte implicitement de son intrication avec l'autre système.

Soit un système $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2\}$ dont l'espace des états est donné par $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. L'opérateur densité du système total ρ permet de définir un opérateur densité associé à l'un des sous-système par trace partielle sur les degrés de libertés des autres sous-systèmes. Pour le premier sous-système, cet *opérateur densité réduit* est donné par

$$\rho^{(1)} = Tr_{\mathcal{H}_2}(\rho) \tag{3.158}$$

Étudions un exemple. Soit deux spins 1/2 dans un état de Bell, $|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. On rappelle qu'un tel état amène une corrélation parfaite entre les résultats d'une mesure faite sur ces spins. Si on ne s'intéresse par exemple qu'à un seul des spins, son état est décrit par un opérateur densité réduit obtenu à partir de la trace partielle sur les degrés de liberté de l'autre spin de l'opérateur densité total ρ défini à l'aide de $|\psi\rangle$. Pour le premier spin,

$$\rho_1 = tr_{\mathcal{H}_2}(\rho) = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) \tag{3.159}$$

On obtient un mélange statistique des états $|0\rangle$ et $|1\rangle$ avec probabilité 1/2 pour les deux. Un point important à remarquer est que les cohérences quantiques sont nulles! Cela se traduit physiquement par le fait que les interférences quantiques entre les deux états sont absentes. Le fait de ne pas tenir compte du second spin et donc de ne capturer qu'une partie de l'information sur l'état du système détruit les effets quantiques.

La décohérence repose sur le même genre d'idée. L'interaction de l'observateur avec son environnement intrique ces deux systèmes. Cependant, lors d'une mesure, il nous est impossible, en pratique, d'observer l'environnement tout entier qui est constitué d'un nombre de degrés de liberté immense. Cette fuite de l'information causée par l'intrication avec l'environnement détruit les effets quantiques et permet de faire émerger la réalité classique. Formellement, nous devons prendre la trace partielle de l'opérateur densité système-observateur-environnement sur les degrés de liberté environnementaux.

3.6.5 Évolution temporelle

Pour déterminer l'évolution temporelle de la matrice densité, on peut par exemple, supposer qu'elle décrit un état pur et dériver l'équation d'évolution à l'aide de l'équation de Schrödinger.

Pour un système fermé décrit par le hamiltonien H , la matrice densité ρ évolue selon ^a

$$i\hbar \frac{d\rho(t)}{dt} = [H(t), \rho(t)] \tag{3.160}$$

a. Attention, cette équation d'évolution est différente de celle du schéma de Heisenberg!

Formellement, nous avons équivalence avec la relation

$$\rho(t) = U(t, t_0)\rho(t_0)U(t_0, t) \tag{3.161}$$

Puisque l'évolution est unitaire, la trace est conservée. Ceci implique donc, en utilisant la caractérisation d'un état pur par la trace, qu'un état pur reste pur au cours du temps. La diagonalisation de $\rho(t)$ dans une base orthonormée permet enfin de conclure à la conservation des probabilités.

On peut remarquer l'analogie formelle avec l'équation de Liouville qui nous donne l'évolution d'une densité de probabilité dans le cadre de la mécanique classique. Ceci permet de donner une autre motivation pour voir l'opérateur densité comme base de la mécanique statistique (quantique mais aussi classique).

3.6.6 Mesures et matrice densité

Reformulons les postulats concernant la mesure en mécanique quantique avec le formalisme de l'opérateur densité.



La probabilité $\text{Prob}(|\chi\rangle)$ de trouver le système quantique dans l'état $|\chi\rangle$ est donnée par $\text{Prob}(|\chi\rangle) = \text{tr}(\rho \mathcal{P}_\chi)$ où $\mathcal{P}_\chi = |\chi\rangle\langle\chi|$ (conséquence de la définition de l'opérateur densité). Si on trouve a_n en mesurant A , alors l'opérateur densité du système juste après la mesure est donnée par

$$\rho' = \frac{\mathcal{P}_n \rho \mathcal{P}_n}{\text{Tr}(\mathcal{P}_n \rho)} \tag{3.162}$$

La forme de l'opérateur densité après une mesure donnée ici est l'analogie du postulat de réduction du paquet d'onde. Dans le cadre de la décohérence, ce postulat n'est plus vraiment nécessaire.

Vérifions ces résultats pour un état pur. Nous avons $\rho = |\psi\rangle\langle\psi|$ avec $|\psi\rangle = \sum_k \langle\psi_k|\psi\rangle |\psi_k\rangle = \sum_k \alpha_k |\psi_k\rangle$. Donc $\rho = \sum_{k, k'} \alpha_k^* \alpha_{k'} |\psi_k\rangle\langle\psi_{k'}|$. La probabilité de trouver α_n est $\text{tr}(\rho \mathcal{P}_n) = \sum_k \langle\psi_k|\rho \mathcal{P}_n|\psi_k\rangle = \langle\psi_n|\rho|\psi_n\rangle = |\langle\psi_n|\psi\rangle|^2$. Enfin, on vérifie facilement que $\rho' = \frac{\mathcal{P}_n \rho \mathcal{P}_n}{\text{tr}(\mathcal{P}_n \rho)} = |\psi_n\rangle\langle\psi_n|$

3.7 Opérateurs de champs

On va dans cette section considérer un système physique à N particules. En physique des particules ou en théorie des champs, le nombre N n'est pas fixé, et il faut donc pouvoir décrire de manière simple les opérations ajoutant ou détruisant des particules. De plus, puisque l'on a des bosons ou des fermions, la fonction d'onde correspondant au système doit posséder des symétries spécifiques : elle doit être symétrique par permutation pour les bosons tandis qu'elle doit être antisymétrique pour les fermions. Cependant, dans le formalisme traditionnel, toutes ces opérations sont difficiles. L'intérêt de l'introduction des opérateurs de champs réside dans le fait que toutes ces contraintes sont directement incorporées dans le formalisme. Avant de donner toutes les définitions, on va s'attarder sur la résolution de l'oscillateur harmonique qui permet d'introduire naturellement ces opérateurs.

3.7.1 Oscillateur harmonique

L'étude de l'oscillateur harmonique est d'une grande importance puisqu'il décrit le comportement d'un système physique autour d'une position d'équilibre stable. En effet, le développement à l'ordre deux d'un potentiel autour d'une position d'équilibre redonne un comportement harmonique. Le hamiltonien de l'oscillateur harmonique est le suivant

$$H = \frac{p^2}{2m} + \frac{1}{2} m \omega^2 q^2 \tag{3.163}$$

où ω est la pulsation de l'oscillateur. On cherchera juste dans la suite à diagonaliser ce hamiltonien et trouver les énergies propres. On va réécrire ce hamiltonien d'une autre manière en définissant de nouveaux opérateurs. Pour l'instant,

$$H = \frac{\hbar\omega}{2} \left(\frac{p^2}{m\hbar\omega} + \frac{m\omega}{\hbar} q^2 \right) \tag{3.164}$$

En posant $P = \frac{p}{(m\hbar\omega)^{1/2}}$ et $Q = q(\frac{m\omega}{\hbar})^{1/2}$, qui vérifient la relation de commutation $[Q, P] = i\hbar$, on adimensionne les opérateurs. Le hamiltonien s'écrit alors

$$H = \frac{\hbar\omega}{2} (P^2 + Q^2) \tag{3.165}$$

En faisant une analogie avec les nombres complexes, on peut voir le terme $\frac{1}{2} (P^2 + Q^2)$ comme résultant du calcul d'un module. Mais attention, ici, on raisonne avec des opérateurs qui ne commutent pas. Néanmoins, cette analogie nous amène à définir les opérateurs dits de *création* et d'*annihilation* par

$$a = \frac{1}{\sqrt{2}} (P + iQ) \tag{3.166a}$$

$$a^\dagger = \frac{1}{\sqrt{2}} (P - iQ) \tag{3.166b}$$



qui vérifient la relation de commutation

$$[a, a^\dagger] = \mathbb{1} \tag{3.167}$$

En utilisant la définition des opérateurs création et annihilation ainsi que la relation de commutation des opérateurs adimensionnés, le hamiltonien s'écrit

$$H = \hbar\omega \left(a^\dagger a + \frac{1}{2} \right) = \hbar\omega \left(N + \frac{1}{2} \right) \tag{3.168}$$

où par définition

$$N = a^\dagger a \tag{3.169}$$

est l'opérateur *nombre de particules*. Ainsi, pour diagonaliser H , il suffit de diagonaliser N , ce qui, par ailleurs, justifiera toutes les dénominations introduites. Diagonalisons N . Soit ν une valeur propre de N associé au vecteur propre $|\nu\rangle$, c'est-à-dire

$$N|\nu\rangle = \nu|\nu\rangle \tag{3.170}$$

Que peut-on dire sur cette valeur propre? Tout d'abord, nous avons

$$0 \leq \|a|\nu\rangle\|^2 = \langle \nu | a^\dagger a | \nu \rangle = \langle \nu | N | \nu \rangle = \nu \langle \nu | \nu \rangle \tag{3.171}$$

Donc $\nu \geq 0$. De plus si $\nu = 0$, $a|\nu\rangle = 0$. On peut montrer que $[N, a] = -a$. Ceci permet de montrer que

$$Na|\nu\rangle = (-a + aN)|\nu\rangle = a(N - 1)|\nu\rangle = (\nu - 1)a|\nu\rangle \tag{3.172}$$

$a|\nu\rangle$ est donc un vecteur propre de N pour la valeur propre $\nu - 1$. C'est pour cela que cet opérateur est appelé opérateur d'annihilation. Nous pouvons faire le même travail avec a^\dagger et montrer qu'il est vecteur propre de N pour la valeur propre $\nu + 1$, d'où son nom d'opérateur création. En itérant l'application de ces opérateurs, nous avons une suite $((a^p|\nu\rangle))_{p \in \mathbb{N}}$ de vecteurs propres de N associée à la suite de valeurs propres $(\nu - p)_{p \in \mathbb{N}}$ et une autre $((a^\dagger)^p|\nu\rangle)_{p \in \mathbb{N}}$ associée à la suite de valeurs propres $(\nu + p)_{p \in \mathbb{N}}$. Or, nous avons la condition supplémentaire que les valeurs propres de N doivent être positives. Ceci implique nécessairement que ses valeurs propres soient entières car sinon la suite de valeurs propres ci-dessus aurait des éléments négatifs. Cela étant dit, nous avons finalement

$$sp(N) = \mathbb{N} \tag{3.173}$$

En notant $|n\rangle$ le vecteur propre de N , aussi vecteur propre de H , associé à la valeur propre entière n , on a

$$H|n\rangle = \hbar\omega \left(N + \frac{1}{2} \right) |n\rangle = \hbar\omega \left(n + \frac{1}{2} \right) |n\rangle \tag{3.174}$$

Les énergies de l'oscillateur harmonique sont donc discrètes et valent

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right) \tag{3.175}$$

Pour établir ce résultat, nous avons supposé l'existence d'une valeur propre, ce qui n'est pas *a priori* acquis. En fait, on peut montrer que l'opérateur N a bien au moins une valeur propre, que les vecteurs propres forment une base orthogonale de l'espace de Hilbert et qu'ils sont non dégénérés. En vertu de ce qui a été établi, nous avons

$$a^\dagger |n\rangle \propto |n + 1\rangle \quad \|a^\dagger |n\rangle\|^2 = n + 1 \tag{3.176}$$

Une fois pour toutes on pose

$$a^\dagger |n\rangle = \sqrt{n + 1} |n + 1\rangle \tag{3.177a}$$

$$a |n\rangle = \sqrt{n} |n - 1\rangle \tag{3.177b}$$

Cette définition assure que a et a^\dagger soient bien adjoints.



3.7.2 Système de N particules identiques

On s'intéresse maintenant à un système de N particules identiques, et on va chercher à donner une formulation adéquate des états, qui rendent compte de toutes les contraintes.

Sur les permutations

L'espace de Hilbert du système est

$$\mathcal{H}^N = \bigotimes_{i=1}^N \mathcal{H} \tag{3.178}$$

Un état est noté

$$|\psi\rangle = \bigotimes_{i=1}^N |\psi_i\rangle = |\psi_1 \cdots \psi_N\rangle \tag{3.179}$$

Développons la manière qu'ont les états de se transformer sous une permutation. Par définition, l'action d'une permutation $\sigma \in \mathfrak{S}_n$ est définie *via* un opérateur noté U_σ tel que

$$U_\sigma |\psi_1 \cdots \psi_n\rangle = |\psi_{\sigma(1)} \cdots \psi_{\sigma(n)}\rangle \tag{3.180}$$

Ces opérateurs sont des opérateurs unitaires vérifiant $U_Q \circ U_P = U_{P \circ Q}$ ainsi que $U_\sigma^{-1} = U_{\sigma^{-1}}$. Quant aux fonctions d'onde en représentation position, elles se transforment comme

$$\begin{aligned} \psi_\sigma(\mathbf{r}_1 \cdots \mathbf{r}_N) &= \langle \mathbf{r}_1 \cdots \mathbf{r}_N | U_\sigma | \psi_1 \cdots \psi_n \rangle \\ &= \langle \mathbf{r}_{\sigma^{-1}(1)} \cdots \mathbf{r}_{\sigma^{-1}(N)} | \psi_1 \cdots \psi_n \rangle \\ &= \psi(\mathbf{r}_{\sigma^{-1}(1)} \cdots \mathbf{r}_{\sigma^{-1}(N)}) \end{aligned} \tag{3.181}$$

Concernant les observables, intéressons nous à ceux qui sont indépendants de l'ordre des états, c'est-à-dire ceux vérifiant

$$\theta = U_\sigma^\dagger \theta U_\sigma \Leftrightarrow [\theta, U_\sigma] = 0 \tag{3.182}$$

Comme le carré de l'opérateur U_σ est égal à l'identité, son spectre est inclus dans l'ensemble $\{-1, 1\}$. Les vecteurs propres associés aux valeurs propres $+1$ et -1 sont respectivement

$$|\psi_S\rangle = \frac{1}{2} (1 + U_\sigma) |\psi\rangle \quad |\psi_A\rangle = \frac{1}{2} (1 - U_\sigma) |\psi\rangle \tag{3.183}$$

On montre facilement en calculant le produit scalaire que ces états sont orthogonaux. Si une observable θ est invariante, on a $\langle \psi_S | \theta | \psi_A \rangle = 0$. Dans le cas particulier où le hamiltonien est invariant, cette égalité permet de dire qu'un état symétrique évolue toujours vers un état symétrique et ne peut se transformer en un état antisymétrique, et vice-versa.

Pour pouvoir symétriser une fonction ou l'antisymétriser, on utilise deux opérateurs, le symétriseur et l'antisymétriseur définis par

$$S_N = \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_N} U_\sigma \quad A_N = \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_N} \epsilon(\sigma) U_\sigma \tag{3.184}$$

Ces opérateurs sont des projecteurs orthogonaux qui commutent. Pour plus de détails sur le groupe des permutations, voir l'appendice B.

Bosons et fermions

Les états quantiques d'un système de N particules identiques sont soit complètement symétriques soit complètement antisymétriques dans l'échange de deux particules. On appelle *boson* une particule dont la fonction d'onde est symétrique et *fermion* une particule dont la fonction d'onde est antisymétrique. On peut montrer en théorie quantique des champs qu'une particule de spin entier est un boson et une particule de spin demi-entier est un fermion.

Représentation en nombre d'occupation

Soit $(e_i)_i$ une base de l'espace de Hilbert à une particule. Dans un système à N particules, on appelle *nombre d'occupation* n_i le nombre de particules dans l'état e_i . Bien sûr, nous avons la condition $\sum_{k=1}^{\infty} n_k = N$. Un des avantages de cette représentation est qu'elle inclut l'indiscernabilité. Un état $|\psi\rangle$, élément de l'espace de Hilbert à N particules \mathcal{H}^N , est représenté de la façon suivante

$$|\psi\rangle = |\underbrace{e_1, \dots, e_1}_{n_1}, \underbrace{e_2, \dots, e_2}_{n_2}, \dots\rangle = |e_1\rangle^{n_1} \dots |e_N\rangle^{n_N} \tag{3.185}$$

Nous savons que suivant le caractère bosonique ou fermionique, nous devons appliquer les opérateurs de symétrisation ou d'antisymétrisation. Pour les bosons, nous avons

$$|n_1, n_2, \dots\rangle = CS_N|\psi\rangle \tag{3.186}$$

où C est une constante de normalisation que l'on détermine de la façon suivante

$$\begin{aligned} 1 &= |C|^2 \langle \psi | S_N^\dagger S_N | \psi \rangle \\ &= |C|^2 \langle \psi | S_N | \psi \rangle \\ &= |C|^2 \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_N} \langle e_1, \dots, e_1, e_2, \dots, e_2, \dots | e_{\sigma(1)}, \dots, e_{\sigma(n_1)}, e_{\sigma(n_1+1)}, \dots, e_{\sigma(n_1+n_2)}, \dots \rangle \\ &= |C|^2 \frac{1}{N!} \prod_{i=1}^{\infty} n_i! \end{aligned} \tag{3.187}$$

ce qui nous donne l'état normalisé pour les bosons

$$|n_1, n_2, \dots\rangle = \sqrt{\frac{N!}{\prod_{i=1}^{\infty} n_i!}} S_N |\psi\rangle \tag{3.188}$$

La même démarche nous donne l'état normalisé pour les fermions, en se rappelant que dans ce cas $n_i \in \{0, 1\}$

$$|n_1, n_2, \dots\rangle = \sqrt{N!} A_N |\psi\rangle \tag{3.189}$$

3.7.3 Particules identiques en nombre indéterminé

On cherche à passer d'un système à n particules à un autre, à $n + 1$ particules, ainsi qu'à déterminer ce qu'est le vide. Ce problème est motivé par la possibilité de création de paires en physique des particules par exemple.

Espace des états

On note \mathcal{H}^0 l'espace de Hilbert sans particules constitué d'un seul état noté $|0\rangle$ que l'on appellera le vide ($\mathcal{H}^0 \cong \mathbb{C}$), ainsi que \mathcal{H}^n un espace d'Hilbert pour un système à n particules. C'est naturellement un produit tensoriel de n espaces de Hilbert. Nous considérons alors un espace de Hilbert que l'on note \mathcal{H}^∞ où le nombre de particules est inconnu et peut varier

$$\mathcal{H}^\infty = \bigotimes_{n=0}^{\infty} \mathcal{H}^n \tag{3.190}$$

Une remarque importante est à faire sur le vide $|0\rangle$. Cet état est bien un état particulier de l'espace de Hilbert et n'est pas égal au zéro de cet espace

$$|0\rangle \neq 0_{\mathcal{H}^\infty} \tag{3.191}$$

Nous pouvons dans cet espace considérer un opérateur N donnant le nombre de particules dans un état donné. Les sous-espaces propres de cet opérateur sont les espaces de Hilbert \mathcal{H}^n qui sont orthogonaux deux à deux. Nous avons donc

$$\forall |\psi\rangle \in \mathcal{H}^n \subset \mathcal{H}^\infty, N|\psi\rangle = n|\psi\rangle \tag{3.192}$$

Dans l'espace construit, nous devons faire la distinction entre bosons et fermions. Il suffit d'appliquer la généralisation des opérateurs de symétrisation et d'antisymétrisation définis comme

$$S = \mathbb{1} \otimes S_1 \otimes \dots \otimes S_n \otimes \dots \tag{3.193a}$$

$$A = \mathbb{1} \otimes A_1 \otimes \dots \otimes A_n \otimes \dots \tag{3.193b}$$

Les sous espaces invariants par ces opérateurs sont les espaces associés respectivement aux bosons noté $\mathcal{H}_{\text{Bosons}}^\infty$ et aux fermions noté $\mathcal{H}_{\text{Fermions}}^\infty$. On les appelle *espaces de Fock* et ce sont les espaces des états physiques pour les bosons et les fermions en nombre indéterminé. Dans la description en nombre d'occupation, nous avons tout de même la restriction $\sum_{i=1}^\infty n_i < \infty$. Un état $|n_1, n_2, \dots\rangle$ est appelé *état de Fock* et l'ensemble de tous ces états forme une base de l'espace. En toute généralité, un état quelconque de l'espace de Fock est combinaison linéaire de ces états.

Opérateurs de création et d'annihilation

Intéressons nous au cas des bosons. Soit un système de bosons identiques en nombre indéterminé. On définit les opérateurs de création et d'annihilation associés à la base $(|u_i\rangle)_{i \in I}$ de \mathcal{H} , espace de Hilbert à une particule

$$a_i^\dagger |n_1, \dots, n_i, \dots\rangle = \sqrt{n_i + 1} |n_1, \dots, n_i + 1, \dots\rangle \tag{3.194a}$$

$$a_i |n_1, \dots, n_i, \dots\rangle = \sqrt{n_i} |n_1, \dots, n_i - 1, \dots\rangle \tag{3.194b}$$

La notation est bien cohérente puisque ces opérateurs sont bien adjoints l'un de l'autre. Physiquement, a_i^\dagger crée une particule dans l'état $|u_i\rangle$, tandis que a_i en détruit une dans cet état. Nous avons les propriétés suivantes

$$[a_i, a_j] = 0 \quad [a_i^\dagger, a_j^\dagger] = 0 \quad [a_i, a_j^\dagger] = \delta_{ij} \tag{3.195}$$

Grâce à ces opérateurs, nous pouvons réécrire l'opérateur associé au nombre de particules. L'opérateur N_i donnant le nombre de particules dans l'état $|u_i\rangle$ vaut $N_i = a_i^\dagger a_i$. Finalement, l'opérateur N donnant le nombre de particules total est

$$N = \sum_{i=1}^\infty N_i = \sum_{i=1}^\infty a_i^\dagger a_i \tag{3.196}$$

En particulier nous avons $N|0\rangle = 0$. Cela étant nous avons les propriétés de commutation suivantes

$$[N, a_i^\dagger] = a_i^\dagger \quad [N, a_i] = -a_i \tag{3.197}$$

Enfin, nous pouvons réécrire un état de Fock comme

$$|n_1, \dots, n_i, \dots\rangle = \frac{1}{\sqrt{\prod_{i=1}^\infty n_i!}} \prod_{i=1}^\infty (a_i^\dagger)^{n_i} |0\rangle \tag{3.198}$$



Construit de cette manière, l'état est nécessairement symétrique comme il se doit pour un système de bosons. Maintenant, comment peut-on écrire l'action des opérateurs de création et d'annihilation pour un état quelconque de l'espace de Fock? Pour arriver au résultat, il faut se rappeler la signification des notations. L'état $|n_1, \dots, n_i + 1, \dots\rangle$ est un produit tensoriel. Explicitement, en utilisant la définition 3.186

$$|n_1, \dots, n_i + 1, \dots\rangle = S(|u_1\rangle^{n_1} \dots |u_i\rangle^{n_i+1} \dots) = S(|u_i\rangle |u_1\rangle^{n_1} \dots |u_i\rangle^{n_i} \dots) = S(|u_i\rangle |n_1, \dots, n_i, \dots\rangle) \tag{3.199}$$

Nous pouvons donc réécrire l'action de l'opérateur création bosonique comme

$$a_i^\dagger |n_1, \dots, n_i, \dots\rangle = \sqrt{n_i + 1} S(|u_i\rangle |n_1, \dots, n_i, \dots\rangle) \tag{3.200}$$

Ainsi, comme les états de Fock forment une base de l'espace de Fock, nous avons

$$\forall |\psi\rangle \in \mathcal{H}_{\text{Bosons}}^\infty, a_i^\dagger |\psi\rangle = \sqrt{n_i + 1} S(|u_i\rangle |\psi\rangle) \tag{3.201}$$

Nous pouvons aller plus loin et généraliser les opérateurs tels que nous les avons définis. Jusqu'alors, les opérateurs que nous avons considéré créent ou annihilent des particules dans l'état $|u_i\rangle$. Nous pouvons, grâce à la relation précédente, définir des opérateurs créant ou annihilant des particules dans un état quelconque $|u\rangle$ que l'on notera $a^\dagger(u)$. Comme l'ensemble des états $(|u_i\rangle)_{i \in I}$ forme une base de l'espace de Hilbert \mathcal{H} à une particule, nous avons directement

$$\forall |\psi\rangle \in \mathcal{H}_{\text{Bosons}}^\infty, a^\dagger(u) |\psi\rangle = \sqrt{N} S(|u\rangle |\psi\rangle) \tag{3.202}$$

Nous pouvons faire le même type de raisonnement pour l'opérateur d'annihilation et arriver à définir un opérateur détruisant une particule dans l'état $|u\rangle$ par produit scalaire partiel, de la façon suivante :

$$\forall |\psi\rangle \in \mathcal{H}_{\text{Bosons}}^\infty, a(u) |\psi\rangle = \langle u | \otimes \sqrt{N} |\psi\rangle \tag{3.203}$$

Nous allons déterminer, comme pour le cas bosonique, les opérateurs de création et d'annihilation pour des fermions. Cependant, nous allons faire le cheminement opposé. Nous partons de la définition des opérateurs créant et détruisant des particules dans un état quelconque $|u\rangle$. Par définition

$$\forall |\psi\rangle \in \mathcal{H}_{\text{Fermions}}^\infty, a^\dagger(u) |\psi\rangle = \sqrt{N} A(|u\rangle |\psi\rangle) \tag{3.204a}$$

$$a(u) |\psi\rangle = \langle u | \otimes \sqrt{N} |\psi\rangle \tag{3.204b}$$

À partir de là, nous pouvons déterminer leur action sur un état de Fock en se rappelant que pour les fermions $n_i \in \{0, 1\}$. Nous avons

$$a_i^\dagger |n_1, \dots, n_i, \dots\rangle = \delta_{n_i 0} \epsilon(\sigma) |n_1, \dots, n_i + 1, \dots\rangle \tag{3.205a}$$

$$a_i |n_1, \dots, n_i, \dots\rangle = \delta_{n_i 1} \epsilon(\sigma) |n_1, \dots, n_i - 1, \dots\rangle \tag{3.205b}$$

où σ est la permutation

$$\sigma = \begin{pmatrix} 1 & \dots & S_i + 1 & S_i + 2 & \dots & n + 1 \\ S_i + 1 & \dots & S_i & S_i + 2 & \dots & n + 1 \end{pmatrix} \tag{3.206}$$

avec S_i le nombre d'états occupés avant celui correspondant à $|u_i\rangle$, à savoir $S_i = \sum_{j=1}^{i-1} n_j$. Les relations vérifiées par ces opérateurs sont

$$\{a_i, a_j\} = 0 \quad \{a_i^\dagger, a_j^\dagger\} = 0 \quad \{a_i, a_j^\dagger\} = \delta_{ij} \tag{3.207}$$

la différence avec le cas bosonique provient de la présence des anticommutateurs. Il intègre le fait que les fermions obéissent au principe de Pauli : nous avons en effet $(a_i^\dagger)^2 = 0$ ce qui signifie bien que nous ne pouvons pas



avoir deux particules dans le même état quantique. Finalement, le comportement des opérateurs vis-à-vis des permutations est contenu dans l’algèbre des opérateurs. Enfin, les relations avec l’opérateur nombre de particules sont

$$[N, a_i^\dagger] = a_i^\dagger \quad [N, a_i] = -a_i \tag{3.208}$$

Résumons tous les résultats. Nous introduisons les notations usuelles qui distinguent les opérateurs bosoniques des opérateurs fermioniques.

Opérateurs de création et d’annihilation

Les opérateurs de création et d’annihilation sont des opérateurs agissant dans l’espace de Fock \mathcal{H}^∞ , espace où le nombre de particules est indéterminé, et permettent de créer ou détruire des particules dans un état donné. L’usage désigne par la lettre b les opérateurs pour les bosons et par c ceux pour les fermions ^a.

$$\{c_i, c_j\} = \{c_i^\dagger, c_j^\dagger\} = 0 \quad \{c_i, c_j^\dagger\} = \delta_{ij} \tag{3.209}$$

$$[b_i, b_j] = [b_i^\dagger, b_j^\dagger] = 0 \quad [b_i, b_j^\dagger] = \delta_{ij} \tag{3.210}$$

Les propriétés de symétrie ou d’antisymétrie sont incorporées dans l’algèbre des opérateurs. Sur un état de Fock, exprimé dans une base donnée de l’espace à une particule, ces opérateurs agissent de la manière suivante

$$b_i^\dagger |n_1, \dots, n_i, \dots\rangle = \sqrt{n_i + 1} |n_1, \dots, n_i + 1, \dots\rangle \tag{3.211a}$$

$$b_i |n_1, \dots, n_i, \dots\rangle = \sqrt{n_i} |n_1, \dots, n_i - 1, \dots\rangle \tag{3.211b}$$

$$c_i^\dagger |n_1, \dots, n_i, \dots\rangle = \delta_{n_i 0} \epsilon(\sigma) |n_1, \dots, n_i + 1, \dots\rangle \tag{3.211c}$$

$$c_i |n_1, \dots, n_i, \dots\rangle = \delta_{n_i 1} \epsilon(\sigma) |n_1, \dots, n_i - 1, \dots\rangle \tag{3.211d}$$

Enfin, on peut définir un opérateur N dont les valeurs propres nous donnent le nombre de particules présent dans un état du système.

$$N = a^\dagger a \tag{3.212}$$

^a. Mais rien ne nous y oblige...

3.7.4 Opérateurs de champs

Dans ce qui précède, nous avons étudié les opérateurs dans une base particulière. Nous pouvons en revanche très bien définir ces mêmes opérateurs dans une base différente. Il est clair qu’ils se transforment de la même manière que les vecteurs de l’espace par un changement de base. Soient $(|u_i\rangle)_{i \in I}$ et $(|v_i\rangle)_{i \in I}$ deux bases de l’espace de Hilbert à une particule. Nous avons donc

$$a^\dagger(v_k) = \sum_i a^\dagger(u_i) \langle u_i | v_k \rangle \tag{3.213a}$$

$$a(v_k) = \sum_i a(u_i) \langle v_k | u_i \rangle \tag{3.213b}$$

La notation avec le signe somme est générique. Nous pourrions tout aussi bien avoir une intégrale qu’une somme.



On appelle *opérateurs de champs* les opérateurs de création ψ^\dagger et d'annihilation ψ exprimés dans la représentation position (dans la base propre associée à l'opérateur position). Nous avons la relation de changement de base suivante

$$\psi^\dagger(\mathbf{x}) = \sum_i a^\dagger(u_i) \langle u_i | \mathbf{x} \rangle \tag{3.214a}$$

$$\psi(\mathbf{x}) = \sum_i a(u_i) \langle \mathbf{x} | u_i \rangle \tag{3.214b}$$

Physiquement, l'opérateur $\psi^\dagger(\mathbf{x})$ crée une particule dans l'état $|\mathbf{x}\rangle$, c'est-à-dire à la position \mathbf{x} . Nous avons toujours bien sûr $\psi(\mathbf{x})|0\rangle = 0$. À partir des relations de commutation des opérateurs bosoniques, nous arrivons à

$$[\psi(\mathbf{x}), \psi(\mathbf{y})] = 0 \quad [\psi^\dagger(\mathbf{x}), \psi^\dagger(\mathbf{y})] = 0 \quad [\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] = \delta(\mathbf{x} - \mathbf{y}) \tag{3.215}$$

Nous pouvons enfin définir l'opérateur *nombre de particules* N par

$$N = \int \underbrace{\psi^\dagger(\mathbf{x})\psi(\mathbf{x})}_{\text{Opérateur densité}} d\mathbf{x} \tag{3.216}$$

Nous pouvons aussi exprimer les opérateurs de champs dans l'espace des impulsions. Il suffit de se rappeler que $\langle \mathbf{k}, \mathbf{x} \rangle = \frac{e^{-i\mathbf{k} \cdot \mathbf{x}}}{(2\pi)^3}$. La formule de changement de base nous donne, par exemple pour l'opérateur de création,

$$\begin{aligned} \psi^\dagger(\mathbf{x}) &= \sum_i a^\dagger(u_i) \langle u_i | \mathbf{x} \rangle \\ &= \sum_i a^\dagger(u_i) \int \langle u_i | \mathbf{k} \rangle \langle \mathbf{k} | \mathbf{x} \rangle d\mathbf{k} \\ &= \int \frac{e^{-i\mathbf{k} \cdot \mathbf{x}}}{(2\pi)^3} \underbrace{\sum_i a^\dagger(u_i) \langle u_i | \mathbf{k} \rangle}_{=a^\dagger(\mathbf{k})} d\mathbf{k} \end{aligned} \tag{3.217}$$

Nous avons donc

$$\psi^\dagger(\mathbf{x}) = \int a^\dagger(\mathbf{k}) \frac{e^{-i\mathbf{k} \cdot \mathbf{x}}}{(2\pi)^3} d\mathbf{k} \tag{3.218a}$$

$$\psi(\mathbf{x}) = \int a(\mathbf{k}) \frac{e^{i\mathbf{k} \cdot \mathbf{x}}}{(2\pi)^3} d\mathbf{k} \tag{3.218b}$$

Les opérateurs $a^\dagger(\mathbf{k})$ et $a(\mathbf{k})$ sont des opérateurs de création et d'annihilation dans l'espace des impulsions. Le premier crée une particule d'impulsion \mathbf{k} tandis que l'autre en détruit une à cette impulsion. Ils vérifient

$$[a(\mathbf{k}), a(\mathbf{k}')] = 0 \quad [a^\dagger(\mathbf{k}), a^\dagger(\mathbf{k}')] = 0 \quad [a(\mathbf{k}), a^\dagger(\mathbf{k}')] = (2\pi)^3 \delta(\mathbf{k} - \mathbf{k}'). \tag{3.219}$$

Enfin, l'opérateur nombre total de particules peut s'écrire

$$N = \int \frac{1}{(2\pi)^3} a^\dagger(\mathbf{k}) a(\mathbf{k}) d\mathbf{k}. \tag{3.220}$$

Observables

Nous cherchons à exprimer dans ce formalisme les observables ainsi que leur évolution. Soit $F^{(n)}$ une observable d'un système à n particules identiques que l'on supposera additive, ce qui signifie qu'elle peut s'écrire comme



somme d'observables à une particule. On a invariance par permutation. Intéressons nous au cas où les particules sont indépendantes, c'est-à-dire sans interaction. Dans une base $(|u_i\rangle)_{i \in I}$ de l'espace de Hilbert à une particule, nous écrivons

$$F^{(1)} = \sum_{ij} F^{ij} |u_i\rangle \langle u_j|. \tag{3.221}$$

Nous avons pour l'observable à n particules

$$F^{(n)} = \sum_{k=1}^n F_k^{(1)} \quad F_k^{(1)} = \mathbb{1} \otimes \dots \otimes \underbrace{F^{(1)}}_{k^{\text{e place}}} \otimes \dots \otimes \mathbb{1}. \tag{3.222}$$

Réécrivons alors cette observable. Nous avons

$$\sum_{\sigma \in \mathfrak{S}_n} U_\sigma^{-1} F_1^{(1)} U_\sigma = \sum_{\sigma \in \mathfrak{S}_n} F_{\sigma^{-1}(1)}^{(1)} = (n-1)! F_1^{(1)} + (n-1)! F_2^{(1)} + \dots = (n-1)! F^{(1)} \tag{3.223}$$

Ce qui donne finalement

$$F^{(n)} = \frac{1}{(n-1)!} \sum_{\sigma \in \mathfrak{S}_n} U_\sigma^{-1} F_1^{(1)} U_\sigma \tag{3.224}$$

Considérons alors un état bosonique à n particules $|\phi\rangle \in \mathcal{H}_{\text{Bosons}}^n$. Nous avons alors

$$F^{(n)} |\phi\rangle = F^{(n)} S_n |\phi\rangle = S_n F^{(n)} |\phi\rangle \tag{3.225}$$

En utilisant la formule précédente, nous arrivons à

$$F^{(n)} |\phi\rangle = \frac{1}{(n-1)!} \sum_{\sigma \in \mathfrak{S}_n} S_n U_\sigma^{-1} F_1^{(1)} U_\sigma |\phi\rangle = \frac{1}{(n-1)!} \sum_{\sigma \in \mathfrak{S}_n} S_n F_1^{(1)} |\phi\rangle = n S_n F_1^{(1)} |\phi\rangle \tag{3.226}$$

Enfin, avec l'écriture dans la base $(|u_i\rangle)_{i \in I}$ de l'opérateur ainsi que l'écriture générale des opérateurs de création et d'annihilation, nous arrivons à

$$F^{(n)} |\phi\rangle = \sqrt{n} S_n \sum_{ij} F^{ij} |u_i\rangle a_j |\phi\rangle = \sum_{ij} F^{ij} \sqrt{n} S_n (|u_i\rangle a_j |\phi\rangle) = \sum_{ij} F^{ij} a_i^\dagger a_j |\phi\rangle \tag{3.227}$$

Par exemple, si l'on veut exprimer une observable en fonction des opérateurs de champs, nous aurons

$$F = \int \psi^\dagger(\mathbf{x}) \langle \mathbf{x} | F | \mathbf{y} \rangle \psi(\mathbf{y}) \, d\mathbf{x} d\mathbf{y} \tag{3.228}$$

Un cas particulier important est celui où l'opérateur est diagonal dans la base considérée. L'opérateur prend la forme simple

$$F^{(n)} |\phi\rangle = \sum_i F^i a_i^\dagger a_i |\phi\rangle \tag{3.229}$$

Hamiltonien sans interaction

En guise d'exemple, nous allons donner dans ce formalisme le hamiltonien sans interaction. Regardons d'abord l'opérateur impulsion. En représentation position, nous avons

$$\langle \mathbf{x} | \mathbf{p} | \mathbf{y} \rangle = -i\hbar \nabla_{\mathbf{x}} \delta(\mathbf{x} - \mathbf{y}) \tag{3.230}$$



Du coup, nous avons

$$\mathbf{p} = -i\hbar \int \psi^\dagger(\mathbf{x}) \nabla \psi(\mathbf{x}) d\mathbf{x} \tag{3.231}$$

Dans l'espace des impulsions, en utilisant 3.218, nous arrivons à

$$\mathbf{p} = \int \hbar \mathbf{k} a^\dagger(\mathbf{k}) a(\mathbf{k}) \frac{d\mathbf{k}}{(2\pi)^3} \tag{3.232}$$

Nous retrouvons alors, dans un cadre plus général qu'en mécanique quantique, que l'impulsion est reliée au vecteur d'onde par la relation $\mathbf{p} = \hbar \mathbf{k}$. Pour pouvoir déterminer la contribution cinétique nous devons trouver comment s'écrit $\frac{\mathbf{p}^2}{2m}$. Un calcul simple montre directement que

$$\frac{\mathbf{p}^2}{2m} = \int \frac{-\hbar^2}{2m} \psi^\dagger(\mathbf{x}) \Delta \psi(\mathbf{x}) d\mathbf{x} = \int \frac{\hbar^2 \mathbf{k}^2}{2m} a^\dagger(\mathbf{k}) a(\mathbf{k}) \frac{d\mathbf{k}}{(2\pi)^3} \tag{3.233}$$

Enfin, si l'on considère un potentiel $V(\mathbf{x})$, nous avons clairement $\langle \mathbf{x} | V(\mathbf{x}) | \mathbf{y} \rangle = V(\mathbf{x}) \delta(\mathbf{x} - \mathbf{y})$. La forme du hamiltonien sans interaction est finalement donnée par

$$H = \int \left[\frac{-\hbar^2}{2m} \psi^\dagger(\mathbf{x}) \Delta \psi(\mathbf{x}) + \psi^\dagger(\mathbf{x}) V(\mathbf{x}) \psi(\mathbf{x}) \right] d\mathbf{x} \tag{3.234}$$

Sur les interactions

Si l'on s'intéresse à une interaction à deux corps, nous aurons une observable de la forme

$$G^{(2)} = \sum_{ijkl} G^{ijkl} |u_i\rangle |u_j\rangle \langle u_k| \langle u_l| \tag{3.235}$$

Cet observable doit être invariant sous la permutation de deux particules. Intuitivement, on comprend bien que cet observable peut s'écrire

$$G^{(2)} = \frac{1}{2} \sum_{ijkl} G^{ijkl} a_i^\dagger a_j^\dagger a_l a_k \tag{3.236}$$

Dynamique

On se place en représentation de Heisenberg. On définit alors $\psi^{(H)}(\mathbf{x}, t) = e^{\frac{iHt}{\hbar}} \psi(\mathbf{x}) e^{-\frac{iHt}{\hbar}}$. La relation de commutation vérifiée par ces opérateurs est la même que celle en représentation de Schrödinger. L'équation d'évolution est donnée dans cette représentation par

$$i\hbar \frac{d\psi^{(H)}(\mathbf{x}, t)}{dt} = \left[\psi^{(H)}(\mathbf{x}, t), H \right] \tag{3.237}$$

Or, avec 3.234,

$$H = \int \left(\psi^{(H)} \right)^\dagger(\mathbf{y}, t) \left[\frac{-\hbar^2}{2m} \Delta_{\mathbf{y}} + V(\mathbf{y}) \right] \psi^{(H)}(\mathbf{y}, t) d\mathbf{y} \tag{3.238}$$

nous avons simplement

$$\left[\psi^{(H)}(\mathbf{x}, t), H \right] = \int \delta(\mathbf{x} - \mathbf{y}) \left[\frac{-\hbar^2}{2m} \Delta_{\mathbf{y}} + V(\mathbf{y}) \right] d\mathbf{y} = \frac{-\hbar^2}{2m} \Delta \psi^{(H)}(\mathbf{x}, t) + V(\mathbf{x}) \psi^{(H)}(\mathbf{x}, t) \tag{3.239}$$

L'équation d'évolution est donc dans ce cas particulier identique à l'équation de Schrödinger

$$i\hbar \frac{\partial \psi^{(H)}(\mathbf{x}, t)}{\partial t} = \frac{-\hbar^2}{2m} \Delta \psi^{(H)}(\mathbf{x}, t) + V(\mathbf{x}) \psi^{(H)}(\mathbf{x}, t) \tag{3.240}$$

Il est à noter que cette équation n'est valable que pour une interaction à un corps. Si, par exemple, on a une interaction à deux corps, un terme non linéaire apparaît dans l'équation d'évolution.

3.8 Généralités sur les symétries

L'étude des symétries est cruciale en physique! L'idée générale derrière une telle étude est de faire un examen des propriétés d'invariance que peut posséder la loi d'évolution et d'en évaluer les conséquences. En général, la présence de symétries restreint drastiquement les solutions d'un problème et permet aussi d'avoir un angle d'approche inédit. De plus, beaucoup de propriétés peuvent être reliées à une symétrie sous-jacente. Par exemple, les lois de conservation telles que celle de l'énergie ou de l'impulsion dérivent de symétries de l'espace-temps, celles d'invariance par translation temporelle et spatiale. On cherchera dans cette section à expliciter le rôle des symétries dans le cas spécifique de la mécanique quantique.

3.8.1 Groupe de transformations

Pour nous, effectuer une transformation sur un système physique signifiera que l'on remplacera chacune de ses variables par une autre, chaque état par un autre tout en conservant les propriétés physiques du système. La première question que l'on se pose est de savoir quelles sont les transformations qui nous intéressent. L'interprétation de la mécanique quantique est basée sur les probabilités. Les points importants sont les suivants. Un état est décrit par un rayon d'un espace de Hilbert, c'est-à-dire par un ensemble du type $\{e^{i\theta}|\psi\rangle, \theta \in \mathbb{R}\}$. Une probabilité de transition est donnée par le module au carré du produit scalaire entre l'état initial et l'état final. Enfin, la probabilité d'obtenir un résultat lors d'une mesure est elle aussi donnée par le module au carré d'un produit scalaire. C'est pourquoi, en mécanique quantique, la classe de transformations qui nous intéresse est celle qui conserve les probabilités.

En général, lorsque l'on considère un ensemble de transformations, celles-ci forment ce que l'on appelle un *groupe*, noté G . La théorie des groupes est exposée rapidement dans l'appendice B. Intuitivement, cette structure mathématique formalise ce que l'on peut attendre d'un ensemble de transformations. Tout d'abord, on dispose d'une transformation qui ne fait rien sur le système appelée l'identité. Ensuite, l'application de deux transformations successives forme aussi une transformation de l'ensemble. Enfin, si l'on se donne une transformation, il existe alors une autre transformation qui a exactement les effets contraires sur le système, ce que l'on appelle l'inverse. Formellement, un groupe G est défini de la manière suivante

Un groupe G est un ensemble non vide muni d'une loi $\star : G \times G \rightarrow G$ tel que

$$\begin{cases} \forall (x, y, z) \in G^3, (x \star y) \star z = x \star (y \star z) & \text{Associativité de la loi} \\ \exists e \in G, \forall x \in G \ x \star e = e \star x = x & \text{Existence d'un élément neutre} \\ \forall x \in G, \exists x^{-1} \in G \ x^{-1} \star x = x \star x^{-1} = e & \text{Existence d'un inverse} \end{cases} \quad (3.241)$$

Si l'ordre n'est pas important lorsqu'on applique la loi de composition, la loi est dite *commutative* et le groupe est dit commutatif ou *abélien*. Le point important à retenir est qu'un groupe est un ensemble d'éléments ayant un lien particulier entre eux : ils se composent d'une manière très particulière. Ce qui importe est la loi de composition \star . C'est elle qui donne toute la structure.

3.8.2 Groupe de Lie

On distingue souvent deux types de groupes de transformations, les *groupes discrets*, groupes ayant un nombre fini ou dénombrable d'éléments comme les groupes de symétries cristallographiques, et les *groupes continus*, groupes dont les éléments peuvent être paramétrés par un ou plusieurs paramètres continus. Par exemple, les rotations dans le plan, formant le groupe $SO(2)$, sont paramétrées par un angle pouvant prendre ses valeurs dans \mathbb{R} .

Beaucoup utilisés en physique, les groupes de Lie G sont des groupes munis d'une structure additionnelle (voir B pour plus de détails). Cette structure a pour conséquence que la connaissance du groupe entier peut se faire en ne regardant que les éléments autour de l'élément neutre. Un développement limité au premier ordre d'un tel élément nous fournit un ensemble d'opérateurs $(t_i)_i$, en nombre fini dans le cas où le groupe est paramétré par un nombre fini de paramètres, appelés *générateurs infinitésimaux*. L'espace vectoriel \mathfrak{g} engendré par ces opérateurs est appelé *algèbre de Lie* associée au groupe de Lie G .

$$\mathfrak{g} = \text{vect} [(t_i)_i] \quad (3.242)$$

La connaissance de l'algèbre suffit à connaître le groupe entier. Nous verrons comment plus en détail dans la suite.

3.8.3 Représentation, théorème de Wigner

Sur les représentations

Un groupe est un objet abstrait. Cependant, il est possible de le représenter avec des objets plus familiers (la théorie de la représentation des groupes est une branche importante de la théorie des groupes, voir B). Il faut savoir qu'un groupe peut être représenté d'une infinité de façons différentes. L'idée est d'associer à chaque élément d'un groupe une matrice inversible¹⁴. *A priori*, rien ne contraint la dimension des matrices que l'on doit choisir et c'est pourquoi nous avons une infinité de représentations. On exige de plus que l'application qui à un élément du groupe associe une matrice soit un morphisme. Cette contrainte est nécessaire si l'on veut préserver la structure de notre groupe. Une dernière remarque : formellement, il n'est pas exigé d'avoir une bijection entre le groupe abstrait et les matrices qu'on lui associe. Cependant, en physique, ce sera souvent le cas.

Une représentation linéaire du groupe G est un morphisme $\rho : G \rightarrow \text{Aut}(X)$ où X est un espace vectoriel de dimension finie. Le morphisme, exigé pour préserver la structure originelle du groupe, dit

$$\forall x, y \in G, \rho(xy) = \rho(x)\rho(y) \tag{3.243}$$

Théorème de Wigner

En mécanique quantique, on travaille dans un espace de Hilbert, un état étant un élément de cet espace sur lequel agit des opérateurs de cet espace. Il serait donc bien de savoir comment on traduit dans ce formalisme la notion de transformation. Pour cela, nous avons le théorème de Wigner.

Pour toute transformation préservant les probabilités, on peut associer un opérateur U agissant sur l'espace de Hilbert tel que soit U est unitaire et linéaire, soit U est anti-linéaire et anti-unitaire.

Les opérateurs anti-unitaires n'interviennent que lorsqu'on considère la symétrie par renversement du temps. On voit déjà arriver l'idée d'une représentation : à une transformation, nous avons associé un opérateur inversible vérifiant certaines propriétés. Il manque encore des propriétés à vérifier pour avoir une représentation correcte.

Intermède mathématique

Soit U une application de \mathcal{H} espace de Hilbert complexe¹⁵. On dit que U est antilinéaire lorsque

$$\forall x, y \in \mathcal{H}, U(x + y) = U(x) + U(y) \tag{3.244a}$$

$$\forall x \in \mathcal{H}, \forall \lambda \in \mathbb{C}, U(\lambda x) = \bar{\lambda}U(x) \tag{3.244b}$$

De manière analogue à la définition de l'adjoint d'un opérateur linéaire, on définit l'adjoint pour un opérateur antilinéaire U^\dagger par la relation

$$\forall x, y \in \mathcal{H}, (x, U(y)) = \overline{(U^\dagger(x), y)} \tag{3.245}$$

Cas d'un groupe de transformations

Le théorème de Wigner permet d'avoir une représentation adéquate d'un groupe pour les besoins de la mécanique quantique. Montrons comment. Soit g une transformation du groupe G . D'après le théorème de Wigner, nous avons

$$|\psi_g\rangle = U(g)|\psi\rangle \tag{3.246}$$

où $U(g)$ désigne l'opérateur que nous fournit le théorème tandis que $|\psi_g\rangle$ désigne l'état que l'on obtient à partir de $|\psi\rangle$ après transformation. Considérons maintenant deux transformations g_1 et g_2 . Posons $g = g_2g_1$. Nous avons alors les deux égalités suivantes

$$|\psi_{g_2g_1}\rangle = U(g_2)U(g_1)|\psi\rangle \tag{3.247a}$$

$$|\psi_g\rangle = U(g)|\psi\rangle \tag{3.247b}$$

¹⁴. Une application linéaire inversible d'un espace vectoriel de dimension finie...

¹⁵. Un espace vectoriel sur \mathbb{C} est suffisant.



Comme ces deux vecteurs représentent le même état, ils sont égaux à une phase près que l'on notera $\theta(g_2, g_1)$. Il est possible de montrer¹⁶ que ce facteur de phase ne dépend pas de l'état $|\psi\rangle$. Nous avons donc en réalité une égalité entre opérateurs

$$U(g) = e^{i\theta(g_2, g_1)} U(g_2) U(g_1) \tag{3.248}$$

Lorsque nous avons une telle égalité, on dit que nous avons une *représentation projective* du groupe G . Nous ne considérerons que deux types de représentations : l'une où le facteur de phase est égal à un, nous avons alors une représentation telle que définie dans l'appendice B, et l'autre où le facteur de phase vaut ± 1 , nous avons alors ce que l'on appelle une *représentation spinorielle*.

En mécanique quantique, nous avons la plupart du temps un isomorphisme entre le groupe abstrait et l'ensemble des opérateurs donnés par le théorème de Wigner. À une transformation est associée un et un seul opérateur. Cependant, il peut arriver à la limite que pour une transformation, on dispose de deux opérateurs opposés l'un de l'autre la représentant. Ceci arrive quand le spin total du système est demi-entier. Dans le cas où le spin du système est entier nous n'avons pas ce problème et il y a bien une bijection.

3.8.4 Groupe et algèbre de Lie

Première approche

Revenons sur les groupes de Lie ainsi que sur leur lien avec leur algèbre. On se donne un groupe de Lie G . On supposera qu'il est paramétré par p variables que l'on notera $(\theta_i)_{i \in [1, p]}$. Le nombre p est appelé *dimension* du groupe. Une transformation T de G est donc déterminée par la donnée de ces variables. On notera finalement de manière abrégée $T(\theta_i)$ les transformations du groupe. On fait l'hypothèse supplémentaire que le paramètre se transforme par composition de la manière suivante

$$T(\alpha_i) T(\beta_j) = T(f^k(\alpha_i, \beta_j)) \tag{3.249}$$

où les f^k sont des fonctions des paramètres régulières. Ces fonctions reflètent la structure du groupe. Maintenant, nous associons à ce groupe l'ensemble des opérateurs donnés par le théorème de Wigner. On suppose que la représentation donnée par le théorème de Wigner est de la forme

$$U[T(\alpha_i) T(\beta_j)] = U[T(\alpha_i)] U[T(\beta_j)] \tag{3.250}$$

(on a une "vraie" représentation) et que nous n'avons pas de symétrie par renversement de temps, ou en d'autres termes, que nous avons des opérateurs unitaires. Maintenant que nous avons fait toutes les hypothèses sur le groupe de transformation que l'on considère ainsi que sur sa représentation, nous pouvons aller plus loin dans la compréhension. Bien que les hypothèses paraissent assez restrictives, le cadre est encore assez général pour englober de nombreux groupes utiles en physique.

Comme mentionné plus haut, pour connaître le groupe de Lie, il suffit d'étudier les éléments autour de l'élément neutre ce qui, pour les opérateurs, revient à faire une étude autour de l'identité. Au deuxième ordre, nous avons alors

$$U[T(\theta_i)] = \mathbb{1} + i \sum_{i=1}^p \theta_i t_i - \sum_{i,j=1}^p \theta_i \theta_j t_{ij} + (\theta^2) \tag{3.251}$$

où les opérateurs hermitiens t_i sont les générateurs infinitésimaux du groupe¹⁷. De manière analogue au théorème de Schwartz, nous avons $t_{ij} = t_{ji}$. Dans la suite, on utilisera la convention d'Einstein. Il existe un lien entre les fonctions f^k et les opérateurs t_i et t_{ij} . En effet, nous avons

$$U[T(\alpha_i)] U[T(\beta_j)] = U[T(\alpha_i) T(\beta_j)] = U[T(f^k(\alpha_i, \beta_j))] \tag{3.252}$$

En écrivant le développement limité de f^k comme

¹⁶. Il suffit de considérer deux états ψ et ϕ , d'écrire que $|\psi_g\rangle = e^{i\alpha} |\psi_{g_2 g_1}\rangle$ et $|\phi_g\rangle = e^{i\beta} |\phi_{g_2 g_1}\rangle$ et de calculer leur produit scalaire en se rappelant que la transformation le préserve.

¹⁷. Le facteur i est mis par commodité.



$$f^k(\alpha_i, \beta_j) \alpha^k + \beta^k + f_{ij}^k \alpha_i \beta_j \tag{3.253}$$

et le développement de la relation 3.252, nous avons la relation générale reliant les opérateurs du premier et deuxième ordre aux fonctions liées à la structure du groupe

$$t_{ij} = -i f_{ij}^k t_k - t_i t_j \tag{3.254}$$

En utilisant la relation de symétrie de ces opérateurs, nous arrivons aux relations de commutation

$$[t_i, t_j] = i c_{ij}^k t_k \tag{3.255}$$

où les coefficients sont appelés *constantes de structure* de l'algèbre de Lie et s'écrivent $c_{ij}^k = -f_{ij}^k + f_{ji}^k$. Ils reflètent la structure du groupe sous-jacent puisqu'ils sont reliés aux fonctions f^k . On le verra mieux dans la suite.

Un exemple

Supposons le groupe G abélien paramétrisé par des paramètres additifs¹⁸. La propriété d'additivité signifie que l'on a $f^k(\alpha_i, \beta_j) = \alpha^k + \beta^k$ ou encore $f_{ij}^k = 0$. Nous avons donc $t_{ij} = -t_i t_j$. De plus, nous avons la propriété de symétrie $t_{ij} = t_{ji}$, ce qui donne $t_i t_j = t_j t_i$,

$$[t_i, t_j] = 0 \tag{3.256}$$

Cela étant, en regardant maintenant le développement de $U[T(\theta_i)]$, on retrouve le développement d'une exponentielle. Nous pouvons donc finalement écrire n'importe quelle transformation du groupe de la manière suivante

$$U[T(\theta_i)] = e^{-i\theta^i t_i} \tag{3.257}$$

Cet exemple illustre pourquoi les opérateurs hermitiens t_i sont appelés les générateurs. Grâce à eux, nous pouvons avoir tous les éléments du groupe. C'est pourquoi aussi nous avons dit que la connaissance de l'algèbre de Lie d'un groupe, espace vectoriel engendré par les générateurs, nous permet de retrouver le groupe tout entier. L'apparition de l'application exponentielle n'est pas un hasard et est en fait très générale : l'exponentielle d'un élément de l'algèbre de Lie \mathfrak{g} nous redonne un élément du groupe G . Pour un peu plus de détails, voir B.

Cadre général

Voyons de manière un peu plus générale la méthode que l'on a employée. On se donne un groupe de Lie G de dimension p dont les paramètres sont notés $(\theta_i)_{i \in [1,p]}$ ¹⁹. On se donne ensuite une représentation du groupe G que l'on notera \mathcal{R} (précédemment, cette représentation était celle donnée par le théorème de Wigner, les opérateurs étant supposés unitaires par commodité). On détermine ensuite les générateurs t_i du groupe par un développement limité ou, c'est équivalent, par la formule

$$t_i = i \frac{\partial \mathcal{R}(\theta_i)}{\partial \theta_i} \Big|_{\theta_i=0} \tag{3.258}$$

Il faut bien comprendre que dans cette approche, la forme des générateurs dépend de la représentation que l'on choisit. Une approche plus intrinsèque est faite dans l'appendice B. Une chose à savoir par contre est que le nombre de générateurs est indépendant de la représentation. L'espace vectoriel engendré par ces générateurs forme l'algèbre de Lie du groupe \mathfrak{g} (on en a en fait une représentation). Sa connaissance nous permet de connaître la représentation du groupe G en entier. En effet, nous avons la formule suivante, liant un opérateur $\mathcal{R}(\theta_i)$ caractérisé par les valeurs des paramètres θ_i aux générateurs du groupe

¹⁸. Lorsque le paramètre est additif, le groupe est forcément abélien.
¹⁹. On n'introduit pas cette fois les fonctions f^k , on n'en a pas besoin.

$$\mathcal{R}(\theta_i) = e^{-i\theta^i t_i} \tag{3.259}$$

Ce que l'on voit ici, c'est qu'un opérateur $\mathcal{R}(\theta_i)$ (qui, en physique, est souvent associé à un et un seul élément du groupe abstrait) et un élément de l'algèbre de Lie $i\theta^i t_i$ sont reliés par l'application exponentielle. La question qui vient alors est de savoir comment et où la structure du groupe apparaît. Pour cela, on doit donc étudier la composition de deux éléments du groupe $\mathcal{R}(\alpha_i) \mathcal{R}(\beta_i)$. Il est possible de montrer la formule dite de *Baker Campbell Hausdorff* qui stipule que

$$e^{i\alpha^i t_i} e^{i\beta^j t_j} = e^{i(\alpha^i t_i + \beta^j t_j - \frac{1}{2} [\alpha^i t_i, \beta^j t_j])} \tag{3.260}$$

Il apparaît donc que les commutateurs des générateurs nous donnent l'information sur la structure sous-jacente du groupe, sur la manière dont les éléments se composent entre eux. On comprend donc mieux leur importance! On peut encore aller plus loin. On rappelle que les générateurs forment une base de l'algèbre de Lie \mathfrak{g} . Ainsi, le commutateur de deux générateurs est lui même un élément de \mathfrak{g} et est donc une combinaison linéaire des éléments de la base. Il existe donc des coefficients c_{ij}^k , appelés constantes de structure, tels que

$$[t_i, t_j] = i c_{ij}^k t_k \tag{3.261}$$

La connaissance de ces coefficients nous permet de connaître tous les commutateurs des générateurs qui en retour nous renseignent sur toute la structure du groupe (d'où leur noms...).

3.8.5 Symétries

Une transformation laissant invariantes les probabilités, représentée par l'opérateur U , est une symétrie d'un système quantique si le hamiltonien est invariant par U . En d'autres termes,

$$[H, U] = 0 \tag{3.262}$$

Une telle définition est justifiée par le fait suivant. Soit un état $|\psi\rangle$. On suppose que le hamiltonien est invariant sous la transformation représentée par l'opérateur U . Alors

$$i\hbar \frac{dU|\psi\rangle}{dt} = HU|\psi\rangle \tag{3.263a}$$

$$i\hbar \frac{d|\psi\rangle}{dt} = H|\psi\rangle \tag{3.263b}$$

Du fait de l'invariance, c'est bien le même hamiltonien qui apparaît dans les deux équations. Nous avons alors directement

$$HU|\psi\rangle = i\hbar \frac{dU|\psi\rangle}{dt} = U \left(i\hbar \frac{d|\psi\rangle}{dt} \right) = UH|\psi\rangle \tag{3.264}$$

Ainsi, lorsqu'un opérateur commute avec le hamiltonien, nous pourrons directement en déduire des informations sur le système de manière générale.

3.9 Rotations et moments angulaires

TEXTE INTRO À FAIRE!

3.9.1 Groupe des rotations

Rotations en trois dimensions

Les rotations dans l'espace euclidien \mathbb{R}^3 forment un groupe appelé *groupe spécial orthogonal* dont la définition est la suivante

$$SO(3) = \{M \in M_3(\mathbb{R}), {}^tMM = \mathbb{1}_3 \text{ et } \det(M) = 1\} \tag{3.265}$$

Le déterminant positif traduit le fait que l'on ne change pas le sens de la base de l'espace lors d'une rotation (on exclut les réflexions). L'une des propriétés intéressantes de ces applications est qu'elles préservent le produit scalaire. Les éléments de ce groupe, les rotations, peuvent être décrites par trois angles (les angles d'Euler). Cela signifie que nous avons trois matrices de base pour décrire notre groupe, correspondant aux rotations autour de trois axes.

$$\mathcal{R}_1(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & -\sin(\phi) \\ 0 & \sin(\phi) & \cos(\phi) \end{pmatrix} \quad \mathcal{R}_2(\psi) = \begin{pmatrix} \cos(\psi) & 0 & \sin(\psi) \\ 0 & 1 & 0 \\ -\sin(\psi) & 0 & \cos(\psi) \end{pmatrix} \quad \mathcal{R}_3(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.266}$$

Il est alors facile de trouver les trois générateurs de $SO(3)$ en dérivant ces matrices. Nous les appellerons J_1, J_2 et J_3

$$J_1 = i \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad J_2 = i \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad J_3 = i \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.267}$$

Une rotation quelconque est alors trouvée en "passant à l'exponentielle". Par exemple, une rotation d'angle α autour de l'axe k s'écrira $\mathcal{R}_k(\alpha) = e^{-i\alpha J_k}$. Le facteur i dans la définition des générateurs est une convention prise en physique. Il arrive parfois qu'il n'y soit pas et, dans ce cas, une rotation quelconque est donné juste en prenant l'exponentielle de l'angle fois le générateur ; il n'y a plus le facteur $-i$.

L'algèbre de Lie est l'espace vectoriel engendré par J_1, J_2 et J_3 . Il se trouve qu'elle correspond à l'ensemble des matrices antisymétriques : cela se voit facilement en se rappelant que l'ensemble des matrices antisymétriques $A_3(\mathbb{R})$ forme un sous-espace vectoriel de $M_3(\mathbb{R})$ de dimension trois²⁰ et que nos trois matrices sont libres.

Le point important concerne les constantes de structure et les relations de commutations. Il apparaît le tenseur complètement antisymétrique de Levi-Cevita.

$$[J_i, J_j] = i\epsilon_{ijk}J_k \tag{3.268}$$

Représentation unitaire

Arrêtons nous un instant sur une représentation particulière du groupe des rotations qu'on appelle représentation unitaire. On comprend bien d'où sort l'idée de l'unitarité, concept central en mécanique quantique.

Soit \mathcal{H} un espace de Hilbert. On appelle *représentation unitaire* de $SO(3)$ sur \mathcal{H} une application

$$\mathcal{D} : \begin{matrix} SO(3) & \rightarrow & U(\mathcal{H}) \\ \mathcal{R} & \mapsto & \mathcal{D}(\mathcal{R}) \end{matrix} \tag{3.269}$$

telle que $\mathcal{D}(\mathcal{R}_1\mathcal{R}_2) = \mathcal{D}(\mathcal{R}_1)\mathcal{D}(\mathcal{R}_2)$.

Cette représentation signifie qu'à chaque rotation on associe un opérateur unitaire d'un espace de Hilbert. En particulier, nous avons $\mathcal{D}(\mathbb{1}) = \mathbb{1}$ et $\mathcal{D}(\mathcal{R})^\dagger = \mathcal{D}(\mathcal{R})^{-1} = \mathcal{D}(\mathcal{R}^{-1})$.

Il est possible de définir une représentation sur l'espace de Hilbert $\mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$, espace associé aux fonctions d'onde. Celle-ci nous donne la manière dont les fonctions d'onde se transforment sous une rotation. On définit l'application qui, cela se montre rapidement, a toutes les propriétés d'une représentation unitaire du groupe des rotations,

$$\mathcal{D} : \begin{matrix} SO(3) & \rightarrow & U(\mathcal{L}^2(\mathbb{R}^3, \mathbb{C})) \\ \mathcal{R} & \mapsto & \mathcal{D}(\mathcal{R}) \end{matrix} \tag{3.270}$$

définie, pour $\mathcal{R} \in SO(3)$,

²⁰. Dans le cas général, $\dim A_n(\mathbb{R}) = \frac{n(n-1)}{2}$.



$$\forall \psi \in \mathcal{L}^2(\mathbb{R}^3, \mathbb{C}), \forall \mathbf{r} \in \mathbb{R}^3, (\mathcal{D}(\mathcal{R})\psi)(\mathbf{r}) = \psi(\mathcal{R}^{-1}\mathbf{r}) \tag{3.271}$$

La manière dont nous avons défini cette représentation nous permet de donner une idée différente de ce qu'est une représentation. On se donne un référentiel R dans lequel on définit nos fonctions d'onde ψ . Maintenant, on effectue un changement de référentiel par rotation spatiale (\mathcal{R} , élément de $SO(3)$) que l'on appelle R' . La question qui vient alors est de savoir comment se transforment les fonctions d'onde que l'on note dans le nouveau référentiel ψ' . Nous avons, avec $\mathbf{r}' = \mathcal{R}\mathbf{r}$,

$$\forall \psi \in \mathcal{L}^2(\mathbb{R}^3, \mathbb{C}), \psi'(\mathbf{r}') = \psi(\mathbf{r}) = \psi(\mathcal{R}^{-1}\mathbf{r}') = (\mathcal{D}(\mathcal{R})\psi)(\mathbf{r}') \tag{3.272}$$

On retombe sur nos pieds. Mais le point de vue est différent dans le sens où nous nous sommes donné une rotation \mathcal{R} puis nous avons regardé l'action de cette rotation sur l'ensemble des fonctions d'onde. L'application $\mathcal{D}(\mathcal{R})$ que l'on définit à l'aide du calcul précédent par

$$\mathcal{D}(\mathcal{R}) : \begin{matrix} \mathcal{L}^2(\mathbb{R}^3, \mathbb{C}) & \rightarrow & \mathcal{L}^2(\mathbb{R}^3, \mathbb{C}) \\ \psi & \mapsto & \psi' \end{matrix} \tag{3.273}$$

nous permet de définir une action de groupe (voir appendice B pour la définition). L'idée derrière la notion d'action de groupe est bien de décrire la manière dont un groupe peut agir sur un ensemble; l'étude des symétries en est à l'origine.

Rotation et moment angulaire

La question qui se pose maintenant, ayant déterminé les générateurs du groupe des rotations ainsi que la manière dont ces dernières agissent sur les fonctions d'onde, est de savoir comment les générateurs sont représentés dans $\mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$, espace des fonctions d'onde. Pour cela, il suffit de regarder au premier ordre l'effet d'un des générateurs des rotations. Par exemple,

$$\psi(\mathcal{R}_z^{-1}\mathbf{r}) = \psi(x + \alpha y, -\alpha x + y, z) = \psi(\mathbf{r}) + \alpha y \frac{\partial \psi}{\partial x}(\mathbf{r}) - \alpha x \frac{\partial \psi}{\partial y}(\mathbf{r}) + o(\alpha^2) \tag{3.274}$$

$$= \psi(\mathbf{r}) - \underbrace{\alpha(x\partial_y - y\partial_x)}_{\frac{i}{\hbar}L_z} \psi(\mathbf{r}) + o(\alpha^2) \tag{3.275}$$

On voit donc que la composante L_z du *moment angulaire orbital* est la représentation d'un des générateurs des rotations pour les fonctions d'onde. On peut montrer facilement que le moment cinétique \mathbf{L} est le générateur (à un facteur près) des rotations dans $\mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$. Une rotation finie autour d'un des axes s'écrit comme l'exponentielle d'un des générateurs $e^{-\frac{i\alpha}{\hbar}L_k}$. Le fait que le moment cinétique soit hermitien correspond à l'unitarité de la représentation. Nous avons toujours les mêmes relations de commutation, à savoir

$$[L_i, L_j] = i\epsilon_{ijk}L_k \tag{3.276}$$

3.9.2 Moment cinétique

Revenons au cas plus général où l'on représente sur un espace de Hilbert \mathcal{H} . Il est possible de montrer que le meilleur moyen de comprendre comment le groupe des rotations agit sur \mathcal{H} , ou de manière équivalente comment est représenté ce groupe, est d'étudier les valeurs propres et vecteurs propres des générateurs dans cette représentation. La seule chose que l'on a besoin de savoir est que les générateurs sont des opérateurs hermitiens vérifiant

$$[J_i, J_j] = i\hbar\epsilon_{ijk}J_k \tag{3.277}$$

L'opérateur vectoriel \mathbf{J} est ce que l'on appelle *un moment cinétique*²¹. Rien que la seule connaissance de ces relations de commutation nous permet de préciser drastiquement ce que doivent être les valeurs propres et les

²¹. Parfois, on prend $\hbar = 1$. Ce n'est pas important.



vecteurs propres de ces opérateurs. Le principe est le même que pour l'étude de l'oscillateur harmonique et est fait en 5.5.1. Ce que l'on doit savoir est la chose suivante. Il est possible de trouver une base orthonormée de vecteurs notés $|j, m\rangle$ dans laquelle J_z et \mathbf{J}^2 sont diagonaux tels que

$$\mathbf{J}^2|j, m\rangle = j(j+1)\hbar^2|j, m\rangle \tag{3.278a}$$

$$J_z|j, m\rangle = m\hbar|j, m\rangle \tag{3.278b}$$

avec les contraintes

$$j \in \mathbb{N}/2 \tag{3.279}$$

$$m \in \llbracket -j, j \rrbracket \tag{3.280}$$

Pour une valeur de j donnée, l'opérateur J_z agit sur un sous-espace de Hilbert de \mathcal{H} noté \mathcal{H}^j de dimension $2j+1$ dont une base orthonormée est donnée par $(|j, m\rangle)_{m \in \llbracket -j, j \rrbracket}$. Il en est de même pour les deux autres opérateurs. Plus explicitement, en notant $J_{\pm} = J_x \pm iJ_y$, nous avons

$$J_z|j, m\rangle = m\hbar|j, m\rangle \tag{3.281a}$$

$$J_{\pm}|j, m\rangle = \hbar\sqrt{j(j+1) - m(m \pm 1)}|j, m \pm 1\rangle \tag{3.281b}$$

Il est intéressant de retrouver pour $j = 1/2$ les matrices de Pauli ou encore pour $j = 1$ les générateurs des matrices de rotation de l'espace ordinaire (à une transformation unitaire près).

Moment angulaire et harmonique sphérique

Retournons un instant dans l'espace des fonctions d'onde $\mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$. Nous avons déjà défini les générateurs des rotations dans cette représentation. D'après ce qui précède, pour mieux les comprendre, il faut regarder les valeurs et vecteurs propres. Ces vecteurs propres sont des fonctions de $\mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$ appelées *harmoniques sphériques* et notées \mathcal{Y}_l^m . En exprimant explicitement les opérateurs J_z et \mathbf{J}^2 , nous arrivons à une équation différentielle vérifiée par ces harmoniques.

Matrices de rotation

Nous avons vu que les rotations peuvent s'écrire en fonction de trois générateurs, et en fonction de trois angles. Géométriquement, une rotation peut être décrite par la succession des rotations suivantes : une rotation d'angle α autour de Oz ($(Ox, Oy, Oz) \rightarrow (Ox_1, Oy_1, Oz_1)$), puis une rotation d'angle β autour de Oy_1 ($(Ox_1, Oy_1, Oz_1) \rightarrow (Ox_2, Oy_2, Oz_2)$) puis enfin une rotation d'angle γ autour de Oz_2 ($(Ox_2, Oy_2, Oz_2) \rightarrow (Ox', Oy', Oz')$). La matrice de rotation $\mathcal{R}(\alpha, \beta, \gamma)$ s'écrit alors, en utilisant la notation standard en mécanique quantique (avec $i\hbar$ dans le commutateur)²²,

$$\mathcal{R}(\alpha, \beta, \gamma) = e^{-\frac{i}{\hbar}\gamma J_{z_2}} e^{-\frac{i}{\hbar}\beta J_{y_1}} e^{-\frac{i}{\hbar}\alpha J_z} \tag{3.283}$$

Pour que cela soit plus commode, il est possible d'exprimer cette rotation uniquement à partir de la base initiale (Ox, Oy, Oz) en utilisant la relation donnant la transformation d'un opérateur sous une rotation, ici d'axe Oz ,

$$J_{y_1} = \mathcal{R}_z(\alpha) J_y \mathcal{R}_z^\dagger(\alpha) = e^{-\frac{i}{\hbar}\alpha J_z} J_y e^{\frac{i}{\hbar}\alpha J_z} \tag{3.284}$$

Ayant une égalité similaire pour le terme en J_{z_2} , nous obtenons après exponentiation,

$$\mathcal{R}(\alpha, \beta, \gamma) = e^{-\frac{i}{\hbar}\alpha J_z} e^{-\frac{i}{\hbar}\beta J_y} e^{-\frac{i}{\hbar}\gamma J_z} \tag{3.285}$$

²². Précédemment, nous avions défini le générateur par $t_i = i \frac{\partial \mathcal{R}(\theta_i)}{\partial \theta_i} |_{\theta_i=0}$ où le facteur i est mis par commodité puisque cela garantit l'hermiticité du générateur. En mécanique quantique, il arrive parfois que l'on sorte un \hbar et donc que l'on définisse le générateur par

$$t_i = \frac{i}{\hbar} \frac{\partial \mathcal{R}(\theta_i)}{\partial \theta_i} |_{\theta_i=0} \tag{3.282}$$

Bien que nous ayons raisonné de manière géométrique, dans \mathbb{R}^3 , on peut se convaincre que cette relation est valable dans n'importe quelle représentation unitaire des rotations. Pour une telle représentation $\mathcal{R}(\alpha, \beta, \gamma)$ sera l'opérateur rotation agissant sur les éléments de l'espace de Hilbert et \mathbf{J} sera le moment cinétique, opérateur hermitien générateur des rotations. Le meilleur moyen d'étudier cet opérateur rotation est encore de le regarder sous forme matricielle. Nous savons que pour j donné, les rotations agissaient sur un sous-espace de Hilbert \mathcal{H}^j de dimension $2j + 1$. Naturellement, nous allons nous placer dans la base des $(|j, m\rangle)_{m \in \llbracket -j, j \rrbracket}$. Ce qui est important ici est de voir le rôle joué par j : c'est ce nombre qui nous donne l'espace dans lequel on représente les rotations ! Pour j donné, la matrice que l'on obtient, appelée *matrice rotation* et notée $D_{mm'}^j$ est

$$D_{mm'}^j(\alpha, \beta, \gamma) = \langle jm | \mathcal{R}(\alpha, \beta, \gamma) | jm \rangle \tag{3.286}$$

C'est donc une matrice de déterminant 1, unitaire et de dimension $2j + 1$. Étant donné que l'opérateur J_z est diagonal dans la base choisie, nous avons la simplification

$$D_{mm'}^j(\alpha, \beta, \gamma) = e^{-(\alpha m + \gamma m')} d_{mm'}^j(\beta) \quad d_{mm'}^j(\beta) = \langle jm | e^{-\frac{i}{\hbar} \beta J_y} | jm \rangle \tag{3.287}$$

Voyons un exemple qui montre une caractéristique importante de ces matrices. Regardons la rotation $\mathcal{R}(0, 0, \gamma) = e^{-\frac{i}{\hbar} \gamma J_z}$. Sur \mathcal{H}^j , nous aurons

$$\mathcal{R}(0, 0, \gamma) |j, m\rangle = e^{-\frac{i}{\hbar} \gamma J_z} |j, m\rangle = e^{-im\gamma} |j, m\rangle \tag{3.288}$$

Quels sont alors les résultats lors d'une rotation de 2π et de 4π ? Si $m \in \mathbb{N}$ les deux opérateurs sont égaux. Cependant, si $m \in \mathbb{N}/2$, les deux opérateurs ne le sont plus et sont opposés l'un de l'autre ! La même rotation est représentée par deux opérateurs différents.

$SU(2)$ et le spin

Un autre groupe que le groupe des rotations de l'espace ordinaire possède des constantes de structure égales à ϵ_{ijk} (voir B). Ce groupe s'appelle le groupe spécial unitaire et est le groupe des rotations de \mathbb{C}^2 .

$$SU(2) = \{M \in \mathcal{M}_2(\mathbb{C}) / M^\dagger M = \mathbb{1}_2, \det(M) = 1\} \tag{3.289}$$

Les générateurs de ce groupe sont proportionnels aux matrices de Pauli ! Une propriété remarquable lie $SU(2)$ et $SO(3)$: il existe une application R de $SU(2)$ dans $SO(3)$ surjective (non injective) telle que $R(AB) = R(A)R(B)$. En fait, pour un élément de $SO(3)$ on dispose de deux matrices de $SU(2)$ qui sont envoyées sur cette rotation. L'application est parfois dite bivaluée.

Le résultat qui nous intéresse vient du lien entre $SU(2)$ et les matrices de rotation définies plus haut. L'ensemble des matrices de rotations $D_{mm'}^j$ donne toujours, pour un j donné, une représentation de $SU(2)$ sur \mathcal{H}^j contrairement au groupe $SO(3)$ où il faut distinguer suivant les valeurs de j . Pour j entier, les matrices de rotations donnent bien une représentation du groupe $SO(3)$ tandis que pour j demi-entier elle est bivaluée.

Le spin est une propriété fondamentale des particules élémentaires. C'est un moment cinétique intrinsèque qui n'a aucun lien avec un quelconque mouvement de la particule. Formellement, le spin est donc un moment cinétique qui commute avec les variables position et impulsion. Il est caractérisé par un unique nombre s pouvant être entier ou demi-entier²³. Ainsi, une particule est décrite, dans le cas le plus général, par un *spineur* qui est un élément de $\mathcal{H}^j \otimes \mathcal{L}^2(\mathbb{R}^3, \mathbb{C})$. Le cas important de l'électron est traité en physique des particules 5.5.1.

23. Nous l'avons vu tout moment cinétique est caractérisé par j que l'on appelle s pour le spin...



Chapitre 4

Physique nucléaire

Sommaire

4.1	Introduction	92
4.2	Premières observations expérimentales	94
4.3	Modèle de la goutte liquide	96
4.4	Étude détaillée de la radioactivité	99
4.5	Structure nucléaire : modèle en couches	104
4.6	Taille et forme des noyaux	108

Ce chapitre sera une brève introduction à la physique nucléaire, où plusieurs modèles théoriques et empiriques du noyau seront étudiés, chacun ayant ses avantages et inconvénients. Le but ici sera de pouvoir décrire raisonnablement des phénomènes nucléaires connus tels que les désintégrations radioactives ou l'existence d'états excités pour les noyaux. Nous commencerons par une description générale des phénomènes nucléaires. Deux modèles seront développés pour comprendre la structure interne d'un noyau : le modèle de la goutte liquide et le modèle en couches. Chacun de ces modèles sera présenté avec les phénomènes dont il permet une description correcte. Le premier de ces modèles permettra de revenir plus en détail sur les réactions nucléaires de désintégrations radioactives, de fission et de fusion. Enfin, nous étudierons les méthodes utilisées pour déterminer la taille d'un noyau.

La référence principale utilisée ici est [18]. En bonus, une courte vidéo sur la physique nucléaire en général, [1]

Ici, aucune section ne peut réellement être passée pour la compréhension du sujet. Il est de plus fortement conseillé d'avoir déjà lu les parties sur l'effet tunnel et sur la diffusion du chapitre de mécanique quantique.

4.1 Introduction

4.1.1 L'expérience de Rutherford

La physique nucléaire a pris naissance au début du XXe siècle, avec l'expérience dite de Rutherford (réalisée par Geiger et Marsden) qui prouva l'existence du noyau atomique (voir figure 4.1). L'idée de cette expérience est simple : une fine feuille d'or est "bombardée" par des particules α , sous-produits de certaines désintégrations radioactives dont nous reparlerons dans ce chapitre. On observe alors sur un écran fluorescent situé après la feuille d'or les points d'impacts des particules traversant la feuille d'or.

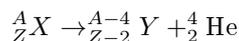
Le modèle atomique en vigueur à l'époque (dit de Thomson) suppose que la charge positive des atomes est répartie sur l'ensemble de l'atome, et que les électrons (qui étaient déjà connus) sont situés à l'intérieur de cette sphère chargée positivement¹. En utilisant ce modèle, le champ électrique résiduel à la surface des atomes est très faible et ne peut pas entraîner une forte diffusion des particules alpha incidente : on devrait observer l'ensemble des points d'impacts dans un très petit angle autour de l'axe correspondant au faisceau incident.

Or, l'expérience montre que la grande majorité des particules sont peu diffusées, mais que certaines d'entre elles subissent des diffusions très importantes, voire repartent vers l'arrière ! Ceci ne peut s'expliquer que si le champ électrique rencontré par les particules α est bien plus important que celui prévu par le modèle de Thomson, ce qui signifie que la charge positive portée par l'atome doit être contenue dans un volume bien plus faible. Rutherford interprète donc cette expérience comme prouvant que les atomes sont constitués principalement de

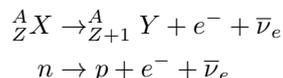
1. L'analogie la plus souvent utilisée pour ce modèle est celle d'un pudding au raisins...



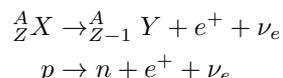
La désintégration α , qui correspond à l'émission d'un noyau d'hélium ${}^4_2\text{He}$ par le noyau. C'est un cas particulier de fission.



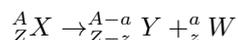
La désintégration β^- , dans laquelle un neutron se transforme en proton en émettant un électron et une antiparticule (un anti-neutrino électronique) qui permet la conservation du moment total.



La désintégration β^+ , qui est la réaction inverse : un proton devient un neutron en émettant un anti-électron et un neutrino électronique.



La fission, dans laquelle le noyau se sépare en deux noyaux plus petits.



Une désintégration est une transition d'un état quantique initial vers un état quantique final, et une étude exhaustive utilisant les règles de la mécanique quantique sera effectuée dans la section 4.4. Le seul élément important est que la probabilité de désintégration par unité de temps est une constante notée λ . De manière simple, exprimer la loi de désintégration radioactive revient à donner la probabilité qu'un noyau existe au bout d'un temps t . Notons $p(t)$ cette probabilité. La probabilité que ce noyau existe au temps $t + dt$ est donc égale à $p(t + dt) = p(t) + \frac{dp}{dt} dt$. Comme λdt est la probabilité que le noyau se désintègre pendant la durée dt , $1 - \lambda dt$ est la probabilité que le noyau existe toujours après l'écoulement d'une durée dt . Ceci permet donc d'écrire que $p(t + dt) = p(t)(1 - \lambda dt)$ et d'obtenir l'équation $\frac{dp}{dt} = -p(t)\lambda$ dont la solution donne la loi recherchée

$$p(t) = e^{-\lambda t} \quad (4.5)$$

On peut l'exprimer autrement en utilisant le temps de vie moyen. La probabilité que le noyau vive à l'insant t puis se soit désintégré au temps $t + dt$ est donnée par $p(t)\lambda dt$. Le *temps de vie moyen* étant défini par

$$\tau = \langle t \rangle = \int t p(t) \lambda dt = \frac{1}{\lambda} \quad (4.6)$$

la loi de désintégration s'exprime comme

$$p(t) = e^{-\lambda t} = e^{-\frac{t}{\tau}} \quad (4.7)$$

Si l'on considère un matériau composé de N_0 noyaux identiques, le nombre de noyaux restant au bout d'un temps t est donné par la loi de désintégration

$$N(t) = N_0 e^{-\lambda t} \quad (4.8)$$

Cela provient du fait que $N(t)$ vérifie clairement la même équation que la probabilité, à savoir

$$\frac{dN(t)}{dt} = -\lambda N(t) \quad (4.9)$$

On définit le temps de demi-vie comme le temps au bout duquel la moitié des noyaux se sont désintégrés. On obtient



$$\frac{N(t_{1/2})}{N_0} = \frac{1}{2} = e^{-\lambda t_{1/2}} \quad (4.10)$$

d'où l'on tire

$$t_{1/2} = \frac{\ln(2)}{\lambda} \quad (4.11)$$

Ce temps de demi-vie peut aller de quelques micro-secondes à plusieurs millions d'années.

On peut définir d'autres quantités, comme l'*activité* qui est le nombre de désintégrations par unité de temps, dont l'unité est le Becquerel (une autre unité qui peut être rencontrée est le Curie : $1 \text{ Ci} = 3,7 \cdot 10^{10} \text{ Bq}$). Par définition, $A(t) = \left| \frac{dN(t)}{dt} \right| = A(0)e^{-\lambda t}$. Si l'on a plusieurs désintégrations, ou même une chaîne de désintégrations, il faut, pour chaque noyau, réécrire l'équation 4.9 en tenant compte de toutes les désintégrations dont il est le père ainsi que de toutes celles dont il est le fils!

Dosimétrie

La dosimétrie correspond à l'étude de l'effet des radiations, notamment sur le corps humain, et de leur dangerosité. Il faut pour cela étudier l'énergie absorbée par unité de masse pour une radiation donnée. Les unités utilisées en dosimétrie sont le Gray ($1 \text{ Gy} = 1 \text{ J/kg}$) et le Sievert (Sv), qui prend en plus en compte la nocivité des différents types de rayonnement pour le corps humain. On multiplie la dose en Gray par le facteur correspondant au type de rayonnement (voir le tableau 4.1) pour avoir la dose équivalente en Sievert. Le tableau 4.2 donne lui un aperçu des effets sur le corps humain de l'absorption d'une dose (en une fois).

Type de radiation	Facteur multiplicatif
X, γ, β	1
Proton	10
α , Fragment de fission	20
Neutron $< 10\text{keV}$	1
Neutron $10\text{keV} - 100\text{keV}$	10
Neutron $100\text{keV} - 2\text{MeV}$	20
Neutron $2\text{MeV} - 20\text{MeV}$	10
Neutron $> 20\text{MeV}$	5

TABLE 4.1 – Tableau donnant les facteurs à prendre en compte pour passer d'une dose en Gray à une dose équivalente en Sievert, c'est-à-dire la dangerosité d'un rayonnement.

Dose (en mSv)	Effet sur le corps
< 250	Aucun
$250 - 1000$	Nausées
$1000 - 2500$	Vomissements, problèmes sanguins, guérison possible
$2500 - 5000$	Conséquences graves, hospitalisation obligatoire
> 5000	Mort

TABLE 4.2 – Tableau présentant l'effet de l'absorption d'une dose en une irradiation homogène du corps entier.

4.3 Modèle de la goutte liquide

On va s'intéresser ici à un premier modèle permettant d'étudier le noyau atomique : le modèle de la goutte liquide. L'idée qui a mené à ce modèle est la suivante : l'interaction nucléaire forte étant une interaction à courte portée saturée, chaque nucléon n'interagit qu'avec ses voisins. Expérimentalement, cela se traduit par le fait que $B(A, Z)/A$ est quasi-constante, à 8 MeV/Nucléon pour des noyaux ayant plus de vingt nucléons. De plus une propriété particulière est l'indépendance de la densité de charge vis-à-vis de la taille des noyaux. Ces propriétés nous poussent à modéliser le noyau comme étant une goutte de "liquide" chargé et incompressible.



Les hypothèses de base du modèle de la goutte liquide sont les suivantes. L'interaction étant à courte portée, elle peut être considérée comme une force de contact. À cause de la répulsion à courte distance, on suppose le noyau incompressible. Enfin, le noyau est supposé sphérique. Le noyau est donc assimilé à une goutte de matière nucléaire incompressible.

4.3.1 Formule de Weizsäcker

Le formule de Weizsäcker, basée sur le modèle de la goutte liquide, est une formule semi-empirique donnant l'énergie de liaison comme une fonction de A et de Z . Tous les coefficients apparaissant dans la formule sont ajustés par l'expérience et sont indépendants, en principe, de A et de Z . Cette formule donne des résultats en accord avec l'expérience pour des noyaux ayant plus de vingt nucléons. Pour de plus petits noyaux, le faible nombre de nucléons ne nous permet pas de voir le noyau comme une goutte de matière nucléaire; les effets de couches sont alors prédominants (voir la section 4.5 pour le modèle en couches du noyau).

Formule de Weizsäcker

$$B(A, Z) = u_v A - u_s A^{2/3} - u_c \frac{Z^2}{A^{1/3}} - u_T \frac{(N - Z)^2}{A} + \delta \quad (4.12)$$

Terme de volume

Le terme de volume correspond au terme $B_v = u_v A$. Ce terme provient du fait que la force nucléaire forte affecte à la fois les neutrons et les protons. Il est proportionnel au nombre de nucléons puisque nous avons la formule empirique pour les rayons des noyaux $R = r_0 A^{1/3}$. Enfin, il reflète le fait que la force est indépendante de la nature des nucléons. Expérimentalement, on trouve $u_v = 15,8$ MeV.

Terme de surface

Le terme de surface correspond au terme $B_s = -u_s A^{2/3}$. Tout comme le terme de volume, la puissance $2/3$ provient de la formule empirique des rayons des noyaux. Le terme de volume suggère que chaque nucléon interagit avec un nombre constant d'autres nucléons. Ceci n'est pas vrai pour les nucléons de surface puisqu'ils ont moins de voisins, justifiant le signe de l'expression. Ce terme peut être pensé comme une tension de surface. Expérimentalement, on trouve $u_s = 18,3$ MeV.

Terme de Coulomb

Le terme de Coulomb correspond au terme $B_c = -u_c \frac{Z^2}{A^{1/3}}$. L'origine de ce terme est dans l'interaction électrostatique entre les protons du noyau. La puissance provient toujours de la formule $R = r_0 A^{1/3}$. Dans un noyau, on compte Z protons. Le terme de Coulomb est donc proportionnel à $Z(Z - 1)$ qui est souvent approximé par Z^2 . Pour déterminer le coefficient de proportionnalité, on peut modéliser le noyau comme une sphère chargée uniformément. On trouve alors $u_c = \frac{2\alpha\hbar c}{5r_0} = 0,72$ MeV qui est proche de la valeur expérimentale $u_c = 0,714$ MeV.

Tous les termes décrits forment la base du modèle en goutte liquide du noyau. Cependant, certains faits expérimentaux ne sont pas pris en compte dans cette formule comme le fait que $N \approx Z$ et que les nucléons ont tendance à former des paires.

Terme d'asymétrie

Ce terme correspond à $B_{\text{asym}} = -u_T \frac{(N - Z)^2}{A}$. Il provient du principe d'exclusion de Pauli. En effet, les protons et les neutrons sont confinés dans le noyau, leurs niveaux d'énergie sont donc discrets. Or, deux fermions ne peuvent pas occuper le même état quantique. C'est pourquoi plus on ajoute de nucléons, plus des niveaux excités seront occupés. Les neutrons et les protons occupent des niveaux distincts. S'il y a beaucoup plus de neutrons que de protons, l'énergie totale du noyau sera plus grande (et l'énergie de liaison plus faible) que si certains des neutrons excités étaient des protons disposés sur des niveaux d'énergie plus bas. La figure 4.3 montre un exemple de noyau dont l'énergie serait plus faible si deux neutrons devenaient des protons. C'est pourquoi l'asymétrie entre le nombre de protons et le nombre de neutrons tend à diminuer l'énergie de liaison. Ce terme peut être dérivé par un modèle type gaz de Fermi. Expérimentalement, on trouve $u_T = 23,2$ MeV.

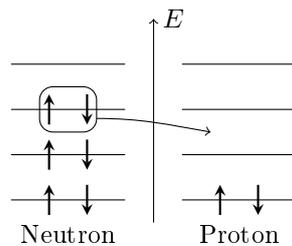


FIGURE 4.3 – Représentation d’un noyau avec trop de neutrons devant son nombre de protons. On voit que l’énergie du noyau serait plus faible, et donc que celui-ci serait plus stable, si une partie des neutrons devenaient des protons. Ceci est à l’origine du terme d’asymétrie dans la formule de Weizsäcker

Terme d’appariement

Ce terme correspond au terme δ . Il est complètement empirique. En fait, nous avons

$$B_{\text{paire}} = \begin{cases} \delta(A) & \text{noyau pair-pair} \\ -\delta(A) & \text{noyau impair-impair} \\ 0 & \text{noyau pair-impair} \end{cases} \quad (4.13)$$

où $\delta(A) = u_p A^{-1/2}$. Les termes pair et impair s’appliquent aux nombres de protons et de neutrons du noyau. Ce terme tend à favoriser les couples de nucléons identiques avec des spins opposés. Expérimentalement, $u_p = 12$ MeV.

4.3.2 Applications

Noyaux stables

La vallée de stabilité correspond à tous les noyaux stables. Pour un nombre de nucléons A donné, un noyau est dit stable lorsque l’énergie de liaison $B(A, Z) = B_A(Z)$ est maximale. Les noyaux stables vérifient donc

$$\frac{\partial B_A(Z)}{\partial Z} = 0 \quad (4.14)$$

Si on utilise la formule de Weizsäcker, l’énergie de liaison est un polynôme du second degré en Z . Le noyau stable, pour un A donné, est donné par

$$Z = \frac{2A}{4 + \frac{u_c}{u_T} A^{2/3}} \quad (4.15)$$

Cette formule nous donne donc le noyau le plus stable pour un nombre de nucléons donné. Cette formule donne les bons résultats expérimentaux : pour $A = 208$, elle donne une valeur théorique de $Z_t = 81.9$. Le nombre de protons étant entier, le Z prédit doit être l’entier le plus proche de Z_t , ici 82. On obtient donc les bons résultats, puisque le noyau de plomb le plus stable est ${}^{208}_{82}Pb$.

Bien que ce modèle permette de retrouver la forme générale de la vallée de stabilité des noyaux, il ne permet pas d’expliquer l’abondance de noyaux possédant un nombre dit *magique* de neutrons ou de protons. Expérimentalement, on a pu montrer des écarts au modèle de la goutte liquide dans le cas de noyaux possédant 2, 8, 20, 28, 50, 82 ou 126 protons ou neutrons. Le but du modèle en couches, exposé dans la section 4.5, sera de retrouver ces nombres.

Autres applications

La formule de Weizsäcker et le modèle de la goutte liquide permettent de mieux comprendre la stabilité des noyaux vis-à-vis des différentes désintégrations radioactives possibles (alpha, beta) et des réactions de fissions, en donnant une possibilité de calculer l’énergie de séparation de manière théorique. Elle permet également de montrer qu’un noyau déformé est plus stable qu’un noyau sphérique.



Nous allons nous intéresser ici à une méthode plus graphique pour déterminer la stabilité d'un noyau donné vis-à-vis des désintégrations beta. On sait que ces réactions ne modifient pas A , et sont donc des réactions *isobariques*. Or, nous avons vu précédemment que, à A donné, l'énergie de liaison suit une loi du second degré en Z . Dès lors, si l'on trace la parabole correspondant à une valeur de A donnée, on peut voir la façon dont un noyau va se désintégrer vers le minimum de la parabole. C'est ce que l'on appelle une *chaîne isobarique* dont les cas A pair et A impair, qui sont à traiter différemment, sont tous deux visibles sur la figure 4.4.

Dans le cas A impair, le terme d'appariement est toujours nul, et on a donc une unique parabole. En notant Z_0 la valeur de Z correspondant au minimum de la parabole, on voit que les noyaux avec $Z < Z_0$ doivent augmenter leur nombre de protons, c'est-à-dire subir des désintégrations β^- . Les noyaux avec $Z > Z_0$ vont eux subir des désintégrations β^+ . On voit alors assez facilement qu'il n'y a qu'un seul noyau stable pour les séries avec A impair.

Dans le cas A pair, on aura un changement de signe du terme d'appariement selon le caractère pair ou impair de Z . En effet, si Z est pair, N l'est également, et le noyau est plus stable. Dans l'autre cas, le terme d'appariement change de signe, et le noyau est alors moins stable. On doit donc tracer deux paraboles, avec un écart de 2δ entre elles. Ceci est bien visible sur la figure 4.4, où l'on voit apparaître deux minima Z_1 et Z_2 possibles pour le système. On en déduit alors que, dans le cas où A est pair, il n'existe pas de noyau impair-impair stable et que, pour une chaîne isobarique donnée, il existe plusieurs noyaux pair-pair stables.

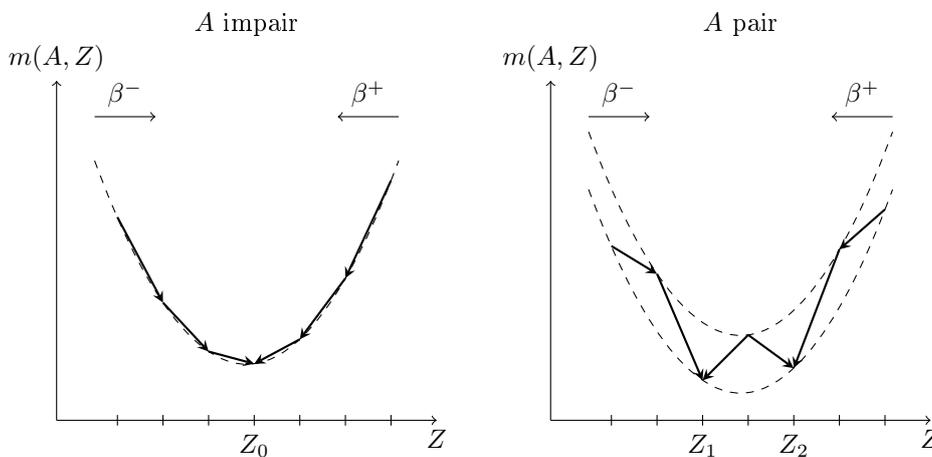


FIGURE 4.4 – Dans le cas A impair, on a un unique noyau stable. Tous les autres atomes de l'isobare vont subir des désintégrations β pour augmenter ou diminuer leur nombre de protons de façon à atteindre ce cas stable. Dans le cas A pair, on a plusieurs noyaux stables possibles, tous avec un nombre de protons pair.

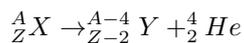
4.4 Étude détaillée de la radioactivité

Maintenant que l'on s'est donné un modèle pour l'étude du noyau stable, il est temps de s'intéresser plus précisément aux différentes désintégrations dont on a déjà parlé.

4.4.1 Désintégration α, β

Désintégration α , modèle de Gamov

On se rappelle que les désintégrations α correspondent à la réaction suivante :



Le modèle que nous allons étudier fut établi par Gamov et est basé sur la traversée d'une barrière de potentiel par effet tunnel (voir la section 3.3.6). Si l'on pense de manière classique, une particule α ne peut pas rentrer ou sortir du noyau. Seule la mécanique quantique permet d'expliquer cette désintégration.

Le modèle de base est le suivant. Le potentiel nucléaire est décrit par un puits carré de profondeur $-V_0$ et de largeur R correspondant à la somme des rayons du noyau fils et du noyau d'hélium (voir figure 4.5). À des distances plus élevées, le potentiel est coulombien. Nous l'écrivons $V(r) = \frac{k}{r}$ avec $k = 2Z'\alpha\hbar c$. Si l'on note E_α l'énergie cinétique finale de la particule α , nous pouvons déterminer la distance à laquelle la particule s'échappe du potentiel par



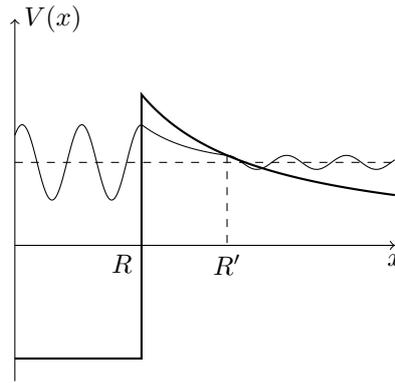


FIGURE 4.5 – Représentation de la désintégration α en termes d’effet tunnel. La particule α présente dans le noyau peut en sortir par effet tunnel.

$$V(R') = E_\alpha \Rightarrow R' = \frac{Z'\alpha\hbar c}{E_\alpha}. \tag{4.16}$$

Maintenant, nous devons donner la probabilité de désintégration λ_α . Gamov supposa la forme suivante

$$\lambda_\alpha = p(\alpha) f T. \tag{4.17}$$

$p(\alpha)$ est la probabilité d’existence de la particule α dans le noyau. Dans ce modèle nous la supposons égale à un. f est la “fréquence de collision” entre la particule α et la barrière de potentiel. Dans une vision classique, nous aurions $f = \frac{v}{2R}$ où v est la vitesse de la particule dans le noyau. En terme d’énergie, cela donne (toujours en termes classiques)

$$f = \frac{1}{R} \sqrt{\frac{E_0}{2m}} \tag{4.18}$$

Enfin, T est la probabilité de transmission par effet tunnel. La détermination de cette probabilité de transmission est faite dans 3.3.6. Dans le cas particulier du potentiel coulombien, nous arrivons à

$$T = e^{-2G} \tag{4.19}$$

où G est la *facteur de Gamov* valant $G = \int_R^{R'} \frac{\sqrt{2m(V(r)-E_\alpha)}}{\hbar} dr$. Le potentiel, d’origine coulombienne, est créé par le noyau fils. Le calcul explicite peut être fait, mais dans la pratique, la distance d’échappement R' est bien plus grande que les dimensions du noyau données par R . Nous avons

$$G = \sqrt{2mc^2}\pi\alpha \frac{Z'}{\sqrt{E_\alpha}}. \tag{4.20}$$

En utilisant tous ces résultats, la probabilité de désintégration alpha est donnée par

$$\lambda_\alpha = \sqrt{\frac{E_\alpha}{2m_\alpha c^2}} \frac{c}{r_0 A^{1/3}} \exp\left(-\pi\alpha\sqrt{8m_\alpha c^2} \frac{Z'}{\sqrt{E_\alpha}}\right). \tag{4.21}$$

À partir de cette formule, nous pouvons donner les lois de Geiger-Nuttal, observées expérimentalement, à savoir que le logarithme de λ_α ainsi que du temps de demi-vie sont des fonctions affines de $\frac{Z'}{\sqrt{E_\alpha}}$. Ceci explique le fait que des isotopes à faible durée de vie émettent des particules alpha plus énergétiques que ceux à durée de vie plus longue.

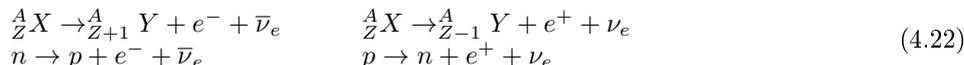
Plusieurs remarques sur ce modèle peuvent être faites. Tout d’abord, il est basé sur des considérations classiques pour la détermination de la fréquence f . Nous avons de plus supposé l’existence d’un état lié alpha dans le noyau,

ce qui n'est pas *a priori* justifié. Nous aurions dû considérer la probabilité d'existence d'un tel état. Pour être encore plus précis, la masse réduite du système peut être introduite de même que des considérations sur le moment angulaire.

Lors d'une désintégration alpha, des règles de sélection doivent être respectées. En effet le moment cinétique total doit être conservé ainsi que la parité.

Désintégration β , modèle de Fermi

Les désintégrations β peuvent être de deux types, qui sont pour rappel :



Le modèle de désintégration beta que nous allons développer est celui de Fermi, visible sur la figure 4.6. Cette désintégration est issue de l'interaction faible, qui sera décrite par une interaction de contact à quatre corps. Dans cette théorie, aucun propagateur n'est présent et son influence est absorbée dans une constante de couplage appelée *constante de Fermi* et valant $G_F = 1,166 \cdot 10^{-5} \text{ GeV}^{-2}$.

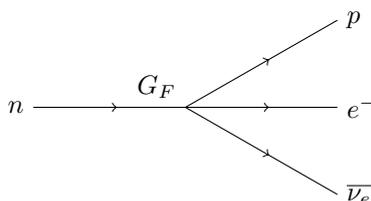


FIGURE 4.6 – Diagramme de Feynman de la désintégration β dans le modèle de Fermi. On suppose dans ce modèle que l'interaction faible peut être modélisée comme une interaction de contact à quatre corps.

Notre but est de donner la probabilité de transition Γ d'une telle réaction. Grâce à la règle d'or de Fermi, qui peut être appliquée car un traitement perturbatif est justifié par la faiblesse réelle de l'interaction faible, nous avons

$$\Gamma = 2\pi |M_{fi}|^2 \rho(E_f). \tag{4.23}$$

Reste maintenant à déterminer tous les termes présents. Commençons par la densité d'états. Pour cela, considérons deux des particules émises. La conservation de l'énergie-impulsion nous donnera les informations sur la troisième. Il faut tout de même garder en mémoire cette contrainte. Choisissons l'électron et le neutrino. Nous avons d'abord

$$d^6 N = \frac{p_\nu^2 dp_\nu d\Omega_\nu}{(2\pi)^3} \frac{p_e^2 dp_e d\Omega_e}{(2\pi)^3}. \tag{4.24}$$

L'énergie totale E_0 est partagée entre celle de l'électron, du neutrino et le recul du noyau. Cette dernière est négligée compte tenu de la masse du noyau comparée à celle des particules produites. Donc $E_0 \simeq E_e + E_\nu$. Pour le neutrino, $E_\nu = p_\nu$. Si l'on se donne une énergie pour l'électron E_e , nous aurons donc $dp_\nu = dE_\nu = dE_0$. D'où

$$\frac{d^6 N}{dE_0} = \frac{d^6 N}{dp_\nu} = \frac{E_\nu^2 p_e^2}{(2\pi)^6} d\Omega_\nu d\Omega_e dp_e. \tag{4.25}$$

En faisant l'hypothèse que la distribution angulaire est isotrope, nous pouvons faire l'intégration sur les variables angulaires. En utilisant la relation sur les énergies nous avons

$$\frac{d^2 N}{dE_0} = \frac{(E_0 - E_e)^2 p_e^2}{4\pi^2} dp_e = d\rho(E_F). \tag{4.26}$$

Nous pourrions utiliser cette formule en étudiant la différentielle de la probabilité de transition $d\Gamma$. Maintenant, déterminons les coefficients de la matrice de transition. Rappelons l'hypothèse de travail, à savoir que l'interaction faible est assimilée à une interaction de contact à quatre corps. Nous avons donc

$$M_{fi} = G_F \int \psi_p^* \psi_e^* \psi_\nu^* \psi_n d^3 \mathbf{r} \quad (4.27)$$

Après la désintégration, l'électron ainsi que le neutrino sont libres, d'où $\psi_e = e^{i\mathbf{p}_e \cdot \mathbf{r}}$ et $\psi_\nu = e^{i\mathbf{p}_\nu \cdot \mathbf{r}}$. Ceci permet d'écrire la forme finale des coefficients comme

$$M_{fi} = G_F \int \psi_p^* \psi_n e^{-i(\mathbf{p}_e + \mathbf{p}_\nu) \cdot \mathbf{r}} d^3 \mathbf{r} = G_F M_{\text{nucl}}. \quad (4.28)$$

Si l'on regroupe tous les résultats ensemble, nous avons après intégration sur l'impulsion électronique la forme finale pour la probabilité de transition

$$\Gamma = \frac{G_F^2}{2\pi^3} |M_{\text{nucl}}|^2 \int_0^{E_0} (E_0 - E_e)^2 p_e^2 dp_e. \quad (4.29)$$

Cette formule permet de rendre compte de différents faits expérimentaux comme la règle de Sargent qui dit que, dans la limite relativiste où $E_e = p_e$, la probabilité de transition est proportionnelle à la puissance cinquième de l'énergie initiale E_0 . Plusieurs améliorations peuvent être faites. Par exemple, l'impulsion de l'électron est modifiée par l'interaction de Coulomb avec le noyau. Cet effet peut être pris en compte par la *fonction de Fermi* $F(Z_Y, p_e)$ que l'on rajoute dans l'intégrale.

Tout comme pour la désintégration alpha, des règles de sélection doivent être satisfaites. De manière qualitative, les transitions permises sont celles où le moment cinétique de la paire électron-neutrino l est nul ainsi que lorsque $e^{-i(\mathbf{p}_e + \mathbf{p}_\nu) \cdot \mathbf{r}} \simeq 1$. Un cas particulier de transition permise, appelé transition super-permise, est celui où $M_{\text{nucl}} \simeq \int \psi_p^* \psi_n d^3 \mathbf{r} \simeq 1$. Ce cas arrive lorsque les fonctions d'onde du proton et du neutron sont identiques.

L'électron et le neutrino étant des particules de spin 1/2, le spin total de la paire ne peut être égal qu'à 0 ou 1. Les transitions sont classées en deux types, les transitions de Fermi sont celles où le spin vaut 0 et les transitions de Gamov-Teller sont celles de spin 1.

D'autres types de réactions similaires à la désintégration beta sont théoriquement possibles. Par exemple, la double désintégration β^- où deux neutrons se désintègrent en protons est à ce jour le phénomène de désintégration le plus rare observé. Une double désintégration β^+ n'a cependant toujours pas été observée. Dans un autre registre, si le neutrino est sa propre antiparticule (il devrait alors avoir une masse), il est alors possible d'envisager une double désintégration sans émission de neutrinos. Son observation fait toujours l'objet de recherches.

4.4.2 Fission et fusion nucléaire

Les noyaux les plus stables sont ceux proches du fer. Pour les noyaux plus lourds, un des modes de désintégration possible est par fission en deux éléments plus légers de taille moyenne. Pour les noyaux légers, la fusion de plusieurs éléments en un de taille plus grosse est privilégiée. Ces deux phénomènes sont d'énormes sources d'énergie, que ce soit sur Terre dans les centrales nucléaires, ou dans les étoiles.

Fission nucléaire

L'origine de la fission provient de la compétition entre l'énergie de surface (dans le modèle de la goutte liquide 4.3) et l'énergie coulombienne. En effet, l'énergie de Coulomb est réduite si le noyau se scinde en deux. Quant bien même l'énergie de surface augmente, la fission est énergétiquement favorable.

Intéressons-nous d'abord au processus de fission spontané. Pour pouvoir se scinder, le noyau doit d'abord se déformer. Pour se donner une idée intuitive, on peut repenser au modèle de la goutte liquide. Au début de la déformation, l'énergie de surface augmente. À une énergie seuil, dite énergie de scission, la répulsion électrostatique prend le dessus et le noyau se sépare en deux fragments. Durant tout ce processus, le système doit traverser une barrière de potentiel, voir la figure 4.7. L'effet tunnel est à nouveau essentiel.

Le modèle le plus simple de déformation consiste à considérer que le noyau sphérique se transforme en un ellipsoïde de révolution. Le demi grand axe a et le demi petit axe b sont donnés par $a = R(1 + \epsilon)$ et $b = R(1 + \epsilon)^{-1/2}$ où ϵ est supposé très petit. L'énergie de liaison B_{def} est alors pour ce noyau déformé égale à



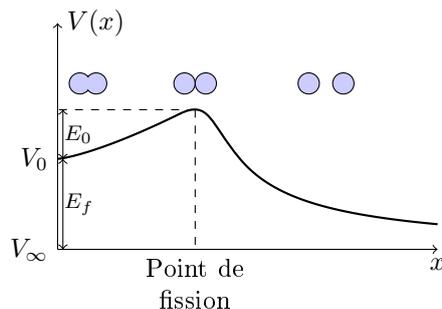


FIGURE 4.7 – Représentation de la fission nucléaire en terme d’effet tunnel. Tout se passe comme si un noyau fils cherchait à “sortir” du noyau initial. Le noyau commence par se déformer avant de se séparer en deux.

$$B_{\text{def}} = B_{\text{sph}} - \underbrace{\frac{2}{5}u_s\epsilon^2 A^{2/3}}_{\text{Différence dans le terme de surface}} + \underbrace{\frac{1}{5}u_c\epsilon^2 \frac{Z^2}{A^{1/3}}}_{\text{Différence dans le terme de Coulomb}} \tag{4.30}$$

Ainsi, un noyau déformé est plus stable qu’un noyau sphérique lorsque

$$B_{\text{def}} > B_{\text{sph}} \Leftrightarrow \frac{Z^2}{A} > \frac{2u_s}{u_c} \simeq 47 \tag{4.31}$$

Cette condition est déjà un critère pour que la fission spontanée puisse se produire. Qualitativement, la probabilité de transmission par effet tunnel dépend du rapport $\frac{Z^2}{A}$ (l’énergie d’activation est donc plus petite pour les noyaux lourds), ainsi que de la racine de la masse du noyau, c’est pourquoi les noyaux lourds ont une probabilité de transmission plus faible. On rappelle que cette probabilité est de la forme $T = e^{-2G}$. Cependant, la probabilité de désintégration par fission spontanée est bien plus faible que pour une désintégration alpha : c’est un mode de désintégration très rare pour les noyaux stables ou quasi-stables.

Intéressons nous maintenant à la fission induite. À faible énergie, des neutrons peuvent être absorbés par le noyau et cela d’autant plus que l’énergie est faible : la section efficace d’absorption est inversement proportionnelle à la vitesse de la particule. Après absorption, le noyau rentre dans un état excité; deux cas peuvent se produire. Soit l’énergie de l’état excité est trop faible, auquel cas une désintégration gamma est la plus probable; soit l’énergie est suffisante pour que la fission soit le mode de désintégration privilégié et se produira rapidement. Ce cas là est ce que l’on appelle la fission induite.

Les noyaux fils riches en neutrons auront une tendance à expulser les neutrons en excès très rapidement. Cependant, ces noyaux peuvent aussi se désintégrer de manière plus lente par émission beta. Ces neutrons peuvent être utilisés de manière convenable pour induire de nouvelles fissions : nous aurons alors affaire à une *réaction en chaîne*. Évidemment, une réaction en chaîne ne peut se produire que si au moins un neutron par fission en induit une autre. Si l’on note k le nombre de neutrons pour une fission induisant une autre fission, alors lorsque $k = 1$ la réaction est dite *critique*, lorsque $k < 1$ elle est dite *sous-critique* et enfin lorsque $k > 1$ elle est dite *sur-critique*. Pour être efficace, le mieux est d’avoir des neutrons lents comme mentionné plus haut.

L’idée de la réaction en chaîne est à la base du fonctionnement des réacteurs nucléaires. Pour pouvoir fonctionner, les réactions doivent être critiques. Un modérateur est nécessaire pour pouvoir ralentir les neutrons rapides et contrôler les réactions. Si le contrôle des neutrons venait à faillir, les réactions en chaîne s’emballeraient. Détaillons un peu. Soit $N(t)$ le nombre de neutrons au temps t . $k - 1$ représente le surplus de neutrons entre deux réactions et l’on définit τ le temps moyen entre deux fissions. Alors, nous avons

$$N(t + dt) = N(t) + (k - 1)N(t) \frac{dt}{\tau} \tag{4.32}$$

Nous obtenons alors une équation différentielle dont les solutions sont données par

$$N(t) = N(0)e^{(k-1)\frac{t}{\tau}} \tag{4.33}$$



Nous avons donc une croissance exponentielle. Dans le cas d'une réaction surcritique, le nombre de neutrons augmente exponentiellement avec le temps. Cependant, le réacteur n'explosera pas. À mesure que la température augmente, la section efficace d'absorption des neutrons diminue. Le réacteur aura tendance à se stabiliser à haute température et à fondre. C'est pourquoi il faut s'arranger pour que le réacteur reste critique.

Fusion nucléaire

L'origine de la fusion nucléaire provient toujours de la compétition entre l'énergie de surface et l'énergie de Coulomb. Contrairement au cas de la fission, les noyaux légers ont une énergie de surface trop importante. Lorsqu'ils fusionnent, l'énergie de surface diminue. Même si dans ce cas l'énergie de Coulomb augmente, elle reste moins importante que l'avantage sur l'énergie de surface.

En revanche, pour pouvoir fusionner, les noyaux doivent surmonter la répulsion coulombienne. Typiquement, pour que deux protons puissent former du deutérium, 1 MeV est nécessaire. Une telle énergie ne peut être fournie thermiquement que si la température est de l'ordre de 10^{10} K (provient de $k_B T = 1$ MeV). La température au centre d'une étoile est de l'ordre de dix millions de Kelvin, ce qui reste trop faible. La fusion ne peut alors être expliquée que si l'on tient compte encore de la possibilité de l'effet tunnel.

L'observation des neutrinos solaires fournit de nombreux tests des modèles stellaires puisque, du fait de leur faible interaction avec la matière, ils permettent de sonder directement le cœur du soleil. Par exemple, le flux mesuré était trois fois plus petit que celui prédit par les modèles. Ce problème fut résolu en admettant le phénomène d'oscillation des neutrinos, c'est-à-dire la possibilité pour les neutrinos de changer de saveur. Ceci impose d'ailleurs que les neutrinos doivent posséder une masse.

4.5 Structure nucléaire : modèle en couches

On va s'intéresser à un autre modèle du noyau, permettant d'expliquer certains faits qui ne peuvent pas être prédits dans le modèle de la goutte liquide. En effet, bien que la formule semi-empirique de Weizsäcker donne des résultats corrects, en accord avec certains faits expérimentaux (prédiction des noyaux stables, des stabilités vis-à-vis des désintégrations radioactives...), elle ne repose sur aucune base théorique et ne permet pas, par exemple, d'expliquer l'existence d'états excités du noyau. L'expérience montre que certains comportements nucléaires sont similaires à ceux de l'atome et permettent de penser que le noyau possède une structure en couches à la manière du nuage électronique.

Dans tous les modèles qui vont suivre, on fera les hypothèses suivantes : on supposera le noyau sphérique, immobile, constitué de nucléons non relativistes indépendants subissant un potentiel moyen représentant l'interaction avec tous les autres nucléons.

4.5.1 Modèle du gaz de Fermi

L'idée ici est de décrire le noyau comme un gaz de fermions dégénéré. On considère donc N fermions dans un volume V à la température T . On supposera enfin une dégénérescence maximale : en d'autres termes, tous les niveaux d'énergie sont remplis jusqu'à un maximum E_F appelé *niveau de Fermi* (un gaz est dit dégénéré si $E_F \gg k_B T$).

Comme vu dans 3.3.5, une particule occupe un volume minimum dans l'espace des phases de h^3 . Le nombre d'état N à l'intérieur d'une sphère dans l'espace des impulsions de rayon p_F vaut³

$$N = \int_0^{p_F} g_s dn = \frac{2V}{h^3} 4\pi \int_0^{p_F} p^2 dp \quad (4.34)$$

On en déduit l'impulsion de Fermi puis l'énergie de Fermi

$$E_F = \frac{h^2}{2M} \left(\frac{3\pi^2 N}{V} \right)^{2/3} \quad (4.35)$$

3. Nous avons $dn = \frac{V 4\pi p^2 dp}{h^3}$. Il faut ajouter à cela la dégénérescence de spin g_s .



Le noyau est décrit par deux gaz de fermions dégénérés, un gaz de protons et un gaz de neutrons enfermés dans un même volume $V = \frac{4}{3}\pi r_0^3 A$. Le principe d'exclusion de Pauli agit de manière indépendante sur les deux gaz. Les énergies associées sont

$$E_F(p) = \frac{h^2}{2M_p} \left(\frac{3\pi^2 Z}{V} \right)^{2/3} \quad E_F(n) = \frac{h^2}{2M_n} \left(\frac{3\pi^2 N}{V} \right)^{2/3} \quad (4.36)$$

On en déduit l'énergie cinétique totale des nucléons

$$\begin{aligned} T(A) &= \int \frac{p^2}{2M_p} dn + \int \frac{p^2}{2M_n} dn \\ &= \int_0^{p_F(p)} \frac{p^2}{2M_p} \left[g_s \frac{V 4\pi p^2}{h^3} \right] dp + \int_0^{p_F(n)} \frac{p^2}{2M_n} \left[g_s \frac{V 4\pi p^2}{h^3} \right] dp \\ &= \frac{3}{5} [N E_F(n) + Z E_F(p)] \end{aligned} \quad (4.37)$$

Supposons que le noyau possède environ le même nombre de protons que de neutrons ($|N - Z| \ll A$) et négligeons leur différence de masse. L'énergie cinétique s'écrit alors, en posant $\epsilon = N - Z$ et en effectuant un développement limité,

$$\begin{aligned} T(A) &= \frac{3}{10} E_F A \left[\left(1 + \frac{\epsilon}{A}\right)^{5/3} + \left(1 - \frac{\epsilon}{A}\right)^{5/3} \right] \\ &= \frac{3}{5} E_F A + \frac{1}{3} E_F \frac{(N - Z)^2}{A} + \dots \end{aligned} \quad (4.38)$$

Ici, E_F correspond à la valeur de l'énergie de Fermi pour un gaz de $A/2$ nucléons de masse M . Le premier terme est une contribution au terme de volume tandis que le second contribue au terme d'asymétrie. Cependant, une évaluation numérique montre que ce modèle n'explique qu'une partie du terme. Des effets de recouvrement de fonctions d'onde sont à prendre en compte.

4.5.2 Potentiel nucléaire moyen

Les nucléons se meuvent dans un potentiel moyen tenant compte des effets de tous les autres nucléons dans le noyau. Dans le potentiel, il faut aussi prendre en compte le fait que l'interaction nucléaire est saturée et à courte portée.

Une bonne modélisation de ce potentiel est le potentiel de Wood-Saxon

$$V_{WS} = -\frac{V_0}{1 + e^{\frac{r-R}{0.228a}}} \quad (4.39)$$

Ce potentiel est visible sur la figure 4.8.

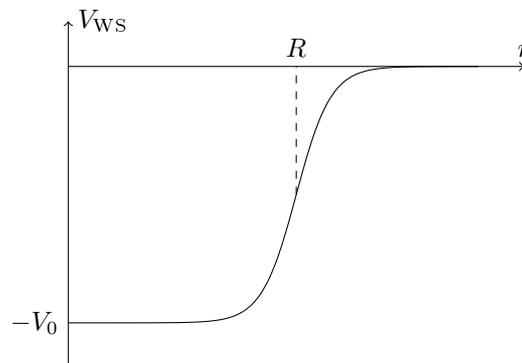


FIGURE 4.8 – Représentation du potentiel moyen de Wood-Saxon

À cela s'ajoute le fait que les protons ressentent l'interaction coulombienne mais aussi un couplage spin-orbite pour tous les nucléons. Le potentiel nucléaire total vaut donc finalement



$$V_{\text{nucl}} = V_{\text{WS}}(r) + V_c(r) + g\mathbf{L} \cdot \mathbf{S} \quad (4.40)$$

Une solution analytique des équations avec un tel potentiel n'est pas possible. Soit on utilise des méthodes perturbatives, soit on approxime le potentiel par d'autres plus simples comme un potentiel carré ou un potentiel harmonique du type

$$V_{\text{harm}} = -V_0 + \frac{1}{2}m\omega_0^2 r^2, \quad \omega_0 = \sqrt{\frac{2V_0}{mR^2}} \quad (4.41)$$

Une correction possible consiste à considérer un terme d'ordre quatre qui contribue au potentiel par un terme en $D\mathbf{L}^2$.

4.5.3 Modèle en couches

Généralités

Compte tenu des hypothèses que nous avons faites précédemment, nous devons donc résoudre l'équation de Schrödinger $H_0\psi = E\psi$ avec le hamiltonien

$$H_0 = \sum_{\alpha=1}^A [T_\alpha + V(r_\alpha) + g\mathbf{L}_\alpha \cdot \mathbf{S}_\alpha] \quad (4.42)$$

Cet hamiltonien est donc une somme de A hamiltonien à un corps. La fonction d'onde ψ_α du nucléon α vérifie l'équation

$$[T_\alpha + V(r_\alpha) + g\mathbf{L}_\alpha \cdot \mathbf{S}_\alpha] \psi_\alpha(\mathbf{r}_\alpha) = \epsilon_\alpha \psi_\alpha(\mathbf{r}_\alpha) \quad (4.43)$$

où ϵ_α est donc l'énergie du nucléon α . L'énergie totale du noyau est alors dans cette approximation la somme des énergies des nucléons individuels

$$E = \sum_{\alpha=1}^A \epsilon_\alpha \quad (4.44)$$

Concernant la fonction d'onde décrivant le noyau, elle se construit à partir des fonctions d'onde $\psi_\alpha(\mathbf{r}_\alpha)$. On rappelle que le noyau est constitué de N neutrons et Z protons qui sont tous des fermions. La fonction d'onde totale doit donc être antisymétrique : il s'agit d'un déterminant de Slater.

$$\psi(\mathbf{r}_1, \dots, \mathbf{r}_A) = \left[\prod_{n=1}^N \psi_n(\mathbf{r}_n) \right]_{\text{Slater}} \left[\prod_{p=1}^Z \psi_p(\mathbf{r}_p) \right]_{\text{Slater}} \quad (4.45)$$

où $[\cdot]_{\text{Slater}}$ désigne le déterminant de Slater. Explicitement, il vaut

$$\left[\prod_{n=1}^N \psi_n(\mathbf{r}_n) \right]_{\text{Slater}} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \cdots & \psi_1(\mathbf{r}_N) \\ \vdots & \ddots & \vdots \\ \psi_N(\mathbf{r}_1) & \cdots & \psi_N(\mathbf{r}_N) \end{vmatrix} \quad (4.46)$$

Solutions pour un potentiel central

Dans le cas d'un potentiel central comme ceux que nous considérons, la solution de l'équation de Schrödinger pour le nucléon α s'écrit de la même manière que la fonction d'onde de l'atome d'hydrogène, à savoir

$$\psi_\alpha(\mathbf{r}_\alpha, \mathbf{s}_\alpha) = R_{nl}(r_\alpha) \mathcal{Y}_l^{m_l}(\theta_\alpha, \phi_\alpha) \chi_s^{m_s}(\mathbf{s}_\alpha) \quad (4.47)$$



où $n \in \mathbb{N}$ est le nombre quantique radial, $l \in \llbracket 0, n-1 \rrbracket$ est le moment angulaire orbital, $m_l \in \llbracket -l, l \rrbracket$ est le moment magnétique orbital, $s = \frac{1}{2}$ est le spin du nucléon et $m_s \in \{-\frac{1}{2}; \frac{1}{2}\}$ est la projection du spin sur l'axe z . Chaque niveau d'énergie dépend *a priori* de n et l . La dégénérescence g_{nl} de chaque niveau d'énergie est alors égale à⁴

$$g_{nl} = (2s + 1)(2l + 1) = 2(2l + 1) \tag{4.48}$$

Dans le cas d'un potentiel harmonique en trois dimensions, l'énergie est donnée par

$$E_N = \left(N + \frac{3}{2}\right) \hbar\omega_0 \tag{4.49}$$

où $N = 2n + l$. On a dans ce cas une dégénérescence en l ce qui conduit, en sommant la relation donnée ci-dessus sur l'intervalle d'existence de l au degré de dégénérescence $g_N = (2s + 1) \frac{(N+1)(N+2)}{2}$. Si l'on rajoute une correction en potentiel, c'est-à-dire si l'on considère

$$V = V_{\text{harm}} + D\mathbf{L}^2 \tag{4.50}$$

la dégénérescence en l est levée et on obtient l'énergie⁵

$$E_{nl} = \left(N + \frac{3}{2}\right) \hbar\omega_0 + D l(l + 1) \hbar^2 \tag{4.51}$$

Cependant, même avec ce potentiel, les résultats théoriques ne reproduisent pas les données expérimentales, notamment *les nombres magiques*, qui donnent des noyaux très stables pour certaines valeurs de Z et de $A - Z$. Ces valeurs sont les mêmes que celles correspondant au nombre d'électrons présent dans une couche électronique pleine.

Interaction spin-orbite

Nous avons vu précédemment qu'une interaction spin-orbite devait être considéré, ce que nous n'avons pas fait pour l'instant. On doit donc tenir compte du terme

$$g\mathbf{L} \cdot \mathbf{S}, \quad g < 0 \tag{4.52}$$

Pour pouvoir déterminer la contribution de ce terme à l'énergie, considérons le moment angulaire total

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \tag{4.53}$$

dont les valeurs propres sont $\hbar^2 j(j + 1)$. Nous avons alors la relation

$$\mathbf{L} \cdot \mathbf{S} = \frac{1}{2} [\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2] \tag{4.54}$$

qui nous donne la contribution recherchée

$$\frac{\hbar^2}{2} [j(j + 1) - l(l + 1) - s(s + 1)] \tag{4.55}$$

Pour un nucléon de spin $\frac{1}{2}$, j ne peut prendre que deux valeurs pour un l donné

$$j = l + \frac{1}{2} \tag{4.56a}$$

$$j = l - \frac{1}{2} \tag{4.56b}$$

4. En effet, pour des valeurs de n et l données, tous les valeurs possibles de m_l redonnent la même énergie. Similairement, la valeur du spin s étant donnée, toutes celles de m_s donnent la même énergie. Compte tenu des bornes, on arrive à la formule donnée.

5. On rappelle que les valeurs propres de \mathbf{L}^2 s'écrivent $l(l + 1)\hbar^2$. Comme \mathbf{L}^2 commute avec l'hamiltonien du système, nous pouvons sommer directement toutes les contributions à l'énergie.

L'interaction spin-orbite divise donc chaque niveau en deux. La contribution de cette interaction est finalement donnée, en utilisant les expressions qui précèdent, par

$$-g \frac{l+1}{2}, \text{ si } j = l - \frac{1}{2} \tag{4.57a}$$

$$g \frac{l}{2}, \text{ si } j = l + \frac{1}{2} \tag{4.57b}$$

En se rappelant que g est négatif, l'écart en énergie est donné par $\Delta E = -g\hbar^2(l + \frac{1}{2})$. L'énergie totale est donc donnée par

$$E_{nl} = \left(2n + l + \frac{3}{2}\right) \hbar\omega_0 + Dl(l+1)\hbar^2 + \begin{cases} -g\hbar^2 \frac{l+1}{2} & \text{si } j = l - \frac{1}{2} \\ g\hbar^2 \frac{l}{2} & \text{si } j = l + \frac{1}{2} \end{cases} \tag{4.58}$$

La dégénérescence est maintenant donnée en étudiant le moment angulaire total \mathbf{J} , qui donne $g_j = 2j + 1$. La prise en compte de cette interaction permet d'expliquer un plus grand nombre de nombres magiques. Cependant, pour des noyaux lourds, les prédictions ne sont plus bonnes. En effet, les effets coulombiens entre protons, que nous avons négligés, prennent de l'importance.

La figure 4.9⁶ montre l'apparition des nombres magiques au fur et à mesure des levées de dégénérescence liées à l'amélioration du potentiel.

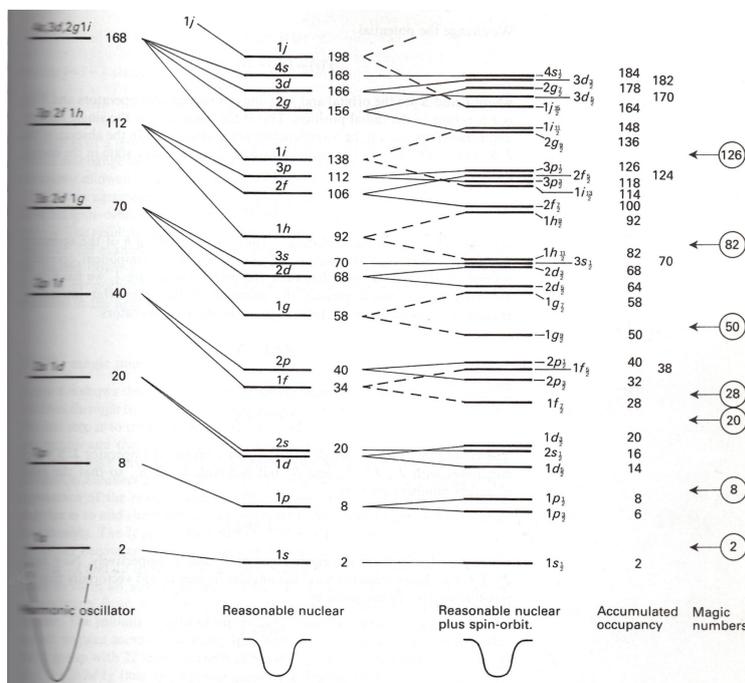


FIGURE 4.9 – Diagramme tiré de [18] montrant les levées de dégénérescence lorsque l'on tient compte des différents termes correctifs du potentiel. On voit bien apparaître les différents nombres magiques comme correspondant à des couches remplies lorsque la dégénérescence est suffisamment levée.

4.6 Taille et forme des noyaux

Après ces deux modèles théoriques permettant de comprendre mieux les noyaux atomiques, il est intéressant de se demander ce qui peut être observé expérimentalement. En effet, la grande variété des noyaux nécessite de se poser la question de leur taille et de leur forme. Comment mesurer ces quantités ? La méthode la plus utilisée pour déterminer la taille d'un noyau repose sur le phénomène de diffusion. On verra ici que la section efficace est directement liée à la distribution de charge du noyau si on décrit le noyau non plus comme ponctuel mais comme un objet étendu.

6. Toutes mes excuses pour la qualité. C. Cabart



4.6.1 Section efficace

Dans un processus de collision, la section efficace est une “aire effective d’interaction”. C’est l’aire fictive minimale que devrait avoir la particule cible pour que la collision ait bien lieu. La section efficace est une mesure de la probabilité de collision. Par définition, la section efficace σ est le rapport entre le taux de diffusion pour une particule cible Γ et le flux de particules incidentes Φ

$$\sigma = \frac{\Gamma}{\Phi} \tag{4.59}$$

Considérons un faisceau de particules incidentes sur une cible d’aire A , d’épaisseur dx et de densité particulaire n . On a donc $N_T = nAdx$ particules cibles contre un flux incident $\phi = \frac{N}{A}$. Par définition de Γ , le nombre de particules diffusées par unité de temps est $N_S = \Gamma N_T$. D’où $\sigma = \frac{\Gamma}{\Phi} = \frac{N_S}{N_T \Phi}$. D’un point de vue expérimental, nous avons donc

$$N_S = \Phi \sigma N_T \tag{4.60}$$

La variation de N est donnée par le nombre de particules diffusées N_S : $N_S = -dN = \Gamma N_T = (\Phi \sigma)(nAdx) = N \sigma n dx$. Nous avons donc $\frac{dN}{dx} = -n \sigma N$. Pour une cible d’épaisseur L ,

$$N = N_0 e^{-\sigma n L} \tag{4.61}$$

4.6.2 Section efficace différentielle

La distribution angulaire des particules diffusées n’est pas nécessairement uniforme. Notons $\delta N_{d\Omega}$ le nombre de particules diffusées dans l’angle solide $d\Omega$. Alors $\delta N_{d\Omega} = \Phi N_T d\sigma$. La section efficace angulaire vaut alors

$$\frac{d\sigma}{d\Omega} = \frac{\delta N_{d\Omega}}{\Phi N_T d\Omega} \tag{4.62}$$

Si l’on intègre sur tous les angles, on obtient la section efficace totale de diffusion σ_s (voir 3.4)

$$\sigma_s = \int \frac{d\sigma}{d\Omega} d\Omega \tag{4.63}$$

4.6.3 Taille et densité des noyaux

Diffusion en mécanique quantique

Pour la théorie derrière cette section, se reporter à 3.4. Ici, seule une dérivation rapide des résultats sera faite. On considère un faisceau de particules incidentes sur un potentiel fixé. On suppose la diffusion élastique, c’est-à-dire que l’énergie de la particule est conservée. La norme de l’impulsion de transfert $\mathbf{q} = \mathbf{p}_i - \mathbf{p}_f$ ⁷, où \mathbf{p}_i est l’impulsion de la particule incidente et \mathbf{p}_f est celle de la particule après diffusion est alors :

$$q = \sqrt{p_i^2 + p_f^2 - 2p_i p_f \cos(\theta)} = 2p \sin \frac{\theta}{2} \tag{4.64}$$

où $p = |p_i| = |p_f|$. La diffusion est caractérisée par la section efficace d’interaction $\sigma = \frac{\Gamma}{\Phi}$. Pour pouvoir mener le calcul, on fait les approximations suivantes : les états initiaux et finals sont décrits par des ondes planes, et on se limite au premier ordre (approximation de Born). On considère donc une onde plane de la forme $\psi = N e^{-i(Et - \mathbf{p} \cdot \mathbf{r})}$. La condition de normalisation de la fonction d’onde dans une boîte de côté L nous donne la constante N et

$$\psi(\mathbf{r}, t) = \frac{1}{L^{\frac{3}{2}}} e^{-i(Et - \mathbf{p} \cdot \mathbf{r})} \tag{4.65}$$

7. On pose $\hbar = 1$, sinon $\mathbf{q} = \frac{\mathbf{p}_i - \mathbf{p}_f}{\hbar}$



Les éléments de la matrice de perturbation s'écrivent, en utilisant leur définition,

$$\begin{aligned}
 M_{if} &= \langle \psi_f | H | \psi_i \rangle \\
 &= \int \psi_f^* H \psi_i d^3\mathbf{r} \\
 &= \frac{1}{L^3} \int e^{-i\mathbf{q}\cdot\mathbf{r}} V(\mathbf{r}) d^3\mathbf{r}
 \end{aligned}
 \tag{4.66}$$

En notant v_i la vitesse des particules incidentes et $n = \frac{1}{L^3}$ leur densité, le flux incident vaut

$$\Phi = \frac{v_i}{L^3}
 \tag{4.67}$$

Pour déterminer la densité d'états, on imagine une boîte de dimension L et on impose des conditions aux limites périodiques. La densité d'états en énergie vaut (voir 3.3.5)

$$\mathcal{D}(E) = \left(\frac{L}{2\pi}\right)^3 p E d\Omega
 \tag{4.68}$$

Finalement, en regroupant tous les termes, la section efficace différentielle s'écrit

$$\frac{d\sigma}{d\Omega} = \frac{1}{(2\pi)^2} \frac{1}{v_i} \left| \int e^{-i\mathbf{q}\cdot\mathbf{r}} V(\mathbf{r}) d^3\mathbf{r} \right|^2 p_f E_f
 \tag{4.69}$$

Dans le cas non relativiste, on retrouve la formule 3.123. Mais, si on considère une diffusion coulombienne sur un noyau ponctuel, les résultats théoriques ne concordent pas avec l'expérience, même en tenant compte de corrections relativistes (formule de Mott).

Diffusion coulombienne sur un noyau étendu

Comme on vient de le dire, un modèle de noyau ponctuel ne redonne pas les observations expérimentales. Cela signifie en fait que le noyau ne peut être considéré comme ponctuel. Pour décrire correctement sa charge, on doit donc utiliser une distribution de charge $\rho(\mathbf{r})$. Cette distribution sera considérée comme normalisée à l'unité, c'est-à-dire que son intégration sur tout l'espace donne 1 et non Ze . On va s'intéresser dans cette partie à la diffusion d'un électron par le noyau. L'énergie potentielle de l'électron à proximité de la distribution de charge est donnée par (intégration de la loi de Coulomb⁸)

$$E(\mathbf{r}) = - \int \frac{Ze^2 \rho(\mathbf{r}') d^3\mathbf{r}'}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|}
 \tag{4.71}$$

Cette formule peut être réécrite comme une convolution du potentiel de Coulomb $\frac{Z\alpha}{r}$ (où $\alpha = \frac{e^2}{4\pi\epsilon_0}$ est la constante de structure fine) et de la distribution de charge $\rho(\mathbf{r})$

$$E(\mathbf{r}) = -Z\alpha \int \frac{\rho(\mathbf{r}') d^3\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}
 \tag{4.72}$$

Pour calculer la section efficace de diffusion dans l'approximation de Born, nous avons vu en 4.69 qu'il faut simplement faire une transformée de Fourier du potentiel. En utilisant le théorème stipulant que la transformée de Fourier d'une convolution n'est que le produit des transformées, nous obtenons

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega}\right)_{\text{Rutherford}} |F(\mathbf{q})|^2
 \tag{4.73a}$$

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{Rutherford}} = \frac{4m^2\alpha^2}{q^4} = \frac{m^2\alpha^2}{4p^4 \sin^4(\theta/2)}
 \tag{4.73b}$$

8. Un élément de charge du noyau est donné par

$$dQ = Ze\rho(\mathbf{r}')d^3\mathbf{r}'
 \tag{4.70}$$



La section efficace différentielle pour un noyau décrit par une distribution de charge n'est autre que le produit de la section efficace de Rutherford avec la transformée de Fourier $F(\mathbf{q})$ de la distribution de charge. $F(\mathbf{q})$ est appelé *facteur de forme*. Dans le cas relativiste, la section efficace de Rutherford est remplacée par celle de Mott.

On voit donc que, grâce à la mesure de la section efficace différentielle, il est possible de remonter à la distribution de charge des noyaux. Cependant, même après une telle correction, la formule n'est pas parfaite et peut être encore améliorée en tenant par exemple compte du recul du noyau ou de son moment magnétique.

Densité de charge

On peut décrire de plusieurs manières la densité de charge du noyau. La plus simple est une distribution uniforme de type sphère dure. Cependant, une distribution plus réaliste est une fonction de type *Fermi – Dirac*

$$\rho(\mathbf{r}) = \frac{\rho_0}{e^{\frac{r-R}{d}} + 1} \tag{4.74}$$

où R est le "rayon" du noyau qui est donné par l'expérience et d est la "largeur" de la surface.

Résumé

Le phénomène de diffusion est efficace pour déterminer les propriétés du noyau *via* la mesure de la section efficace différentielle de diffusion. Dans le cadre de l'approximation de Born, celle-ci est donnée par

$$\frac{d\sigma}{d\Omega} = \frac{1}{(2\pi)^2} \frac{1}{v_i} \left| \int e^{-i\mathbf{q}\cdot\mathbf{r}} V(\mathbf{r}) d^3\mathbf{r} \right|^2 p_f E_f \tag{4.75}$$

où $V(\mathbf{r})$ est le potentiel diffuseur. Lorsqu'on modélise le noyau par une distribution de charge étendue, la section efficace est le produit de deux termes

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega} \right)_{\text{Rutherford}} |F(\mathbf{q})|^2. \tag{4.76}$$

Le premier correspond à la diffusion par un potentiel coulombien classique^a. Le second, $F(\mathbf{q})$, appelé *facteur de forme*, est la transformée de Fourier de la distribution. Une modélisation réaliste de la distribution de charge est une fonction de type Fermi-Dirac

$$\rho(\mathbf{r}) = \frac{\rho_0}{e^{\frac{r-R}{d}} + 1} \tag{4.77}$$

où les paramètres sont donnés par l'expérience. Le rayon du noyau est donné par la formule $R = r_0 A^{\frac{1}{3}}$, avec $r_0 = 1.4$ fm et A le nombre atomique. La distance $d = 2.5$ fm est gouvernée par la portée de l'interaction nucléaire forte tandis que la densité $\rho_0 = 0.08$ fm⁻³ est constante pour tous les noyaux.

a. Ici dans le cas non relativiste, mais peut être remplacé sans problème par la formule de Mott

4.6.4 Moments nucléaires

Afin de caractériser un peu plus finement la manière dont les courants et les charges sont distribués dans le noyau, on va utiliser un développement multipolaire du potentiel créé par le noyau. Le développement multipolaire se fait en effectuant un développement de la fonction $\frac{1}{r-r'}$. Nous avons

$$V(\mathbf{r}) = \frac{Ze}{4\pi r} \left[1 + \frac{1}{r} \int z\rho(\mathbf{r}')d^3\mathbf{r}' + \frac{1}{2r} \int (3z^2 - r'^2)\rho(\mathbf{r}')d^3\mathbf{r}' + \dots \right] \tag{4.78}$$

$\rho(\mathbf{r})$ est ici défini comme précédemment, c'est-à-dire normalisé à l'unité. Les éléments qui nous intéressent le plus dans ce développement sont le moment dipolaire magnétique et le moment quadripolaire électrique. À la limite quantique, $\rho(\mathbf{r}') = |\psi(\mathbf{r}')|^2$. La charge est alors donnée par $Ze \int \psi(\mathbf{r}')\psi^*(\mathbf{r}')d^3\mathbf{r}'$, le moment dipolaire magnétique par $Ze \int \psi(\mathbf{r}')z\psi^*(\mathbf{r}')d^3\mathbf{r}'$ et le moment quadripolaire électrique par $Ze \int \psi(\mathbf{r}')(3z^2 - r'^2)\psi^*(\mathbf{r}')d^3\mathbf{r}'$.



La parité de la fonction d'onde fait que le moment dipolaire électrique est nul. Lorsque le moment quadrupolaire est nul, le noyau a la symétrie sphérique alors que dans le cas contraire, le noyau est un ellipsoïde de révolution. Le signe du moment quadrupolaire donne la direction du grand axe. Ce moment peut être mesuré *via* le spectre hyperfin d'un atome.

UNE CONCLUSION GÉNÉRALE DU CHAPITRE ?



SCORE : 0112

Chapitre 5

Physique des particules

Sommaire

5.1	Observation du monde des particules élémentaires	113
5.2	Particules et interactions	118
5.3	Introduction aux théories de jauge	118
5.4	Équation de Klein-Gordon	121
5.5	Équation de Dirac	123
5.6	Interaction électromagnétique et diagramme de Feynman	132

Cette partie est une introduction élémentaire à la mécanique quantique relativiste appliquée principalement à la physique des particules. L'idée est d'introduire des concepts tels que l'invariance de jauge qui sont à la base de la description de toutes les interactions fondamentales, hormis la gravitation. Cependant, pour pouvoir marier complètement les idées de la théorie quantique et de la relativité restreinte, il faudrait adopter le point de vue des théories quantiques des champs. Dans ce formalisme, le nombre de particules n'est plus fixé et nous raisonnons avec des opérateurs de création et d'annihilation. Ce point de vue n'est pas adopté dans la suite ; on restera dans le cadre plus restreint où le nombre de particules est fixé. Seules les équations de Dirac et de Klein-Gordon seront développées. Elles généralisent l'équation de Schrödinger au cadre relativiste sans pour autant fournir une description complète de tous les processus élémentaires possibles tel que la création et l'annihilation de paires de particules. Enfin, la représentation graphique dite "diagramme de Feynman" sera expliquée.

Pour cette partie, les références que nous avons utilisées sont [12], [10], [8] et [9].

Comme en physique nucléaire, il est difficile de voir quelles parties peuvent être passées ici, mais les parties les plus importantes sont sans doute celles sur l'équation de Dirac et sur les diagrammes de Feynman.

5.1 Observation du monde des particules élémentaires

Les particules élémentaires ne peuvent être observées directement et des détecteurs toujours plus sophistiqués doivent donc être construits. La plupart du temps, la détection des particules requiert une action sur elle (absorption d'énergie, modification de la trajectoire...); on fait interagir la particule avec de la matière. De nombreux processus physiques rentrent en jeu. Généralement, il résulte de ces interactions une perte d'énergie de la particule. Deux processus dominant : les collisions inélastiques avec les atomes du matériau et les collisions élastiques avec les noyaux. D'autres effets, comme l'émission Cerenkov, rayonnement émis lorsqu'une particule dans un matériau va plus vite que la vitesse de la lumière dans ce matériau, doivent être considérés. En ce qui concerne les photons, nous avons l'effet photo-électrique (émission d'un électron d'un atome par absorption du photon), la diffusion Compton (diffusion inélastique d'un photon sur un électron d'un atome, résultant en une modification de la longueur d'onde du photon) ou encore la production de paire.

5.2 Particules et interactions

5.2.1 Introduction aux monde des particules

La nature est constituée de nombreuses particules qui peuvent être classées de différentes manières. Si l'on fait une classification selon le spin, nous avons deux grandes familles de particules : les *fermions* qui sont des



particules de spin demi-entier (elles obéissent à la statistique de Fermi-Dirac) et les *bosons* qui sont des particules de spin entier (et obéissent à la statistique de Bose-Einstein).

À partir de là, nous pouvons encore affiner la classification en classant les particules selon les forces auxquelles elles sont sensibles. Parmi les fermions, on distingue deux familles : les *leptons*, particules insensibles à l'interaction forte et les *quarks*, particules sensibles à cette interaction. Les leptons sont toutes des particules de spin $\frac{1}{2}$. Trois d'entre elles sont chargées, l'électron e^- , le muon μ^- et le τ^- tandis que les trois autres sont neutres et se nomment neutrino électronique ν_e , neutrino muonique ν_μ et neutrino tauique ν_τ . Finalement, nous avons six *saveurs* différentes de leptons. Les quarks, particules de spin $\frac{1}{2}$, ont aussi six saveurs différentes nommées *up*, *down*, *charm*, *strange*, *top* et *bottom*. Ces particules ont des charges fractionnaires. Les quarks sont dotés d'une "couleur" rouge, bleue ou verte¹.

Toutes les particules élémentaires peuvent interagir entre elles *via* différentes interactions. Les forces proviennent de l'échange de particules virtuelles, médiatrices des interactions, appelées *bosons de jauge*. Par exemple, le boson de jauge de la force électromagnétique est le photon. Tous les bosons de jauge connus sont des bosons de spin 1. La masse d'un boson est le reflet de la portée de l'interaction qu'il transmet. En effet, nous avons approximativement $\Delta E \Delta t \sim \hbar$. Avec $E = mc^2$, la portée de l'interaction R vaut

$$R \sim c \Delta t \sim \frac{\hbar}{mc} \tag{5.1}$$

On voit donc que plus la particule est massive, plus l'interaction est de courte portée. Dans le cas extrême où le boson de jauge est de masse nulle, comme le photon, la portée de l'interaction est infinie. L'interaction faible possède trois bosons de jauge massifs, W^-, W^+ et Z^0 . Les *gluons* sont les bosons médiateurs de l'interaction forte. Le tableau 5.1 résume l'ensemble des particules élémentaires connues actuellement.

Fermions (spin demi-entier)						Bosons (spin entier)					
Leptons (spin= 1/2)			Quarks (spin= 1/2)			Interaction électrofaible			Interaction forte		
Saveur	Masse	Charge	Saveur	Masse	Charge	Nom	Masse	Charge	Nom	Masse	Charge
ν_e	$< 10^{-8}$	0	u	0.003	2/3	γ	0	0	g	0	0
e	0.000511	-1	d	0.006	-1/3	photon	0	0	gluon	0	0
ν_μ	< 0.0002	0	c	1.3	2/3	W^-	80.4	-1			
μ	0.106	-1	s	0.1	-1/3	W^+	80.4	+1			
ν_τ	< 0.02	0	t	175	2/3	Z^0	91.187	0			
τ	1.7771	-1	b	4.3	-1/3						

TABLE 5.1 – Tableau présentant les différentes particules élémentaires connues à ce jour. Les fermions sont les constituants de la matière, tandis que les bosons sont les médiateurs d'une interaction. Les masses sont données en GeV/c², et la charge correspond à la charge *électrique* de la particule.

Les quarks ne peuvent être observés individuellement et font toujours partie d'un état lié appelé *hadron* dont la couleur résultante doit être "blanche". Deux types de hadrons existent : ceux composés de trois quarks appelés *baryons* et ceux composés de deux quarks (un quark et son antiparticule associée), les *mésons*.

5.2.2 Nombre leptonique, nombre de saveur

À chaque type de particule, on va associer de plus un nombre : le *nombre leptonique* pour les leptons, et le *nombre de saveur* pour les quarks. Ces nombres sont résumés dans les tableaux 5.2 et 5.3. Enfin, on définit le *nombre baryonique*, qui vaut 1/3 pour un quark et -1/3 pour un anti-quark. L'intérêt de ces nombres est de permettre de voir rapidement quelles réactions sont autorisées en physique des particules. **Seules les réactions conservant les différents nombres leptoniques et le nombre baryonique sont possibles. De plus, parmi celles-ci, on a conservation des nombres de saveurs pour toutes celles qui ne font pas intervenir l'interaction faible.**

5.2.3 L'isospin

Plusieurs faits expérimentaux ont fait émerger l'idée que certaines particules étaient en fait différentes facettes d'un même objet. L'un d'eux concerne l'apparente "symétrie" de l'interaction forte vis-à-vis de la charge électrique, les interactions proton-proton, proton-neutron et neutron-neutron étant identiques. L'égalité presque

¹. Attention, ceci est juste une dénomination de la "charge" associée à l'interaction forte, de la même manière que la "charge" associée à l'interaction électromagnétique est la charge électrique !

	e^-	μ^-	τ^-	e^+	μ^+	τ^+
	ν_e	ν_μ	ν_τ	$\bar{\nu}_e$	$\bar{\nu}_\mu$	$\bar{\nu}_\tau$
L_e	1	0	0	-1	0	0
L_μ	0	1	0	0	-1	0
L_τ	0	0	1	0	0	-1

TABLE 5.2 – Les différents nombres leptoniques associés aux leptons. Le nombre leptonique est conservé par toutes les interactions lors d’une réaction en physique des particules.

	u	\bar{u}	d	\bar{d}	c	\bar{c}	s	\bar{s}	t	\bar{t}	b	\bar{b}
S_u	1	-1	0	0	0	0	0	0	0	0	0	0
S_d	0	0	1	-1	0	0	0	0	0	0	0	0
S_c	0	0	0	0	1	-1	0	0	0	0	0	0
S_s	0	0	0	0	0	0	1	-1	0	0	0	0
S_t	0	0	0	0	0	0	0	0	1	-1	0	0
S_b	0	0	0	0	0	0	0	0	0	0	1	-1
B	1/3	-1/3	1/3	-1/3	1/3	-1/3	1/3	-1/3	1/3	-1/3	1/3	-1/3

TABLE 5.3 – Les différents nombres de saveurs associés aux quarks, ainsi que le nombre baryonique. Une réaction dans laquelle un nombre de saveur n’est pas conservé implique nécessairement l’interaction faible. Enfin, le nombre baryonique total est conservé par toutes les interactions.

parfaite entre les masses de ces particules en est un autre. C’est pourquoi Heisenberg eut l’idée de considérer le neutron et le proton comme étant deux états d’une même particule, le nucléon. On peut faire une analogie avec le spin. Par exemple, les électrons sont des particules de spin 1/2 qui peuvent être classées en deux groupes, ceux ayant une projection du spin sur l’axe de quantification L_z égale à +1/2 et les autres, dont la projection du spin est égale à -1/2. La description mathématique de l’isospin est identique à celle du spin (d’où son nom...).

La symétrie d’isospin est considérée comme exacte pour l’interaction forte bien qu’en réalité elle ne le soit pas. On introduit donc un opérateur \mathbf{I} appelé *isospin*. Dans une approche formelle, on considère un espace abstrait, l’espace d’isospin, dans lequel l’interaction forte est invariante par rotation. La quantité conservée, donnée par le théorème de Noether, est l’isospin. Mathématiquement, l’opérateur \mathbf{I} vérifie les mêmes propriétés que le spin, à savoir

$$[I_i, I_j] = \epsilon_{ijk} I_k \tag{5.2a}$$

$$\mathbf{I}^2 |I, I_3\rangle = I(I + 1) |I, I_3\rangle \tag{5.2b}$$

$$I_3 |I, I_3\rangle = I_3 |I, I_3\rangle \tag{5.2c}$$

Dans le cas d’un isospin 1/2 nous avons le doublet

$$|p\rangle = \left| I = \frac{1}{2}, I_3 = \frac{1}{2} \right\rangle = \left| \frac{1}{2}, \frac{1}{2} \right\rangle \quad |n\rangle = \left| I = \frac{1}{2}, I_3 = -\frac{1}{2} \right\rangle = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \tag{5.3}$$

5.2.4 Sur les interactions fondamentales

Description sommaire

Toutes les particules élémentaires peuvent interagir entre elles *via* différentes interactions. Les forces proviennent de l’échange de particules virtuelles, médiatrices des interactions, appelées *bosons de jauge*. Par exemple, le boson de jauge de la force électromagnétique est le photon. Tous les bosons de jauge connus sont des bosons de spin 1. La masse d’un boson est le reflet de la portée de l’interaction qu’il transmet. En effet, nous avons approximativement $\Delta E \Delta t \sim \hbar$. Avec $E = mc^2$, la portée de l’interaction R vaut

$$R \sim c \Delta t \sim \frac{\hbar}{mc} \tag{5.4}$$

On voit donc que plus la particule est massive, plus l’interaction est de courte portée. Dans le cas extrême où le boson de jauge est de masse nulle, comme le photon, la portée de l’interaction est infinie. L’interaction faible

possède trois bosons de jauge massifs, W^- , W^+ et Z^0 . Les *gluons* sont les bosons médiateurs de l'interaction forte.

Enfin, chaque interaction est caractérisée par une constante de couplage qui donne "l'intensité" de l'interaction ressentie par la particule. En ce qui concerne par exemple l'électromagnétisme, la charge électrique joue le rôle de constante de couplage. Cependant, on considère en général des constantes sans dimension. Pour l'électromagnétisme, ce serait la constante de structure fine $\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c}$.

Propagateurs

On appelle *propagateur* une fonction de Green de l'équation d'évolution d'une particule libre. C'est une distribution.

La plupart du temps on utilisera le propagateur dans l'espace des impulsions, qui est plus simple². Considérons par exemple l'équation de Klein-Gordon, que nous étudierons par la suite,

$$\left[\square + \frac{m^2 c^2}{\hbar^2} \right] \psi(x^\mu) = 0 \tag{5.5}$$

Par définition, une fonction de Green $G(x^\mu, y^\mu)$ est solution de l'équation au sens des distributions³

$$\left[\square_{x^\mu} + \frac{m^2 c^2}{\hbar^2} \right] G(x^\mu, y^\mu) = -\delta(x^\mu - y^\mu) \tag{5.6}$$

Pour pouvoir déterminer complètement la solution, il faudra ajouter des conditions aux limites. Dans l'espace des impulsions, l'équation vérifiée par une fonction de Green est

$$\left[-(p^\mu)^2 + \frac{m^2 c^2}{\hbar^2} \right] G(p^\mu) = -1 \tag{5.7}$$

dont une solution est donnée par

$$G(p^\mu) = \frac{1}{(p^\mu)^2 - \frac{m^2 c^2}{\hbar^2}} \tag{5.8}$$

Conventionnellement, en physique des particules, on choisit des unités telles que $\hbar = c = 1$. De plus, pour les calculs de diagrammes de Feynman, on rajoute un facteur i au propagateur précédent. Nous aurons donc en général sur un diagramme tel que celui de la figure 5.1 un propagateur qui s'écrit

$$G(p) = \frac{i}{p^2 - m^2} \tag{5.9}$$

Les propagateurs sont en fait d'une grande importance en mécanique quantique. Ils sont à la base de la formulation de la mécanique quantique par Feynman. Pour comprendre un peu à quoi correspond le propagateur, regardons le cas non relativiste et l'équation de Schrödinger. Une fonction de Green $K(\mathbf{x}', t'; \mathbf{x}, t)$ sera donc par définition une solution de l'équation

$$\left[-H_x + i\hbar \frac{\partial}{\partial t} \right] K(\mathbf{x}', t'; \mathbf{x}, t) = i\hbar \delta(\mathbf{x} - \mathbf{x}') \delta(t - t') \tag{5.10}$$

Pour pouvoir définir complètement la solution, on impose la condition aux limites suivante

$$K(\mathbf{x}', t'; \mathbf{x}, t) = 0, \quad t' < t \tag{5.11}$$

Le propagateur ainsi défini est une fonction de Green retardée. Cela étant, nous pouvons écrire la fonction d'onde solution de l'équation de Schrödinger, sous forme intégrale, de la façon suivante

2. Pour le déterminer, il faudra donc faire une transformée de Fourier.

3. Le signe $-$ est mis par commodité.

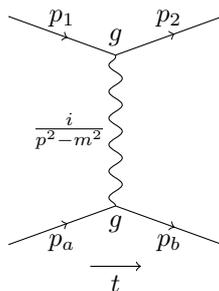


FIGURE 5.1 – Diagramme de Feynman représentant une interaction entre deux particules. Elle se fait par échange d’un boson de spin 0, le propagateur. g représente un couplage entre particule et propagateur. La création et la signification réelle d’un tel diagramme seront discutées plus loin.

$$\psi(\mathbf{x}', t') = \int K(\mathbf{x}', t'; \mathbf{x}, t) \psi(\mathbf{x}, t) d\mathbf{x} dt \tag{5.12}$$

On peut comprendre que le propagateur nous donne l’amplitude de probabilité pour qu’une particule, se trouvant à la position \mathbf{x} au temps t , se retrouve, à l’instant t' , à la position \mathbf{x}' . La formulation de Feynman de la mécanique quantique est entièrement basée sur ces amplitudes de probabilité. Feynman postule qu’une telle amplitude est la somme sur “tous les chemins γ possibles” reliant l’événement \mathbf{x}, t à l’événement \mathbf{x}', t' que l’on note par

$$K(\mathbf{x}', t'; \mathbf{x}, t) = \int e^{\frac{i}{\hbar} S[\gamma]} \mathcal{D}\gamma \tag{5.13}$$

où $S[\gamma]$ est l’action classique calculée sur le chemin γ . Cette intégrale est une intégrale fonctionnelle (voir A.2). L’intérêt d’une telle approche réside dans le fait que l’on raisonne directement dans l’espace-temps et qu’elle se prête donc facilement à une généralisation relativiste. Pour plus de détails, l’approche de Feynman de la mécanique quantique est discutée dans la section 3.1.

Diagramme de Feynman

Les diagrammes de Feynman sont une représentation schématique des phénomènes d’interaction de particules élémentaires. Une chose à comprendre est qu’ils ne représentent pas vraiment la réalité physique. Ils sont simplement un moyen commode de voir les choses, de représenter les amplitudes de transitions. À chaque interaction est associé un ou plusieurs diagrammes de bases que l’on appelle *les vertex*. Ce sont les briques élémentaires permettant de construire un diagramme complet d’interaction. À un processus physique donné est associé une infinité de diagrammes possibles. Cependant, une amplitude de transition peut être décomposée en plusieurs contributions de différents

ordres, chacune étant représentée par un diagramme particulier. Les diagrammes de Feynman associés à un processus sont en fait des représentations du développement perturbatif de l’amplitude de transition. On s’arrêtera dans la suite au premier ordre et nous ne donnerons donc toujours que ces diagrammes.

La figure 5.2 correspond au vertex de base de l’interaction électromagnétique. En effet, nous savons que le champ électromagnétique est véhiculé par le photon et qu’une particule chargée interagit *via* ce champ en émettant ou en absorbant un photon virtuel.

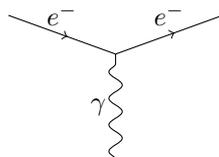


FIGURE 5.2 – Vertex de l’électromagnétisme.

L’interaction forte, quant à elle, est plus délicate. Il existe huit charges de couleur pour les bosons de jauge de cette interaction, les gluons. Cependant, l’idée est toujours la même au niveau du diagramme. La figure 5.3 montre un diagramme complet où deux quarks u colorés interagissent par l’échange d’un gluon.

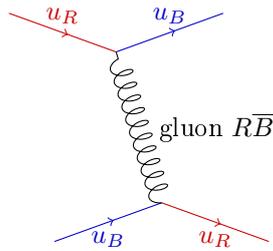


FIGURE 5.3 – Interaction forte.

Enfin, l'interaction faible possède trois bosons de jauge. Tous les vertex possibles pour cette interaction sont représentés sur la figure 5.4.

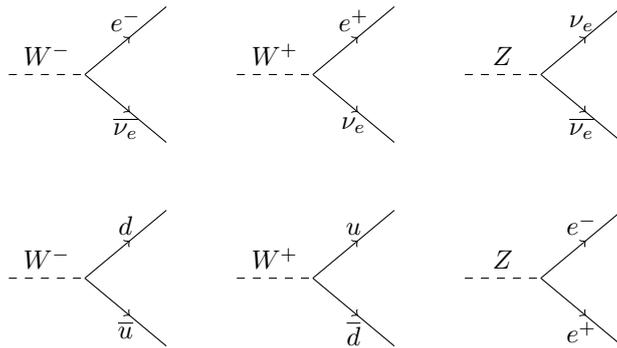


FIGURE 5.4 – Vertex de l'interaction faible.

5.3 Introduction aux théories de jauge

En mécanique classique, l'invariance du lagrangien par translation spatiale, temporelle ou par rotation conduisent à des quantités conservées données par le théorème de Noether. Toutes ces symétries sont des symétries d'espace-temps. Cependant, le lagrangien peut aussi être invariant par d'autres types de symétrie. Par exemple, en mécanique quantique non relativiste, la description d'une particule libre est invariante par la transformation

$$\psi'(\mathbf{x}, t) = e^{i\theta} \psi(\mathbf{x}, t). \tag{5.14}$$

Cette invariance particulière est une *invariance de jauge globale* par le groupe $U(1)$ qui est par définition le groupe des complexes de module unité. On verra dans la suite comment cette idée peut être généralisée et pourquoi on considère de telles transformations via une phase. Avant cela, faisons un retour sur l'invariance de jauge en électromagnétisme classique et sur ses conséquences dans une description quantique.

5.3.1 La quantification canonique

Ce que l'on appelle la quantification canonique est la procédé habituel qui permet de trouver un opérateur A correspondant à une grandeur physique \mathcal{A} définie classiquement (c'est une fonction de la position et de l'impulsion). Le procédé est simple, on remplace la position et l'impulsion dans l'expression de la grandeur par les opérateurs correspondants, tout en symétrisant l'expression (ce point est important puisqu'en mécanique classique les observables commutent alors que ce n'est pas le cas en mécanique quantique). L'énergie quant à elle est remplacée par l'opérateur dérivation temporelle. Pour résumer, nous avons

$$\begin{aligned} E &\rightarrow i\hbar \frac{\partial}{\partial t} \\ \mathbf{p} &\rightarrow -i\hbar \nabla \\ \mathbf{r} &\rightarrow \mathbf{r} \end{aligned} \tag{5.15}$$

Le principe de correspondance, énoncé par Bohr, permet de caractériser des observables à partir de leur grandeur classique associée. Partant de la formulation **hamiltonienne** de la mécanique classique, le principe de correspondance relie la relation de commutation de deux observables A et B aux crochets de Poisson des grandeurs classiques \mathcal{A} et \mathcal{B} par :

$$[A, B] = i\hbar\{\mathcal{A}, \mathcal{B}\} \quad (5.16)$$

5.3.2 Théorie de Maxwell et invariance de jauge

On rappelle que la théorie de Maxwell est invariante par la transformation sur les potentiels suivante

$$A'^{\mu}(x^{\mu}) = A^{\mu}(x^{\mu}) - \partial^{\mu}\chi(x^{\mu}) \quad (5.17)$$

où $\chi(x^{\mu})$ est une fonction régulière. En effet, les équations de Maxwell sont données en fonction du tenseur électromagnétique, c'est pourquoi nous avons, grâce au théorème de Schwartz,

$$F'^{\mu\nu} = \partial^{\mu}A'^{\nu} - \partial^{\nu}A'^{\mu} \quad (5.18)$$

$$\begin{aligned} &= \partial^{\mu}(A^{\nu} - \partial^{\nu}\chi) - \partial^{\nu}(A^{\mu} - \partial^{\mu}\chi) \\ &= \partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu} \\ &= F^{\mu\nu} \end{aligned} \quad (5.19)$$

Le hamiltonien en théorie classique des champs s'écrit

$$H = \frac{1}{2m} (\mathbf{p} - q\mathbf{A}(\mathbf{r}, t))^2 + qV(\mathbf{r}, t) \quad (5.20)$$

En appliquant la quantification canonique, nous obtenons l'équation de Schrödinger en présence d'un champ électromagnétique

$$i\hbar \frac{\partial\psi(\mathbf{r}, t)}{\partial t} = \left[\frac{1}{2m} [-i\hbar\nabla - q\mathbf{A}(\mathbf{r}, t)]^2 + qV(\mathbf{r}, t) \right] \psi(\mathbf{r}, t). \quad (5.21)$$

Le hamiltonien que l'on peut associer à cette équation est appelé *hamiltonien de Pauli*. Une première remarque à faire est la suivante. Si l'on fait un parallèle avec l'équation de Schrödinger libre, on voit qu'il suffit, pour obtenir l'hamiltonien en présence d'un champ, de faire les substitutions suivantes dans l'équation de Schrödinger libre

$$\begin{aligned} \nabla &\rightarrow \nabla - i\frac{q}{\hbar}\mathbf{A} \\ \frac{\partial}{\partial t} &\rightarrow \frac{\partial}{\partial t} + i\frac{q}{\hbar}V \end{aligned} \quad (5.22)$$

Cela étant, la question qui se pose maintenant est de savoir qu'elle est l'influence de la liberté de jauge. En effet, les potentiels ne sont pas uniques, nous en avons une infinité ce qui donne donc une infinité de hamiltoniens possibles pour décrire notre système. Si l'on opère une transformation de jauge, est-ce que les solutions de la nouvelle équation prédisent la même physique? Par une transformation de jauge, l'équation de Schrödinger s'écrit simplement

$$i\hbar \frac{\partial\psi'(\mathbf{r}, t)}{\partial t} = \left[\frac{1}{2m} [-i\hbar\nabla' - q\mathbf{A}'(\mathbf{r}, t)]^2 + qV'(\mathbf{r}, t) \right] \psi'(\mathbf{r}, t). \quad (5.23)$$

On peut montrer, en effectuant le calcul explicite, que la solution de cette équation s'écrit simplement en fonction de celle de départ

$$\psi'(\mathbf{r}, t) = e^{iq\chi(\mathbf{r}, t)}\psi(\mathbf{r}, t). \quad (5.24)$$

Nous voyons donc que la densité de probabilité est la même dans les deux cas. Par contre, si on regarde des quantités dépendantes des dérivées, il faut bien faire attention à substituer les dérivées de la manière indiquée



précédemment si l'on veut conserver l'invariance de jauge de la théorie de Maxwell en mécanique quantique. Par exemple, si l'on regarde la densité de courant de probabilité $\mathbf{j}(\mathbf{r}, t)$, on montre directement qu'elle n'est pas invariante de jauge. La bonne quantité invariante de jauge à regarder est celle où nous avons fait la substitution

$$\partial^\mu \rightarrow D^\mu = \partial^\mu + i\frac{q}{\hbar}A^\mu. \tag{5.25}$$

Pour résumer, en mécanique quantique, l'invariance de jauge de la théorie de Maxwell est conservée si l'on fait les transformations suivantes

$$\begin{aligned} A^\mu(x^\mu) &\rightarrow A'^\mu(x^\mu) = A^\mu(x^\mu) - \partial^\mu\chi(x^\mu) \\ \psi(\mathbf{r}, t) &\rightarrow \psi'(\mathbf{r}, t) = e^{iq\chi(\mathbf{r}, t)}\psi(\mathbf{r}, t) \\ \partial^\mu &\rightarrow D^\mu = \partial^\mu + i\frac{q}{\hbar}A^\mu \end{aligned} \tag{5.26}$$

5.3.3 Invariance de jauge locale

L'idée de l'invariance de jauge locale est de renverser ce qui a été développé ci-dessus. Nous avons supposé d'abord le hamiltonien de l'électromagnétisme connu et nous avons dérivé ensuite les transformations nécessaires pour que l'invariance de jauge de la théorie de Maxwell soit conservée en mécanique quantique. Maintenant, nous ne supposons rien de connu mis à part l'équation d'évolution de la particule libre (dans notre cas particulier l'équation de Schrödinger libre). Nous savons que la description reste inchangée si l'on effectue la transformation de jauge globale

$$\psi'(\mathbf{r}, t) = e^{i\theta}\psi(\mathbf{r}, t). \tag{5.27}$$

Nous imposerons dorénavant que la théorie soit *invariante de jauge locale* par le groupe $U(1)$, c'est à dire invariante par la transformation

$$\psi'(\mathbf{r}, t) = e^{i\frac{q}{\hbar}\theta(\mathbf{r}, t)}\psi(\mathbf{r}, t). \tag{5.28}$$

Le paramètre de phase dépend maintenant de la position de la particule dans l'espace-temps⁴. Voyons ce que donne cette fonction d'onde en la réinjectant dans l'équation de Schrödinger. On arrive alors à l'équation suivante

$$i\hbar\frac{\partial\psi'(\mathbf{r}, t)}{\partial t} = \left[\frac{1}{2m}(\mathbf{p} - q\nabla\theta)^2 + V - q\frac{\partial\theta}{\partial t} \right] \psi'(\mathbf{r}, t) \tag{5.29}$$

Cette équation n'est pas identique à l'équation de Schrödinger. Le problème vient des opérateurs de dérivations. Pour assurer l'invariance de jauge, l'idée est de définir de nouvelles dérivées qui soient covariantes par rapport à la transformation de jauge. En d'autres termes, on cherche des opérateurs de dérivations qui se transforment de la manière suivante

$$\partial^\mu\psi(\mathbf{r}, t) \rightarrow \partial'^\mu\psi'(\mathbf{r}, t) = e^{i\frac{q}{\hbar}\theta(\mathbf{r}, t)}\partial^\mu\psi(\mathbf{r}, t) \tag{5.30}$$

Pour cela, on introduit la dérivation suivante, dite *dérivée covariante*,

$$\partial^\mu \rightarrow D^\mu = \partial^\mu + i\frac{q}{\hbar}A^\mu \tag{5.31}$$

où le potentiel A^μ vérifie la condition de transformation

$$A^\mu(x^\mu) \rightarrow A'^\mu(x^\mu) = A^\mu(x^\mu) - \partial^\mu\theta(x^\mu) \tag{5.32}$$

Cela étant, nous pouvons, au prix de l'introduction d'un champ A^μ , mettre l'équation de Schrödinger sous une forme invariante par transformation de jauge locale sous le groupe $U(1)$ en l'écrivant

4. Le préfacteur $\frac{q}{\hbar}$ est choisi par commodité.



$$i\hbar D_0\psi(\mathbf{r}, t) = -\frac{\hbar^2}{2m}(\mathbf{D})^2\psi(\mathbf{r}, t) \quad (5.33)$$

qui, si on explicite chaque terme, redonne l'équation de Schrödinger en présence d'un champ électromagnétique. Finalement, en imposant à la théorie d'être invariante sous une transformation de jauge locale par le groupe $U(1)$, nous avons été amené à introduire un champ qui n'est rien d'autre que le champ électromagnétique et nous avons obtenu le hamiltonien du système.

5.3.4 Modèle standard

Les interactions fondamentales (faible, forte et électromagnétique) sont toutes construites via une symétrie de jauge locale. La différence entre ces interactions provient du groupe que l'on utilise, $U(1)$ pour l'électromagnétisme, $SU(2)$ pour l'interaction faible et $SU(3)$ pour l'interaction forte (voir **B**, complément sur la théorie des groupes). Ces invariances conduisent à l'introduction de nouveaux champs, appelés *champs de jauge*, qui sont véhiculés par des bosons dits *bosons de jauge* dont les caractéristiques sont déterminées par le groupe de symétrie sous-jacent.

5.4 Équation de Klein-Gordon

L'équation de Klein-Gordon est l'une des premières équation d'onde relativiste qui fut proposée. Ses défauts ont permis plus tard à Dirac de proposer une nouvelle équation dont les succès sont aujourd'hui nombreux. Cependant, l'équation de Klein-Gordon n'est pas sans intérêt puisqu'il s'est avéré qu'elle donnait une description correcte des particules de spin 0. Nous allons en faire une étude rapide.

5.4.1 Dérivation de l'équation de Klein-Gordon

L'idée est de partir de la formule donnant l'énergie en relativité restreinte et d'appliquer la quantification canonique. À première vue il semble difficile, à cause de la racine, d'utiliser la formule suivante

$$E = \sqrt{\mathbf{p}^2 c^2 + m^2 c^4} \quad (5.34)$$

Partons alors de la formule donnant la norme du quadrivecteur impulsion

$$p^\mu p_\mu = m^2 c^2 \quad (5.35)$$

Ceci nous donne l'équation dite de *Klein-Gordon* sous forme covariante

$$-\hbar^2 \partial^\mu \partial_\mu \psi(x^\nu) = m^2 c^2 \psi(x^\nu) \quad (5.36)$$

qui se réécrit explicitement

$$\left(\square + \frac{m^2 c^2}{\hbar^2} \right) \psi(\mathbf{r}, t) = 0 \quad (5.37)$$

L'équation de Klein-Gordon est l'équation relativiste d'évolution permettant de décrire des particules de spin 0. Regardons par exemple une solution en onde plane de la forme

$$\psi(x^\mu) = \mathcal{N} e^{\frac{i}{\hbar} p^\mu x_\mu} \quad (5.38)$$

De telles solutions doivent satisfaire à la relation

$$p^\mu p_\mu = m^2 c^2 \quad (5.39)$$

qui est ce que l'on attend pour une particule libre relativiste. Pourtant, cela implique que pour une impulsion \mathbf{p} donnée, deux solutions pour l'énergie sont possibles. Comment alors leur donner un sens ?



5.4.2 Interprétation probabiliste

Pour pouvoir donner une interprétation probabiliste de la théorie, il est nécessaire d'avoir au moins une équation de conservation des probabilités. Pour l'obtenir, on procède de la même manière que dans le cas non relativiste. En multipliant l'équation 5.37 par $\psi^*(x^\mu)$ d'une part, l'équation conjuguée par $\psi(x^\mu)$ d'autre part et en soustrayant les équations obtenues nous arrivons à l'équation de conservation

$$\partial_\mu j^\mu = 0, \text{ où } j^\mu(x^\nu) = \frac{i\hbar}{2m} [\psi^*(x^\nu) \partial^\mu \psi(x^\nu) - \psi(x^\nu) \partial^\mu \psi^*(x^\nu)] \quad (5.40)$$

La première composante de ce quadrivecteur nous donne ce qui était interprété dans le cas non relativiste par la densité de probabilité

$$\rho(\mathbf{r}, t) = \frac{i\hbar}{2mc^2} \left(\psi^*(\mathbf{r}, t) \frac{\partial \psi(\mathbf{r}, t)}{\partial t} - \psi(\mathbf{r}, t) \frac{\partial \psi^*(\mathbf{r}, t)}{\partial t} \right) \quad (5.41)$$

La différence avec la densité de probabilité non relativiste est qu'ici elle dépend des dérivées temporelles. Cela a pour conséquence que ρ n'est plus toujours à valeurs positives. Pour pouvoir voir cela plus explicitement, il suffit de calculer ρ dans le cas d'une onde plane et on voit qu'elle est proportionnelle à l'énergie, dont les valeurs ne sont pas restreinte à être positives dans cette équation.

La solution envisagée par Pauli et Weisskopf est de considérer non plus j^μ mais qj^μ qui est alors une densité de courant. L'équation de conservation est alors une équation de conservation de la charge. La théorie est donc vue comme une théorie à une charge et non comme une théorie à une particule.

5.4.3 Formulation lagrangienne et invariance de jauge

L'équation précédente n'est pas invariante de jauge. Si on impose cette invariance, nous serons amenés à introduire un champ, le champ électromagnétique, et à modifier l'équation d'évolution. Nous allons faire cela via une formulation lagrangienne. Considérons la densité de lagrangien suivant

$$\mathcal{L}(\partial^\mu \psi, \psi, t) = \partial_\mu \psi^\dagger \partial^\mu \psi - \frac{m^2 c^2}{\hbar^2} \psi^\dagger \psi \quad (5.42)$$

En appliquant l'équation d'Euler-Lagrange, nous arrivons de nouveau à l'équation de Klein-Gordon. Maintenant, nous imposons l'invariance de jauge locale par le groupe $U(1)$. Nous avons vu dans la section 5.3 qu'il suffisait d'introduire un champ de jauge A^μ se transformant de la manière suivante

$$A^\mu(x^\mu) \rightarrow A'^\mu(x^\mu) = A^\mu(x^\mu) - \partial^\mu \chi(x^\mu) \quad (5.43)$$

ainsi que des dérivées covariantes telles que

$$\partial^\mu \rightarrow D^\mu = \partial^\mu + i \frac{q}{\hbar} A^\mu \quad (5.44)$$

Le lagrangien devient alors

$$\mathcal{L}(\partial^\mu \psi, \psi, t) = D_\mu \psi^\dagger D^\mu \psi - \frac{m^2 c^2}{\hbar^2} \psi^\dagger \psi \quad (5.45)$$

$$= \partial_\mu \psi^\dagger \partial^\mu \psi - \frac{m^2 c^2}{\hbar^2} \psi^\dagger \psi - i \frac{q}{\hbar} [A^\mu \psi^\dagger \partial_\mu \psi - \partial_\mu \psi^\dagger A^\mu \psi] + \frac{q^2}{\hbar^2} A_\mu \psi^\dagger A^\mu \psi \quad (5.46)$$

Pour pouvoir déterminer complètement le champ, il faudrait adjoindre les équations reliant le champ aux sources ainsi que la dynamique propre aux champs. Le lagrangien complet de l'électrodynamique quantique pour une particule de spin 0 est finalement

$$\mathcal{L}(\partial^\mu \psi, \psi, t) = \partial_\mu \psi^\dagger \partial^\mu \psi - \frac{m^2 c^2}{\hbar^2} \psi^\dagger \psi - i \frac{q}{\hbar} [A^\mu \psi^\dagger \partial_\mu \psi - \partial_\mu \psi^\dagger A^\mu \psi] + \frac{q^2}{\hbar^2} A_\mu \psi^\dagger A^\mu \psi - \frac{1}{4\mu_0} F^{\mu\nu} F_{\mu\nu} - J^\mu A_\mu \quad (5.47)$$



où J^μ désigne le quadri-vecteur courant de la théorie de Maxwell. On voit bien que la dynamique du champ électromagnétique est découplée de celle du champ de Klein-Gordon. Le lagrangien se décompose en trois termes, l'un correspondant à la particule libre, l'un à l'interaction entre la particule et le champ et le dernier à la dynamique du champ en présence de sources. Nous dériverons dans la suite l'équation de Klein-Gordon pour ce lagrangien et nous en déduirons le potentiel créé par l'interaction électromagnétique. Nous verrons qu'il y aura deux types de contribution, l'une linéaire et l'autre quadratique.

5.4.4 Résumé

Équation de Klein-Gordon

$$-\hbar^2 \partial^\mu \partial_\mu \psi(x^\nu) = m^2 c^2 \psi(x^\nu) \quad \left(\square + \frac{m^2 c^2}{\hbar^2} \right) \psi(\mathbf{r}, t) = 0 \quad (5.48)$$

En formulation lagrangienne, nous avons la densité de lagrangien de Klein-Gordon

$$\mathcal{L}(\partial^\mu \psi, \psi, t) = \partial_\mu \psi^\dagger \partial^\mu \psi - \frac{m^2 c^2}{\hbar^2} \psi^\dagger \psi \quad (5.49)$$

À partir de ces résultats, il est possible de retrouver assez rapidement les autres quantités comme l'équation de conservation ou le lagrangien invariant de jauge. Retenir la méthode de dérivation est plus efficace que les formules explicites.

5.5 Équation de Dirac

L'équation de Dirac est une équation d'onde relativiste qui cherche à palier les problèmes rencontrés par celle de Klein-Gordon. Nous verrons que cette équation nous permet de décrire toutes les particules de spin 1/2 telles que l'électron. En plus de pouvoir lui donner un sens probabiliste, cette équation prédit l'existence des antiparticules, non encore découverte à l'époque de sa publication par Dirac. Un des autres succès de cette équation est d'incorporer directement le spin d'une particule. La structure fine des raies de l'atome d'hydrogène est en conséquence expliquée sans besoin de rajouter de manière *ad hoc* des interactions dans le hamiltonien du système.

5.5.1 Moment cinétique, spin et matrices de Pauli

Moment cinétique

On définit le moment cinétique comme un opérateur vectoriel \mathbf{J} dont les composantes sont des opérateurs hermitiens vérifiant les relations de commutation

$$[J_i, J_j] = i\hbar \epsilon_{ijk} J_k \quad (5.50)$$

Cette relation impose des restrictions sur les valeurs propres du moment cinétique. Pour cela, considérons l'opérateur \mathbf{J}^2 défini par

$$\mathbf{J}^2 = J_x^2 + J_y^2 + J_z^2. \quad (5.51)$$

Grâce aux relations de commutation, on peut montrer que

$$[J_i, \mathbf{J}^2] = 0 \quad (5.52)$$

que l'on note souvent $[\mathbf{J}, \mathbf{J}^2] = 0$. Cela étant, il est possible de former un ensemble complet d'opérateurs qui commutent en considérant \mathbf{J}^2 et une des composantes, par exemple J^z . Comme \mathbf{J}^2 est un opérateur hermitique,



Moment cinétique

Un opérateur vectoriel \mathbf{J} est appelé *moment cinétique* lorsque ses composantes sont des opérateurs hermitiques vérifiant les relations de commutation

$$[J_i, J_j] = i\hbar\epsilon_{ijk}J_k \tag{5.57}$$

L'opérateur $\mathbf{J}^2 = J_x^2 + J_y^2 + J_z^2$ commute avec \mathbf{J} . \mathbf{J}^2 et J_z peuvent former un ensemble complet d'opérateurs qui commutent. On note $|j, m\rangle$ les vecteurs propres communs et

$$\mathbf{J}^2|j, m\rangle = j(j+1)\hbar^2|j, m\rangle \tag{5.58a}$$

$$J_z|j, m\rangle = m\hbar|j, m\rangle \tag{5.58b}$$

Nous pouvons obtenir, grâce aux relations de commutation et aux opérateurs

$$J_+ = J_x + iJ_y \quad J_- = J_x - iJ_y \tag{5.59}$$

les restrictions

$$j \in \mathbb{N}/2 \tag{5.60a}$$

$$m \in \llbracket -j, j \rrbracket \tag{5.60b}$$

Le sous espace propre associé au moment cinétique j est de dimension $2j + 1$.

Spin et matrices de Pauli

Pour pouvoir interpréter les expériences, comprendre les spectres atomiques, l'hypothèse que les particules possèdent un moment cinétique intrinsèque appelé *le spin*, \mathbf{s} , fut formulée. Au spin, on associe un moment magnétique $\boldsymbol{\mu}_s$ tel que

$$\boldsymbol{\mu}_s = g_s \frac{e}{2mc} \mathbf{s} \tag{5.61}$$

Le facteur g_s appelé facteur gyromagnétique est égal, pour l'électron, particule de spin 1/2, à deux. La théorie relativiste de l'électron, en d'autres termes l'équation de Dirac, explique la valeur de ce terme. Spécifions un peu le spin 1/2. D'après ce qui précède, l'opérateur \mathbf{s}^2 n'a qu'une seule valeur propre valant $3\hbar^2/4$. Les opérateurs agissent sur un espace de dimension deux. On peut montrer que $\mathbf{s}_+^2 = \mathbf{s}_-^2 = 0$, ce qui donne, en écrivant explicitement les opérateurs via leur définition, la relation d'anticommutation

$$\{s_i, s_j\} = \frac{\hbar^2}{2} \delta_{ij} \tag{5.62}$$

Introduisons directement les *matrices de Pauli*, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ ainsi que les relations qu'elles vérifient

$$\mathbf{s} = \frac{\hbar}{2} \boldsymbol{\sigma} \tag{5.63a}$$

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij} \tag{5.63b}$$

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \tag{5.63c}$$

Donnons une représentation explicite de ces matrices. Nous devons exhiber deux vecteurs formant une base. Comme $\sigma_z^2 = 1$ et que la matrice est hermitique, ses valeurs propres sont ± 1 . Soit $|+\rangle$ le vecteur propre de σ_z associé à la valeur propre $+1$. Posons $|-\rangle = \sigma_x|+\rangle$. Par anticommution avec σ_z , nous avons $\sigma_z|-\rangle = -|-\rangle$. Nous avons ainsi nos vecteurs de base. Par ailleurs, $\sigma_x|\pm\rangle = |\mp\rangle$. Pour finir, avec les relations de commutation et d'anticommutation, nous avons $\sigma_z\sigma_x = i\sigma_y$, ce qui permet de déterminer la matrice de l'opérateur σ_y .



Entre autres propriétés, les matrices de Pauli sont toutes de trace nulle et de déterminant -1 , ce qui peut se montrer directement à partir des relations algébriques. On peut calculer explicitement l'opérateur de rotation associé à ce moment dans l'espace des spins qui s'écrit de manière générale

$$\mathcal{R}_{\mathbf{n}}(\theta) = e^{-\frac{i}{2}\theta\boldsymbol{\sigma}\cdot\mathbf{n}} \tag{5.64}$$

Résumons les propriétés principales des matrices de Pauli.

Matrices de Pauli

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij} \quad [\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \tag{5.65}$$

$$\text{Tr}(\sigma_i) = 0 \quad \det(\sigma_i) = -1 \tag{5.66}$$

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{5.67}$$

$$\mathcal{R}_{\mathbf{n}}(\theta) = e^{-\frac{i}{2}\theta\boldsymbol{\sigma}\cdot\mathbf{n}} = \cos\left(\frac{\theta}{2}\right) - i\boldsymbol{\sigma}\cdot\mathbf{n}\sin\left(\frac{\theta}{2}\right), \quad \mathbf{n} \text{ vecteur unitaire.} \tag{5.68}$$

Spineur de Pauli

De par sa définition, le spin commute avec les opérateurs impulsion et position. Nous pouvons donc décrire l'espace des états comme un produit tensoriel. Dans le cas d'une particule de spin $1/2$, l'espace des états de spin \mathcal{H}_s est un espace de Hilbert complexe de dimension deux. Nous avons donc l'isomorphisme

$$\mathcal{H}_s \cong \mathbb{C}^2 \tag{5.69}$$

En représentation position, la particule à l'instant t est donc décrite par

$$\psi(\mathbf{r}) = \begin{pmatrix} \psi_{+1/2}(\mathbf{r}) \\ \psi_{-1/2}(\mathbf{r}) \end{pmatrix} \in \mathcal{L}^2(\mathbb{R}^3) \otimes \mathbb{C}^2 \tag{5.70}$$

Un tel élément est appelé *spineur de Pauli*. Une particule de spin $1/2$ est donc décrite au cours du temps par un champ de spineur de Pauli. Un tel élément a des propriétés de transformation particulières, différentes de celles d'un vecteur classique. Si l'on considère une rotation \mathcal{R} du spineur, celle-ci peut se décomposer en un produit de deux rotations qui commutent agissant chacune sur l'un des espaces formant le produit tensoriel. On aura

$$\mathcal{R}(\alpha, \beta, \gamma) = \underbrace{\mathcal{R}_s(\alpha, \beta, \gamma)}_{\text{agit sur } \mathbb{C}^2} \underbrace{\mathcal{R}_0(\alpha, \beta, \gamma)}_{\text{agit sur } \mathcal{L}^2(\mathbb{R}^3)} \tag{5.71}$$

Le spineur se transformera alors de la manière suivante

$$\psi'(\mathbf{r}) = [\mathcal{R}(\alpha, \beta, \gamma)\psi](\mathbf{r}) = \mathcal{R}_s(\alpha, \beta, \gamma) \begin{pmatrix} \psi_{+1/2}(\mathcal{R}_0^{-1}(\alpha, \beta, \gamma)\mathbf{r}) \\ \psi_{-1/2}(\mathcal{R}_0^{-1}(\alpha, \beta, \gamma)\mathbf{r}) \end{pmatrix} \tag{5.72}$$



5.5.2 L'équation de Dirac

Le problème majeur de l'équation de Klein-Gordon était qu'il était impossible de donner une interprétation probabiliste correcte des solutions. La cause était que l'équation était du second ordre. L'équation proposée par Dirac cherche à éliminer ce problème. En plus de l'invariance relativiste, Dirac imposa à l'équation d'évolution d'être du premier ordre et imposa également que les solutions vérifient l'équation de Klein-Gordon (ce qui permet d'assurer que la relation relativiste de l'énergie soit satisfaite). On peut comprendre que pour garder la symétrie, requise en relativité, entre les coordonnées d'espace et de temps, nous devons considérer des dérivées spatiales du premier ordre. À cela s'ajoute le fait que lorsque la particule est au repos son énergie est égale à mc^2 . Nous devons, dans la limite non relativiste, retrouver le hamiltonien de Schrödinger, voire celui de Pauli 5.21 quand on tiendra compte de l'invariance de jauge. Le hamiltonien a donc la forme générale suivante

$$i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} = H_D \psi(\mathbf{r}, t) = (c\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc^2) \psi(\mathbf{r}, t) \tag{5.73}$$

Pour rester le plus général possible, nous avons *a priori*

$$\psi(\mathbf{r}, t) = \begin{pmatrix} \psi_1(\mathbf{r}, t) \\ \vdots \\ \psi_n(\mathbf{r}, t) \end{pmatrix} \in \mathcal{L}^2(\mathbb{R}^3) \otimes \mathbb{C}^n \tag{5.74}$$

c'est pourquoi $\boldsymbol{\alpha}$ est un vecteur de matrices et β une matrice. On impose que les composantes soient de carré intégrable pour l'interprétation probabiliste. À noter que nous nous sommes placés dans une représentation particulière de l'espace de Hilbert produit tensoriel. Il reste tout de même à préciser les espaces qui le composent, surtout celui isomorphe à \mathbb{C}^n . On remarque l'analogie avec les spineurs de Pauli. Le problème est donc de déterminer les opérateurs que nous avons introduit. Réécrivons l'équation comme

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \boldsymbol{\alpha} \cdot \nabla + i \frac{mc}{\hbar} \beta \right) \psi(\mathbf{r}, t) = 0 \tag{5.75}$$

et en la multipliant par l'opérateur $(\frac{1}{c} \frac{\partial}{\partial t} - \boldsymbol{\alpha} \cdot \nabla - i \frac{mc}{\hbar} \beta)$ nous pouvons arriver à

$$\left[\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \frac{1}{2} \sum_{i,j} (\alpha_i \alpha_j + \alpha_j \alpha_i) \frac{\partial^2}{\partial x_i \partial x_j} - i \frac{mc}{\hbar} \sum_k (\alpha_k \beta + \beta \alpha_k) \frac{\partial}{\partial t} + \left(\frac{mc}{\hbar} \right)^2 \beta^2 \right] \psi(\mathbf{r}, t) = 0 \tag{5.76}$$

Pour que cette équation se réduise à celle de Klein-Gordon, nous devons imposer les relations suivantes

$$\{\alpha_i, \alpha_j\} = 2\delta_{ij} \quad \{\alpha_i, \beta\} = 0 \quad \beta^2 = \mathbb{1} \tag{5.77}$$

Reste maintenant à déterminer l'espace sur lequel agissent les opérateurs α_i et β . On prendra l'espace le plus simple dont une base est formée de ces opérateurs. On peut montrer en utilisant une méthode similaire à celle qui a permis de construire les matrices de Pauli, que l'espace le plus simple est un espace de dimension quatre et que les opérateurs peuvent s'écrire

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix} \quad \beta = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{pmatrix} \tag{5.78}$$

où $\boldsymbol{\sigma}$ est le vecteur de matrices formé des matrices de Pauli. Finalement, l'équation de Dirac s'écrit sous forme matricielle

$$i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} = \begin{pmatrix} mc^2 \mathbb{1}_2 & c\boldsymbol{\sigma} \cdot \mathbf{p} \\ c\boldsymbol{\sigma} \cdot \mathbf{p} & -mc^2 \mathbb{1}_2 \end{pmatrix} \psi(\mathbf{r}, t). \tag{5.79}$$

On remarque que le carré du hamiltonien de Dirac est $(\mathbf{p}^2 c^2 + m^2 c^4) \mathbb{1}_4$, ce qui est cohérent avec le fait que nous avons imposé que les solutions de cette équation soient solutions de l'équation de Klein-Gordon, elle-même

dérivée du prérequis de retrouver l'énergie relativiste. Donnons encore une autre écriture de l'équation pour que son invariance relativiste soit plus manifeste.

$$\begin{pmatrix} i\hbar\partial_t & i\hbar\boldsymbol{\sigma}\cdot\nabla \\ -i\hbar\boldsymbol{\sigma}\cdot\nabla & -i\hbar\partial_t \end{pmatrix} \psi(\mathbf{r}, t) = mc^2\psi(\mathbf{r}, t) \Leftrightarrow i\hbar(\gamma^0\partial_0 + \gamma^i\partial_i) = mc\psi(\mathbf{r}, t) \quad (5.80)$$

d'où finalement l'équation sous forme covariante

$$(i\hbar\gamma^\mu\partial_\mu - mc)\psi(\mathbf{r}, t) = 0 \quad (5.81)$$

Explicitement, les matrices γ^μ s'écrivent

$$\gamma^0 = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{pmatrix} \quad \gamma^i = \begin{pmatrix} 0 & \boldsymbol{\sigma}^i \\ -\boldsymbol{\sigma}^i & 0 \end{pmatrix} \quad (5.82)$$

On peut vérifier que ces matrices ont les propriétés suivantes

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbb{1}_4 \quad (5.83a)$$

$$(\gamma^0)^2 = \mathbb{1}_4 \quad (5.83b)$$

$$(\gamma^i)^2 = -\mathbb{1}_4 \quad (5.83c)$$

$$\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0 \quad (5.83d)$$

Dans la suite, nous allons donner une autre représentation de l'équation de Dirac qui va nous permettre de donner une interprétation probabiliste de la fonction d'onde, ou en d'autres termes, trouver une équation de conservation des probabilités. Avant faisons un point. En physique des particules, nous prenons le plus souvent $\hbar = c = 1$ ⁵.

Équation de Dirac

$$(i\gamma^\mu\partial_\mu - m)\psi(\mathbf{r}, t) = 0 \quad \psi(\mathbf{r}, t) \in \mathcal{L}^2(\mathbb{R}^3) \otimes \mathbb{C}^4 \quad (5.84)$$

$\psi(\mathbf{r}, t)$ est appelé *spineur de Dirac* et les matrices γ^μ sont définies par et ont les propriétés suivantes

$$\gamma^0 = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & -\mathbb{1}_2 \end{pmatrix} \quad \gamma^i = \begin{pmatrix} 0 & \boldsymbol{\sigma}^i \\ -\boldsymbol{\sigma}^i & 0 \end{pmatrix} \quad (5.85)$$

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu} \quad (5.86a)$$

$$(\gamma^0)^2 = \mathbb{1}_4 \quad (5.86b)$$

$$(\gamma^i)^2 = -\mathbb{1}_4 \quad (5.86c)$$

$$\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0. \quad (5.86d)$$

5.5.3 Interprétation probabiliste

Équation adjointe

Développons l'équation de Dirac et prenons-en le conjugué

$$-i\partial_\mu\psi^\dagger(\mathbf{r}, t)(\gamma^\mu)^\dagger - m\psi^\dagger(\mathbf{r}, t) = 0 \quad (5.87)$$

En utilisant les propriétés des matrices γ^μ , nous arrivons à l'équation *adjointe* de Dirac

⁵. Nous les utiliserons dans toute la suite!



$$i\partial_\mu \bar{\psi}(\mathbf{r}, t) \gamma^\mu + m\bar{\psi}(\mathbf{r}, t) = 0 \quad (5.88)$$

où par définition $\bar{\psi}$ est l'adjoint de ψ : $\bar{\psi} = \psi\gamma^0$

Quadrivecteur densité de courant

En multipliant l'équation de Dirac par $\bar{\psi}$ et l'équation adjointe par ψ , on arrive à une équation de conservation sous forme covariante

$$\partial_\mu (\bar{\psi}(\mathbf{r}, t) \gamma^\mu \psi(\mathbf{r}, t)) = 0. \quad (5.89)$$

On définit alors le quadrivecteur densité de courant de probabilité j^μ par

$$j^\mu = \bar{\psi} \gamma^\mu \psi \quad (5.90)$$

5.5.4 États stationnaires

On cherche à trouver les états stationnaires en ondes planes de l'équation de Dirac, autrement dit, on cherche des solutions de la forme

$$\psi(x^\mu) = \mathcal{N} \omega(p^\mu) e^{-ip^\mu x_\mu} \quad (5.91)$$

où $\omega(p^\mu)$ est un spineur à quatre composantes que l'on écrira en fait dans la suite comme un spineur à deux composantes de spineur, soit

$$\omega = \begin{pmatrix} \phi \\ \chi \end{pmatrix} \quad (5.92)$$

En explicitant l'équation de Dirac ainsi que les quadrivecteurs, nous arrivons à l'équation suivante, qui n'est rien d'autre que l'équation aux valeurs propres du hamiltonien de Dirac

$$H_D \omega = E \omega \Leftrightarrow \begin{pmatrix} m \mathbb{1}_2 & \boldsymbol{\sigma} \cdot \mathbf{p} \\ \boldsymbol{\sigma} \cdot \mathbf{p} & -m \mathbb{1}_2 \end{pmatrix} \begin{pmatrix} \phi \\ \chi \end{pmatrix} = E \begin{pmatrix} \phi \\ \chi \end{pmatrix} \quad (5.93)$$

La seconde ligne de cette équation nous permet de donner χ comme fonction de ϕ par ⁶ $\chi = \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m} \phi$. La forme générale du spineur ω est donc

$$\omega(p^\mu) = \begin{pmatrix} \phi \\ \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m} \phi \end{pmatrix} \quad (5.94)$$

Mais quelles sont les valeurs propres, les énergies E autorisées ? En multipliant l'équation aux valeurs propres 5.93 par H_D (on se rappelle que les solutions de l'équation de Dirac vérifient l'équation de Klein-Gordon), nous arrivons après quelques calculs à la relation relativiste de l'énergie. Les énergies sont donc telles que

$$E = \pm \sqrt{\mathbf{p}^2 + m^2} \quad (5.95)$$

Les solutions d'énergie négative sont donc toujours autorisées dans la théorie de Dirac. Cependant, nous verrons dans la suite que nous pouvons en donner une interprétation et nous distinguerons les solutions positives et négatives de l'équation de Dirac. On peut déjà faire une première remarque. L'ensemble des énergies positives est minoré par mc^2 (considérer une particule au repos dans un référentiel) tandis que l'ensemble des énergies négatives est majoré par $-mc^2$. Il existe donc un "gap" en énergie de $2mc^2$ entre les états d'énergie positive et négative.

⁶. On rappelle que l'on raisonne ici avec des matrices, il faut donc vérifier qu'elles sont bien inversibles si on veut faire des manipulations algébriques correctes. Ici, c'est évident.



5.5.5 Interprétation de Feynman

Les équations d'onde relativistes ont été dérivées à partir de l'expression de l'énergie relativiste. L'équation de Klein-Gordon prédisait l'existence de probabilités négatives ainsi que des énergies négatives, ce qui rendait impossible une interprétation probabiliste correcte. Celle de Dirac au contraire donne un sens à cette interprétation, les probabilités étant cette fois positives, mais l'existence d'énergies négatives demeure.

L'interprétation de Dirac repose sur l'idée que le nombre de particules est indéterminé à cause de la possibilité de création de paires. Ceci montre qu'en fait le cadre de la théorie des champs est indispensable. Dirac suppose que l'ensemble des états d'énergies négatives est rempli par une infinité de particules. L'énergie maximale de cette mer de Dirac est de $-mc^2$ et est séparé par un gap en énergie de $2mc^2$ avec le niveau d'énergie positif le plus bas. Le système composé de la mer de Dirac remplie constitue le vide électronique, son énergie est prise égale à zéro tout comme sa charge. Un électron est alors interprété comme une excitation au-dessus de cette mer. Le principe de Pauli exclut la possibilité qu'un électron puisse tomber dans la mer de Dirac. Pour exciter la mer de Dirac, il faut fournir une énergie au moins égale à celle du gap (on peut faire un parallèle avec l'énergie minimale de création d'une paire). Cela étant, nous avons un électron d'une part et une lacune dans la mer. Nous pouvons attribuer à ce trou une charge positive, qui correspond alors au *positron*, antiparticule de l'électron. Cette nouvelle particule fut prédite par Dirac en 1928.

Bien que forte de plusieurs succès, l'idée de Dirac a plusieurs limites. D'abord, son interprétation repose sur un nombre de particules indéterminés alors que la théorie prétend ne décrire qu'une seule particule. De plus, cette interprétation ne peut être faite pour les solutions de l'équation de Klein-Gordon puisqu'elles décrivent des bosons de spin 0. Pour s'affranchir de ces difficultés, Feynman donna l'interprétation suivante

Des particules d'énergie négative se propageant dans le sens opposé à la flèche du temps sont équivalentes à des antiparticules d'énergie positive se propageant dans le sens du temps.

En conséquence, l'émission (l'absorption) d'une antiparticule de quadrivecteur impulsion p^μ est physiquement équivalente à l'absorption (l'émission) d'une particule de quadrivecteur impulsion $-p^\mu$. Dans le cas des fermions, il est de plus nécessaire de changer le signe du spin⁷.

5.5.6 Spineurs et énergie négative

Nous allons maintenant analyser plus précisément les spineurs de Dirac en fonction du signe de l'énergie et utiliser l'interprétation de Feynman pour en donner la forme conventionnelle. Nous ne nous intéressons pas tout de suite à la normalisation que l'on reporte à une section suivante. On regarde donc simplement le spineur ω . Il vérifie l'équation aux valeurs propres 5.93 écrite de manière explicite

$$p^0\phi = m\phi + \boldsymbol{\sigma} \cdot \mathbf{p}\chi \quad p^0\chi = \boldsymbol{\sigma} \cdot \mathbf{p}\phi - m\chi \quad (5.96)$$

Supposons d'abord l'énergie de l'électron E positive, l'impulsion s'écrivant alors $p^\mu(E, \mathbf{p})$. Dans le référentiel où l'électron est au repos, nous avons alors $\mathbf{p} = 0$ et $E = m$. Ceci implique donc que $\chi = 0$. Dans ce référentiel particulier, nous avons donc

$$\omega = \begin{pmatrix} \phi \\ 0 \end{pmatrix} \quad (5.97)$$

Ce spineur peut s'écrire comme une combinaison linéaire de deux spineurs, l'un où $\phi = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et l'autre où $\phi = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Ces deux spineurs sont généralement notés respectivement ϕ^1 et ϕ^2 . Maintenant, si l'on se place dans un référentiel quelconque, on peut trouver la forme du spineur en éliminant naturellement χ des équations. Finalement,

$$\omega^{1,2} = \begin{pmatrix} \phi^{1,2} \\ \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m} \phi^{1,2} \end{pmatrix} \quad \phi^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \phi^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (5.98)$$

Regardons maintenant les états d'énergie négative. On écrira alors $p^\mu(-E, \mathbf{p})$ ⁸. En suivant la même démarche, en se plaçant dans le référentiel au repos, nous avons $E = m$ donc $p^0 = -m$. Ceci implique cette fois $\phi = 0$. Tout

⁷. Pour les fermions, le signe de l'amplitude de transition est changé lorsqu'on passe d'un point de vue à l'autre. Ceci n'est pas forcément important puisqu'on regarde le module au carré.

⁸. Ainsi, partout où apparaît l'énergie dans les spineurs, c'est le module de l'énergie que l'on considère, et que l'on note E . Le signe est incorporé dans les expressions.



comme précédemment, le spineur ω est combinaison de deux spineurs de base. Dans un référentiel quelconque, en éliminant cette fois naturellement ϕ , nous avons

$$\omega = \begin{pmatrix} -\frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m} \chi \\ \chi \end{pmatrix} \tag{5.99}$$

On va maintenant utiliser l'interprétation de Feynman. Un fermion d'énergie négative d'impulsion p^μ dont l'orientation du spin est donnée est alors vu comme un anti-fermion d'énergie positive d'impulsion $-p^\mu$ dont l'orientation du spin est opposée. Nous définissons alors les spineurs associés aux énergies négatives, $s \in \{1, 2\}$

$$\omega^s = \begin{pmatrix} \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m} \chi^s \\ \chi^s \end{pmatrix} \quad \chi^1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \chi^2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{5.100}$$

Le fait que pour une énergie donnée, la particule puisse être décrite par deux spineurs différents donne bien l'idée que l'équation de Dirac décrit bien les particules de spin 1/2.

5.5.7 Le spin

Jusqu'alors nous avons affirmé que l'équation de Dirac décrivait des particules de spin 1/2, avons parlé d'électron et de positron. Bien que nous puissions nous en convaincre de différentes façon (par les spineurs, la présence des matrices de Pauli), ceci peut parfaitement bien se montrer à partir de l'équation elle-même. Une manière de voir est de montrer que le vecteur

$$\mathbf{J} = \mathbf{L} + \frac{\hbar}{2} \boldsymbol{\Sigma} \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \tag{5.101}$$

commute avec le hamiltonien de Dirac H_D . Ce vecteur peut être interprété comme un moment cinétique vu qu'il contient déjà le vecteur \mathbf{L} . Ceci se fait assez facilement à partir de l'écriture explicite de H_D . Ainsi $\mathbf{S} = \frac{\hbar}{2} \boldsymbol{\Sigma}$ peut s'interpréter comme un moment cinétique intrinsèque, *le spin*, qui, ajouté au moment cinétique orbital \mathbf{L} , forme le moment cinétique total \mathbf{J} qui est une constante du mouvement. On peut de plus vérifier à partir des propriétés des matrices de Pauli que ce moment de spin vérifie toutes les propriétés d'un moment cinétique défini dans la section 5.5.1.

5.5.8 Formulation lagrangienne et invariance de jauge

On peut donner une formulation lagrangienne de l'équation de Dirac. Il suffit pour cela de considérer la densité de lagrangien suivante

$$\mathcal{L}(\partial^\mu \psi, \psi, t) = \bar{\psi} (i\gamma^\mu \partial_\mu - m) \psi \tag{5.102}$$

Si l'on impose l'invariance de jauge locale par le groupe $U(1)$, nous tenons compte du champ électromagnétique. Nous avons alors

$$\mathcal{L}(\partial^\mu \psi, \psi, t) = \bar{\psi} (i\gamma^\mu D_\mu - m) \psi \tag{5.103}$$

En développant la dérivée, nous voyons directement arriver le terme supplémentaire correspondant à l'interaction entre le champ et les particules. Explicitement

$$\mathcal{L}(\partial^\mu \psi, \psi, t) = \bar{\psi} (i\gamma^\mu \partial_\mu - m) \psi - e \underbrace{\bar{\psi} \gamma^\mu \psi}_{j^\mu} A_\mu \tag{5.104}$$

Le terme d'interaction est donc proportionnel au courant j^μ . Si l'on calcule les équations d'Euler Lagrange et que nous déterminons le potentiel créé par l'interaction électromagnétique, on voit que, contrairement au cas des bosons de spin 0, le potentiel A^μ n'intervient que de façon linéaire. Pour avoir le lagrangien complet, il faut ajouter la dynamique propre au champ ainsi qu'un éventuel terme de source. Le lagrangien complet de l'électrodynamique quantique pour une particule de spin 1/2 en présence d'un champ est donc

$$\mathcal{L}(\partial^\mu \psi, \psi, t) = \bar{\psi} (i\gamma^\mu D_\mu - m) \psi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - J^\mu A_\mu \tag{5.105}$$



5.6 Interaction électromagnétique et diagramme de Feynman

Nous allons voir dans ce qui suit les bases permettant de faire des calculs en électrodynamique quantique. Comme précédemment, nous n'utiliserons pas le formalisme de la théorie des champs. Les méthodes perturbatives sont efficaces pour cette théorie du fait de la petite valeur de la constante de structure fine $\alpha = 1/137$ et, en ce qui nous concerne, nous effectuerons des calculs au premier ordre. En fait, nous verrons que ces calculs peuvent être représentés sous forme de diagrammes qui seront à la base de notre approche de l'électrodynamique quantique.

5.6.1 Sur la section efficace

En électrodynamique quantique, nous savons que l'interaction entre particules est modélisée par un échange de bosons virtuels appelés bosons de jauge ou photons. Nous aurons donc en quelque sorte à étudier des processus de diffusion. Il est donc naturel de commencer à étudier plus en détail la section efficace différentielle.

Normalisation des états

Pour discuter de la section efficace, nous devons d'abord préciser la forme des solutions libres des équations d'évolution. Dans le cas non relativiste, la normalisation est choisie pour que l'intégrale sur le volume de la densité d'états soit égale à l'unité

$$\int_V \psi^*(\mathbf{r}, t) \psi(\mathbf{r}, t) d\mathbf{r} = \int_V \rho(\mathbf{r}, t) d\mathbf{r} = 1 \tag{5.106}$$

Dans le domaine relativiste tel n'est pas le cas. En général, on choisit la normalisation telle que l'intégrale sur le volume soit égale à $2E$. Cette normalisation est appelée normalisation covariante.

$$\int_V \rho(\mathbf{r}, t) d\mathbf{r} = 2E \tag{5.107}$$

Dans le cas particulier d'une onde plane, la normalisation \mathcal{N} vaut alors

$$\mathcal{N} = \frac{1}{\sqrt{V}} \tag{5.108}$$

Expression générale de la section

Considérons le processus général de diffusion suivant dont on cherche à exprimer la section efficace.

$$a + b \rightarrow 1 + \dots + n \tag{5.109}$$

Tout d'abord, on peut montrer que le taux de transition Γ_{if} par unité de volume entre un état i et un état f est donné par l'expression

$$\Gamma_{if} = \frac{(2\pi)^4 \delta^4(p_f - p_i)}{V^{n+2}} |T_{fi}|^2 \tag{5.110}$$

où T_{fi} représente la matrice de transition entre l'état final f et l'état initial i . Le sens de la moyenne que nous effectuons sera donné plus tard et est une simplification demandée pour les expériences.

Cependant, pour pouvoir avoir une quantité qui puisse être comparée quelle que soit l'expérience, nous devons supprimer la dépendance en ϕ , flux incident, et en N_s , nombre de particules cibles. Il est utile de se placer dans le référentiel où la particule b est au repos. Exprimons le flux incident. Par définition, c'est le nombre de particules du faisceau par unité d'aire et de temps. Rappelons qu'à cause de la normalisation que nous avons choisie, la densité de particules incidentes est égale à $\frac{2E_a}{V}$. Le nombre de particules traversant la surface A orthogonale au faisceau pendant le temps t est alors donné par $\frac{2E_a}{V} \cdot vtA$ où v est le module de la vitesse relative. Le flux vaut donc

$$\phi = \frac{2E_a}{V} \cdot \|\mathbf{v}_a - \mathbf{v}_b\| \tag{5.111}$$



Enfin, le nombre de particules cibles par unité de volume vaut $\frac{2E_b}{V}$. Au final, nous devons donc diviser par la quantité $\frac{2E_b 2E_a}{V^2} \|\mathbf{v}_a - \mathbf{v}_b\|$. Lors d'un processus physique de diffusion, nous ne savons pas quels sont les états finals. Nous devons donc sommer sur toutes les possibilités. L'élément différentiel dans l'espace des phases est, compte tenu de la normalisation que nous avons choisi,

$$\prod_{k=1}^n \frac{V}{(2\pi)^3 2E_k} d\mathbf{p}_k \tag{5.112}$$

Maintenant, si l'on met tout bout à bout, nous avons l'expression de la section efficace différentielle $d\sigma$

$$d\sigma = \frac{1}{2E_b 2E_a \|\mathbf{v}_a - \mathbf{v}_b\|} \int (2\pi)^4 \delta^4(f - p_i) \overline{|T_{fi}|^2} \prod_{k=1}^n \frac{d\mathbf{p}_k}{(2\pi)^3 2E_k} \tag{5.113}$$

La présence de la distribution de Dirac assure la conservation de l'impulsion. Le problème est maintenant de déterminer les coefficients de la matrice de transition.

5.6.2 Interaction pour un boson de spin 0

Amplitude de transition au premier ordre

On s'intéresse en premier lieu à la diffusion de bosons de spin 0. Nous utiliserons donc l'équation de Klein-Gordon. Nous nous arrêterons toujours dans ce qui suit à des calculs au premier ordre. Nous supposons, mais cela peut se justifier, que l'amplitude de probabilité de transition s'écrit de la même manière qu'en mécanique quantique non relativiste, à savoir

$$T_{fi} = -i \langle \psi_f | V | \psi_i \rangle = -i \int \psi_f^*(x^\mu) V(x^\mu) \psi_i(x^\mu) d^4x \tag{5.114}$$

Maintenant, quel est donc le potentiel V ? Pour le déterminer, il faut retourner au lagrangien de Klein-Gordon invariant de jauge 5.45. En écrivant les équations d'Euler Lagrange, nous arrivons à l'équation d'évolution d'un boson de spin 0 en présence d'un champ

$$(\square + m^2) \psi(\mathbf{r}, t) = -V(\mathbf{r}, t) \psi(\mathbf{r}, t) \quad V(\mathbf{r}, t) = ie(\partial_\mu A^\mu + A^\mu \partial_\mu) - e^2 A_\mu A^\mu \tag{5.115}$$

Si l'on s'arrête au premier ordre dans le développement perturbatif, il suffit de ne pas tenir compte du second terme dans le potentiel. Le potentiel électromagnétique pour un boson de spin 0 est donc

$$V(\mathbf{r}, t) = ie(\partial_\mu A^\mu + A^\mu \partial_\mu) \tag{5.116}$$

Maintenant, si l'on injecte ce potentiel dans T_{fi} et que l'on effectue une intégration par parties, nous arrivons à la forme commode pour l'amplitude de transition

$$T_{fi} = -i \int A^\mu \underbrace{(-ie) [\psi_f^* \partial_\mu \psi_i - \psi_i \partial_\mu \psi_f^*]}_{j_\mu^{fi}} d^4x \tag{5.117}$$

j_μ^{fi} peut être vu comme un courant de transition électromagnétique analogue au simple courant de probabilité. Si on considère que les états initial et final décrivent une particule libre, l'état s'écrit comme $\psi_{i,f} = \mathcal{N}_{i,f} e^{-ip_{i,f}^\mu x_\mu}$. Ceci donne donc

$$T_{fi} = -i \int A^\mu \underbrace{(-e) \mathcal{N}_i \mathcal{N}_f (p_i + p_f)_\nu}_{j_\mu^{fi}} e^{-i(p_i - p_f)^\mu x_\mu} d^4x \tag{5.118}$$

L'inconnue qu'il nous reste est le quadripotentiel A^μ que nous allons maintenant étudier.



Le quadripotentiel

Spécifions un peu la diffusion que l'on veut étudier, soit

$$a + 1 \rightarrow b + 2 \tag{5.119}$$

Du point de vue de la particule passant de l'état initial a à l'état final b, le champ A^μ qu'elle perçoit n'est plus un champ libre mais un champ créé par le courant j_μ^{12} . Nous devons donc résoudre les équations de Maxwell en présence d'une source. En jauge de Lorentz, le quadripotentiel est solution de l'équation

$$\partial^\nu \partial_\nu A^\mu(x) = j_\mu^{12}(x) \tag{5.120}$$

On rappelle qu'une telle équation peut se résoudre facilement à l'aide des fonctions de Green. Celle-ci est ici simplement égale à $-\frac{1}{q^2}$. La solution générale s'écrit comme la convolution de la fonction de Green par la source. Dans le cas particulier qui nous intéresse ici, où l'expression du courant $j_\mu^{12}(x)$ est une onde plane, $j_\mu^{12}(x) = -e\mathcal{N}_i\mathcal{N}_f(p_i + p_f)_\mu e^{-i(p_i - p_f)^\mu x}$, nous avons une solution particulière simple

$$A^\mu(x) = -\frac{g^{\mu\nu}}{q^2} j_\nu^{12}(x) \tag{5.121}$$

où $q = p_i - p_f$ est le moment relatif. Le tenseur métrique est introduit pour plus de commodité dans la suite. Au final, nous avons l'expression de l'élément de matrice de transition comme

$$T_{fi} = -i \int j_\mu^{ab}(x) \frac{-g^{\mu\nu}}{q^2} j_\nu^{12} d^4x \tag{5.122}$$

On voit que cette expression est symétrique par rapport aux courants. Le fait que l'on aie choisi le point de vue de la particule "ab" et non l'autre est sans importance. Nous pouvons injecter les expressions des courants et effectuer l'intégration. Cela donne au final

$$T_{fi} = -i(-e)^2 \mathcal{N}_a\mathcal{N}_b\mathcal{N}_1\mathcal{N}_2 (p_a + p_b)_\mu \frac{-g^{\mu\nu}}{q^2} (p_1 + p_2)_\nu \cdot (2\pi)^4 \delta^4(p_f - p_i - k) \tag{5.123}$$

Ceci nous fait apparaître la distribution de Dirac qui assure la conservation de l'énergie impulsion. En général ce terme est implicite dans les calculs puisqu'il suffit juste de se rappeler qu'il assure la loi de conservation. On le fait apparaître en fait directement dans la section efficace (terme $(2\pi)^4 \delta^4(p_f - p_i)$)⁹. L'amplitude que l'on obtient en oubliant ces termes ainsi que toutes les constantes de normalisation est parfois appelée *amplitude de Feynman*. C'est celle que l'on trouvera dans la suite avec les diagrammes.

Un point important à retenir est l'apparition du propagateur du photon $\frac{-g^{\mu\nu}}{q^2}$. C'est ce propagateur qui nous permet de dire que l'interaction se fait via l'échange d'un photon ! Enfin, il faut retenir qu'un terme en $-ie(p_a + p_b)_\mu$ apparaît lors d'une émission ou d'une absorption. Cette amplitude peut être représentée par un diagramme comme sur la figure 5.5. Les détails de leur construction seront donnés plus loin.

5.6.3 Interaction pour un fermion de spin 1/2

Amplitude de transition au premier ordre

Qu'en est-il des fermions ? On se contente toujours de faire des calculs au premier ordre. En utilisant le lagrangien de Dirac 5.104, nous pouvons dériver le potentiel électromagnétique pour une particule de charge $-e$. Une différence majeure est que ce potentiel apparaît comme une matrice de dimension quatre.

$$V(\mathbf{r}, t) = -e \begin{pmatrix} A^0 & -\boldsymbol{\sigma} \cdot \mathbf{A} \\ -\boldsymbol{\sigma} \cdot \mathbf{A} & A^0 \end{pmatrix} \tag{5.124}$$

L'amplitude de probabilité s'écrit en fonction des spineurs de Dirac comme

⁹. Le problème de l'apparition de la distribution de Dirac vient du fait que nous avons décrit nos particules par des ondes planes. De manière plus correcte, il aurait fallu les décrire par des paquets d'ondes et la formule de la section efficace serait tombée plus rigoureusement.



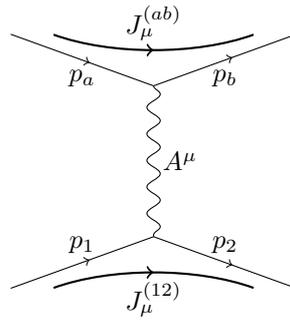


FIGURE 5.5 – Diagramme de Feynman générique représentant l'interaction électromagnétique entre deux particules. Elle représente de manière graphique l'amplitude de transition en oubliant quelques termes...

$$T_{fi} = -i \int \bar{\psi}_f^\dagger(k'^\mu, s') V(x^\mu) \psi_i(k^\mu, s) d^4x \quad (5.125)$$

où s et s' désignent le spin des particules, dont la présence est obligatoire puisqu'une interaction peut les changer. Une attention toute particulière a été observée pour bien obtenir un scalaire!¹⁰. Cette amplitude peut être écrite sous une forme covariante en utilisant les matrices γ^μ ainsi que l'adjoint d'un spineur

$$T_{fi} = -i \int \bar{\psi}_f(k'^\mu, s') (-e\gamma^\mu A_\mu) \psi_i(k^\mu, s) d^4x \quad (5.126)$$

De manière analogue au cas bosonique, nous pouvons définir un courant de transition électromagnétique j^μ pour les fermions analogue au courant de probabilité issu de l'équation de Dirac. L'amplitude de transition s'écrit alors sous la forme compacte et similaire à celle des bosons.

$$T_{fi} = -i \int j_{fi}^\mu A_\mu d^4x \quad j_{fi}^\mu = -e\bar{\psi}_f \gamma^\mu \psi_i \quad (5.127)$$

Normalisation des spineurs de Dirac

Nous avons choisi précédemment une normalisation particulière de la densité de probabilité sur le volume égale à deux fois l'énergie. Une onde plane de Dirac d'énergie positive s'écrit de manière générale comme

$$\psi = \mathcal{N}\omega e^{-ip^\mu x_\mu} = \mathcal{N} \begin{pmatrix} \phi \\ \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m} \phi \end{pmatrix} e^{-ip^\mu x_\mu} \quad (5.128)$$

Le calcul explicite de la densité de probabilité donne pour la normalisation

$$\mathcal{N}^2 V \left(\frac{(\boldsymbol{\sigma} \cdot \mathbf{p})(\boldsymbol{\sigma} \cdot \mathbf{p})^*}{E+m} + 1 \right) = 2E \Leftrightarrow \mathcal{N}^2 V \left(\frac{\mathbf{p}^2}{(E+m)^2} + 1 \right) = 2E \Leftrightarrow \mathcal{N} = \sqrt{\frac{E+m}{V}} \quad (5.129)$$

Il est alors commun de définir les spineurs $u(\mathbf{p}, s)$ pour les énergies positives (électrons) et $v(\mathbf{p}, s)$ pour celles négatives (positrons) incorporant le facteur $\sqrt{E+m}$ pour que l'onde plane soit analogue au cas Klein-Gordon. Soit donc

$$u(\mathbf{p}, s) = \sqrt{E+m} \begin{pmatrix} \phi^s \\ \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m} \phi^s \end{pmatrix} \quad v(\mathbf{p}, s) = \sqrt{E+m} \begin{pmatrix} \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{E+m} \chi^s \\ \chi^s \end{pmatrix} \quad (5.130)$$

Comme pour tous les spineurs, nous pouvons définir les spineurs adjoints à $u(\mathbf{p}, s)$ et $v(\mathbf{p}, s)$.

¹⁰. La transposée d'un spineur, qui peut être vu comme un vecteur colonne, est un vecteur ligne.



Retour sur l'amplitude de probabilité

Comme dans le cas bosonique, on calcule ce courant à l'aide des ondes planes définies ci-dessus. De plus, le calcul du potentiel A^μ est analogue à celui déjà établi. Au final, l'amplitude de transition a exactement la même forme que ce que nous avons précédemment, à savoir

$$T_{fi} = -i \int j_\mu^{ab}(x) \frac{-g^{\mu\nu}}{q^2} j_\nu^{12} d^4x \quad j_\mu^{fi} = -e (\bar{\psi}_f \gamma_\mu \psi_i) \tag{5.131}$$

Pour donner un exemple, dans le cas particulier où nous avons toujours un fermion entre l'état initial et final, le courant s'écrit comme

$$j_\mu^{fi} = -e (\bar{u}_f \gamma_\mu u_i) e^{i(p_f - p_i)^\nu x_\nu} \tag{5.132}$$

Cette expression peut être comparée à celle que nous avons trouvée dans le cas des bosons dans la section 5.6.2. Un point délicat concerne les spineurs que l'on doit utiliser. Il faut bien faire attention au type de particule qui rentre en jeu dans le processus. Ceci va être détaillé dans la suite.

5.6.4 Diagramme de Feynman

Les diagrammes de Feynman sont des moyens efficaces pour représenter les termes du développement perturbatif de l'amplitude de transition. La construction de tels diagrammes repose sur les calculs qui ont été faits précédemment, et principalement sur la forme des courants électromagnétiques. Les *règles de Feynman* nous permettent de donner directement, à partir d'un diagramme, l'amplitude de transition déjà intégrée (comme dit plus haut, le facteur $(2\pi)^4 \delta^4(p_f - p_i)$ est omis tout comme les facteurs de normalisation).

Les lignes externes correspondent à la propagation libre des différentes particules que l'on considère. Les fermions et les bosons sont décrits par des objets différents. Les fermions sont décrits par les spineurs $u(\mathbf{p}, s)$, les anti-fermions par $v(\mathbf{p}, s)$. D'après la forme du courant électromagnétique pour les fermions, on voit que, suivant si la particule est entrante ou sortante, il faut considérer soit le spineur soit son adjoint. La figure 5.6 détaille tous les cas.

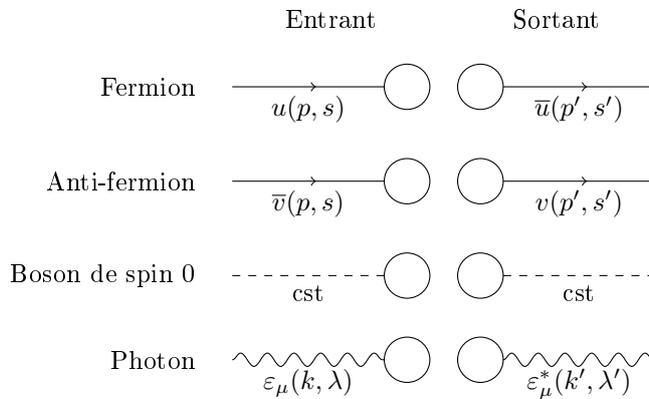


FIGURE 5.6 – Règles de Feynman concernant les lignes externes d'un diagramme. Ces règles proviennent de la forme des différents courants électromagnétiques que nous avons défini.

Nous savons que les interactions fondamentales sont médiées par des bosons virtuels. Pour l'interaction électromagnétique, on rappelle que le boson de jauge est un photon. Dans le calcul des amplitudes de transition, ils apparaissent *via* la présence des propagateurs des équations d'onde. Nous avons déjà fait le calcul du propagateur de l'équation de Klein-Gordon dans 5.2.4. Il est facile de faire de même avec l'équation de Dirac. La figure 5.7 donne les propagateurs de l'équation de Klein-Gordon et de Dirac ainsi que celui associé au photon, le plus utilisé en électrodynamique. Nous avons utilisé la notation de Feynman \not{p} qui ne signifie rien d'autre que $\gamma^\mu p_\mu$. Les dernières règles concernent l'absorption/émission du boson virtuel. On ne s'intéresse qu'au cas de l'électromagnétisme. L'idée est simple et provient toujours de la forme des courants électromagnétiques. À chaque nœud d'un vertex, on place $ie\gamma^\mu$ si l'on a émission/absorption d'un photon par un fermion ou on place $-ie(p + p')^\mu$ si nous avons un boson. La charge e est considérée comme positive; le signe est donc à changer suivant la charge réelle de la particule. Par contre, il n'est pas utile de se soucier du signe dans le cas d'une création/annihilation, cela ne change rien dans l'amplitude que l'on écrira au final.



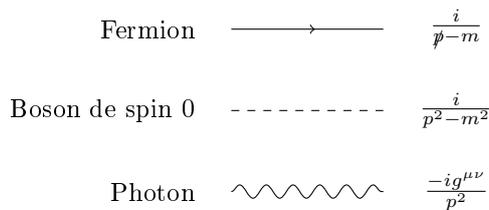


FIGURE 5.7 – Propagateurs pour les fermions de spin 1/2, les bosons de spin 0 et les photons. La notation de Feynman est utilisée, $\not{p} = \gamma^\mu p_\mu$.

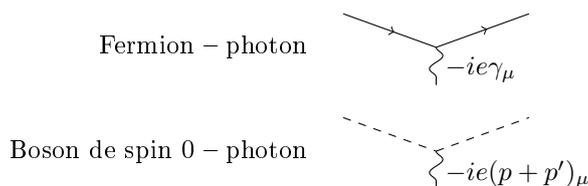


FIGURE 5.8 – Vertex associé à l’émission/absorption d’un photon par un boson ou un fermion. Le signe de la charge est important. Ici, $e > 0$.

5.6.5 Diffusion électron-positron

Sur les variables de Mandelstam

Souvent, dans les calculs que nous allons faire, il est utile d’utiliser un jeu de variables dites variables de Mandelstam, qui permettent d’exprimer simplement les expressions. La figure 5.9 montre les diagrammes de base servant à leur définition. La variable s , égale à la somme des impulsions au carré est associée à l’impulsion totale du système qui est conservée. Nous avons donc $s = (k + p)^2 = (k' + p')^2$. Ce diagramme est typique d’une annihilation/création. Les variables t et u sont quant à elles associées à l’impulsion transférée lors de la diffusion. C’est pourquoi nous avons $t = (k - k')^2$. Bien que les deux diagrammes associés puissent paraître identiques, il correspondent aux deux situations que l’on peut rencontrer lorsqu’on discute d’une diffusion de deux particules identiques! Lors d’une telle diffusion, il nous est impossible de savoir quel est le chemin exact suivi par les particules initiales, c’est pourquoi nous devons considérer deux diagrammes où les impulsions finales sont inversées.

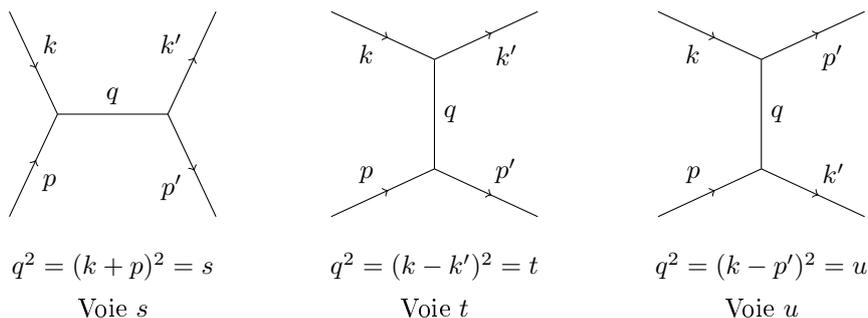


FIGURE 5.9 – Diagrammes associés aux variables de Mandelstam. La variable s est associée à l’impulsion totale du système qui est le cas typique d’une annihilation et d’une création. Les variables t et u sont associées à l’impulsion transférée lors d’une diffusion. Le dernier diagramme est nécessaire si l’on considère une diffusion de deux particules identiques.

Étude de la diffusion

Nous allons appliquer les règles de Feynman pour étudier le processus de diffusion entre un électron et un positron. La première chose à faire est de tracer tous les diagrammes de Feynman possibles pour une telle diffusion. On se limitera à des faibles énergies. En d’autres termes, les effets de l’interaction faible sont négligés

et seule l'interaction électromagnétique est présente. Seuls deux diagrammes sont possibles et sont représentés sur la figure 5.10.

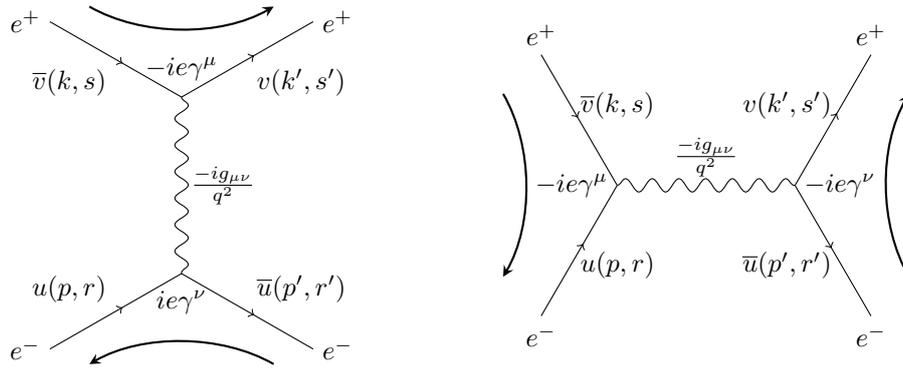


FIGURE 5.10 – Diagrammes de Feynman associés à la diffusion d'un électron et d'un positron. Deux situations se présentent. L'une, le diagramme de gauche, correspond à une diffusion classique, tandis que l'autre correspond à l'annihilation de la paire puis à sa création. À basse énergie, seuls ces processus rentrent en jeu.

Analysons en détail l'un des diagrammes, celui de gauche par exemple, qui représente une diffusion au sens classique. D'abord regardons ce qui se passe au début du processus. D'après les règles, le positron, antiparticule de l'électron, doit être représenté par le spineur $\bar{v}(k, s)$ tandis que l'électron doit être représenté par $u(p, r)$ (r et s sont les variables de spin). Ensuite vient l'interaction, l'échange d'un photon virtuel. À ce photon est associé un propagateur $\frac{-ig_{\mu\nu}}{q^2}$, où q est l'impulsion transférée. De plus, puisque nous avons deux fermions, $-ie\gamma^\mu$ représente l'émission/absorption du photon par la particule. Remarquons le changement de signe pour l'électron, particule de charge négative. Ces contributions sont placées aux nœuds des vertex. Finalement, l'état final de l'électron est donné par le spineur $\bar{u}(p', r')$ tandis que celui du positron par $v(k', s')$. Ici, rien de plus que l'application stricte des règles de Feynman n'a été fait ! Avec cela, nous pouvons donner directement l'amplitude de transition associé à ce diagramme.

Le sens des flèches sur les lignes n'a pas d'importance. Nous avons choisi de les prendre toutes dans la direction d'évolution temporelle. Les grandes flèches sur les diagrammes ne sont là qu'à titre indicatif. Elles sont là pour rappeler dans quel "sens" va le courant électromagnétique. La chose importante à se rappeler est qu'au final nous devons toujours avoir un scalaire pour chaque courant que l'on considère. En d'autres termes, un spineur adjoint est toujours à gauche d'une matrice γ^μ tandis qu'un spineur est toujours à sa droite. Finalement, l'amplitude de transition pour le diagramme de diffusion classique T_d est la suivante

$$T_d = \underbrace{\bar{v}(k, s) (-ie\gamma^\mu) v(k', s')}_{\text{Courant du positron}} \underbrace{\frac{-ig_{\mu\nu}}{q^2}}_{\text{Propagateur}} \underbrace{\bar{u}(p', r') (ie\gamma^\nu) u(p, r)}_{\text{Courant de l'électron}} \tag{5.133}$$

$$= -\frac{ie^2}{s} \bar{v}(k, s) \gamma^\mu v(k', s') \bar{u}(p', r') \gamma_\mu u(p, r) \tag{5.134}$$

où nous avons utilisé la variable s de Mandelstam et contracté le tenseur métrique. De la même manière, nous pouvons donner l'amplitude de transition T_a associé au processus d'annihilation et de création

$$T_a = \underbrace{\bar{v}(k, s) (-ie\gamma^\mu) u(p, r)}_{\text{Courant d'annihilation}} \underbrace{\frac{-ig_{\mu\nu}}{q^2}}_{\text{Propagateur}} \underbrace{\bar{u}(p', r') (-ie\gamma^\nu) v(k', s')}_{\text{Courant de création}} \tag{5.135}$$

$$= \frac{ie^2}{s} \bar{v}(k, s) \gamma^\mu u(p, r) \bar{u}(p', r') \gamma_\nu v(k', s') \tag{5.136}$$

On remarque bien que chaque matrice γ^μ est entourée par un spineur adjoint à sa gauche et un spineur à sa droite, ce qui donne bien au final un scalaire. L'amplitude totale de diffusion T est la somme de ces deux contributions, puisque nous ne pouvons pas savoir laquelle a réellement lieu,

$$T = T_a - T_d \tag{5.137}$$

Pourquoi un signe négatif? Ceci vient en fait de la nature fermionique du système; pour des fermions, on sait que nous devons avoir une antisymétrie de la fonction d'onde (principe de Pauli) tandis que pour des bosons, la fonction d'onde doit être symétrique. Ces propriétés se traduisent de manière simple lorsqu'on discute d'amplitude de probabilité : pour un système bosonique, nous devons faire la somme des amplitudes tandis que pour un système fermionique nous devons en faire la différence.¹¹

Expérimentalement, ce que l'on mesure correspond à la section efficace différentielle où apparaît le module au carré de l'amplitude T . Ainsi, on voit bien qu'il va y avoir trois contributions, l'une correspondant à la diffusion classique $|T_d|^2$, une autre au processus de création/annihilation $|T_a|^2$ et la dernière à l'interférence entre les deux $T_a^*T_d + T_aT_d^*$.

En revanche, ce n'est pas réellement l'amplitude de diffusion telle quelle que l'on doit considérer d'un point de vue expérimental. En effet, en général, les états de spin ne sont pas pris en compte, on ne s'y intéresse pas. C'est pourquoi la quantité qui est intéressante expérimentalement est la moyenne sur tous les spins initiaux et finals possibles des modules des amplitudes. On fait bien la moyenne sur les modules des amplitudes et non sur les amplitudes elles-mêmes car, en principe, il nous est possible de déterminer les états de spin. Dans ce cas, les postulats de la mécanique quantique nous obligent à sommer les modules et non les amplitudes, contrairement à ce que nous avons fait ci-dessus pour déterminer l'amplitude totale.

Ainsi, si l'on note $d\sigma_{s'r'sr}$ la section efficace associé à un processus ayant les spins s' , r' , s et r , la section efficace qui nous intéresse est celle où l'on moyenne sur toutes les possibilités de ces différents spins. Compte tenu de l'expression de la section efficace, cela revient à faire la moyenne directement sur le module de l'amplitude, ce que l'on note $\overline{|T|^2}$.

Cette complexité supplémentaire simplifie en fait grandement les calculs.

$$d\sigma = \frac{1}{2s+1} \frac{1}{2r+1} \sum_{s'r'sr} d\sigma_{s'r'sr} = \frac{1}{4} \sum_{s'r'sr} d\sigma_{s'r'sr} \quad (5.138)$$

Technique des traces

Cette moyenne sur les spins que l'on s'impose permet de simplifier les expressions des modules des amplitudes grâce aux propriétés des matrices de Dirac. On ne va s'intéresser qu'à la contribution de la diffusion classique et calculer $|T_d|^2$. Calculons d'abord le module au carré.

$$\begin{aligned} |T_d|^2 &= T_d^* T_d & (5.139) \\ &= \frac{e^4}{s^2} [\bar{v}(k, s) \gamma^\mu v(k', s') \bar{u}(p', r') \gamma_\mu u(p, r)]^* \bar{v}(k, s) \gamma^\nu v(k', s') \bar{u}(p', r') \gamma_\nu u(p, r) \\ &= \frac{e^4}{s^2} [\bar{v}(k, s) \gamma^\mu v(k', s') \bar{u}(p', r') \gamma_\mu u(p, r)]^\dagger \bar{v}(k, s) \gamma^\nu v(k', s') \bar{u}(p', r') \gamma_\nu u(p, r) \\ &= \frac{e^4}{s^2} [\bar{u}(p', r') \gamma_\mu u(p, r)]^\dagger [\bar{v}(k, s) \gamma^\mu v(k', s')]^\dagger \bar{v}(k, s) \gamma^\nu v(k', s') \bar{u}(p', r') \gamma_\nu u(p, r) \\ &= \frac{e^4}{s^2} [\bar{u}(p', r') \gamma_\mu u(p, r)]^\dagger \bar{u}(p', r') \gamma_\nu u(p, r) [\bar{v}(k, s) \gamma^\mu v(k', s')]^\dagger \bar{v}(k, s) \gamma^\nu v(k', s') \end{aligned} \quad (5.140)$$

Il nous faut donc calculer le conjugué du courant. Pour cela, nous allons utiliser les propriétés des matrices de Dirac

$$\begin{aligned} [\bar{u}(p', r') \gamma_\mu u(p, r)]^\dagger &= [u(p, r)]^\dagger \gamma_\mu^\dagger [\bar{u}(p', r')]^\dagger \\ &= [u(p, r)]^\dagger \gamma_0 \gamma_\mu \gamma_0 [u^\dagger(p', r') \gamma_0]^\dagger \\ &= \underbrace{[u(p, r)]^\dagger \gamma_0}_{=\bar{u}(p, r)} \underbrace{\gamma_\mu \gamma_0 \gamma_0}_{=1} u(p', r') \\ &= \bar{u}(p, r) \gamma_\mu u(p', r') \end{aligned} \quad (5.141)$$

D'où le module de l'amplitude de transition

$$|T_d|^2 = \frac{e^4}{s^2} \bar{u}(p, r) \gamma_\mu u(p', r') \bar{u}(p', r') \gamma_\nu u(p, r) \bar{v}(k', s') \gamma^\mu v(k, s) \bar{v}(k, s) \gamma^\nu v(k', s') \quad (5.142)$$

¹¹. Ceci est particulièrement bien expliqué dans *La Bible*.



Maintenant, nous devons effectuer la moyenne sur tous les spins. Nous allons séparer les sommes en deux pour faire apparaître deux tenseurs, appelés tenseurs leptoniques, pour écrire l'amplitude moyennée de façon compacte.

$$\begin{aligned}
 \overline{|T_d|^2} &= \frac{e^4}{s^2} \frac{1}{4} \sum_{s'r'sr} \bar{u}(p, r) \gamma_\mu u(p', r') \bar{u}(p', r') \gamma_\nu u(p, r) \bar{v}(k', s') \gamma^\mu v(k, s) \bar{v}(k, s) \gamma^\nu v(k', s') \\
 &= \frac{e^4}{s^2} \left[\frac{1}{2} \sum_{r'r} \bar{u}(p, r) \gamma_\mu u(p', r') \bar{u}(p', r') \gamma_\nu u(p, r) \right] \left[\frac{1}{2} \sum_{s's} \bar{v}(k', s') \gamma^\mu v(k, s) \bar{v}(k, s) \gamma^\nu v(k', s') \right] \\
 &= \frac{e^4}{s^2} L_{\mu\nu}(e^-) L^{\mu\nu}(e^+)
 \end{aligned} \tag{5.143}$$

où nous avons défini les deux tenseurs leptoniques de manière évidente. Reste maintenant à calculer ces tenseurs. Puisqu'ils ont la même forme, on va se concentrer sur celui associé à l'électron. Pour pouvoir avancer, il va falloir écrire explicitement les produits matriciels. En effet, le but est de simplifier l'expression. Or vu la forme qu'elle a, il serait judicieux de faire apparaître la relation de fermeture des spineurs ¹².

$$\begin{aligned}
 L^{\mu\nu}(e^-) &= \frac{1}{2} \sum_{r'r} \bar{u}(p, r) \gamma^\mu u(p', r') \bar{u}(p', r') \gamma^\nu u(p, r) \\
 &= \frac{1}{2} \sum_{r'r} \sum_{ijkl} \bar{u}_i(p, r) \gamma_{ij}^\mu u_j(p', r') \bar{u}_k(p', r') \gamma_{kl}^\nu u_l(p, r) \\
 &= \frac{1}{2} \sum_{ijkl} \underbrace{u_l(p, r) \bar{u}_i(p, r)}_{(\not{p} + m_e)_{li}} \gamma_{ij}^\mu \underbrace{u_j(p', r') \bar{u}_k(p', r')}_{(\not{p}' + m_e)_{jk}} \gamma_{kl}^\nu \\
 &= \frac{1}{2} \sum_{ijkl} (\not{p} + m_e)_{li} \gamma_{ij}^\mu (\not{p}' + m_e)_{jk} \gamma_{kl}^\nu \\
 &= \frac{1}{2} \sum_l [(\not{p} + m_e) \gamma^\mu (\not{p}' + m_e) \gamma^\nu]_{ll} \\
 &= \frac{1}{2} \text{Tr} [(\not{p} + m_e) \gamma^\mu (\not{p}' + m_e) \gamma^\nu]
 \end{aligned} \tag{5.145}$$

Sous cette forme, il est facile de simplifier ce tenseur. En effet, grâce aux propriétés des matrices de Dirac, il existe plusieurs relations intéressantes sur la trace du produit de ces matrices. Cela simplifie considérablement les expressions. Pour l'instant, nous sommes arrivés à une expression simple de la moyenne sur les spins du module de l'amplitude de transition. Nous pouvons faire le même genre de calculs pour les autres contributions à la section efficace. En fait, cette méthode est générale et s'applique pour de nombreux diagrammes de Feynman. Lorsqu'il apparaît un autre type de particule, nous pouvons définir de la même manière un tenseur, par exemple un tenseur hadronique, pionique... On peut aussi ne pas passer par ces tenseurs. Dans ce problème de diffusion particulier, on peut tout aussi bien partir de la formule générale pour $\overline{|T_d|^2}$ et faire diverses combinaisons pour donner des relations de fermeture qui au final nous permettent d'écrire cette moyenne comme un produit de trace de tenseur différents de ceux que nous avons définis.

¹². Pour les spineurs $u(p, s)$ et $v(p, s)$, nous avons

$$\sum_s u(p, s) \bar{u}(p, s) = \not{p} + m \quad \sum_s v(p, s) \bar{v}(p, s) = \not{p} - m \tag{5.144}$$



Chapitre 6

Physique statistique

Sommaire

6.1	Thermodynamique	141
6.2	Fondements de la physique statistique	151
6.3	Physique statistique classique	158
6.4	Statistiques quantiques	160
6.5	Gaz réels et liquides	162
6.6	Théorie cinétique des gaz	166

Comprendre la physique à l'échelle macroscopique à partir de la dynamique microscopique donnée par les lois de la mécanique et les interactions fondamentales est une tâche quasi impossible de manière analytique et même numérique du fait du nombre astronomique de degrés de liberté en présence. Cependant, et de façon assez remarquable, il est tout de même possible de comprendre la plupart des phénomènes physiques macroscopiques à partir d'une description statistique en terme d'un jeu de variables très restreint.

À la genèse de la physique statistique se trouve la thermodynamique avec tous les concepts et résultats qu'elle a su apporter depuis sa construction au *XIX*^{ième} siècle. Le but premier de la physique statistique est d'éclaircir les concepts thermodynamiques en lui donnant une assise solide et elle permet d'aller plus loin dans les prédictions que l'on peut faire sur la dynamique des systèmes macroscopiques.

Le but de ce chapitre est donc d'introduire les bases de la physique statistique. Il commence par un gros rappel de thermodynamique,¹ pour ensuite aborder les fondements d'une description statistique à partir de la mécanique quantique et classique. On termine ensuite par différentes applications, pour les systèmes quantiques avec les statistiques d'un ensemble de particules indiscernables, une discussion des gaz réels et des liquides et enfin un petit aperçu de la théorie cinétique des gaz qui ouvre la voie vers la physique statistique hors-équilibre.

Les références pour ce chapitre sont [15], [4], [3] et [2]. Le Dieu de thermodynamique est un excellent livre de thermodynamique, très complet et extrêmement précis.

6.1 Thermodynamique

6.1.1 Introduction et définitions

La *thermodynamique* est une théorie phénoménologique des systèmes macroscopiques basée sur des concepts issus de l'expérience. Sa portée est très générale puisqu'elle peut être appliquée *a priori* à tous les systèmes physiques ayant un nombre de degrés de liberté suffisamment grand pour que l'on puisse négliger les fluctuations des grandeurs physiques.

Un système thermodynamique est génériquement n'importe quel système macroscopique. Son *état thermodynamique* est la donnée d'un ensemble de paramètres mesurables caractérisant le système. Leur origine peut être purement mécanique mais aussi, plus subtilement, d'origine statistique comme c'est le cas pour l'entropie. La thermodynamique que l'on expose dans la suite étudie principalement les états d'équilibre et les transitions entre eux. Un état d'équilibre est un état qui, à conditions extérieurs invariables, est lui aussi invariable. Ceci est à bien distinguer des états stationnaires qui sont par définition des états indépendants du temps mais où l'existence de flux entre le système et son environnement sont autorisés.

1. Il est honnêtement inutile d'essayer de faire de la physique statistique sans une connaissance solide de la thermodynamique.



Un système à l'équilibre thermodynamique peut être amené, *via* des transformations, à un autre état d'équilibre. La thermodynamique ne nous dit rien sur la manière dont le système atteint ce nouvel état d'équilibre mais, en spécifiant certaines caractéristiques de ces transformations, elle nous renseigne sur les grandeurs thermodynamiques à l'état final.

C'est ce dernier point qui fait toute la force et l'étrangeté de la thermodynamique. La description microscopique de la matière est complètement ignorée ici et on se focalise simplement sur les relations entre des grandeurs macroscopiques à l'équilibre. C'est d'ailleurs l'objet de la physique statistique que de comprendre l'origine microscopique des concepts thermodynamiques.

Il est très important en thermodynamique d'être très précis sur le système que l'on étudie ainsi que le référentiel dans lequel on l'étudie !

À un *système thermodynamique* donné, nous pouvons associer différentes grandeurs appelées *grandeurs d'état* caractérisant l'état du système. Parmi celles-ci, le volume V , le nombre de particules N ainsi que l'énergie U du système sont définies sans ambiguïtés. Ces grandeurs peuvent être définies quelque soit l'état du système et sont qualifiées *variables d'états primitives*. Cette qualification vient du fait qu'elles serviront de base à la formulation du postulat fondamental, et surtout qu'elles prennent des valeurs bien définies pour un système isolé. Un autre exemple d'une telle variable est la charge q pour un système électrostatique. De façon générique, on notera x l'ensemble des variables primitives servant à caractériser le système en plus de celles définies plus haut. On verra dans la suite que l'on peut définir de nouvelles grandeurs d'états dites conjuguées à partir de la fonction entropie.

Toute fonction des grandeurs d'états caractérisant complètement l'état d'équilibre du système est appelé *fonction d'état*. Cette notion n'est pas à confondre avec celle de potentiel thermodynamique. On discutera de cela dans la suite.

Pour rappel, on dit qu'un système est *isolé* lorsqu'il ne peut ni échanger de la matière ni de l'énergie avec son environnement. Il sera qualifié de *fermé* s'il peut seulement échanger de l'énergie mais pas de matière.

Toutes les grandeurs thermodynamiques ne sont pas toutes indépendantes les unes des autres. Elles sont reliées par des relations fonctionnelles appelées *équations d'état* et dépendent du modèle considéré. Par conséquent, le nombre de paramètres indépendants du système est réduit. Typiquement, nous avons

$$f(V, N, U) = 0 \quad (6.1)$$

Concernant les transformations thermodynamiques, on qualifie une transformation de *quasi-statique* si elle passe par une suite infinitésimale d'états d'équilibre. Elle sera qualifiée de *réversible* si elle est quasi-statique et si la transformation inverse nous permet de retourner à l'état d'équilibre initial. Par exemple, étirer un chewing-gum tout doucement est une transformation quasi-statique mais clairement pas réversible. On définira dans la suite d'autres types de transformations.

6.1.2 Le postulat fondamental et l'entropie

On va adapter un point de vue de la thermodynamique proche de la physique statistique dans le sens où nous allons la construire en postulant l'existence d'une unique fonction d'état contenant toute l'information sur le système qui n'est autre que l'entropie S . Mais avant d'énoncer ce postulat, nous allons préciser certaines idées.

Énergie interne

La notion d'énergie est d'origine mécanique et prend son fondement théorique dans l'étude de la symétrie par translation temporelle. D'ailleurs lorsque le système est invariant sous translation temporelle, l'énergie est une quantité conservée. C'est toujours le cas par exemple pour un système isolé. Précisons un peu les choses sur la notion d'énergie en thermodynamique.

On considère un système isolé constitué de N particules en interaction avec un potentiel invariant par translation à l'équilibre dans son référentiel barycentrique. Nous verrons dans la suite que seul un mouvement de rotation uniforme est possible pour un système à l'équilibre, en plus d'une éventuelle translation uniforme. L'énergie de ses constituants est appelée *énergie interne* et est de la forme

$$U = \sum_{i=1}^N \frac{1}{2} m_i \mathbf{v}_i^2 + V(\mathbf{r}_i - \mathbf{r}_j) \quad (6.2)$$



Systèmes faiblement couplés

Pour deux systèmes isolés, l'énergie totale est la somme des énergies des sous-systèmes. Ceci n'est plus *a priori* le cas lorsqu'ils sont en interaction du fait de l'énergie potentielle d'interaction. L'énergie n'est donc plus une grandeur extensive. Cependant, pour des systèmes faiblement couplés, cette même énergie d'interaction peut être négligée devant les énergies propres des sous-systèmes. L'énergie est alors de nouveau la somme des énergies des sous-systèmes : on parle d'extensivité approchée. La plupart des grandeurs qualifiées d'extensives ne le sont qu'en fait de manière approchée. On parlera du coup dans la suite d'extensivité en sous-entendant cette subtilité.

Énoncé du postulat

POSTULAT FONDAMENTAL : Il existe une fonction S appelée *entropie* des variables primitives U, V, N, x , contenant toute l'information sur le système, qui vérifie les propriétés suivantes :

- $S(U, V, N, x)$ est une fonction d'état ;
- S est extensive pour des systèmes faiblement couplés ;
- un état d'équilibre correspond à un maximum de S compatible avec les contraintes ;
- S est positive ($S \geq 0$), fonction croissante de l'énergie ($\frac{\partial S}{\partial U} \geq 0$) et s'annule et a une dérivée par rapport à l'énergie infinie au minimum d'énergie ($S(U_{\min}) = 0, \frac{\partial S}{\partial U}(U_{\min}) = \infty$).

6.1.3 Variables conjuguées, état d'équilibre

Variables conjuguées, réservoirs et identités thermodynamiques

Précédemment, nous avons introduit des variables primitives avec lesquelles le postulat fondamental a été défini. Il est alors possible de définir pour chaque variable primitive une *variable conjuguée* à partir de la fonction entropie.

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_{V,x} \quad \frac{P}{T} = \left(\frac{\partial S}{\partial V} \right)_{U,x} \quad \frac{X}{T} = \left(\frac{\partial S}{\partial x} \right)_{V,U} \quad (6.3)$$

La grandeur conjuguée à l'énergie interne correspond à la *température* T . D'après le postulat fondamental, la température est toujours positive. La grandeur P appelée *pression thermodynamique* permet de définir la grandeur conjuguée au volume. La mécanique donne une définition différente de la pression dite dans ce cas cinétique. On verra que ces deux définitions *a priori* différentes sont en fait identiques. La dernière définition est la définition générique de la variable X conjuguée à la variable primitive x .

Ces variables permettent de caractériser les échanges d'une grandeur thermodynamique, comme par exemple des échanges d'énergie, de volume, de particules... entre un système et un réservoir. Un réservoir est un système pouvant échanger avec le système avec lequel il est en contact une grandeur x sans que cela altère sa valeur de la variable conjuguée correspondante X . Par exemple, un thermostat est un réservoir d'énergie, il peut donc échanger de l'énergie avec le système sans pour autant que cela affecte sa température.

Les définitions ci-dessus permettent d'exprimer la différentielle de l'entropie en terme des variables internes et conjuguées.

$$dS = \frac{1}{T}dU + \frac{P}{T}dV + \frac{X}{T}dx \quad (6.4)$$

Énergie interne comme fonction d'état

Le postulat fondamental stipule que l'entropie est une fonction strictement croissante de l'énergie interne. Ainsi, il est possible d'inverser la relation entre ces deux grandeurs et de définir la fonction d'état énergie interne $U(S, V, x)$ qui sera (entre autres) fonction de l'entropie. La différentielle de cette fonction, parfois dénommée *première identité thermodynamique* s'écrit

$$dU = TdS - PdV - Xdx \quad (6.5)$$

Les termes du membre de droite ne peuvent être identifiés au travail et à la chaleur que dans le cas d'une transformation réversible. Il est clair que les dérivées partielles peuvent être reliées à la température et à la pression.



Équilibre thermique et mécanique

On va maintenant expliciter ce que nous dit le postulat sur l'état d'équilibre d'un système. On rappelle qu'à l'équilibre l'entropie est à son maximum compatible avec les contraintes imposées.

Soit un système isolé à l'équilibre thermique composé de deux parties 1 et 2 faiblement couplées. Ici, les deux sous-parties peuvent échanger de l'énergie et rien d'autre. L'énergie interne U du système étant extensive $U = U_1 + U_2$. Cette énergie est une constante, le système étant isolé. L'entropie, elle aussi extensive, vaut $S(U) = S_1(U_1) + S_2(U_2)$. La fonction d'état entropie $S(U_1, U_2, V_1, V_2, x)$ doit être maximale à l'équilibre. Les volumes et toutes les autres variables primitives x sont ici imposées. Pour déterminer l'état d'équilibre, on définit le potentiel thermodynamique $S(U_1; V_1, V_2, x)$ ² de la seule variable d'énergie U_1 à partir de la fonction d'état entropie et le maximum de ce potentiel nous déterminera l'état d'équilibre. La distinction entre fonction thermodynamique et potentiel thermodynamique sera discutée plus loin. Au maximum, nous avons donc

$$\frac{dS}{dU_1} = \frac{dS_1}{dU_1} + \frac{dS_2}{dU_1} = \frac{dS_1}{dU_1} + \frac{dU_2}{dU_1} \frac{dS_2}{dU_2} = \frac{dS_1}{dU_1} - \frac{dS_2}{dU_2} = 0 \quad (6.6)$$

Comme $\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_{V,x}$ on conclut à l'égalité des températures des sous-systèmes à l'équilibre thermique.

Considérons maintenant le même système mais en plus à l'équilibre mécanique. Les énergies internes des sous-systèmes ainsi que leur volume ne sont plus imposés et deviennent des variables internes. Seuls sont donnés le volume total et l'énergie totale. On peut alors de la même manière que plus haut définir un nouveau potentiel thermodynamique (différent du précédent), toujours grâce à la fonction d'état entropie, et chercher son maximum compatible avec les contraintes précédemment citées. En étudiant la dérivée du potentiel par rapport au volume, la condition d'équilibre mécanique nous donne

$$\frac{P_1}{T_1} = \frac{P_2}{T_2} \quad (6.7)$$

Si de plus nous avons l'équilibre thermique, nous avons en plus l'égalité des températures. Des systèmes à la fois en équilibre thermique et mécanique auront donc même température et même pression.

Mouvement macroscopique

Analysons la possibilité de mouvements macroscopiques de parties d'un système isolé à l'équilibre. On se place dans un référentiel galiléen. L'entropie d'une partie (macroscopique) du système est notée S_i , son énergie totale E_i , son impulsion \mathbf{P}_i et sa masse m_i . L'entropie est une fonction de l'énergie interne, c'est-à-dire de $E_i - \frac{\mathbf{P}_i^2}{2m_i}$. L'entropie totale du système s'écrit par extensivité

$$S = \sum_i S_i \left(E_i - \frac{\mathbf{P}_i^2}{2m_i} \right) \quad (6.8)$$

Le système total étant isolé, l'énergie, l'impulsion et le moment cinétique sont des constantes du mouvement

$$\sum_i \mathbf{P}_i = \text{constante} \quad \sum_i \mathbf{r}_i \times \mathbf{P}_i = \text{constante} \quad (6.9)$$

La maximisation de l'entropie par rapport aux impulsions, requise pour un état d'équilibre, doit tenir compte de ces contraintes. Le théorème des extrema liés nous permet de faire cela en introduisant des multiplicateurs de Lagrange. Nous devons donc trouver la condition d'extremum sur le potentiel suivant

$$\sum_i S_i \left(E_i - \frac{\mathbf{P}_i^2}{2m_i} \right) + \lambda \cdot \mathbf{P}_i + \mu \cdot (\mathbf{r}_i \times \mathbf{P}_i) \quad (6.10)$$

Les énergies et les impulsions sont ici les variables internes. On omet l'écriture explicite de toutes les autres variables imposées x . La condition d'extrémalité sur les énergies internes permet de conclure à l'égalité des températures des sous-systèmes. Enfin, grâce à

². On ne fait pas une distinction notationnelle claire entre la fonction et le potentiel thermodynamiques ici. Notons aussi que le potentiel n'est fonction que de U_1 et non aussi de U_2 puisque $U_2 = U - U_1$.



$$\frac{\partial S_i}{\partial \mathbf{P}_i} \left(E_i - \frac{\mathbf{P}_i^2}{2m_i} \right) = -\frac{\mathbf{P}_i}{m_i} \frac{\partial S_i}{\partial E_i} = -\frac{\mathbf{P}_i}{m_i T} = -\frac{\mathbf{v}_i}{T} \quad (6.11)$$

nous arrivons à la condition d'extrémalité sur les impulsions

$$\mathbf{v}_i = \mathbf{v} + \mathbf{r}_i \times \boldsymbol{\Omega} \quad (6.12)$$

où \mathbf{v} et $\boldsymbol{\Omega}$ ne dépendent pas de l'état des sous-systèmes (et sont directement reliés aux multiplicateurs de Lagrange). Cette condition exprime le fait qu'à l'équilibre il ne peut exister de mouvements macroscopiques internes hormis une translation ou une rotation uniforme globale du système! Le système se comporte comme un solide.

Potentiel chimique

Jusqu'ici, nous avons pas explicité le rôle du nombre de particules. Or, celui-ci joue un rôle crucial dans de nombreux phénomènes physiques, comme par exemple lors de réactions chimiques, ou dans des systèmes quantiques.

Les fonctions et potentiels thermodynamiques sont alors des fonctions du nombre de particules N . On définit alors le *potentiel chimique* μ comme variable conjuguée à N (attention au signe)

$$-\frac{\mu}{T} = \left(\frac{\partial S}{\partial N} \right)_{U,V,x} \quad (6.13)$$

Si l'on prend par exemple l'énergie interne $U(S, V, N, x)$ comme fonction thermodynamique, nous avons

$$dU = TdS - PdV + \mu dN - Xdx \quad \mu = \left(\frac{\partial U}{\partial N} \right)_{S,V,x} \quad (6.14)$$

Le potentiel chimique s'interprète comme étant l'énergie nécessaire pour extraire une particule du système.

Si le système est composé de différents types de particules, un potentiel chimique est associé à chacun de ces types. Nous avons alors la généralisation immédiate sur l'exemple de l'énergie interne $U(S, V, N_i, x)$

$$dU = TdS - PdV + \sum_i \mu_i dN_i - Xdx \quad (6.15)$$

De part la définition des fonctions thermodynamiques comme transformée de Legendre sur les variables autres que N , prendre en compte la variation du nombre de particules revient à simplement rajouter un terme μdN aux différentielles.

Regardons qu'elles sont les conditions pour qu'un système isolé soit à l'équilibre. Séparons le système en deux parties 1 et 2 composé respectivement de N_1 et N_2 particules. L'entropie est une fonction du nombre de particules N telle que $S(N) = S_1(N_1) + S_2(N_2)$. Notons la contrainte $N = N_1 + N_2$. Une condition nécessaire à l'équilibre est que l'entropie soit maximale par rapport à N_1 . Du coup

$$\frac{\partial S}{\partial N_1} = \frac{\partial S_1}{\partial N_1} + \frac{\partial S_2}{\partial N_1} = \frac{\partial S_1}{\partial N_1} + \frac{\partial N_2}{\partial N_1} \frac{\partial S_2}{\partial N_2} = \frac{\partial S_1}{\partial N_1} - \frac{\partial S_2}{\partial N_2} = 0 \quad (6.16)$$

Avec $\frac{\partial S}{\partial N} = -\frac{\mu}{T}$, et

$$\frac{\mu_1}{T_1} = \frac{\mu_2}{T_2} \quad (6.17)$$

Comme à l'équilibre, les températures des sous-systèmes sont égales, une condition nécessaire à l'équilibre du systèmes est l'égalité des potentiels chimiques

$$\mu_1 = \mu_2 \quad (6.18)$$

Un tel résultat peut s'appliquer pour les mélanges à l'équilibre comme entre deux phases (voir le chapitre 8) ou encore pour un système dans un champ extérieur. Cette égalité se comprend naturellement si l'on se rappelle de son interprétation physique, à savoir que le potentiel chimique correspond à l'énergie d'extraction d'une particule du système. Pour deux systèmes, comme deux phases par exemple, qui sont à l'équilibre, aucun sens de transfert de particules ne doit être énergétiquement privilégié car sinon leur coexistence serait impossible.



6.1.4 Coefficients calorimétriques

La calorimétrie a pour but l'étude de la mesure des quantités de chaleur. Elle se focalise en particulier sur la détermination de la *chaleur spécifique à volume constant* ou *capacité thermique à volume constant* notée C_V qui caractérise l'évolution de l'énergie du système lors d'un changement de température. C'est un cas particulier de coefficient de réponse. Par définition,

$$C_V = \left(\frac{\partial U}{\partial T} \right)_V \quad (6.19)$$

Le calcul de cette quantité n'est pas toujours évident. Pour un solide par exemple, une description quantique de sa structure est nécessaire pour être en accord avec les faits expérimentaux (voir le chapitre 7 de matière condensée). Sa détermination expérimentale est en revanche assez simple et sera discutée dans la partie suivante après l'introduction du premier principe.

6.1.5 Les principes

Le postulat de la thermodynamique tel qu'énoncé permet en principe de déterminer toutes les propriétés d'équilibre du système. En revanche une telle formulation n'est pas toujours des plus transparentes du point de vue physique. Les principes de la thermodynamique sont une reformulation du postulat qui extrait de façon un peu plus explicite son contenu physique. C'est d'ailleurs *via* ces principes que la thermodynamique fut historiquement formulée.

Travail, chaleur et premier principe

Le premier principe se focalise sur l'énergie d'un système \mathcal{S} et son comportement entre différents états d'équilibres. Le schéma physique général que l'on considère est celui du système physique \mathcal{S} qui intéresse un opérateur \mathcal{O} en présence d'un environnement \mathcal{E} qui contient tous les autres degrés de libertés en interaction avec \mathcal{S} .

La plupart du temps, cette loi est écrite en explicitant différentes contributions à l'énergie totale du système. Le système est pris généralement au repos ce qui permet d'annuler les contributions cinétique et potentiel macroscopiques. L'énergie totale est alors assimilée à l'énergie interne du système U qui correspond à toutes les contributions énergétiques de ses constituants. Lors d'une transformation de ce système par un opérateur, l'énergie va évoluer et se transformer. En considérant le système global $\mathcal{S}\mathcal{O}\mathcal{E}$ qui est par conséquent isolé, l'énergie est conservée. Les forces déployées par l'opérateur sont caractérisées par leur *travail* W tandis que la variation d'énergie interne du système est caractérisée par ΔU . Le reste de l'énergie provenant de l'interaction avec l'environnement \mathcal{E} est appelée *chaleur* et est notée Q . Le premier principe s'énonce finalement comme

PREMIER PRINCIPE DE LA THERMODYNAMIQUE : L'énergie interne U est une fonction d'état du système qui vérifie la loi de conservation suivante

$$\Delta U = Q + W \quad (6.20)$$

où Q et W sont respectivement la chaleur et le travail.

Une formulation pour des transformations infinitésimales s'en déduit immédiatement. Suivant les situations physiques que l'on souhaite analyser et les modèles choisis, ce principe permet d'accéder à une information sur les variations d'énergie (de température...) du système, sur la dissipation...

Traisons un exemple simple qui illustre d'ailleurs un procédé utile pour la calorimétrie. Soit deux systèmes aux températures respectives T_1 et T_2 caractérisés par des capacités thermiques C_{V1} et C_{V2} (ce sont bien les capacités thermiques à volume constant qui sont pertinentes ici vu que le volume ne changera pas) qu'on supposera constantes dans la gamme de température explorée. On met en contact ces deux systèmes (non calorifugés sinon c'est sans intérêt). Quelle est la température finale T du système total à l'équilibre?

D'après le premier principe qui stipule entre autre la conservation de l'énergie, avec $dU = C_V dT$, on a

$$C_{V1}T + C_{V2}T = C_{V1}T_1 + C_{V2}T_2 \Leftrightarrow T = \frac{C_{V1}T_1 + C_{V2}T_2}{C_{V1} + C_{V2}} \quad (6.21)$$

La température est donc le barycentre des températures initiales pondérées par les capacités thermiques à volume constant des deux systèmes. Mettre en contact deux systèmes à des températures différentes permet donc bien de faire de la calorimétrie, c'est-à-dire faire des mesures de capacité thermique. Connaissant les températures initiales et la température finale ainsi qu'une des deux capacités, on peut en déduire la capacité inconnue. Bien sûr cela présuppose d'avoir une autre méthode de mesure pour déterminer au préalable une des capacités.



Le second principe : énoncés historiques

La seconde loi de la thermodynamique provient de la constatation expérimentale que certains processus, même s'ils respectent le premier principe, ne se produisent pas. Les deux énoncés historiques qui suivent reflètent ces faits expérimentaux.

ÉNONCÉ DE KELVIN : Il n'existe pas de transformation thermodynamique dont le *seul* effet est de convertir entièrement en travail une quantité de chaleur extraite d'une unique source de chaleur.

ÉNONCÉ DE CLAUSIUS : Il n'existe pas de transformation thermodynamique dont le *seul* effet est de transférer une quantité de chaleur d'une source froide à une source chaude.

Ces deux énoncés du second principe de la thermodynamique sont en fait équivalents. Supposons que l'énoncé de Kelvin soit faux : il existe une transformation convertissant entièrement en travail une quantité de chaleur extraite d'une source de chaleur à la température T_1 . Ce travail peut être transformé en chaleur et peut être fourni à une source à la température $T_2 > T_1$. Nous avons donc un processus dont le seul effet est de transférer une quantité de chaleur d'une source froide à une source chaude : l'énoncé de Clausius est donc faux. Il est possible de démontrer l'autre implication de la même manière par contraposée en définissant une bonne machine thermique.

Cycle réversible, cycle de Carnot

Un *cycle* est un ensemble de transformations où l'état initial et l'état final sont les mêmes. Le *cycle de Carnot* est un cycle réversible qui nous permet de concevoir en principe l'existence d'une machine réversible. Ce cycle est constitué de quatre transformations, une détente isotherme, une détente adiabatique, une compression isotherme et enfin une compression adiabatique FIG. Toutes ces transformations peuvent être considérées comme réversibles si elles sont faites assez lentement.

Ayant justifié la possible existence d'une machine réversible, nous allons établir un des résultats majeurs de la thermodynamique, à savoir qu'aucune machine ne peut fournir plus de travail à partir d'une quantité de chaleur donnée qu'une machine réversible. Supposons cela faux et considérons une machine réversible R ainsi qu'une machine quelconque M fonctionnant entre deux sources de chaleur à la température T_1 et $T_2 < T_1$ FIG. Supposons que ces deux machines peuvent extraire la même quantité de chaleur Q_1 à la source T_1 . La machine M fournit un travail W plus grand que la machine réversible R noté W_R et transfère une quantité de chaleur à la source froide T_2 . Comme $W > W_R$, on garde une partie du travail fourni par la machine M pour faire fonctionner la machine réversible à l'envers, la différence $W - W_R$ étant gardée comme source de travail effectif. La machine réversible R extrait alors une quantité de chaleur à la source froide et fournit Q_1 à la source chaude T_1 .

Faisons un bilan. Au final, la source froide revient à son état de départ et le seul effet de cette machine est de fournir un travail $W - W_R$ à partir d'une unique source, la source froide T_2 . Or ceci n'est pas possible d'après le second principe de la thermodynamique. Ainsi, il n'existe pas de machine pouvant fournir plus de travail à partir d'une quantité de chaleur donnée qu'une machine réversible.

En appliquant cet énoncé à deux machines réversibles, nous avons le résultat remarquable que le travail fourni pour une quantité de chaleur donnée est le même quelque soit la machine réversible de l'on considère ! Ce travail est une propriété universelle indépendante de la substance. Entre autres, si on trouve un résultat pour une machine réversible donnée, nous savons alors qu'il sera vrai pour toutes les autres.

L'entropie

Historiquement, le second principe de la thermodynamique tel que nous l'avons énoncé a permis de définir une nouvelle fonction d'état du système à l'équilibre appelée *entropie*. À partir du résultat de Carnot, il est possible de définir une échelle absolue de température, l'échelle *Kelvin*, puis de montrer

THÉORÈME DE CLAUSIUS : Quelque soit le cycle (quasi-statique) γ , nous avons

$$\int_{\gamma} \frac{\delta Q}{T} \leq 0 \quad (6.22)$$

et l'égalité n'est vérifiée que lors d'un cycle réversible.

Un corollaire de ce résultat est que l'intégrale de $\frac{\delta Q}{T}$ sur une transformation réversible ne dépend pas du chemin suivi. Cette intégrale nous permet de définir une nouvelle fonction d'état, l'entropie, à une constante additive près. De par sa définition, l'entropie est une fonction différentiable de différentielle



$$dS = \frac{dQ}{T} \quad (6.23)$$

Explicitons un peu plus les implications du théorème de Carnot sur l'entropie. Pour une transformation quelconque γ ,

$$\int_{\gamma} \frac{\delta Q}{T} \leq \Delta S \quad (6.24)$$

En particulier, pour un système thermiquement isolé, la variation d'entropie est positive. En d'autres termes, l'état d'équilibre d'un tel système est caractérisé par un maximum de son entropie.

Le cas d'égalité du théorème de Carnot se traduit simplement par la constance de l'entropie pour une transformation réversible. En clair, l'entropie caractérise la *réversibilité* d'une transformation

$$\Delta S = 0 \text{ Transformation réversible} \quad (6.25)$$

$$\Delta S \neq 0 \text{ Transformation irréversible} \quad (6.26)$$

Il est en fait tout à fait possible d'énoncer le second principe de la thermodynamique directement en postulant l'existence de l'entropie³ et de montrer qu'il est équivalent aux énoncés de Kelvin et de Clausius. Comme l'entropie est globalement utilisée pour exprimer concrètement le second principe, il a pu être ré-énoncé ainsi

SECOND PRINCIPE DE LA THERMODYNAMIQUE : Il existe une fonction d'état appelée *entropie* caractérisant l'équilibre du système et vérifiant pour un cycle γ (quasi-statique) quelconque

$$\int_{\gamma} \frac{\delta Q}{T} \leq \Delta S \quad (6.27)$$

où le cas d'égalité est atteint pour un cycle réversible.

La notion d'entropie exposée sous cette forme, dérivée à partir du second principe sous ses formes historiques, apparaît très abstraite. En fait, c'est un concept purement statistique et la mécanique statistique donne un éclairage nouveau sur l'information qu'elle véhicule (comme nous le verrons dans la section 6.2.2). Telle quelle, c'est une fonction d'état qui nous permet de déterminer l'état d'équilibre en cherchant ses maximums.

6.1.6 Fonctions et potentiels thermodynamiques

Suivant les systèmes que l'on étudie et l'information que l'on souhaite en extraire, il peut être plus utile de travailler avec d'autres grandeurs thermodynamiques que l'entropie ou l'énergie interne qui sont, en plus de cela, difficiles d'accès pour l'expérimentateur. L'énergie libre, l'enthalpie ou l'énergie libre de Gibbs sont les plus utilisées et dépendent chacune de variables différentes. Telles qu'elles sont construites, ces grandeurs contiennent exactement la même information que l'entropie. Elles déterminent donc complètement l'état d'équilibre du système et sont appelées *fonctions thermodynamiques*.

D'après le postulat fondamental, l'état d'équilibre est caractérisé par un maximum de l'entropie compatible avec les contraintes imposées au système. Différents exemples ci-dessus ont permis d'illustrer la méthode pour obtenir les valeurs des variables internes (non fixées par l'expérimentateur) en définissant un *potentiel thermodynamique* adapté à la situation étudiée et en regardant ses extremum. Mais, comme on l'a remarqué précédemment, l'entropie n'est pas toujours la fonction la plus facile à manipuler. Les fonctions thermodynamiques les plus adaptées au problème envisagé permettent à leur tour de définir des potentiels plus pertinents pour répondre aux questions posées.

Construction des fonctions thermodynamiques

La méthode qui permet de construire toutes les fonctions thermodynamiques est basée sur la transformation de Legendre. Au lieu de discuter mathématiquement cette transformation, dont on supposera les hypothèses d'applicabilités vérifiées pour l'entropie (et les fonctions qu'on en déduira), nous allons illustrer la méthode en partant de l'énergie interne.

3. C'est d'ailleurs ce que nous avons fait précédemment !



Par construction, l'énergie est une fonction naturelle de $U(y, x)$ où y est l'ensemble des variables primitives que l'on souhaite conserver (donc en général contrôlées par l'expérimentateur) et x sont les variables que l'on souhaite changer par des grandeurs que l'on contrôle mieux, typiquement par leur variable conjuguée X . Un exemple d'une telle situation que l'on décrira dans la suite est celle où l'énergie interne, fonction de $U(V, N, S)$ est remplacée par une fonction des variables (V, N, T) plus accessibles expérimentalement.

L'idée est donc de construire une nouvelle fonction d'état \mathcal{F} des variables (y, X) contenant la même quantité d'information que l'énergie interne (ou l'entropie ou tout autre fonction thermodynamique) définie comme

$$\mathcal{F}(y, X) = U(y, x(X)) - Xx \quad d\mathcal{F} = dU - x dX - X dx \quad (6.28)$$

Le point de vue adopté ici à partir de l'énergie interne est historique. Un point de vue plus proche de la physique statistique serait de partir directement de l'entropie elle-même et de procéder de la même façon. Ce sont bien sûr deux approches équivalentes.

Qu'est ce qu'un potentiel thermodynamique ?

En thermodynamique, l'état d'équilibre est déterminé par un maximum de l'entropie (second principe). L'idée des potentiels thermodynamiques est une généralisation de ce que l'on connaît en mécanique où l'état d'équilibre d'un système et sa stabilité peuvent être connus à l'aide d'un potentiel (pour des forces conservatives) et de ses extremum. À partir des contraintes expérimentales que l'on impose au système thermodynamique, l'équilibre est déterminé par un minimum d'un potentiel thermodynamique fonction des variables internes du système.

Énergie libre

Beaucoup de systèmes expérimentaux sont en contact avec un thermostat qui fixe leur température. La raison en est que la température est beaucoup plus facilement contrôlable que l'entropie ou l'énergie interne. Quelle est donc la fonction thermodynamique adaptée aux variables T, V, x ? D'après la méthode explicitée plus haut, la fonction thermodynamique que l'on cherche, appelée *énergie libre* $F(T, V, x)$ est définie par

$$F(T, V, x) = U(T, V, x) - TS(T, V, x) \quad dF = -S(T, V, x)dT - P(T, V, x)dV - X(T, V, x)dx \quad (6.29)$$

Nous avons bien mis en évidence toutes les dépendances pour bien souligner que pour connaître l'énergie libre il est nécessaire de connaître soit l'entropie et l'énergie interne comme fonctions de (T, V, x) soit l'entropie, la pression et toutes les variables conjuguées X comme fonctions de (T, V, x) .

On pourra retenir en particulier les formules

$$P = - \left. \frac{\partial F}{\partial V} \right|_{T, x} \quad S = - \left. \frac{\partial F}{\partial T} \right|_{V, x} \quad X = - \left. \frac{\partial F}{\partial x} \right|_{T, V} \quad (6.30)$$

Voyons maintenant l'aspect potentiel. On considère une transformation réversible *isotherme* à la température T_0 ⁴. Nous avons alors

$$dU = T_0 dS + \delta W \Leftrightarrow d(U - T_0 S) = \delta W \quad (6.31)$$

Notons pour l'instant $F_0(V, x; T_0) = U(V, x; T_0) - T_0 S(V, x; T_0)$. C'est une fonction des variables internes (V, x) , la température T_0 étant fixée par un thermostat.

Considérons maintenant de façon plus générale une transformation γ isotherme quelconque. Nous avons alors grâce au second principe ainsi qu'au premier principe

$$\int_{\gamma} \frac{\delta Q}{T_0} = \frac{Q}{T_0} \leq \Delta S \Rightarrow W \geq \Delta U - T_0 \Delta S = \Delta F_0 \quad (6.32)$$

L'égalité est atteinte lors d'une transformation réversible comme vu plus haut. Ainsi, le travail maximal que l'on peut extraire lors d'une transformation isotherme quelconque est au plus égal à la variation d'énergie libre. En particulier, pour un système isolé, $W = 0$. L'énergie libre ne peut que diminuer et est donc minimale

4. Une transformation où la température *extérieure* est constante est dite *monotherme*.



à l'équilibre. On en déduit que pour un système en contact avec un thermostat à la température T_0 , l'état d'équilibre est un état d'énergie libre $F_0(V, x; T_0)$ minimal. Autrement dit, $F_0(V, x; T_0)$ est un bon potentiel thermodynamique pour ce type de système.

L'énergie libre permet par exemple d'étudier les solides élastiques en nous permettant de déterminer le tenseur des contraintes à l'équilibre en fonctions des déformations, les transitions de phases, etc. . .

Enthalpie

Quelle est la fonction thermodynamique adaptée aux variables (S, P, x) ? Encore une fois on la définit par transformation de Legendre

$$H(S, P, x) = U(S, P, x) + PV(S, P, x) \quad (6.33)$$

Regardons l'aspect potentiel. Nous pouvons nous demander ce qui se passe pour une transformation à pression constante P_0 appelée *isobare*⁵ (comme la pression est toujours définie, cette transformation doit forcément être quasi-statique). Nous avons alors

$$dU = dQ - P_0 dV \Leftrightarrow d(U + P_0 V) = \delta Q \quad (6.34)$$

On définit le potentiel *enthalpie* H_0 comme $H_0(S, x; P_0) = U(S, x; P_0) + P_0 V(S, x; P_0)$. La quantité de chaleur échangée lors d'une transformation isobare est donnée par la variation d'enthalpie. Ses dérivées partielles sont reliées à la température et au volume.

Enthalpie libre

En chimie, les systèmes étudiés sont en général à température et pression fixées par l'expérimentateur. *L'enthalpie libre*, ou énergie libre de Gibbs, est la fonction thermodynamique adaptée à ces variables

$$G(T, P, x) = H(T, P, x) - TS(T, P, x) \quad (6.35)$$

Regardons encore une fois l'aspect potentiel. Considérons pour cela une transformation isotherme à la température T_0 et isobare à la pression P_0 . Le premier principe nous dit alors

$$dU = T_0 dS - P_0 dV \Leftrightarrow d(U + P_0 V - T_0 S) = 0 \quad (6.36)$$

On définit alors le potentiel associé à l'enthalpie libre $G_0(T, P, x) = U(T, P, x) + P_0 V(T, P, x) - T_0 S(T, P, x)$. Allons plus loin. Pour une transformation isotherme, nous avons l'inégalité 6.32 entre l'énergie libre F_0 et le travail. Or, pour une transformation isobare, $\delta W = -P_0 dV$, ce qui au final nous donne

$$\Delta G_0 \leq 0 \quad (6.37)$$

En d'autres termes, l'état d'équilibre d'un système ayant subi une transformation à température et pression constante est caractérisé par un minimum de l'enthalpie libre.

Grand potentiel

On peut aller encore plus loin et considérer que le système est en contact avec un réservoir de particules. Le potentiel chimique est alors une variable pertinente alors que dans les fonctions définies plus haut c'est le nombre N de particules qui était choisi comme variable (implicitement contenue dans x). Quel est donc le potentiel thermodynamique adapté aux variables (T, V, μ, x) ? On définit alors la fonction thermodynamique *grand potentiel* J (attention au nom) par

$$J(T, V, \mu, x) = U(T, V, \mu, x) - TS(T, V, \mu, x) + \mu N(T, V, \mu, x) \quad (6.38)$$

5. Une transformation où la pression *extérieure* est constante est dite *monobare*.



6.1.7 Le gaz parfait

Le modèle du *gaz parfait* est le modèle de base de la thermodynamique. Bien qu'il soit une idéalisation de la réalité, tous les gaz se comportent dans une certaine limite (gaz dilué) de manière similaire à un gaz parfait, ce qui justifie son étude.

Considérons un ensemble de N particules identiques de masse m supposées ponctuelles, contenues dans un récipient de volume V , sans interactions hormis les collisions binaires et sur la paroi. Un tel gaz est qualifié de parfait.

Les gaz parfaits suivent deux lois de comportement appelées *première et seconde loi de Joule* qui stipulent respectivement que l'énergie interne et l'enthalpie ne dépendent que de la température. Nous avons donc

$$dU = C_V dT \quad dH = C_P dT \quad (6.39)$$

où les capacités sont indépendantes de la température. L'expérience permettant de montrer qu'un gaz suit la première loi de Joule est l'expérience de Joule-Gay-Lussac.

Soit un gaz de N particules contenu dans un récipient calorifugé (transformation adiabatique) de volume V_1 , à la pression P_1 et à la température T_1 . On met en contact ce récipient avec un autre de volume V_2 totalement vide et on laisse le gaz se détendre. La conservation de l'énergie nous permet d'écrire $dU = \delta W + \delta Q = 0$ et donc que pour un gaz parfait $dT = 0$: la température est donc constante pour un gaz parfait lors de cette transformation.

Regardons la variation d'entropie. Grâce à l'identité thermodynamique $dU = TdS - PdV$, et à la loi du gaz parfait, nous avons

$$dS = \frac{Nk_B T}{TV} dV = Nk_B \frac{dV}{V} \Rightarrow \Delta S = nk_B \ln \left(\frac{V_1 + V_2}{V_2} \right) > 0 \quad (6.40)$$

La variation d'entropie est strictement positive, signe que la détente de Joule-Gay-Lussac est irréversible.

La définition donnée ici d'un gaz parfait est une définition purement thermodynamique et assez abstraite. Nous verrons plus loin une approche plus microscopique de ce modèle qui permettra de préciser son domaine de validité.

6.2 Fondements de la physique statistique

Dans la suite, nous allons construire une description statistique d'un système physique à partir de la mécanique quantique. De cette construction, il est possible de retrouver la description statistique classique faite entre autres par Boltzmann, Maxwell et Gibbs. Cette approche est plus fondamentale et permet par exemple de comprendre certains faits et ajouts *ad hoc* nécessaire à la cohérence de la description classique (indiscernabilité par exemple).

6.2.1 Description statistique

Un système quantique est décrit de la manière la plus générale par un opérateur densité. Cette description permet de tenir compte de notre connaissance partielle de l'état du système lors de sa préparation.

L'état d'un système parfaitement connu est donné par un *état pur* $|\psi\rangle$. Dans le contexte statistique, on parlera de *micro-état*. Au contraire, lorsque notre connaissance n'est que partielle, l'opérateur densité est la description adaptée et peut être vu comme un *mélange statistique* d'état. On parlera alors de *macro-état*.

Intuitivement, un macro-état est la donnée d'un nombre réduit de variables décrivant un système. On comprend bien que plusieurs micro-états peuvent conduire à un même macro-état, c'est-à-dire aux mêmes valeurs des quelques variables choisies. Dans l'approche de Gibbs, la collection des micro-états réalisant un macro-état donné est appelé *ensemble statistique*.

Dans la suite, nous étudierons trois de ces ensembles, celui décrivant un système isolé, celui décrivant un système en contact avec un thermostat et celui en plus en contact avec un réservoir de particules. Mais avant cela, il est nécessaire de disposer d'une manière de pouvoir quantifier l'information ou plutôt le manque d'information que nous avons sur le système. Un tel objet est en fait donné par *l'entropie*. C'est elle qui nous permettra de déterminer complètement l'opérateur densité décrivant un ensemble statistique.

6.2.2 Entropie statistique

On a déjà vu apparaître, dans le cadre de la thermodynamique, la notion d'entropie. Cette notion sera fondamentale pour l'étude de tout système statistique. Cependant, elle peut être définie de manière détachée de tout système physique grâce à la théorie de l'information. C'est ce que nous allons voir dans cette partie, avant d'étudier comment cette notion se transpose à la physique.

L'entropie comme mesure de notre ignorance

On va étudier ici un ensemble d'événements⁶ auxquels sont associés des probabilités de réalisation. Par exemple, on peut considérer l'ensemble des lettres dans un message écrit, ainsi que leur probabilité d'apparition, ou de manière plus simple l'ensemble des tirages successifs d'un jeu de pile ou face. Si chaque événement i de notre ensemble possède une probabilité p_i d'avoir lieu, on va définir l'entropie associée à notre ensemble d'événements de la manière suivante :

$$S = -k \sum_i p_i \ln(p_i), \quad (6.41)$$

k étant un facteur "arbitraire", qui sera discuté plus en détail dans la suite. Cette notion peut s'interpréter comme étant l'information contenue par un événement de notre système. En effet, si on étudie une pièce de monnaie pipée, dont la probabilité de donner le résultat "face" est de 1, on aura une entropie associée à notre système nulle. Cela revient à dire que l'événement qui va avoir lieu lors d'un lancer de cette pièce ne nous apporte aucune information, puisque l'on sait que l'on obtiendra le résultat "face" avant le lancer. À l'inverse, un système dans lequel on a de nombreux événements possibles, chacun avec une probabilité faible de se réaliser, aura une entropie élevée, ce qui traduit le fait que l'événement obtenu nous apporte beaucoup d'information (obtenir un événement particulier est en quelque sorte un "scoop").

Cette définition de l'entropie vérifie également plusieurs résultats utiles que nous allons présenter ici, et qui coïncident avec ce qui est attendu pour une notion d'information⁷.

$$S \geq 0 \quad (6.42)$$

Ce résultat traduit le fait qu'on ne peut pas avoir moins d'information sur un système après qu'un événement s'est produit qu'avant (puisque l'on sait quel événement s'est produit).

$$S = 0 \leftrightarrow p_i = \delta_{ij} \quad (6.43)$$

Le seul moyen de n'avoir gagné absolument aucune information après avoir effectué le tirage d'un événement dans notre ensemble est que le résultat du tirage soit connu avec certitude à l'avance, c'est-à-dire qu'un unique événement a une probabilité 1 de se produire (ici l'événement j).

Imaginons que l'on divise notre ensemble d'événements en deux ensembles \mathcal{A} et \mathcal{B} , tels que \mathcal{A} contienne tous les événements $i \in \llbracket 1, n \rrbracket$ et \mathcal{B} tous les autres. La probabilité d'obtenir un événement de \mathcal{A} (respectivement \mathcal{B}) est notée q_A (respectivement q_B). On a alors le résultat suivant :

$$S(p_1, \dots, p_n, p_{n+1}, \dots, p_N) = S(q_A, q_B) + q_A S(p_1, \dots, p_n) + q_B S(p_{n+1}, \dots, p_N). \quad (6.44)$$

Comment interpréter ce résultat ? On sait qu'un événement nous apporte plusieurs informations lorsqu'il a lieu : son information personnelle ainsi que l'ensemble \mathcal{A} ou \mathcal{B} auquel il appartient. Dès lors, l'entropie totale, qui mesure l'information de tout notre ensemble d'événements, peut être vue comme la somme d'une entropie associée au groupe \mathcal{A} et d'une entropie associée au groupe \mathcal{B} , chacune multipliée par la probabilité de se trouver dans chacun de ces groupes, auxquelles on doit rajouter l'entropie associée à l'ensemble des deux événements \mathcal{A} et \mathcal{B} . L'information totale est donc tout d'abord l'information sur le groupe d'événements auquel un message donné appartient, puis l'information qu'il contient au sein de ce groupe.

$$S \leq k \ln(N) \quad (6.45)$$

La valeur maximale pour l'entropie correspond au cas où tous les événements sont équiprobables ($p_i = \frac{1}{N}$). En effet, le cas équiprobable correspond au cas où l'on a le moins d'information (ignorance maximale) sur notre système initialement, et où chaque événement apporte la même quantité d'information. Dans un cas non équiprobable, l'événement qui a le plus de chance de se produire apporte moins d'information que les autres, puisqu'il s'agit d'un résultat plus "attendu".

6. On parlera également dans la suite de messages.

7. On peut même montrer, ce qui ne sera pas fait ici, qu'il s'agit de la seule fonction possible vérifiant ces différentes propriétés.



$$S_{\mathcal{A}+\mathcal{B}} \leq S_{\mathcal{A}} + S_{\mathcal{B}} \quad (6.46)$$

L'entropie totale associée à deux sources est plus faible que la somme des entropies de ces sources. On a égalité dans le cas de deux sources indépendantes. Cela peut se comprendre aisément : si une information sur les messages de \mathcal{B} peut être obtenue à partir de ceux de \mathcal{A} (voire, dans le cas extrême, si les messages de \mathcal{B} dépendent entièrement de ceux de \mathcal{A}), alors la mesure d'un événement de \mathcal{A} apporte une information sur les messages possible de \mathcal{B} , et diminue donc l'information qui sera apportée par ce deuxième message. L'information totale est donc plus faible que l'information des deux messages pris séparément.

L'entropie comme borne de compression

L'entropie d'un ensemble d'événement donné mesure l'information apportée par chaque événement de cet ensemble. Elle permet donc de connaître les ressources à mettre en œuvre pour conserver une trace de cet événement. Une façon simple de comprendre ceci consiste à étudier un jeu de pile ou face. Dans le cas où les résultats sont équiprobables, $p_{\text{pile}} = p_{\text{face}} = \frac{1}{2}$, on trouve une entropie $S = \ln(2)k$. k mesure alors l'unité de stockage d'information. Si on peut écrire le résultat sur n valeurs, on a $k = \frac{1}{\ln(n)}$. Ainsi, en binaire, on note une information avec les valeurs 0 ou 1. Dans ce cas, on a $k = \frac{1}{\ln(2)}$ et $S = 1$. Cela signifie qu'on a besoin au minimum d'une unité binaire pour stocker chaque résultat. Par exemple, on peut noter 0 les résultats "pile" et 1 les résultats "face", et noter les résultats des tirages dans l'ordre avec une suite de 0 et de 1.

Étudions désormais le cas d'une pièce pipée, avec $p_{\text{pile}} = \frac{1}{15}$, $p_{\text{face}} = \frac{14}{15}$. On a alors $S = 0.35$. Cela signifie qu'on a besoin au minimum de 0.35 unités binaires pour noter chaque résultat. Attention, il n'existe pas nécessairement de codage permettant d'obtenir ce résultat, mais on peut chercher à s'en approcher. Par exemple on peut décider que, la pièce ayant de grandes chances de tomber sur "face", on se contentera de noter la position des résultats "pile" au cours de la série, et on saura que tous les autres résultats étaient des "face". Si on effectue des séries de 15 lancers consécutifs, on peut noter les positions des "pile" avec 4 unités binaires, ainsi que la séparation entre séries avec une des combinaisons sur 4 bits (par exemple 0000). En moyenne, sur 15 lancers, on a un unique résultat "pile", et on a donc 8 bits utilisés par série (4 pour le pile et 4 pour le séparateur). On a alors 0.53 bits utilisés par résultat, ce qui n'est pas le minimum mais reste meilleur que la valeur de 1 bit par résultat du cas équiprobable.

L'entropie comme mesure de vraisemblance

Imaginons un ensemble d'événements liés par des contraintes, et dont on cherche à connaître la distribution de probabilité des événements. Dans le cas général, si on n'a que peu de contraintes (par exemple savoir que la somme de tous les p_i doit donner 1, qui est une contrainte toujours vérifiée) il existe une infinité de distributions de probabilité. En l'absence d'informations supplémentaires, l'entropie nous permet de savoir quelle distribution est la plus vraisemblable : il s'agit de celle qui maximise l'entropie compte tenu des contraintes. En effet, c'est cette distribution qui correspond au plus faible nombre de suppositions faites sur le système, c'est-à-dire celle dans laquelle on a le moins d'information sur le système *a priori*.

Lien avec la physique statistique

Pour se ramener au cas qui nous intéresse ici, c'est-à-dire celui de la mécanique statistique, nous allons effectuer une légère avancée sur ce qui va suivre en nous plaçant dès maintenant dans le cadre de l'ensemble microcanonique. Comme nous allons le voir dans la suite, cet ensemble est adapté à la description d'un système isolé, et chaque état du système est alors considéré comme équiprobable. Dans ce cadre, on a vu ici que l'entropie d'un ensemble d'événements équiprobables était de la forme

$$S = k \ln(N) \quad (6.47)$$

où N est le nombre d'événements accessibles. La mesure de l'entropie statistique d'un système physique va donc se faire en "comptant" tout simplement le nombre d'états accessibles pour ce système. De plus, le facteur multiplicatif utilisé en physique statistique est la constante de Boltzmann, k_B .

6.2.3 Ensemble microcanonique

Le point de départ de la physique statistique est l'étude d'un système isolé, c'est-à-dire où l'énergie E , le volume V et le nombre de particules N sont fixés. Lorsque notre connaissance de l'état du système n'est que partielle, une description *via* l'opérateur densité est la plus commode.

La description statistique d'un système isolé, appelé *ensemble microcanonique*, est basée sur le postulat suivant



Tous les états accessibles par le système sont *équiprobables* et *incohérents* entre eux.

Ces deux postulats permettent d'affirmer que l'opérateur densité est proportionnel à l'identité. En effet, l'hypothèse d'incohérence, ou en d'autres termes que les interférences quantiques entre les états sont nulles, nous donne un opérateur diagonal. L'hypothèse d'équiprobabilité rend égaux les termes diagonaux puisque étant effectivement égaux aux probabilités d'obtenir un état particulier. Si l'espace de Hilbert décrivant le système est de dimension W (cet espace rend compte de toutes les informations certaines sur le système), nous aurons

$$\rho = \frac{1}{W} \mathbb{1} \quad (6.48)$$

L'entropie statistique est facilement calculable dans ce cas. En fait, on est dans le pire des cas possibles où l'entropie est maximale : aucun état n'est privilégié, notre méconnaissance du système est maximale.

$$S = -k_B \sum_i p_i \ln(p_i) = k_B \ln(W) \quad (6.49)$$

On retrouve alors la formule de Boltzmann ayant permis une compréhension de l'origine microscopique de l'entropie. Elle nous dit qu'à l'équilibre, le nombre de micro-états donné par W associés à un macro-état est maximal. L'hypothèse que l'entropie est maximale à l'équilibre va être à la base de toutes les dérivations à venir et est une généralisation de l'hypothèse d'équiprobabilité

Parmi toutes les descriptions statistiques possibles d'un système satisfaisant aux conditions imposées, celle décrivant le système correspond à un maximum d'entropie.

Retour à la thermodynamique

Le système étant isolé, son énergie E est conservée. On définit de manière formelle la température microcanonique T à partir de l'entropie comme

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_x \quad (6.50)$$

où x représente l'ensemble de toutes les autres variables caractérisant le système. À partir de là, nous pouvons par exemple déterminer quelle est la température de deux systèmes en contact thermique à l'équilibre en déterminant le nombre d'états à une énergie donnée. On définit de même toutes les autres grandeurs thermodynamique à partir de l'entropie, comme la pression P et le potentiel chimique μ

$$\frac{P}{T} = \left. \frac{\partial S}{\partial V} \right|_x \quad \frac{\mu}{T} = - \left. \frac{\partial S}{\partial N} \right|_x \quad (6.51)$$

6.2.4 Ensemble canonique

Opérateur densité canonique

Les systèmes que l'on rencontre les plus souvent ne sont pas isolés mais sont en interaction avec un environnement. La situation où l'énergie n'est plus une donnée exacte du problème mais seulement donnée en moyenne est la plus fréquente. Physiquement, cela revient à considérer que le système est en contact avec un thermostat. De tel systèmes forment *l'ensemble canonique*. En clair, le volume V ainsi que le nombre de particules N sont fixés et l'énergie interne du système U n'est donnée qu'en moyenne avec

$$U = \text{Tr}(\rho H) = \langle H \rangle_\rho \quad (6.52)$$

où H désigne le hamiltonien du système et ρ son opérateur densité qui est à déterminer. Pour cela, on utilise le postulat de maximum d'entropie à l'équilibre sous les contraintes 6.52 et trace unité! La résolution se fait à l'aide de la méthode des multiplicateurs de Lagrange, soit

$$0 = \text{Tr}(\delta\rho(\ln(\rho) + 1)) + \beta \text{Tr}(\delta\rho H) + \lambda \text{Tr}(\delta\rho) \Leftrightarrow \quad (6.53)$$

$$0 = \text{Tr}(\delta\rho(\ln(\rho) + \beta H + \lambda + 1)) \quad (6.54)$$



Ceci étant valable quelle que soit la variation $\delta\rho$, on en déduit l'opérateur densité canonique à l'équilibre

$$\rho = \frac{1}{Z} e^{-\beta H} \quad (6.55)$$

Le facteur de normalisation Z , appelé *fonction de partition*, est la quantité fondamentale en physique statistique : sa connaissance permet à elle seule de déterminer toutes les grandeurs thermodynamiques du système ! Explicitement, en utilisant le fait que la trace de l'opérateur densité vaut un, nous avons

$$Z = \text{Tr} (e^{-\beta H}) \quad (6.56)$$

De manière plus développée, dans la base des états propres du Hamiltonien $(|\phi_k\rangle)_k$ que l'on supposera de cardinal N , la probabilité p_k que le système ait l'énergie E_k ainsi que la fonction de partition sont donnés par ⁸

$$p_k = \text{Tr} (\rho |\phi_k\rangle \langle \phi_k|) = \frac{e^{-\beta E_k}}{Z} \quad Z = \sum_{k=1}^N e^{-\beta E_k} \quad (6.57)$$

La chose importante à voir pour des systèmes en équilibre avec un thermostat est que tous les micro-états ne sont plus équiprobables. Leur influence dans la statistique du système est dépendante de leur énergie. Nous discuterons cela dans la suite de manière plus en rapport avec la thermodynamique.

Retour à la thermodynamique

Une fois que la fonction de partition est connue, toutes les quantités thermodynamiques peuvent être trouvées. On définit l'énergie libre de Helmholtz F par

$$Z = \sum_{k=1}^N e^{-\frac{E_k}{k_B T}} = e^{-\beta F} \quad (6.58)$$

ou de manière équivalente

$$F = -\frac{1}{\beta} \ln(Z) = -k_B T \ln(Z) \quad (6.59)$$

On peut alors exprimer l'entropie en fonction de l'énergie libre, comme on l'avait déjà vu dans la section de thermodynamique 6.1.6, par

$$S = -\frac{\partial F}{\partial T} \quad (6.60)$$

On peut ensuite retrouver simplement, grâce à la fonction de partition ou l'énergie libre, plusieurs quantités physiques comme l'énergie moyenne $\langle U \rangle$ qui vaut

$$\langle U \rangle = \sum_k p_k E_k = \sum_k E_k \frac{e^{-\frac{E_k}{k_B T}}}{Z} = -\frac{\partial \ln Z}{\partial \beta} \quad (6.61)$$

La capacité thermique quant à elle vaut alors

$$C_V = \frac{\partial U}{\partial T} = -T \frac{\partial^2 F}{\partial T^2} \quad (6.62)$$

⁸. Attention, chaque énergie apparaît autant de fois que sa dégénérescence $g(E_l)$; une écriture plus correcte de la fonction de partition serait $Z = \sum_{E_l} g(E_l) e^{-\beta E_l}$



Du microcanonique au canonique

L'ensemble canonique décrit un système physique \mathcal{S} en contact avec un thermostat \mathcal{T} qui fixe en moyenne l'énergie du système (sa température). En considérant le système et le thermostat comme un unique système isolé, nous pouvons dériver, à partir de la description microcanonique, l'opérateur densité canonique. Les états de $\mathcal{S} + \mathcal{T}$ sont équiprobables. En notant, E_{Tot} l'énergie de l'ensemble (qui est donc fixée puisque le système est isolé), la probabilité $P(E)$ que \mathcal{S} ait l'énergie E est donnée par

$$P(E) = \frac{\Omega_{\mathcal{S}}(E_{\text{Tot}} - E)}{\Omega_{\mathcal{S}+\mathcal{T}}(E_{\text{Tot}})} \quad (6.63)$$

où Ω désigne le nombre d'état d'un ensemble à l'énergie E . En oubliant la normalisation qui peut être remise à la fin, cette probabilité est une fonction de l'entropie de Boltzmann,

$$P(E) \propto \exp \left[\frac{S_{\mathcal{S}}(E_{\text{Tot}} - E)}{k_B} \right] \quad (6.64)$$

Maintenant, le fait de voir le thermostat comme un réservoir d'énergie qui fixe la température du système \mathcal{S} pré-suppose que $E \ll E_{\text{Tot}}$. Du coup, nous pouvons approximer l'argument de l'exponentielle par un développement limité de l'entropie autour de l'énergie totale,

$$S_{\mathcal{S}(E_{\text{Tot}}-E)} = S_{\mathcal{S}(E_{\text{Tot}})} - \underbrace{\frac{\partial S_{\mathcal{S}}}{\partial E}(E_{\text{Tot}})}_{\frac{1}{T_{\mathcal{T}}}} E + \frac{E^2}{2} \underbrace{\frac{\partial^2 S_{\mathcal{S}}}{\partial E^2}(E_{\text{Tot}})}_{-\frac{1}{T_{\mathcal{T}}^2 C_{V,\mathcal{T}}}} \quad (6.65)$$

Du coup,

$$P(E) \propto \exp \left[-\frac{E}{k_B T_{\mathcal{T}}} \left(1 + \frac{E}{2C_{V,\mathcal{T}} T_{\mathcal{T}}} \right) \right] \simeq \exp \left[-\frac{E}{k_B T_{\mathcal{T}}} \right] \quad (6.66)$$

dans le cas où $E \ll C_{V,\mathcal{T}} T_{\mathcal{T}}$, c'est-à-dire lorsque E est très petit devant l'énergie stockée dans le thermostat. En normalisant par Z , fonction de partition, nous retrouvons finalement la distribution de Boltzmann avec directement $\beta = \frac{1}{k_B T_{\mathcal{T}}}$.

6.2.5 Ensemble grand-canonique

Opérateur densité grand-canonique

L'ensemble grand-canonique permet de donner une description statistique d'un système où l'énergie U ainsi que le nombre de particules N ne sont donnés qu'en moyenne. On peut imaginer le système comme étant en équilibre avec un réservoir avec qui l'échange d'énergie et de particules se fait. Dans cet ensemble, les particules sont considérées comme indiscernables.

La dérivation de l'opérateur densité statistique grand-canonique repose sur la même méthode que pour l'ensemble canonique. On introduit de plus le potentiel chimique qui représente l'énergie que l'on doit fournir pour qu'une particule aille du système au réservoir; c'est le multiplicateur de Lagrange associé à l'opérateur N nombre de particules.

$$\rho = \frac{1}{\Xi} e^{-\beta H + \mu N} \quad (6.67)$$

Le facteur de normalisation ξ est cette fois appelé *fonction de partition grand-canonique*. De manière similaire au cas canonique,

$$\Xi = \text{Tr} (e^{-\beta H + \mu N}) \quad (6.68)$$

Regardons l'écriture de ces quantités de manière plus explicite; on en donnera deux écritures différentes. Le nombre de particules n'étant pas fixé, on introduit le nombre d'occupation n_a donnant le nombre de particules à l'énergie ϵ_a . La fonction de partition grand-canonique Ξ s'écrit finalement



$$\begin{aligned}\Xi &= \sum_{\{n_i\}} e^{-\beta[\sum_a n_a(\epsilon_a - \mu)]} \\ \Xi &= \sum_s e^{-\beta[E(s) - \mu N(s)]}\end{aligned}\quad (6.69)$$

où s désigne un état particulier du système N_s particules à l'énergie E_s . On peut écrire la fonction de partition grand canonique de différentes façons plus commode pour certains calculs

$$\begin{aligned}\Xi &= \sum_{\{n_i\}} e^{-\beta[\sum_a n_a(\epsilon_a - \mu)]} = \sum_{\{n_i\}} \prod_a e^{-\beta n_a(\epsilon_a - \mu)} = \prod_a \sum_{n_a} e^{-\beta n_a(\epsilon_a - \mu)} \\ \Xi &= \sum_s e^{-\beta[E(s) - \mu N(s)]} = \sum_{N_l=1}^{\infty} e^{\beta \mu N_l} Z_{\beta, N_l}\end{aligned}\quad (6.70)$$

où Z_{β, N_l} désigne la fonction de partition canonique pour un système à N_l particules tandis que $e^{\beta \mu N_l}$ est appelé *fugacité*. Cette forme particulière de la fonction de partition grand-canonique permet de la voir comme la transformée de Laplace sur N_l de la fonction de partition canonique.

Retour à la thermodynamique

Comme dans le cas canonique, on introduit le *grand potentiel* J défini par

$$J = -k_B T \ln(\Xi) = -\frac{1}{\beta} \ln(\Xi) \quad (6.71)$$

On peut montrer que le nombre moyen de particules dans le système est

$$\langle N \rangle = \frac{1}{\beta} \frac{\partial \ln(\Xi)}{\partial \mu} \quad (6.72)$$

$$= -\frac{\partial J}{\partial \mu} \quad (6.73)$$

De la même manière, le nombre moyen de particules à l'énergie E_a vaut

$$\langle n_a \rangle = \frac{\partial J}{\partial \epsilon_a} \quad (6.74)$$

L'énergie moyenne s'écrit

$$\langle U \rangle = \frac{1}{\Xi} \sum_{\{n_i\}} \left(\sum_a n_a \epsilon_a \right) e^{-\beta[\sum_a n_a(\epsilon_a - \mu)]} = \frac{\partial \beta J}{\partial \beta} + \mu \langle N \rangle \quad (6.75)$$

la pression comme

$$P = \frac{1}{\Xi} \sum_{\{n_i\}} \left(-\frac{\partial \sum_a n_a \epsilon_a}{\partial V} \right) e^{-\beta[\sum_a n_a(\epsilon_a - \mu)]} = -\frac{\partial J}{\partial V} \quad (6.76)$$

et enfin l'entropie

$$S = -\frac{\partial J}{\partial T} \quad (6.77)$$



Équivalence entre les ensembles

A priori, les différents ensembles que nous avons défini jusqu'à présent ne sont pas sensés donner les mêmes résultats pour les grandeurs thermodynamiques que l'on peut calculer (par exemple, rien ne dit que la température microcanonique s'accorde avec la canonique). Cependant on peut montrer que les prédictions faites dans différents ensembles sont identiques à la limite thermodynamique (pour les systèmes envisagés jusqu'alors).

L'idée derrière ce résultat provient de l'analyse des fluctuations autour de la valeur moyenne \bar{X} d'une grandeur thermodynamique X dans les différents ensembles. On peut alors montrer que les fluctuations relatives $\Delta X/\bar{X}$ varient en $1/\sqrt{N}$ où N est le nombre de particules. Ainsi, à la limite thermodynamique, la valeur à l'équilibre de X est égale à sa valeur moyenne \bar{X} .

6.3 Physique statistique classique

6.3.1 Description statistique

La construction de la physique statistique exposée plus haut repose sur une description quantique du système thermodynamique. En fait, on peut tout aussi bien la construire à partir simplement de la physique classique formulée dans l'espace des phases comme ce fut le cas historiquement. Il s'avère cependant qu'une formulation trop simple (trop "classique", au sens non quantique) amène à des paradoxes ou des prédictions en désaccord avec l'expérience (paradoxe de Gibbs par exemple) et de petites corrections doivent être apportées pour obtenir des bons résultats.

En mécanique, un système à N degrés de liberté est décrit par un point dans l'espace des phases qui évolue selon les équations du mouvement de Hamilton. Pour un très grand nombre de degrés de liberté, une description statistique est plus pertinente et se fait *via* l'introduction d'une densité de probabilité sur l'espace des phases $f(q, p, t)$ qui est fonction de toutes les positions et moments ainsi que du temps. L'étude complète de cette distribution de probabilité, en particulier son comportement temporel, fait l'objet d'un domaine à part entière qui est celui de la physique statistique hors-équilibre (équation de Boltzmann, distribution à une particule...).

À l'équilibre, la distribution de probabilité est, comme nous l'avons déjà vu, proportionnelle au poids de Boltzmann et permet de définir les moyennes d'ensemble ou les fonctions de partition. Par exemple, pour l'ensemble canonique

$$f(p, q) = \frac{1}{Z} e^{-\beta H(p, q)} \quad (6.78)$$

$$Z = \frac{1}{h^{3N} N!} \int e^{-\beta H(p, q)} dp dq \quad (6.79)$$

Les préfacteurs présents dans la fonction de partition se justifient correctement par une approche quantique prenant en compte l'indiscernabilité et le principe d'incertitude. Non présent à l'origine de la construction de la physique statistique, ce sont eux qui amènent à des paradoxes s'ils sont omis.

Hormis le passage au continu pour le calcul des valeurs moyennes et de la fonction de partition, peu de choses changent comparées au cas quantique. On ne va donc plus s'attarder à distinguer les différentes formulations dans la suite.

6.3.2 Du quantique au classique : formulation de Feynman-Kac

Nous avons établi dans la section précédente la forme de l'opérateur densité statistique pour différentes situations physiques. Notre but maintenant est de trouver la limite classique de cette description statistique quantique. On se restreindra ici uniquement à l'ensemble canonique. Une manière commode de voir est d'utiliser une description de l'opérateur densité par une intégrale de chemin. Cette formulation dite de Feynman-Kac permet entre autres d'avoir une vision complémentaire de certaines situations statistiques.

Considérons N particules indiscernables que l'on supposera dans un premier temps bosoniques (on verra qu'à la limite classique cette propriété n'a pas d'influence) décrites par le hamiltonien H_N . On rappelle que la fonction de partition canonique s'écrit

$$Z_N = \text{Tr} (e^{-\beta H_N}) \quad (6.80)$$

Une base de l'espace des états bosoniques est formée des fonctions d'onde symétriques $|\{r_i\}\rangle$



$$|\{\mathbf{r}_i\}\rangle = \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_n} |\mathbf{r}_{\sigma(1)} \cdots \mathbf{r}_{\sigma(N)}\rangle \tag{6.81}$$

Du coup, la fonction de partition s'écrit dans cette base comme

$$\begin{aligned} Z_N &= \frac{1}{N!^2} \sum_{\sigma, \sigma' \in \mathfrak{S}_n} \int \langle \mathbf{r}_{\sigma'(1)} \cdots \mathbf{r}_{\sigma'(N)} | e^{-\beta H_N} | \mathbf{r}_{\sigma(1)} \cdots \mathbf{r}_{\sigma(N)} \rangle d\mathbf{r}_1 \cdots d\mathbf{r}_N \\ &= \frac{1}{N!} \sum_{\sigma \in \mathfrak{S}_n} \int \langle \mathbf{r}_1 \cdots \mathbf{r}_N | e^{-\beta H_N} | \mathbf{r}_{\sigma(1)} \cdots \mathbf{r}_{\sigma(N)} \rangle d\mathbf{r}_1 \cdots d\mathbf{r}_N \end{aligned} \tag{6.82}$$

Tout le problème est donc d'abord de calculer l'intégrande pour une permutation donnée. Toute la statistique est contenue dans ces $Z(\sigma)$. On remarque que leur forme est très semblable à celle de l'opérateur d'évolution en mécanique quantique exprimé dans la base des positions. En fait, si on regarde l'équation en β vérifiée par l'opérateur densité statistique $\rho = e^{-\beta H}$ (on oublie la normalisation ici), nous avons

$$\frac{\partial \rho}{\partial \beta} = -H\rho \tag{6.83}$$

Si l'on pose sans scrupule $t = -i\beta\hbar$, on retrouve l'équation de Schrödinger ! On peut alors avoir l'idée d'introduire un opérateur d'évolution $U(\beta\hbar, 0)$ et l'exprimer à l'aide d'une intégrale de chemin. On fait ici le chemin inverse de celui exposé en mécanique quantique.

La dérivation de la fonction de partition classique pour N particules requiert un peu de discussion. On commencera d'abord, pour avoir une idée de comment les choses se passent, par une unique particule. On donne ici directement le résultat à une dimension, les détails étant exposés dans l'annexe A.2,

$$\langle x' | e^{-\beta H} | x \rangle = \rho(x, x'; \beta) = \int_{\substack{\gamma(0)=x \\ \gamma(\beta)=x'}}^{\beta\hbar} \exp\left(-\frac{1}{\hbar} \int_0^{\beta\hbar} H(\gamma(u)) du\right) \mathcal{D}\gamma \tag{6.84}$$

La limite classique, c'est-à-dire la limite haute température, est facilement accessible avec cette formulation. Le calcul de la fonction de partition requiert juste de connaître la partie diagonale de l'opérateur densité, $x = x'$, dans ce cas précis. À haute température, $\beta\hbar$ est proche de zéro, et l'intégrale du potentiel peut être approximée par $\beta\hbar V(x)$. Du coup,

$$\rho(x, x; \beta) = e^{-\beta V(x)} \int_{\substack{\gamma(0)=x \\ \gamma(\beta)=x}}^{\beta\hbar} \exp\left(-\frac{1}{\hbar} \int_0^{\beta\hbar} \frac{m\dot{\gamma}^2(u)}{2} du\right) \mathcal{D}\gamma = e^{-\beta V(x)} \rho_{\text{libre}}(x, x; \beta) \tag{6.85}$$

où $\rho_{\text{libre}}(x, x'; \beta)$ est l'opérateur densité obtenu pour un système sans potentiel extérieur. Pour expliciter cet opérateur, on peut résoudre l'équation différentielle 6.83 par la méthode des fonctions de Green par exemple. La solution est la même qu'en mécanique quantique avec le changement $t = -i\beta\hbar$. Du coup,

$$\rho_{\text{libre}}(x, x'; \beta) = \sqrt{\frac{m}{2\pi\hbar^2\beta}} e^{-m\frac{(x-x')^2}{2\hbar^2\beta}} \tag{6.86}$$

La généralisation à trois dimensions est tout aussi facile. La fonction de partition est alors donnée, en introduisant la longueur d'onde thermique de de Broglie $\lambda_T = \sqrt{\frac{2\pi\hbar^2\beta}{m}}$, par

$$Z = \int \langle \mathbf{r} | e^{-\beta H} | \mathbf{r} \rangle d\mathbf{r} = \frac{1}{\lambda_T^3} \int e^{-\beta V(\mathbf{r})} d\mathbf{r} \tag{6.87}$$

Ceci est la fonction de partition classique pour une particule de masse m à la température T dans un potentiel extérieur V . En fait, nous pouvons la mettre une forme plus simple faisant apparaître le poids de Boltzmann,

$$Z = \frac{1}{h^3} \int e^{-\beta H(\mathbf{r}, \mathbf{p})} d\mathbf{r} d\mathbf{p} \quad (6.88)$$

Maintenant, si on retourne à notre problème initial, à savoir un système composé de N particules bosoniques, le problème va venir du potentiel : on ne pourra pas *a priori* séparer facilement le potentiel de la partie cinétique dans l'intégrale de chemin. Dans ce cas, regardons d'abord le cas libre. La fonction de partition 6.82 est alors facilement exprimable avec le travail que nous déjà fait. Pour une permutation σ donnée nous avons,

$$Z(\sigma) = \int \langle \mathbf{r}_1 \cdots \mathbf{r}_N | e^{-\beta H_N} | \mathbf{r}_{\sigma(1)} \cdots \mathbf{r}_{\sigma(N)} \rangle d\mathbf{r}_1 \cdots d\mathbf{r}_N = \int e^{-\sum_{i=1}^N \frac{m(\mathbf{r}_i - \mathbf{r}_{\sigma(i)})^2}{2\hbar^2\beta}} \prod_{i=1}^N \frac{d\mathbf{r}_i}{\lambda_N^3} \quad (6.89)$$

La fonction de partition totale est la somme sur toutes les permutations de ces fonctions. Or, pour une permutation différente de l'identité, l'argument de l'exponentielle est non nul tandis qu'à la limite haute température, $\beta \ll 1$ ce qui rend la contribution à la fonction de partition totale de cette permutation négligeable. Si maintenant on considère de nouveau que le système est soumis à un potentiel V et que nous pouvons effectuer la séparation sans soucis, alors

$$Z_N \simeq \frac{1}{\lambda_N^3 N!} \int \langle \mathbf{r}_1 \cdots \mathbf{r}_N | e^{-\beta V(\mathbf{r}_1 \cdots \mathbf{r}_N)} | \mathbf{r}_1 \cdots \mathbf{r}_N \rangle d\mathbf{r}_1 \cdots d\mathbf{r}_N \quad (6.90)$$

Comme précédemment, la forme de la fonction de partition classique est donnée en faisant apparaître les poids de Boltzmann,

$$Z_N = \frac{1}{h^3 N!} \int e^{-\beta H(\mathbf{r}_1 \cdots \mathbf{r}_N, \mathbf{p}_1 \cdots \mathbf{p}_N)} \prod_{i=1}^N d\mathbf{r}_i d\mathbf{p}_i \quad (6.91)$$

Ceci est finalement la fonction de partition d'un système classique de N particules à la température T dans un potentiel V . Notons que la dérivation que nous avons faites partait de l'hypothèse bosonique. Si on avait considéré des fermions, la signature de la permutation serait apparue devant $Z(\sigma)$. En revanche, comme seule l'identité contribue significativement à la limite classique, 6.91 nous donne bien la description classique générale d'un système.

Les préfacteurs qui apparaissent dans cette expression n'étaient pas compris à l'époque de Boltzmann puisque leur origine est purement quantique. En effet, $N!$ provient de l'indiscernabilité des particules quantiques, chose qui en mécanique classique n'existe pas puisque nous pouvons étiqueter chaque particule du système. L'apparition de la constante de Planck h^3 peut être vue comme la conséquence du principe d'incertitude : une particule quantique n'est pas décrite par un point de l'espace des phases ! Les relations d'incertitude nous disent qu'en gros la particule occupe un volume h^3 de l'espace des phases ce qui réduit le nombre des états accessibles. Bien que la prise en compte de ces facteurs ne change pas les équations d'états, ils changent par contre certaines fonctions comme l'entropie : les oublier serait à l'origine de certains paradoxes comme celui de Gibbs.

La forme exacte de la fonction de partition 6.82 devient utile lorsqu'on regarde des systèmes à basse température. Elle permet d'avoir une autre vision de certains systèmes physiques. En jouant avec les permutations, nous pouvons par exemple décrire un gaz parfait de bosons comme un gaz de cycle et voir la condensation de Bose comme l'explosion de la longueur moyenne de ces cycles.

6.4 Statistiques quantiques

L'une des grandes particularités d'un système quantique de particules indiscernables provient de la statistique qu'elles doivent vérifier suivant la nature de leur spin. Pour des particules fermioniques (de spin demi-entier), le principe de Pauli interdit qu'un état quantique soit peuplé par plus d'une seule particule tandis qu'aucune restriction n'existe pour les particules bosoniques (spin entier). Ce lien entre spin et statistique est un résultat majeur des théories quantiques des champs relativistes. On va donc chercher dans ce qui suit à élucider la forme des nombres d'occupation moyen (la statistique) dans ces différents cas.

6.4.1 Statistique de Bose-Einstein

L'étude de systèmes de bosons comme l'Hélium 4, le rayonnement électromagnétique ou les vibrations d'un réseau dans un régime où les effets quantiques doivent être pris en compte (souvent cela signifie être à des basses températures) nécessite une approche statistique.

Les bosons peuvent occuper un même état en nombre arbitrairement grand. Cela revient à dire, dans le formalisme grand-canonique, que les nombres d'occupations peuvent prendre toutes les valeurs entières possibles. En utilisant 6.70, on obtient la fonction de partition grand-canonique pour les bosons

$$\Xi = \prod_i \sum_{n_i=0}^{\infty} e^{-\beta n_i(\epsilon_i - \mu)} = \prod_i \frac{1}{1 - e^{-\beta(\epsilon_i - \mu)}} \tag{6.92}$$

Ce que l'on cherche est la distribution statistique des bosons à une énergie donnée, en d'autres termes $\langle n_i \rangle$. Pour cela, le plus simple est de calculer le grand potentiel J qui vaut

$$J = \frac{1}{\beta} \sum_i \ln \left(1 - e^{-\beta(\epsilon_i - \mu)} \right) \tag{6.93}$$

Ainsi, d'après ??, on en déduit la distribution de *Bose-Einstein*.

Distribution de Bose-Einstein

$$\langle n_i \rangle = \frac{1}{e^{\beta(\epsilon_i - \mu)} - 1} \tag{6.94}$$

Les conséquences physiques de cette distribution sont nombreuses. Outre le fait qu'elle permet d'expliquer le comportement de grandeurs thermodynamiques comme la capacité thermique en fonction de la température comme on le verra en matière condensée, elle prédit le phénomène spectaculaire de la condensation de Bose-Einstein dont l'idée est simplement qu'en deçà d'une certaine température ou au-delà d'un certain nombre de particules, tous les états excités sont saturés et tout l'excédent d'atomes se retrouve dans l'état fondamental.

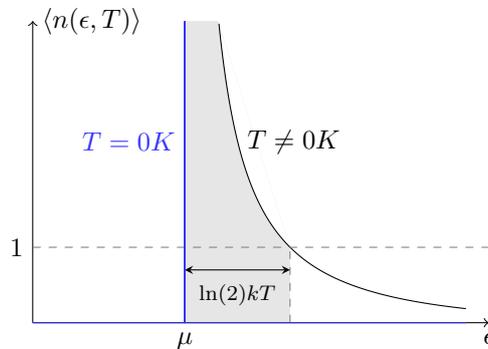


FIGURE 6.1 – Statistique de Bose-Einstein

6.4.2 Statistique de Fermi-Dirac

Pour des systèmes fermioniques comme les électrons dans un métal ou l'Hélium 3, la statistique est drastiquement changée. Les fermions, particules de spin demi-entier, sont sujets au principe de Pauli qui stipule qu'un état quantique ne peut être occupé, au plus, que par une seule particule. Cela se traduit, en grand-canonique, par le fait que $n_i \in \{0, 1\}$. La fonction de partition grand-canonique vaut alors

$$\Xi = \prod_i \left(1 + e^{-\beta(\epsilon_i - \mu)} \right) \tag{6.95}$$

Pour déterminer la distribution de *Fermi-Dirac*, on doit une fois encore calculer le grand potentiel et utiliser la formule ??. Nous discuterons plus en détails le contenu physique de cette distribution pour le gaz de fermion libre, modèle utile pour modéliser les électrons dans un métal en matière condensée.

Distribution de Fermi-Dirac

$$\langle n_i \rangle = \frac{1}{e^{\beta(\epsilon_i - \mu)} + 1} \tag{6.96}$$

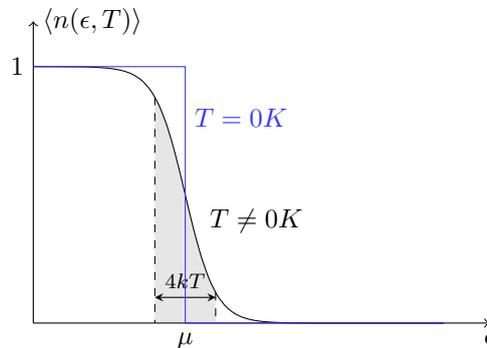


FIGURE 6.2 – Statistique de Fermi-Dirac

6.5 Gaz réels et liquides

6.5.1 Introduction

Le comportement des gaz, et même des solides, est quelque chose de bien connu. Chacun de ces états de la matière peut en effet être étudié comme un cas parfait, auquel seront rajoutées ensuite des variations. Ainsi, pour les gaz, on commence par l'étude du gaz parfait, sans interaction, et de dilution infinie, avant de passer à des modélisations plus réalistes. Pour les solides, on part d'un modèle de solide cristallin infini, auquel on rajoutera ensuite des défauts pour modéliser correctement un solide réel. Néanmoins, pour les liquides, nous ne savons que très peu de choses, et les modélisations théoriques ne sont pas si simples. Intuitivement, on sait que liquides et gaz sont très liés, puisque lorsque la température est au-dessus de la température critique nous ne savons plus les distinguer (voir la vidéo 8.1. Mais les liquides restent des états de la matière difficiles à décrire, et ce surtout d'un point de vue calculatoire. En effet, la théorie des liquides est fondée sur les mêmes principes de physique statistique que celle des gaz, mais les calculs y sont généralement bien plus compliqués, du fait de la densité et des interactions beaucoup plus importantes. Nous nous attacherons ici à faire une brève description statistique des fluides de manière générale.

6.5.2 Calculs préliminaires

Dans toute la suite, on se placera dans le cadre de l'ensemble canonique, c'est-à-dire à nombre de particules fixé. Pour un fluide avec interactions constitué de N particules, le hamiltonien s'écrit

$$H = \sum_{n=1}^N \frac{p_n^2}{2m} + H_{\text{int}}(\mathbf{r}_1, \dots, \mathbf{r}_N) \tag{6.97}$$

où H_{int} correspond à un terme d'interaction entre les différents atomes ou molécules constituant le fluide. La fonction de partition canonique de ce fluide est alors

$$\begin{aligned} Z_c &= \frac{1}{h^{3N} N!} \int d^{3N} \mathbf{r} d^{3N} \mathbf{p} e^{-\beta H} \\ &= \frac{1}{h^{3N} N!} \int d^{3N} \mathbf{r} e^{-\beta H_{\text{int}}(\mathbf{r}_1, \dots, \mathbf{r}_N)} \int d^{3N} \mathbf{p} e^{-\beta \sum_{n=1}^N \frac{p_n^2}{2m}} \\ &= \frac{1}{h^{3N} N!} (2\pi m k_B T)^{\frac{3N}{2}} \int d^{3N} \mathbf{r} e^{-\beta H_{\text{int}}} \end{aligned} \tag{6.98}$$

La dernière ligne s'obtient en effectuant l'intégration des différentes gaussiennes. On pose alors la longueur d'onde de de Broglie associée au fluide, λ_{Th} , et l'intégrale de configuration Q_N , définies par

$$\lambda_{\text{Th}} = \frac{h}{\sqrt{2\pi mk_B T}} \tag{6.99a}$$

$$Q_N = \int d^{3N} \mathbf{r} e^{-\beta H_{\text{int}}} \tag{6.99b}$$

On a alors

$$Z_c = \frac{Q_N}{N! \lambda_{\text{Th}}^{3N}} \tag{6.100}$$

La longueur d'onde λ_{Th} contient l'information sur le caractère quantique ou classique du fluide considéré. Si ρ est la densité du fluide, de telle sorte que $\rho^{-\frac{1}{3}}$ donne l'ordre de grandeur de la distance moyenne entre deux particules, le fluide est classique dans le cas $\lambda_{\text{Th}} \rho^{\frac{1}{3}} \ll 1$. On voit qu'à température suffisamment basse, un fluide doit être traité quantiquement⁹. Enfin, le but de toute étude d'un fluide réel est de trouver Q_N . Dans le cas général, on n'en connaît que des approximations.

Gaz parfait

Pour un gaz parfait, on n'a aucune interaction entre particules : $H_{\text{int}} = 0$. On alors dans ce cas particulier

$$Q_N = V^N \tag{6.101a}$$

$$Z_c = \frac{1}{N!} \left(\frac{V}{\lambda_{\text{Th}}^3} \right)^N \tag{6.101b}$$

6.5.3 Développement du Viriel

Forme du hamiltonien d'interaction

On va supposer que les termes d'interactions proviennent uniquement d'interactions de paires de particules. Ceci est vrai dans le cas de fluides non polaires. En effet, pour un fluide dipolaire, on aurait une induction mutuelle des différents dipôles qui créerait des interactions à plus de deux particules. Ici, on écrira donc

$$H_{\text{int}}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{j < k} v(|\mathbf{r}_j - \mathbf{r}_k|) = \frac{1}{2} \sum_{j \neq k} v(|\mathbf{r}_j - \mathbf{r}_k|) \tag{6.102}$$

Le potentiel v d'interaction peut être choisi de différentes façons. Néanmoins, il doit toujours être attractif à distance suffisante, tendre vers 0 pour $r \rightarrow \infty$, et être répulsif à courte distance. Les approximations les plus fréquentes sont

$$v(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad \text{Potentiel de Lennard-Jones} \tag{6.103a}$$

$$v(r) = \begin{cases} \infty & \text{si } r < \sigma \\ -\epsilon & \text{si } \sigma < r < r_c \\ 0 & \text{si } r > r_c \end{cases} \quad \text{Approximation des sphères dures} \tag{6.103b}$$

$$v(r) = \epsilon \left(e^{-2b(r-a)} - 2e^{-b(r-a)} \right) \quad \text{Potentiel de Morse} \tag{6.103c}$$

Ces différents potentiels sont visibles sur la figure 6.3.

Les méthodes développées dans la suite ne dépendront pas du choix du potentiel.

⁹. Ceci aurait dû faire l'objet d'un futur chapitre, sur les liquides quantiques, mais l'idée a été un peu laissée de côté...

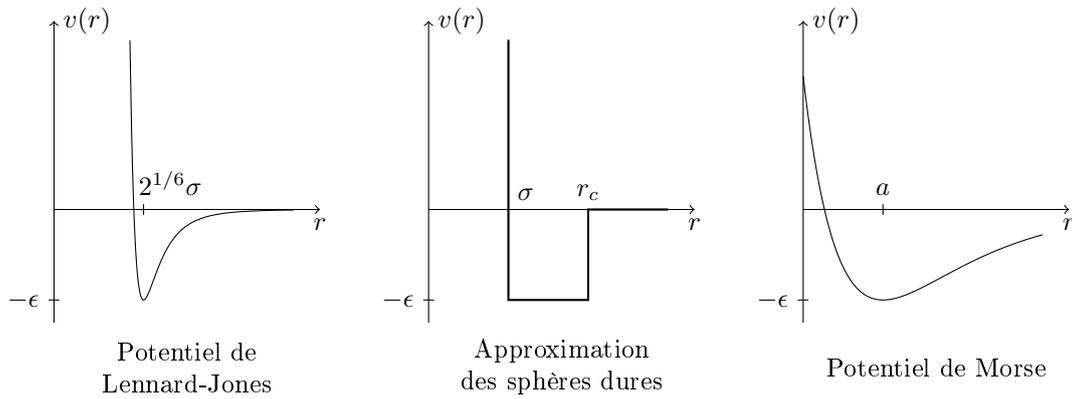


FIGURE 6.3 – Représentation des différents potentiels souvent utilisés pour approximer le potentiel d’interaction réel entre deux particules d’un gaz.

Développement du Viriel

On va chercher ici à faire un développement en puissances de la densité ρ de l’équation d’état. Bien entendu, le calcul de tous les termes de ce développement est extrêmement long, et on ne fera ici que présenter la méthode et le calcul des premiers termes. Repartons de la fonction Q_N introduite précédemment.

$$\begin{aligned}
 Q_N &= \int d^{3N} \mathbf{r} e^{-\beta H_{\text{int}}} \\
 &= \int d^{3N} \mathbf{r} e^{-\beta \sum_{j < k} v(|\mathbf{r}_j - \mathbf{r}_k|)} \\
 &= \int d^{3N} \mathbf{r} e^{-\beta \sum_{j < k} v(|\mathbf{r}_j - \mathbf{r}_k|)} \\
 &= \int d^{3N} \mathbf{r} \prod_{1 \leq j < k \leq N} e^{-\beta v(|\mathbf{r}_j - \mathbf{r}_k|)} \\
 &= \int d^{3N} \mathbf{r} \prod_{1 \leq j < k \leq N} (1 + f_{jk}) \tag{6.104}
 \end{aligned}$$

On a introduit ici la fonction f ¹⁰, telle que $f(r) = e^{-\beta v(r)} - 1$, et on a posé $f_{jk} = f(|\mathbf{r}_j - \mathbf{r}_k|)$.

Le produit 6.104 peut alors être développé en ordonnant les termes suivant les puissances croissantes de f , sous la forme de la série formelle

$$Q_N = \int d^{3N} \mathbf{r} \left[1 + \sum_{1 \leq j < k \leq N} f_{jk} + \sum_{1 \leq j < k \leq l < m \leq N} f_{jk} f_{lm} + \dots \right] \tag{6.105}$$

On peut alors écrire Z_c sous la forme d’une somme de plusieurs termes :

$$\begin{aligned}
 Z_c &= \frac{Q_N}{N! \lambda_{\text{Th}}^{3N}} \\
 &= Z_0 + Z_1 + Z_2 + \dots \tag{6.106}
 \end{aligned}$$

¹⁰. Cette fonction est commode à utiliser car elle est égale à 0 en l’absence d’interactions, strictement positive pour une interaction uniquement attractive, et strictement négative pour une interaction purement répulsive.

Et donc

$$Z_0 = \frac{1}{N! \lambda_{\text{Th}}^{3N}} \int d^{3N} \mathbf{r} = \frac{1}{N!} \left(\frac{V}{\lambda_{\text{Th}}^3} \right)^N \quad (6.107a)$$

$$\begin{aligned} Z_1 &= \frac{1}{N! \lambda_{\text{Th}}^{3N}} \int d^{3N} \mathbf{r} \sum_{1 \leq j < k \leq N} f_{jk} \\ &= \frac{1}{N! \lambda_{\text{Th}}^{3N}} \left(\frac{N(N-1)}{2} V^{N-2} \right) \int d^3 \mathbf{r}_1 d^3 \mathbf{r}_2 f(|\mathbf{r}_1 - \mathbf{r}_2|) \end{aligned} \quad (6.107b)$$

On voit que Z_0 correspond au cas du gaz parfait, sans interaction, ce qui était attendu. En revenant à $F = -k_B T \ln(Z)$, on peut faire le développement

$$P = N \frac{k_B T}{V} [1 - B(T)\rho - C(T)\rho^2 + \dots] \quad (6.108)$$

Les coefficients $B(T)$, $C(T)$ et autres sont appelés coefficients du Viriel. Ils se calculent à partir des intégrales apparaissant dans les Z_k :

$$B(T) = \frac{1}{2} \int d^3 \mathbf{r} f(r) \quad (6.109a)$$

$$C(T) = \frac{1}{3} \iint d^3 \mathbf{r}_1 d^3 \mathbf{r}_2 f(r_{12}) f(r_1) f(r_2) f(r_{12}) \quad (6.109b)$$

Cette méthode de développement a l'avantage d'être systématique et de donner de bons résultats pour les gaz à densité "pas trop forte", mais son coût en temps de calcul et en complexité la rend difficile à mettre en œuvre. Ainsi, dans le cas simple des sphères dures, le développement du Viriel n'a été mené que jusqu'à l'ordre 7. De plus, ce développement apporte toujours assez peu d'informations pour les liquides.

6.5.4 Corrélations

Pour un système canonique, la densité de probabilité de trouver les particules du système aux positions $(\mathbf{r}_i)_i$, est donnée par

$$P_N(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{Q_N} e^{-\beta H_{\text{int}}(\mathbf{r}_1, \dots, \mathbf{r}_N)} \quad (6.110)$$

Cette probabilité nous apporte beaucoup trop d'informations sur le système. On peut définir une densité réduite à s états par

$$P_s(\mathbf{r}_1, \dots, \mathbf{r}_s) = \frac{1}{Q_N} \int d^3 \mathbf{r}_{s+1} \dots \int d^3 \mathbf{r}_N e^{-\beta H_{\text{int}}(\mathbf{r}_1, \dots, \mathbf{r}_N)} \quad (6.111)$$

Ceci décrit la probabilité de trouver s particules aux positions $\mathbf{r}_1, \dots, \mathbf{r}_s$, indépendamment des positions des autres particules. Deux densités réduites vont être particulièrement utiles pour nous : $P_1(\mathbf{r})$ et $P_2(\mathbf{r}, \mathbf{r}')$. On peut montrer assez facilement la relation

$$P_1(\mathbf{r}) = \frac{1}{N} \rho(\mathbf{r}) \quad (6.112)$$

où $\rho(\mathbf{r})$ est la moyenne de la densité ordinaire à une particule. Dans le cas d'un fluide homogène et isotrope, on a tout simplement $P_1(\mathbf{r}) = \frac{1}{V}$. La fonction de corrélation à deux particules donne la probabilité de trouver une particule en \mathbf{r}' , sachant qu'il y en a une en \mathbf{r} . Pour un fluide parfait (sans corrélations), on aurait $P_2(\mathbf{r}, \mathbf{r}') = P_1(\mathbf{r})P_1(\mathbf{r}')$. Dans le cas général, le calcul est beaucoup plus compliqué. Néanmoins, $P_2(\mathbf{r}, \mathbf{r}')$ nous apporte beaucoup d'informations sur la nature du fluide. Généralement, on introduit la fonction sans dimension

$$g(|\mathbf{r} - \mathbf{r}'|) = V^2 P_2(\mathbf{r}, \mathbf{r}') \quad (6.113)$$

Par exemple, on peut démontrer la relation

$$\frac{P}{k_B T} = \rho - \frac{1}{6} \rho \int d^3 \mathbf{r} r g(r) v(r) \quad (6.114)$$

Enfin, g apporte une vision de la structure du fluide, comme cela est visible sur la figure 6.4.



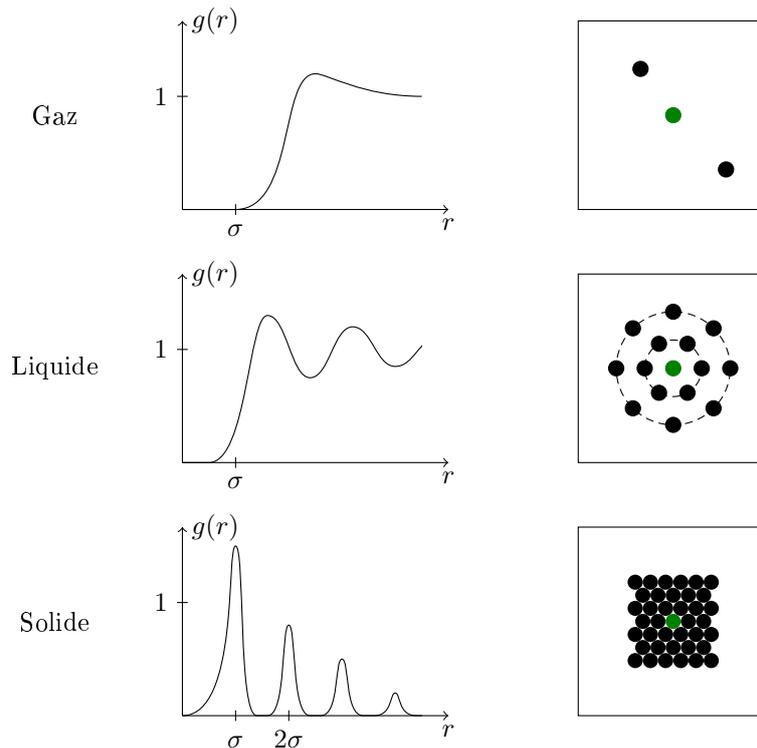


FIGURE 6.4 – Représentation de la fonction de corrélation g dans les différents états de la matière. Dans le cas des gaz, g est sensiblement constant passée une certaine distance, et nul ou presque auparavant. Ceci montre que la probabilité de trouver une particule proche de la particule déjà présente (en vert sur le schéma de droite) est faible aux petites distances, puis est la même ou presque dans tout l'espace. Pour les liquides, on voit apparaître des creux et des bosses dans la corrélation : certaines positions sont clairement plus probables que d'autres pour les particules voisines de la particule verte. On a alors une structure faiblement liée. Enfin, pour les solides, la fonction de corrélation est piquée en certains points : les particules ne peuvent se placer que sur les nœuds d'un réseau.

6.6 Théorie cinétique des gaz

La théorie cinétique des gaz, développée principalement par Boltzmann, est une première approche de la physique statistique hors équilibre. Bien que nous n'étudions pas ce domaine, cette excursion met en perspective le fait que la physique statistique (à l'équilibre) n'est pas la fin de l'histoire. Elle permet en outre de retrouver à partir de calculs mécaniques microscopiques des résultats tels que la loi du gaz parfait et de donner une définition microscopique de la notion de pression. Toutes ces notions sont *a priori* différentes de celles définies en thermodynamique/physique statistique mais on peut montrer qu'elles sont en fait égales.

On s'intéresse donc à la pression exercée par un gaz sur paroi. L'origine microscopique de cette pression vient des collisions entre le gaz et la paroi. Il s'agit donc de faire un bilan de quantité de mouvement.

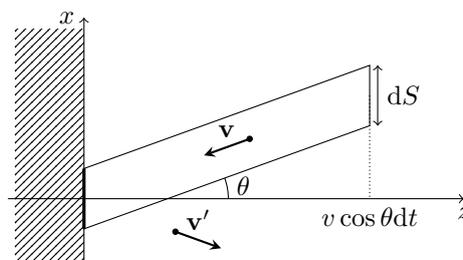


FIGURE 6.5 – Une particule qui arrive avec une vitesse \mathbf{v} sur la paroi repart avec la vitesse \mathbf{v}' , de composante z opposée. Toutes les particules avec une vitesse \mathbf{v} donnée et qui entrent en collision avec l'élément de paroi dS entre t et $t + dt$ sont situées dans le cylindre de volume $v \cos \theta dt dS$.

On va considérer que les collisions entre les particules, toutes de masses m , et la paroi sont élastiques. De plus, l'interaction entre le mur et la paroi fait qu'une particule qui arrive à vitesse \mathbf{v} sur la paroi, repart avec une vitesse \mathbf{v}' , où $v'_x = v_x$, $v'_y = v_y$ et $v'_z = -v_z$ (voir figure 6.5). Dans ces conditions, lorsqu'une particule rebondit sur la paroi, la paroi absorbe une impulsion $-2mv_z\mathbf{e}_z = -2mv \cos\theta\mathbf{e}_z$.

Pour obtenir la pression, on va s'intéresser à la quantité de mouvement absorbée par la paroi sur une surface dS , pendant un temps dt . Si on considère toutes les particules d'une vitesse \mathbf{v} donnée, l'ensemble des particules qui vont entrer en collision avec la paroi pendant le temps dt sont dans le cylindre dessinée en figure 6.5. La quantité de mouvement transférée par l'ensemble de particules de vitesse \mathbf{v} est alors donnée par :

$$d\mathbf{p}_v = -2mnv^2 \cos^2\theta dS dt \mathbf{e}_z \quad (6.115)$$

où n est la densité particulaire. La pression qui s'exerce sur la paroi est alors juste la force par unité de surface. On peut écrire, pour une vitesse :

$$dP_v = 2mnv^2 \cos^2\theta \quad (6.116)$$

Il reste alors à faire la moyenne sur l'ensemble des vitesses dans le demi-plan $z > 0$. Pour cela, on va se placer en coordonnées sphérique, et utiliser le fait que la distribution de probabilité des vitesses est isotrope ($f(\mathbf{v}) = g(v)$). On a alors :

$$P = 2mn \int g(v)v^4 \cos^2\theta \sin\theta dv d\theta d\phi = \frac{4\pi}{3}mn \int_0^\infty g(v)v^4 dv \quad (6.117)$$

On reconnaît la vitesse quadratique moyenne :

$$\langle v^2 \rangle = \int g(v)v^4 \sin\theta dv d\theta d\phi = 4\pi \int_0^\infty g(v)v^4 dv \quad (6.118)$$

Ce qui donne :

$$P = \frac{1}{3}mn \langle v^2 \rangle \quad (6.119)$$

Puisque l'on a des particules ponctuelles, le théorème de l'équipartition de l'énergie (FAUDRA EN PARLER) nous assure que $\frac{1}{2}m\langle v^2 \rangle = \frac{3}{2}k_B T$. Si l'on utilise alors le fait que la densité de particule s'écrit simplement $n = N/V$ où N est le nombre total de particules, on retrouve la loi des gaz parfaits :

$$PV = Nk_B T \quad (6.120)$$

On peut alors se demander ce qui se passe lorsque l'on a un capteur réel, qui mesure la pression comme la force par unité de surface qui s'exerce sur celui-ci. A priori, la surface \mathcal{S} d'un tel capteur n'est pas absolument plane. On va alors supposer que la pression est uniforme sur le capteur et que la force s'exerce selon la normale à celle-ci (voir figure 6.6, à gauche). De plus, le contour de cette surface est inclus dans le plan $z = 0$. La force totale qui s'exerce sur le capteur est alors :

$$\mathbf{f} = \int_{\mathcal{S}} P d\mathbf{S} \quad (6.121)$$

La pression étant uniforme, P est constant sur la surface. Il reste donc à calculer l'intégrale de $d\mathbf{S}$. En fait, on peut réécrire :

$$d\mathbf{S} = (\mathbf{e}_x \cdot d\mathbf{S})\mathbf{e}_x + (\mathbf{e}_y \cdot d\mathbf{S})\mathbf{e}_y + (\mathbf{e}_z \cdot d\mathbf{S})\mathbf{e}_z \quad (6.122)$$

On a alors trois intégrales qui sont des flux de champs de vecteur constant. On peut donc déformer continûment la surface, sans que cela ne change le résultat de l'intégrale. On va donc déformer la surface \mathcal{S} en \mathcal{S}_0 , qui possède le même contour, et qui est incluse dans le plan $z = 0$ (voir figure 6.6, à droite). On a donc

$$\mathbf{f} = \int_{\mathcal{S}} P d\mathbf{S} = P \int_{\mathcal{S}_0} d\mathbf{S} = -PS_0\mathbf{e}_z \quad (6.123)$$

où S_0 est l'aire de la surface \mathcal{S}_0 . La forme de la surface n'influe donc pas sur la valeur de la pression mesurée par le capteur.



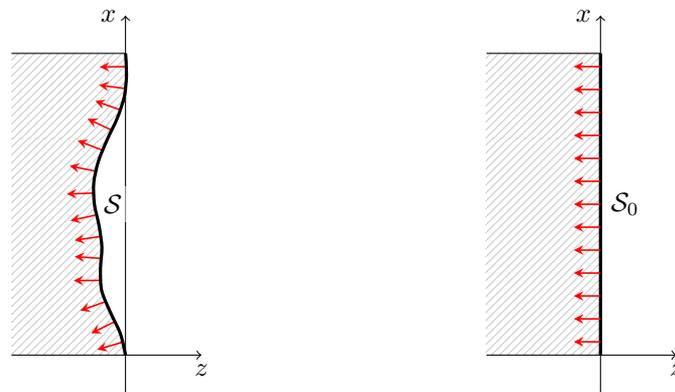


FIGURE 6.6 – La surface \mathcal{S} d'un capteur n'est a priori pas parfaitement plane. La pression qui s'exerce dessus est uniforme et normale à la surface. En fait, si l'on déforme la surface continûment en la surface \mathcal{S}_0 , qui est incluse dans le plan $z = 0$, on ne change pas la valeur de la force qui s'exerce sur le capteur.

Chapitre 7

Matière condensée

Sommaire

7.1	Sur la conduction dans les métaux	169
7.2	Structure cristalline	171
7.3	Dynamique d'un réseau cristallin	175
7.4	Structure électronique	181
7.5	Sur la conduction dans les métaux : approche semi-classique	189
7.6	Énergie de cohésion	192

Jusqu'ici, dans ce Need to Know, nous ne sommes intéressés qu'à des situations "simples" de la physique : particules uniques, peu ou pas d'interactions... Bien que ces situations permettent la description de pans entiers de la physique, elles échouent à décrire des situations complexes dans lequel on va avoir un très grand nombre de degrés de liberté, et/ou de fortes interactions entre composants. Nous avons déjà en partie parlé de ce type de problèmes lorsque nous nous sommes intéressés au cas des gaz réels et des liquides à la fin du chapitre précédent, et nous allons nous baser sur les notions de physique statistique déjà développées pour effectuer une étude d'un état de la matière qui nous entoure communément : la matière condensée. Ce domaine regroupe l'étude de toutes les propriétés des solides ou des liquides, et a donné lieu à des avancées technologiques considérables au cours du dernier siècle (développement des semi-conducteurs et explosion de l'informatique, découverte de nombreuses phases exotiques de la matière telles que la supraconductivité, la superfluidité, le régime d'effet Hall quantique...).

Au cours de ce chapitre, on verra comment le difficile problème à N-corps¹ en interactions que constitue l'étude d'une phase condensée peut être traité à travers certains modèles pour extraire des propriétés d'intérêt. On s'intéressera ici plus spécifiquement aux propriétés de conduction thermique (et notamment de chaleur spécifique) des solides et à leurs propriétés de conduction électrique.

D'un point de vue plus "linéaire", dans ce chapitre nous commencerons par un modèle classique de conduction dans les métaux, qui nous permettra de voir qu'une description quantique des phénomènes en jeu en matière condensée est nécessaire. Nous ferons ensuite une brève étude de la structure cristalline des solides, car les notions qui y seront développées serviront par la suite. Puis seront étudiés la dynamique du réseau et sa structure électronique afin de trouver un modèle raisonnable expliquant les observations expérimentales sur la chaleur spécifique des métaux et des isolants. Un retour sur la conduction des métaux sera alors effectué au vu des résultats obtenus sur leur structure électronique, avant une partie plus qualitative sur l'énergie de cohésion.

La principale référence utilisée ici est [17].

7.1 Sur la conduction dans les métaux

7.1.1 Modèle de Drude

Introduction

Le modèle de Drude vise à expliquer de manière simple certaines propriétés des métaux, comme leur conductivité électrique, à travers un modèle microscopique du mouvement des électrons dans le métal. Le métal est constitué d'atomes supposés fixes. Les électrons composant un atome sont de deux types : certains, appelés *électrons de cœur* sont fortement liés au noyau, tandis que d'autres y sont faiblement attachés.

1. Avec N de l'ordre du nombre d'Avogadro!

Ces derniers sont appelés *électrons de conduction*, ce sont eux qui participent à la conduction électronique. L'idée de Drude fut de les considérer comme un "gaz" d'électrons confinés dans la structure d'ions métalliques. Le modèle de Drude repose sur plusieurs hypothèses simplificatrices. Tout d'abord, on néglige les interactions électrons-électrons, *approximation des électrons indépendants*, et électrons-ions, *approximation des électrons libres*. Les collisions prépondérantes sont celles entre un électron et un ion²; elles sont supposées instantanées. On suppose que la probabilité de collision par unité de temps s'écrit $\frac{1}{\tau}$, où τ est appelé *temps de relaxation ou temps de libre parcours moyen*. Enfin, on suppose que les électrons sont en équilibre thermique et ce *via* les collisions. On peut montrer que le modèle de Drude se réduit à l'introduction d'une force de freinage de la forme $\frac{1}{\tau}\mathbf{p}(t)$ agissant sur les électrons de conduction.

Conductivité électrique d'un métal

Il est possible, grâce au modèle de Drude, de donner une expression de la conductivité électrique d'un métal en fonction de ses caractéristiques microscopiques. Pour cela, considérons un électron de conduction, de charge q , dans un champ électrique \mathbf{E} . Le métal est caractérisé par sa densité électronique n , nombre d'électrons par unité de volume. L'équation de la dynamique de l'électron est donc, dans le modèle de Drude,

$$m_e \frac{d\mathbf{v}(t)}{dt} = -\frac{m_e \mathbf{v}(t)}{\tau} + q\mathbf{E} \quad (7.1)$$

Attention, dans cette équation, on considère en réalité une dynamique moyenne des électrons dans le système. La quantité \mathbf{v} est donc à comprendre comme étant la vitesse moyenne d'ensemble des électrons, $\mathbf{v} = \frac{1}{N} \sum_{i=1}^N \mathbf{v}_i$. Cette vision moyenne permet de comprendre plus naturellement la force visqueuse utilisée pour modéliser les collisions. À l'équilibre, nous avons donc la relation

$$\mathbf{v}_{\text{eq}} = \frac{q\tau\mathbf{E}}{m_e} \quad (7.2)$$

La densité de courant s'écrit, quant à elle,

$$\mathbf{j}(t) = \sum_i n_i q_i \mathbf{v}_i = qn\mathbf{v} \quad (7.3)$$

d'où à l'équilibre

$$\mathbf{j}(t) = \frac{nq^2\tau\mathbf{E}}{m_e}. \quad (7.4)$$

Par définition de la conductivité électrique, $\mathbf{j}(t) = \sigma\mathbf{E}$, nous avons finalement

$$\sigma = \frac{nq^2\tau}{m_e}. \quad (7.5)$$

Bien que le modèle de Drude donne des résultats plausibles, il s'avère qu'il est loin d'être en accord avec l'expérience. Plusieurs objections au modèle peuvent être faites. Un point important est que, dans ce modèle, les électrons sont traités de manière classique. Or, nous savons que la mécanique quantique a un rôle à jouer important : les électrons sont des fermions et doivent donc obéir à la statistique de Fermi-Dirac et non à celle de Boltzmann. Dans le calcul de la conductivité électrique, la statistique n'intervient pas. Prendre en compte le caractère fermionique des électrons ne changera pas ce résultat. Cependant, d'autres applications du modèle de Drude seront modifiées.

7.1.2 Effet Hall

Un des effets qui ne peuvent pas être parfaitement compris par un modèle classique de la conduction est l'effet Hall. On considère un conducteur parallélépipédique plongé à la fois dans un champ électrique dirigé selon l'axe x et un champ magnétique dirigé selon l'axe y . À cause de la présence du champ magnétique uniforme, les électrons sont déviés dans la direction y . Ne pouvant s'échapper, ils s'accumulent sur une face du conducteur

². On verra en fait vers la fin de ce chapitre que cela est faux !



tandis qu'une charge positive apparaît sur l'autre face. À l'équilibre, le champ transverse E_y créé, appelé champ de Hall, compense la force de Lorentz. Le courant ne circulera alors que dans la direction x . On définit alors le *coefficient de Hall* R_H comme

$$R_H = \frac{E_y}{j_x B} \quad (7.6)$$

qui donne une estimation du champ de Hall. Ce que la théorie classique ne peut pas expliquer est le fait qu'il est possible que cette quantité puisse devenir positive, indiquant qu'il existerait des métaux avec des porteurs de charges positifs. Seul un traitement quantique peut l'expliquer. En effet, donnons-en une expression plus explicite. Dans un conducteur, nous avons la relation

$$\mathbf{j}(\mathbf{r}) = \boldsymbol{\sigma} \mathbf{E}(\mathbf{r}) \quad (7.7)$$

où $\boldsymbol{\sigma}$ est le tenseur de conductivité. En écrivant l'équation de Newton à l'équilibre, nous pouvons facilement arriver à l'expression

$$R_H = \frac{1}{ne} \quad (7.8)$$

où n est la densité de porteurs de charges. Ainsi, le coefficient de Hall ne dépend que de la densité de porteurs de charges du métal et de rien d'autre. En revanche, on voit expérimentalement que R_H dépend de la température ou du champ magnétique, ce qu'un calcul classique de n n'explique pas. De plus, on trouve parfois une valeur de R_H négative! On verra au cours de ce chapitre que cela est lié au fait que les porteurs de charge peuvent être des "trous" de charge positive dans certains systèmes. La valeur trouvée par un calcul *via* le modèle de Drude n'est donc retrouvée que dans des conditions particulières, ce qui indique une fois de plus que ce modèle n'est pas suffisant.

7.2 Structure cristalline

Dans cette section, on va étudier la structure de la matière condensée. Nous allons d'abord voir qu'un solide présente très souvent une structure cristalline et qu'il peut dès lors être décrit par un réseau de mailles élémentaires.

7.2.1 Cristal et réseau

Depuis le début du XXe siècle, on a eu des preuves de la périodicité des structures atomiques dans les solides. L'idée est alors de décrire le cristal périodique par un réseau infini. Un réseau peut être décrit par deux quantités : la *maille* du réseau, liée aux *vecteurs* du réseau, et le *motif* de chaque maille.

Quelques définitions

Un *réseau de Bravais* est un ensemble de points arrangés de telle façon qu'en tout point du réseau, cet arrangement apparaisse comme absolument identique. De façon équivalente, les points d'un réseau de Bravais ont un vecteur position \mathbf{R} de la forme

$$\mathbf{R} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad n_1, n_2, n_3 \in \mathbb{Z} \quad (7.9)$$

où les vecteurs \mathbf{a}_1 , \mathbf{a}_2 et \mathbf{a}_3 sont appelés *vecteurs primitifs* du réseau de Bravais, et définissent la *maille* du réseau. On définit de plus le motif du réseau comme étant la position des atomes réels dans une maille. Si on a un seul atome par maille, la maille est dite *primitive*. Si n atomes sont présents dans une maille, on définit le *motif* de la maille comme étant la position des différents atomes dans une maille par rapport au nœud, \mathbf{r}_j , avec $j \in \llbracket 1, n \rrbracket$. Ainsi, la position d'un atome est donnée par \mathbf{R} si les atomes sont les nœuds du réseau, par $\mathbf{R} + \mathbf{r}_j$ sinon. Enfin, on définit la *cellule de Wigner-Seitz* d'un réseau de Bravais comme contenant l'ensemble des points plus proches d'un nœud du réseau que d'un autre. Les différentes notions introduites ici sont toutes visibles sur la figure 7.1.



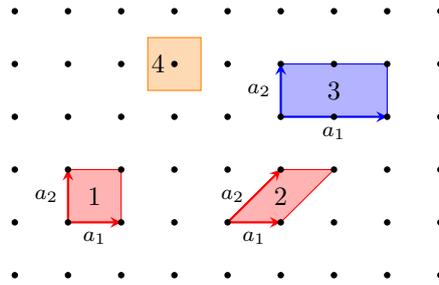


FIGURE 7.1 – Différentes mailles possibles pour décrire un réseau carré à deux dimensions. Les mailles 1 et 2 sont dites primitives, car elles ne contiennent qu'un atome. À l'inverse, la maille 3 a un motif constitué de deux atomes, l'un en (0,0) et l'autre en (0,1/2), dans la base $\mathbf{a}_1, \mathbf{a}_2$ liée à cette maille. Enfin, la maille 4 est une cellule de Wigner-Seitz du réseau : elle contient tous les points plus proches du nœud central que d'un autre.

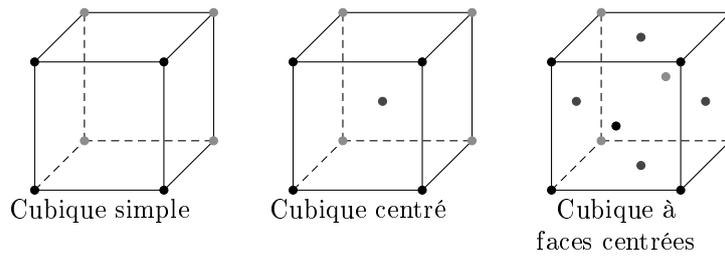


FIGURE 7.2 – Trois cas fréquemment rencontrés de réseaux cubiques. Dans le cubique simple, on a un atome en chaque nœud du réseau. La maille cubique est une maille primitive dans ce cas. Pour le cubique centré, la maille n'est plus primitive : le motif est constitué d'un atome en l'origine et d'un au centre du cube. Pour finir, la maille cubique à faces centrées a quatre atomes dans son motif : un à l'origine, et un au centre de chacune des faces liées au point d'origine.

Exemples élémentaires

La figure 7.2 présente quelques cas fréquents de réseaux de Bravais cubiques.

7.2.2 Réseau réciproque et diffraction

Diffraction d'un rayonnement par une structure périodique

On examine une situation de diffusion où une onde plane de vecteur d'onde \mathbf{k}_0 arrive sur deux points du réseau de Bravais, puis est diffusée en une onde plane de vecteur d'onde \mathbf{k}_1 . La différence de chemin optique est alors $\delta = \mathbf{R} \cdot (\mathbf{k}_1 - \mathbf{k}_0) \frac{\lambda}{2\pi}$. On veut avoir des interférences constructives, c'est-à-dire $\delta = n\lambda$. On a alors la condition de diffraction de Laue, où $\mathbf{k} = \mathbf{k}_1 - \mathbf{k}_0$:

Condition de diffraction de Laue

$$\mathbf{R} \cdot \mathbf{k} = 2\pi n \Leftrightarrow e^{i\mathbf{k}\mathbf{R}} = 1 \tag{7.10}$$

\mathbf{R} peut être n'importe quel vecteur du réseau réel. On peut alors associer à l'ensemble des \mathbf{R} un ensemble de \mathbf{k} . On va définir les vecteurs \mathbf{k} comme appartenant au *réseau réciproque*.

Réseau réciproque

De la même façon que $\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$, pour le réseau réel, un vecteur du réseau réciproque est de la forme :

$$\mathbf{k} = h\mathbf{a}_1^* + k\mathbf{a}_2^* + l\mathbf{a}_3^* \quad h, k, l \in \mathbb{Z} \tag{7.11}$$

Pour que la condition de Laue soit vérifiée, pour un \mathbf{k} donné, pour tout \mathbf{R} , on doit nécessairement avoir

$$\mathbf{a}_i \cdot \mathbf{a}_j^* = 2\pi\delta_{ij} \tag{7.12}$$



Une possibilité est de définir :

$$\mathbf{a}_1^* = 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} \quad (7.13)$$

ainsi que \mathbf{a}_2^* et \mathbf{a}_3^* par permutations circulaires des indices. On appelle *zone de Brillouin* la cellule de Wigner-Seitz du réseau réciproque. Une propriété importante est que tout vecteur diffusé \mathbf{k}_1 étant en limite de zone de Brillouin sera bien dans une condition de diffraction.

7.2.3 Détermination d'une structure cristalline par diffusion

Construction d'Ewald

Cette construction permet de déterminer très simplement si, avec une orientation du cristal donnée et un vecteur d'onde incident \mathbf{k}_0 donné, on peut se trouver dans un cas de diffusion. Si oui, elle donne également le vecteur d'onde \mathbf{k}_1 de l'onde diffusée. La construction est visible sur la figure 7.3. Elle se fait de la manière suivante : on dessine tout d'abord le réseau réciproque du cristal, puis on place le vecteur \mathbf{k}_0 à partir d'un des points du réseau. Si la sphère de rayon $|\mathbf{k}_0|$ et de centre la "pointe" de \mathbf{k}_0 touche un autre point du réseau réciproque que le point d'origine, alors on est dans un cas de diffusion, et le vecteur d'onde de l'onde diffusée est donné par la figure. Si elle ne touche pas d'autre point que l'origine, deux solutions sont possibles expérimentalement pour se placer dans un cas de diffusion : on peut changer la norme de \mathbf{k}_0 , ou on peut faire tourner le réseau réciproque en tournant le cristal dans l'espace réel (c'est souvent la solution la plus simple).

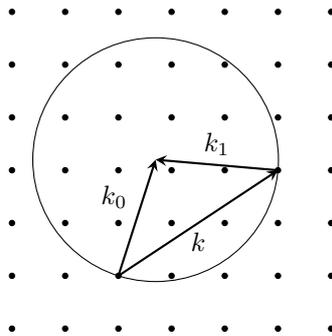


FIGURE 7.3 – La construction d'Ewald permet de trouver le vecteur d'onde \mathbf{k}_1 de l'onde diffusée à partir d'une onde incidente de vecteur d'onde \mathbf{k}_0 . Si le cercle tracé à partir d'un point du réseau réciproque et de \mathbf{k}_0 comme sur la figure touche un autre point du réseau réciproque, alors \mathbf{k}_1 est donné par la construction.

Amplitude diffusée, facteurs de forme et de structure

La construction d'Ewald est efficace lorsque l'on connaît déjà le réseau du cristal, et que l'on cherche uniquement à caractériser ses axes, c'est-à-dire à trouver son orientation. Mais cette construction ne nous apprend rien sur l'intensité de l'onde diffusée, qui dépend elle du réseau de Bravais sous-jacent et des atomes présents dans la maille. Nous allons voir ici comment.

On peut montrer tout d'abord que l'intensité diffusée est proportionnelle à la transformée de Fourier du réseau diffractant. Dans le cas général, une maille contient un motif complexe, constitué de plusieurs atomes différents aux positions \mathbf{r}_j . La densité de charge de la maille est donc

$$\rho_{\text{maille}}(\mathbf{r}) = \sum_{\mathbf{r}_j} \rho_j(\mathbf{r} - \mathbf{r}_j) \quad (7.14)$$

Dans cette équation, $\rho_j(\mathbf{r})$ est la densité de charge de l'atome situé en \mathbf{r}_j dans la maille. De plus, le réseau est périodique, et on a une maille située en chaque point d'un peigne de Dirac à trois dimensions, noté $\mathbb{III}_{\mathbf{R}}$, de telle sorte que la densité de charge totale est la convolution du peigne de Dirac et de la densité de charge de la maille :

$$\rho(\mathbf{r}) = (\mathbb{III}_{\mathbf{R}} * \rho_{\text{maille}}(\mathbf{r})) \quad (7.15)$$



La transformée de Fourier de la densité de charge, $S_{\mathbf{k}}$ est définie par

$$S_{\mathbf{k}} = \int \rho(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}} d^3\mathbf{r} \quad (7.16)$$

Puisque la transformée de Fourier d'un produit de convolution est égale au produit des transformées de Fourier des fonctions, on a

$$S_{\mathbf{k}} = \left(\int \text{III}_{\mathbf{R}}(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}} d^3\mathbf{r} \right) \times \left(\int \rho_{\text{maille}}(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}} d^3\mathbf{r} \right) \quad (7.17)$$

Dans le cas de diffusion où l'on se place, \mathbf{k} appartient au réseau réciproque, et donc $e^{-i\mathbf{k} \cdot \mathbf{R}} = 1$ (c'est-à-dire pour tous les \mathbf{R} gardés par le peigne de Dirac). La première partie de $S_{\mathbf{k}}$, appelée *facteur de structure du réseau* est donc un facteur constant, qui sera omis dans la suite. On a alors

$$\begin{aligned} S_{\mathbf{k}} &= \int \rho_{\text{maille}}(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}} d^3\mathbf{r} \\ &= \int \sum_{\mathbf{r}_j} \rho_j(\mathbf{r} - \mathbf{r}_j) e^{-i\mathbf{k} \cdot \mathbf{r}} d^3\mathbf{r} \\ &= \sum_{\mathbf{r}_j} e^{-i\mathbf{k} \cdot \mathbf{r}_j} f_j(\mathbf{k}) \end{aligned} \quad (7.18)$$

Où l'on définit le *facteur de forme atomique* f_j comme étant la transformée de Fourier de la distribution de charge atomique de l'atome en \mathbf{r}_j .

$$f_j(\mathbf{k}) = \int \rho_j(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}} d^3\mathbf{r} \quad (7.19)$$

Exemple du réseau cubique à face centrée

Le réseau de Bravais de ce réseau est un réseau cubique, on peut alors montrer facilement que son réseau réciproque est également un réseau cubique. Le vecteur \mathbf{k} considéré est donc de la forme

$$\mathbf{k} = \frac{2\pi}{a} (h\mathbf{x} + k\mathbf{y} + l\mathbf{z}) \quad (7.20)$$

avec a le pas du réseau cubique réel. Le motif de ce réseau est un ensemble de quatre atomes (visible sur la figure 7.4) situés en $\{1, 2, 3, 4\} = \frac{a}{2} \{0, \mathbf{x} + \mathbf{y}, \mathbf{y} + \mathbf{z}, \mathbf{x} + \mathbf{z}\}$.

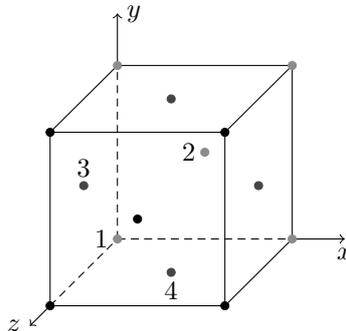


FIGURE 7.4 – La maille considérée habituellement dans le cas d'un réseau cubique à faces centrées. Le motif est alors constitué de quatre atomes situés en $\{1, 2, 3, 4\} = \frac{a}{2} \{0, \mathbf{x} + \mathbf{y}, \mathbf{y} + \mathbf{z}, \mathbf{x} + \mathbf{z}\}$.

Le facteur de structure est donc

$$S_{\mathbf{k}} = f_1 + f_2 e^{i\pi(h+k)} + f_3 e^{i\pi(k+l)} + f_4 e^{i\pi(h+l)} \quad (7.21)$$



Plusieurs cas peuvent être alors envisagés, qui ne seront pas développés ici : atomes identiques, où tous les f_i sont égaux et peuvent donc être factorisés ; atomes différents ordonnés, où l'on connaît avec certitude l'atome présent au site i ; et enfin atomes différents désordonnés, où chaque site peut être vu comme possédant un facteur de forme effectif, qui est la moyenne des facteurs de forme atomique des atomes de la maille. Ce dernier cas est donc équivalent au cas d'atomes identiques. Dans chaque cas, on aura un certain nombre de pics de diffraction possibles qui seront en fait annulés. Ainsi, si tous les f_i sont égaux, seuls les cas où h , k , et l seront de même parité apparaîtront dans la figure de diffraction finale.

7.3 Dynamique d'un réseau cristallin

Considérer des ions statiques formant un réseau est une approximation qui n'est valable qu'aux faibles températures. En effet, à cause de la température, chaque ion possède une énergie thermique et donc un mouvement autour de sa position d'équilibre. Même à température nulle, dans le cadre de la théorie quantique, les ions ne peuvent être considérés comme statiques à cause du principe d'incertitude.

Le comportement de la chaleur spécifique en fonction de la température ne peut être correctement expliqué qu'en tenant compte de la vibration du réseau. À haute température, $c_v \approx 3R$: c'est la loi de Dulong et Petit. À très basse température cependant, c_v varie linéairement pour les métaux alors que pour les isolants, elle varie comme T^3 . D'autres propriétés des matériaux sont influencées par les vibrations du réseau comme les propriétés de transport ou l'interaction avec un rayonnement.

7.3.1 Modèle classique du cristal harmonique

Hypothèses du modèle

Dans le modèle du réseau dynamique, nous faisons deux hypothèses importantes.

On suppose que les positions moyennes des différents ions forment toujours un réseau de Bravais. À chaque ion, on associe un site particulier du réseau \mathbf{R} . C'est la position d'équilibre autour de laquelle l'ion est en mouvement. On note a la distance interatomique. On suppose que l'amplitude du mouvement des ions autour de leur position d'équilibre est petite devant a . Cette hypothèse conduira à l'approximation harmonique.

Finalement, la position $\mathbf{r}_{\mathbf{R}}(t)$ de chaque ion disposé sur le réseau de Bravais des positions moyennes est donnée par

$$\mathbf{r}_{\mathbf{R}}(t) = \mathbf{R} + \mathbf{u}_{\mathbf{R}}(t) \quad (7.22)$$

où $\mathbf{u}_{\mathbf{R}}(t)$ donne la déviation par rapport à l'équilibre et vérifie $\mathbf{u}_{\mathbf{R}}(t) \ll a$, a distance interatomique.

Dynamique d'un cristal unidimensionnel monoatomique. Approximation harmonique

On considère un cristal unidimensionnel monoatomique. On a N atomes de masse m , chacun étant caractérisé par sa position relativement à l'équilibre u_n . L'énergie du système est donnée par

$$E(\{u_n\}) = E(0) + \sum_n \frac{\partial E}{\partial u_n} u_n + \frac{1}{2} \sum_{n,n'} \frac{\partial^2 E}{\partial u_n \partial u_{n'}} u_n u_{n'} + \dots \quad (7.23)$$

où $E(0)$ correspond à l'énergie de cohésion, proche de 0 (voir la section 7.6). Si l'on se place à l'équilibre, le second terme du membre de droite s'annule. Dans le cadre des hypothèses du modèle, nous avons $u_n \ll a$. L'approximation harmonique consiste à ne garder que les termes jusqu'à l'ordre deux dans le développement de l'énergie. On peut alors en déduire la force s'exerçant sur chaque ion par dérivation de l'énergie.

L'énergie du réseau à l'équilibre est donnée dans le cadre de l'approximation harmonique par

$$E(\{u_n\}) = \frac{1}{2} \sum_{n,n'} \frac{\partial^2 E}{\partial u_n \partial u_{n'}} u_n u_{n'} \quad (7.24)$$

Les dérivées secondes de l'énergie, notées $D_{nn'} = \frac{\partial^2 E}{\partial u_n \partial u_{n'}}$, sont les constantes de force. La force s'exerçant sur un ion du réseau est alors donnée par

$$F_n = -\frac{\partial E}{\partial u_n} = -\sum_{n'} D_{nn'} u_{n'} \quad (7.25)$$

La dynamique du système est finalement donnée par l'équation de Newton

$$m\ddot{u}_n = -\sum_{n'} D_{nn'} u_{n'} \quad (7.26)$$

La plupart du temps on cherche des solutions en ondes planes progressives de la forme

$$u_n(t) = Ae^{i(qna - \omega t)} \quad (7.27)$$

où $q = \frac{2\pi}{\lambda}$. Si l'on injecte cette forme dans l'équation de la dynamique, nous obtenons une relation de dispersion $\omega(q)$.

La fonction $\omega(q)$ est appelée courbe de dispersion de phonons et vérifie

$$m\omega^2(q) = \sum_{n'} D_{nn'} e^{iqa(n-n')} \quad (7.28)$$

On appelle *phonon* un mode propre de vibration collectif caractérisé par la pulsation $\omega(q)$.

Plusieurs remarques peuvent être faites. Tout d'abord, ayant N équations, on doit avoir *a priori* N valeurs possibles de pulsations. Pour plus de simplicité, on impose la *condition aux limites de Born-Von Karman* : $\forall n, u_n = u_{n+N}$. Ceci impose $e^{iqNa} = 1$, sachant que $L = (N-1)a \approx Na$. Nous avons la propriété supplémentaire suivante : la solution est identique si on remplace q par $q' = q + \frac{2\pi n'}{a}$. Les valeurs de q peuvent donc être restreinte à un intervalle de longueur $\frac{2\pi}{a}$. Finalement, q est donné par un ensemble de valeurs discrètes

$$q = n \frac{2\pi}{L}, n \in \llbracket 1, N \rrbracket \quad (7.29)$$

Chaîne monoatomique avec interaction des premiers voisins

On étudie un cas particulier du cas précédent général où on ne considère que les interactions entre premiers voisins. On peut modéliser cette situation en considérant notre système comme une chaîne d'oscillateurs harmoniques couplés de raideur k . Ce système est visible sur la figure 7.5.

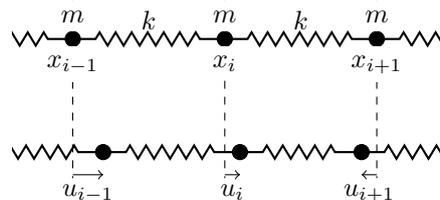


FIGURE 7.5 – Cristal unidimensionnel monoatomique, modélisé par une chaîne d'oscillateurs harmoniques couplés. x_i est la position moyenne de l'atome i , u_i représente son déplacement par rapport à l'équilibre.

La dynamique est alors donnée par l'équation de Newton

$$m\ddot{u}_n = -k(2u_n - u_{n+1} - u_{n-1}) \tag{7.30}$$

Si on cherche des solutions en ondes progressives de la forme $u_n(t) = Ae^{i(qna - \omega t)}$, nous obtenons la relation de dispersion

$$\omega(q) = \sqrt{\frac{4k}{m}} \left| \sin\left(\frac{qa}{2}\right) \right| \tag{7.31}$$

Cette relation de dispersion, présentée sur la figure 7.6 montre qu'il existe une pulsation maximale $\omega_{max} = \sqrt{\frac{4k}{m}}$. Dans l'approximation des milieux continus $qa \ll 1$, $\omega(q) = a\sqrt{\frac{k}{m}}q$. Les vitesses de groupe et de phase sont égales, et donnent la vitesse de propagation des ondes sonores dans le milieu. C'est pourquoi on l'appelle *branche acoustique*.

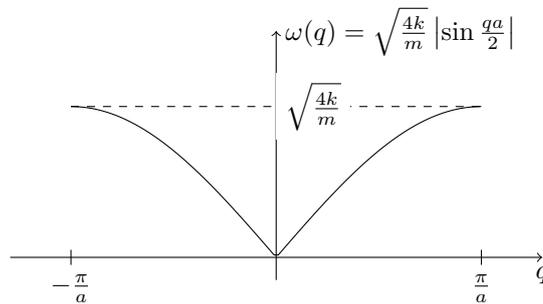


FIGURE 7.6 – Relation de dispersion pour une chaîne monoatomique d'oscillateurs harmoniques couplés. On voit apparaître une pulsation maximale. De plus, à la limite des milieux continus, la relation de dispersion peut être approximée par une relation linéaire qui nous donne la vitesse du son dans le milieu. On parle pour ce type de branches de branches acoustiques.

Chaîne diatomique avec interaction des premiers voisins

On considère cette fois un réseau de Bravais unidimensionnel avec deux ions par maille élémentaire et interaction avec les premiers voisins. On modélise encore l'interaction par un ressort de raideur k . On notera la distance entre deux ions identiques par a_0 . La figure 7.7 représente ce nouveau système.

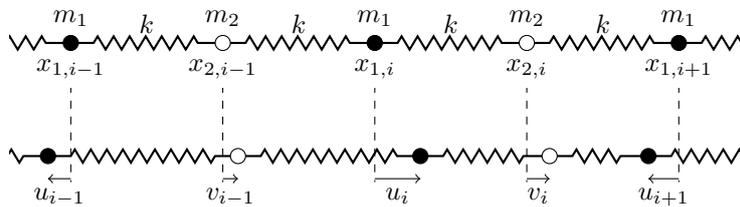


FIGURE 7.7 – Chaîne diatomique d'oscillateurs harmoniques couplés. $x_{1,i}$ est la position moyenne de l'atome i de type 1, de même pour ceux de type 2, et u_i et v_i représentent un déplacement par rapport à l'équilibre.

La dynamique est donnée par les équations

$$m_1\ddot{u}_n = -k(2u_n - v_n - v_{n-1}) \tag{7.32a}$$

$$m_2\ddot{v}_n = -k(2v_n - u_n - u_{n+1}) \tag{7.32b}$$

Pour déterminer la relation de dispersion, on considère des solutions en ondes planes progressives de la forme $u_n(t) = A_1e^{i(qna_0 - \omega t)}$ et $v_n(t) = A_2e^{i(qn(a_0 + \frac{a_0}{2}) - \omega t)}$. Les amplitudes A_1 et A_2 sont à déterminer. Pour ne pas



considérer des solutions triviales, on doit supposer que le déterminant du système d'inconnues A_1 et A_2 , obtenu en réinjectant les fonctions dans les équations ci-dessus, est égal à zéro. Ceci nous donne directement la relation de dispersion du système.

$$\omega^2(q) = k \left(\frac{1}{m_1} + \frac{1}{m_2} \right) \pm k \left[\left(\frac{1}{m_1} + \frac{1}{m_2} \right)^2 - \frac{4}{m_1 m_2} \sin^2 \left(\frac{qa_0}{2} \right) \right]^{\frac{1}{2}} \tag{7.33}$$

On rappelle que si l'on suppose les conditions aux limites de Born-Von Karman, la valeur de q peut prendre N valeurs pour une chaîne de N atomes. Pour une chaîne diatomique, nous avons donc $2N$ modes de vibrations. La relation de dispersion 7.33 possède deux branches, visibles sur la figure 7.8. L'une, similaire à celle trouvée dans le cas monoatomique, est linéaire pour des valeurs de q faibles, tout comme pour des ondes sonores. C'est pourquoi cette branche est appelée elle aussi *branche acoustique*. L'autre branche est appelée *branche optique* car ce sont ses modes qui interagissent avec le champ électromagnétique et qui sont responsables des propriétés optiques du matériau.

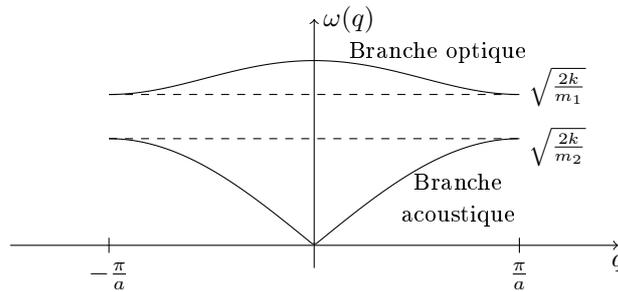


FIGURE 7.8 – Relation de dispersion pour une chaîne diatomique. On a deux branches : l'une acoustique, similaire au cas monoatomique, et l'autre est appelée branche optique. On peut remarquer que les valeurs possibles pour ω sont restreintes à certains intervalles.

Dynamique d'un cristal 3D

On considère le cas général d'un cristal à trois dimensions avec maille. On suppose que le cristal possède N mailles élémentaires, chacune possédant n atomes. Une maille est caractérisée par un vecteur de translation \mathbf{t}_k . La position moyenne d'un atome d'une maille est donné par le vecteur \mathbf{j}_i tandis que son mouvement autour de sa position d'équilibre est donné toujours par $\mathbf{u}_{n_i}(t)$. Finalement, la position d'un atome est donnée par le vecteur $\mathbf{t}_k + \mathbf{j}_i + \mathbf{u}_{n_i}(t)$. L'énergie du système est donné par

$$E(\{\mathbf{u}_{n_i}\}) = E_0 + \frac{1}{2} \sum_{n_{i\mu} n'_{i'\nu}} D_{n_{i\mu} n'_{i'\nu}} u_{n_{i\mu}} u_{n'_{i'\nu}} \tag{7.34}$$

La sommation parcourt toutes les mailles, tous les atomes par maille ainsi que les trois directions d'espace (μ). La dynamique est donnée par

$$m_i \ddot{u}_{n_{i\mu}} = - \sum_{n'_{i'\nu}} D_{n_{i\mu} n'_{i'\nu}} u_{n_{i\mu}} u_{n'_{i'\nu}} \tag{7.35}$$

On cherche des solutions en ondes planes progressives afin d'obtenir la relation de dispersion du réseau. On pose donc $\mathbf{u}_i(t) = \mathbf{A}_i(\mathbf{q}, \omega) e^{\mathbf{q} \cdot \mathbf{t}_n - \omega t}$. Le vecteur $\mathbf{A}_i(\mathbf{q}, \omega)$ est appelé *vecteur polarisation*. On utilise toujours les conditions aux limites périodiques de Born-Von Karman. En réinjectant l'expression et en supposant l'existence de solutions non triviales (déterminant du système nul) on obtient la courbe de dispersion. On obtient finalement l'existence de $3n$ branches et $3nN$ pulsations possibles. Trois de ces branches sont des branches acoustiques tandis que toutes les autres sont des branches optiques. Les $3nN$ pulsations trouvées correspondent à l'existence de $3nN$ degrés de liberté pour le système, avec 3 dimensions, n atomes par maille et N mailles.

On peut montrer que ce modèle redonne la loi de Dulong et Petit pour la chaleur spécifique. Mais ici, la chaleur spécifique serait indépendante de la température, ce qui est démenti par l'expérience.

7.3.2 Modèle quantique du cristal harmonique

Le modèle classique du cristal harmonique n'est pas capable de rendre compte de toutes les propriétés d'un cristal. Une description correcte repose alors sur la mécanique quantique.

Modèle

La description quantique du cristal harmonique se fait en se donnant un hamiltonien et en étudiant ses énergies propres. Le hamiltonien décrivant le système se trouve, dans le cadre de l'approximation harmonique, en quantifiant l'énergie 7.34. Le hamiltonien est donc

$$H = \sum_{n_i} \frac{p_{n_i}^2}{2m} + \frac{1}{2} \sum_{n_{i\mu} n'_{i'\nu}} u_{n_{i\mu}} D_{n_{i\mu} n'_{i'\nu}} u_{n'_{i'\nu}} \quad (7.36)$$

où cette fois p_{n_i} et $u_{n_{i\mu}}$ sont des opérateurs. Il est possible de diagonaliser cet hamiltonien en utilisant des opérateurs de création et d'annihilation. La conclusion est que cet hamiltonien n'est rien d'autre que la somme de $3N$ oscillateurs harmoniques indépendants. L'énergie du cristal vaut alors

$$E = \sum_{\mathbf{k}, s} \left(n_{\mathbf{k}, s} + \frac{1}{2} \right) \hbar \omega_s(\mathbf{k}) \quad (7.37)$$

où \mathbf{k} représente le vecteur d'onde et s la polarisation.

Phonons

Ce qui précède permet de décrire le cristal en terme de modes normaux de vibration harmoniques indépendants. Tout comme pour le champ électromagnétique, qui peut être décrit de la même manière par des oscillateurs harmoniques indépendants, où au lieu de parler de modes on parle de photons, nous parlerons en terme de *phonons*. On dira alors qu'il y a $n_{\mathbf{k}s}$ phonons de polarisation s et de vecteur d'onde \mathbf{k} . À chaque phonon est associé une énergie $\hbar \omega_s(\mathbf{k})$ et une impulsion $\mathbf{p} = \hbar \mathbf{k}$ ³. Ils obéissent à la statistique de Bose-Einstein. Cependant, il faut faire attention au fait que le phonon n'est pas une particule réelle mais bien un mode de vibration du cristal. On dira que le phonon est une quasi-particule.

Phonon

Le phonon est un quanta d'excitation des modes normaux de vibration du cristal. À cette quasi-particule, on associe une énergie $\hbar \omega_s(\mathbf{k})$ et une impulsion $\mathbf{p} = \hbar \mathbf{k}$. Il obéit à la statistique de Bose-Einstein.

Chaleur spécifique : formule générale

On cherche maintenant à déterminer la contribution des vibrations du réseau à la chaleur spécifique du cristal. Exprimons d'abord l'énergie moyenne de vibration par unité de volume u . Nous avons

$$u = -\frac{1}{V} \frac{\partial \ln Z}{\partial \beta} \quad (7.38)$$

où Z est la fonction de partition. Nous pouvons effectuer le calcul directement en utilisant la forme de l'énergie donnée ci-dessus. Un moyen plus rapide d'arriver au résultat est d'utiliser le fait que les phonons obéissent à la statistique de Bose-Einstein. Chaque phonon a une énergie valant $\hbar \omega_s(\mathbf{k})$. Reste maintenant à savoir le nombre moyen de phonons ayant cette énergie. Or, ceci est donné par le facteur de Bose-Einstein

$$\langle n_{\mathbf{k}, s} \rangle = \frac{1}{e^{\beta \hbar \omega_s(\mathbf{k})} - 1} \quad (7.39)$$

L'énergie moyenne rapportée au volume est alors la somme des toutes ces contributions plus celle de l'énergie du vide

³. Attention, ce n'est pas une impulsion conventionnelle !



$$u = \frac{1}{V} \sum_{\mathbf{k},s} \hbar\omega_s(\mathbf{k}) \left(\langle n_{\mathbf{k},s} \rangle + \frac{1}{2} \right) = \frac{1}{V} \sum_{\mathbf{k},s} \frac{\hbar\omega_s(\mathbf{k})}{e^{\beta\hbar\omega_s(\mathbf{k})} - 1} + \frac{1}{V} \sum_{\mathbf{k},s} \frac{1}{2} \hbar\omega_s \quad (7.40)$$

On ne tient pas compte de l'énergie d'équilibre. Finalement, on trouve la contribution des vibrations à la chaleur spécifique suivante

$$c_v = \frac{\partial u}{\partial T} = \frac{1}{V} \sum_{\mathbf{k},s} \frac{\partial}{\partial T} \frac{\hbar\omega_s(\mathbf{k})}{e^{\beta\hbar\omega_s(\mathbf{k})} - 1} \quad (7.41)$$

Contrairement au traitement classique, la chaleur spécifique n'est plus indépendante de la température et dépend essentiellement des fréquences des modes normaux.

Chaleur spécifique à haute température

Supposons la température du système assez élevée pour que l'énergie thermique domine devant toutes les énergies associées aux phonons. Nous pouvons alors effectuer un développement limité du dénominateur avec $\frac{1}{e^x - 1} = \frac{1}{x} \left[1 - \frac{x}{2} + \frac{x^2}{12} + O(x^3) \right]$. Au premier ordre, nous avons

$$\begin{aligned} c_v &= \frac{1}{V} \sum_{\mathbf{k},s} \frac{\partial}{\partial T} \frac{\hbar\omega_s(\mathbf{k})}{\beta\hbar\omega_s(\mathbf{k})} \\ &= \sum_{\mathbf{k},s} \frac{k_B}{V} \\ &= \frac{3Nk_B}{V} \end{aligned} \quad (7.42)$$

Ici, la somme se réduit bien à $3N$, nombre de modes possibles. À haute température, nous retrouvons la loi de Dulong et Petit. Cette loi prédit l'indépendance de la chaleur spécifique vis-à-vis de la température, la constante étant précisément celle trouvée ici. Nous voyons que dans le modèle quantique, ceci n'est vérifié qu'à haute température.

Chaleur spécifique à basse température

Pour déterminer le comportement à basse température, supposons d'abord le cristal assez grand, de telle sorte que le nombre de modes soit suffisant pour que l'on puisse approximer la somme donnant la chaleur spécifique par une intégrale

$$c_v = \frac{\partial}{\partial T} \sum_s \int \frac{\hbar\omega_s(\mathbf{k})}{e^{\beta\hbar\omega_s(\mathbf{k})} - 1} \frac{d\mathbf{k}}{(2\pi)^3} \quad (7.43)$$

Maintenant, si l'on se place dans la limite des très faibles températures, les modes tels que $\beta\hbar\omega_s(\mathbf{k}) \gg 1$ ne vont pas contribuer du fait de la décroissance exponentielle. Seuls les modes acoustiques ne rentrent pas dans ce cadre puisque $\omega_s(\mathbf{k})$ peut s'annuler. Faisons de plus l'approximation consistant à remplacer la relation de dispersion par sa forme aux grandes longueurs d'onde

$$\omega_s(\mathbf{k}) = c_s(\mathbf{k})k \quad (7.44)$$

où k est la norme du vecteur d'onde. Nous avons donc la forme suivante

$$c_v = \frac{\partial}{\partial T} \sum_s \int \frac{\hbar c_s(\mathbf{k})k}{e^{\beta\hbar c_s(\mathbf{k})k} - 1} \frac{d\mathbf{k}}{(2\pi)^3} \quad (7.45)$$

En intégrant sur les angles et en faisant un changement de variable,



$$c_v = \frac{\partial}{\partial T} \frac{(k_B T)^4}{(\hbar c)^3} \frac{3}{2\pi^2} \int_0^\infty \frac{x^3}{e^x - 1} dx \quad \frac{1}{c^3} = \frac{1}{3} \sum_s \int \frac{1}{c_s(\mathbf{k})^3} \frac{d\Omega}{4\pi} \quad (7.46)$$

Nous avons finalement la forme de la chaleur spécifique à basse température

$$c_v = \frac{2\pi^2}{5} k_B \left(\frac{k_B T}{\hbar c} \right)^3 \sim T^3 \quad (7.47)$$

Nous retrouvons donc un résultat expérimental que la théorie classique ne pouvait pas expliquer, à savoir que la chaleur spécifique a un comportement en T^3 à basse température.

Aparté historique : le modèle de Debye

Historiquement, Debye avait trouvé ce comportement en T^3 avec une relation de dispersion simplifiée (visible sur la figure 7.9). L'idée de Debye était de quantifier les ondes de vibration dans le solide de la même manière que Planck avait quantifié les ondes électromagnétiques. On prend donc la relation de dispersion simplifiée suivante :

$$\omega(\mathbf{k}) = \begin{cases} v_D |\mathbf{k}| & \text{si } v_D |\mathbf{k}| \leq \omega_D \\ \omega_D & \text{sinon} \end{cases} \quad (7.48)$$

La nécessité de couper à un certain ω_D permet de retrouver la loi de Dulong et Petit à haute température. La valeur de cette pulsation de coupure se trouve en imposant que le nombre total de modes doit être celui attendu ($3N$ modes pour un cristal 3D).

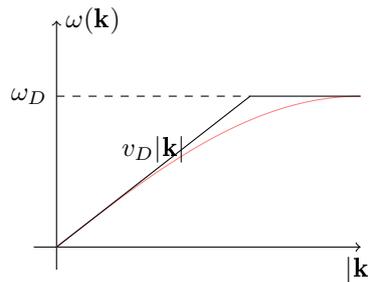


FIGURE 7.9 – Relation de dispersion simplifiée imaginée par Debye en quantifiant les modes de vibration comme Planck l’avait fait pour les photons, c’est-à-dire avec une loi linéaire (coupée à une certaine énergie limite afin de retrouver la loi de Dulong et Petit à haute température). On voit bien que cette loi capture l’essentiel du comportement réel pour une branche acoustique (en rouge).

7.4 Structure électronique

Après nous être intéressés au réseau cristallin lui-même, on s’intéresse ici à la structure électronique. Cette structure existe car le réseau est formé d’atomes pouvant fournir une partie de leurs électrons au réseau. Nous allons voir ici que ces électrons auront une influence sur le comportement du cristal.

7.4.1 Modèle des électrons libres

Dans un atome, nous avons deux “types” d’électrons, les électrons de valence qui participent aux liaisons ainsi que des électrons de cœur qui génèrent un potentiel $V(\mathbf{r})$. Le modèle des électrons libres consiste à considérer que les électrons de valence, ou de conduction, forment un gaz de fermions libres. On doit alors résoudre l’équation de Schrödinger pour la fonction d’onde des électrons libres. Pour cela, on impose des conditions aux limites périodiques associées au confinement des électrons. On peut alors déterminer la densité d’énergie de la même manière que dans 3.3.5. Nous avons

$$\mathcal{D}(E) = \frac{Vm}{2\pi^2 \hbar^3} \sqrt{2mE} \quad (7.49)$$



Gaz d'électrons à température nulle

Un gaz de fermions est décrit par la statistique de Fermi-Dirac

$$f(E, T) = \frac{1}{1 + e^{\frac{E-\mu}{k_B T}}} \tag{7.50}$$

où l'énergie μ est le potentiel chimique. À température nulle, cette fonction est une marche. L'énergie du dernier niveau occupé E_F est appelée *énergie de Fermi* et est donc égale dans ce cas au potentiel chimique. Dans le cas général, E_F et μ sont deux grandeurs différentes. Pourtant, elles sont le plus souvent considérées comme égales. Nous avons tout de même la condition $\lim_{T \rightarrow 0} \mu = E_F$, voir la figure 7.10.

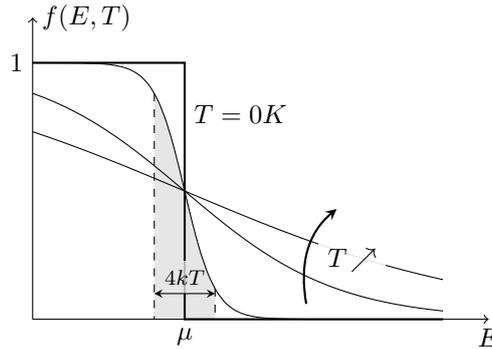


FIGURE 7.10 – Statistique de Fermi-Dirac, pour différentes températures.

Plaçons nous à température nulle. Le nombre N d'électrons est alors donné par

$$N = \int_{\mathbb{R}} \mathcal{D}(E) f(E, 0) dE = \int_0^{E_F} \frac{Vm}{2\pi^2 \hbar^3} \sqrt{2mE} dE = \frac{V}{3\pi^2 \hbar^3} (2mE_F)^{3/2}. \tag{7.51}$$

L'énergie de Fermi s'écrit donc

$$E_F = \frac{\hbar^2}{2m} \left(\frac{3\pi^2 N}{V} \right)^{2/3} \tag{7.52}$$

À partir de là, on définit le vecteur d'onde de Fermi k_F

$$k_F = \left(\frac{3\pi^2 N}{V} \right)^{1/3} \tag{7.53}$$

La surface de Fermi correspond, dans l'espace des impulsions, à l'ensemble des vecteurs d'onde dont l'énergie associée est l'énergie de Fermi : c'est une sphère de rayon k_F . Finalement, on définit la température de Fermi par

$$T_F = \frac{E_F}{k_B} \tag{7.54}$$

Pour un métal, la température de Fermi est en général très supérieure à sa température de fusion.

Gaz d'électrons à température non nulle et chaleur spécifique

On se place à température non nulle. On cherche à déterminer la contribution électronique à la chaleur spécifique du cristal. La distribution de Fermi ne diffère alors de la marche qu'autour de l'énergie de Fermi et cette différence s'étale sur un intervalle de largeur typique de quelques $k_B T$. La densité électronique nous dit de plus qu'une partie des électrons qui étaient auparavant dans la mer de Fermi ont maintenant une énergie supérieure à celle de Fermi ; ils ont été excités thermiquement (figure 7.11).

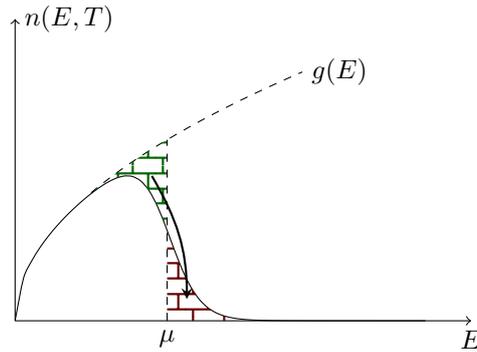


FIGURE 7.11 – Densité électronique à $T \neq 0$. Certains électrons sont excités à une énergie supérieure à celle qu'ils auraient eu dans le cas $T = 0$.

Raisonnons de manière qualitative pour obtenir la contribution que l'on recherche. Le nombre d'états excités vaut $n_e = \frac{1}{2}2k_B T \frac{1}{2}\mathcal{D}(E_F) = \frac{1}{2}k_B T \mathcal{D}(E_F)$ de manière approximative (aire du triangle rouge sur la figure 7.11). En moyenne, les électrons gagnent une énergie $2k_B T$, ce qui nous donne une différence totale en énergie de $(k_B T)^2 \mathcal{D}(E_F)$. D'où la forme de la contribution électronique

$$c_v = 2k_B^2 T \mathcal{D}(E_F) \quad (7.55)$$

Reste maintenant à déterminer la densité d'état à l'énergie de Fermi. Dans un métal, si $T \ll T_F$ nous aurons alors $f(E, T) \simeq 1$. En utilisant la formule 7.49 donnant la densité d'états ainsi que celle donnant le nombre d'électrons 7.51, nous arrivons à la forme suivante pour la densité électronique

$$n = \frac{2}{3}\mathcal{D}(E_F)E_F \quad \text{où} \quad \mathcal{D}(E_F) = \frac{3n}{2E_F} \quad (7.56)$$

Finalement, nous avons

$$c_v = 3nk_B \frac{T}{T_F} \quad (7.57)$$

Un calcul plus exact avec le développement de Sommerfeld donne la même formule à un facteur $\frac{\pi^2}{6}$ près. La contribution électronique à la chaleur spécifique a une dépendance linéaire en la température du système. L'intérêt de ce raisonnement est de voir que la dépendance en température ne provient que des électrons excités au-dessus du niveau de Fermi.

7.4.2 Théorème de Bloch

Le modèle précédent n'est pas entièrement satisfaisant. Pour aller plus loin, nous allons devoir complexifier le modèle. Un théorème général concernant les solutions des équations de Schrödinger dans un potentiel périodique est celui de Bloch, qui est énoncé ici.

Énoncé

Les électrons dans un réseau cristallin subissent un potentiel périodique dont la périodicité est celle du réseau de Bravais sous-jacent. Le hamiltonien total est un hamiltonien à N corps, ce qui rend difficile la résolution exacte. Si l'on fait l'approximation des électrons indépendants, ces électrons, dans un potentiel périodique, sont appelés électrons de Bloch. Comme conséquence de cette périodicité, les états stationnaires des électrons de Bloch ont la propriété qui suit.

Soit un état propre ψ du hamiltonien dans l'approximation des électrons indépendants pour un potentiel périodique tel que

$$\forall \mathbf{R} \in \mathcal{R}, V(\mathbf{r} + \mathbf{R}) = V(\mathbf{r}) \quad (7.58)$$

où \mathcal{R} représente le réseau de Bravais. Alors

$$\exists \mathbf{k} \in \mathbb{R}^3 \text{ tel que } \forall \mathbf{R} \in \mathcal{R}, \psi(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k} \cdot \mathbf{R}} \psi(\mathbf{r}) \quad (7.59)$$

On peut écrire la fonction d'onde de la manière suivante

$$\psi(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}} u(\mathbf{r}) \quad (7.60)$$

où $u(\mathbf{r})$ est une fonction ayant la même périodicité que le réseau.

Conditions aux limites de Born-Von Karman

On note $(\mathbf{a}_i)_i$ les vecteurs primitifs du réseau de Bravais. On rappelle l'hypothèse supplémentaire, appelée conditions aux limites périodiques de Born-Von Karman, sur la fonction d'onde :

$$\psi(\mathbf{r} + N_i \mathbf{a}_i) = \psi(\mathbf{r}) \quad (7.61)$$

où N_i est le nombre de cellules primitives du réseau. En utilisant le théorème de Bloch, on arrive à déterminer la forme du vecteur \mathbf{k} . En effet,

$$\psi(\mathbf{r} + N_i \mathbf{a}_i) = \psi(\mathbf{r}) = e^{i\mathbf{k} \cdot N_i \mathbf{a}_i} \psi(\mathbf{r}) \quad (7.62)$$

Ceci équivaut finalement à dire que

$$e^{i\mathbf{k} \cdot N_i \mathbf{a}_i} = 1 \Leftrightarrow k_i = \frac{2\pi n}{N_i}, n \in \mathbb{Z} \quad (7.63)$$

Remarques

Le vecteur d'onde de Bloch \mathbf{k} n'est pas relié à l'impulsion des électrons par la relation traditionnelle $\mathbf{p} = \hbar \mathbf{k}$. Les états de Bloch ne sont pas des états propres de l'impulsion comme on peut le vérifier facilement.

Le vecteur d'onde \mathbf{k} peut être restreint à la première zone de Brillouin.

Pour un vecteur d'onde de Bloch donné, il existe plusieurs (une infinité de) solutions à l'équation de Schrödinger pour un électron indépendant.

7.4.3 Électrons quasi-libres

Le modèle des électrons quasi-libres modélise les électrons de conduction comme un gaz d'électrons faiblement perturbé par un potentiel périodique. Ce point de vue est justifié pour certains métaux. Par exemple, le principe de Pauli interdit aux électrons de conduction d'être dans un voisinage proche des ions constituant le réseau puisque les électrons de cœur occupent déjà tous les niveaux. De plus, là où les électrons de conduction peuvent exister, du fait de leur mobilité, des effets d'écran masquent le champ créé par les ions. Ce modèle permet de décrire des électrons de couches s et p .

Équation d'évolution dans l'espace réciproque

On se place dans le cas général à trois dimensions. On notera \mathcal{K} le réseau réciproque associé au réseau de Bravais \mathcal{R} . Une fonction d'onde $\psi(\mathbf{r})$ associée à un électron de conduction peut se mettre sous la forme

$$\psi(\mathbf{r}) = \sum_{\mathbf{q} \in \mathcal{K}} c_{\mathbf{q}} e^{i\mathbf{q} \cdot \mathbf{r}} \quad (7.64)$$

Le potentiel ressenti par les électrons est un potentiel périodique ayant la même périodicité que le réseau de Bravais. Nous avons alors la décomposition

$$V(\mathbf{r}) = \sum_{\mathbf{K} \in \mathcal{K}} v_{\mathbf{K}} e^{i\mathbf{K} \cdot \mathbf{r}} \tag{7.65}$$

On rappelle que les coefficients de Fourier peuvent être calculé par $v_{\mathbf{K}} = \frac{1}{V} \int_{\text{Cellule}} V(\mathbf{r}) e^{-i\mathbf{K} \cdot \mathbf{r}} d\mathbf{r}$. On choisit l'énergie à l'origine nulle. Sachant que le potentiel est réel, nous avons $v_{-\mathbf{K}} = v_{\mathbf{K}}^*$. L'équation de Schrödinger se reformule de la façon suivante

$$\sum_{\mathbf{q} \in \mathcal{K}} e^{i\mathbf{q} \cdot \mathbf{r}} \left[(E_{\mathbf{q}}^0 - E) c_{\mathbf{q}} + \sum_{\mathbf{K}' \in \mathcal{K}} v_{\mathbf{K}'} c_{\mathbf{q}-\mathbf{K}'} \right] = 0 \tag{7.66}$$

où par définition $E_{\mathbf{q}}^0 = \frac{\hbar^2 \mathbf{q}^2}{2m}$ est l'énergie d'un électron libre. Chaque terme de la somme doit être nul ce qui nous donne

$$(E_{\mathbf{q}}^0 - E) c_{\mathbf{q}} + \sum_{\mathbf{K}' \in \mathcal{K}} v_{\mathbf{K}'} c_{\mathbf{q}-\mathbf{K}'} = 0 \tag{7.67}$$

Ramenons cette équation à la première zone de Brillouin. On sait qu'il existe \mathbf{k} vecteur du réseau réciproque tel que $\mathbf{q} = \mathbf{k} - \mathbf{K}$. Cela correspond à faire la construction visible sur la figure 7.12.

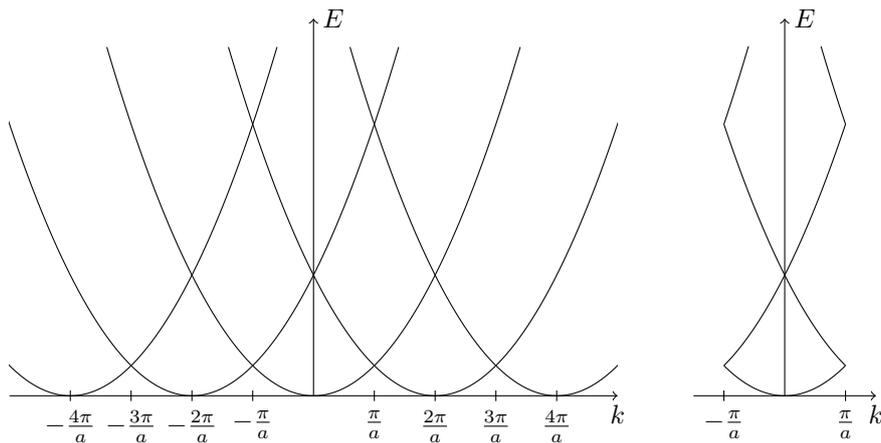


FIGURE 7.12 – Construction permettant de ramener la description du système à une description dans la première zone de Brillouin uniquement. En effet, les paraboles liées aux électrons libres suivants du réseau “croisent” la première zone de Brillouin à haute énergie, et on peut donc s’intéresser uniquement à cette zone pour décrire tout le système.

Finalement, nous avons

$$(E_{\mathbf{k}-\mathbf{K}}^0 - E) c_{\mathbf{k}-\mathbf{K}} + \sum_{\mathbf{K}' \in \mathcal{K}} v_{\mathbf{K}'} c_{\mathbf{k}-\mathbf{K}-\mathbf{K}'} = 0 \tag{7.68}$$

Approximation du potentiel faible

L'hypothèse de base du modèle des électrons quasi-libres est de les considérer comme interagissant très faiblement. Analysons d'abord le cas où les électrons sont libres, c'est-à-dire lorsque le potentiel est nul. Nous avons alors

$$(E_{\mathbf{k}-\mathbf{K}}^0 - E) c_{\mathbf{k}-\mathbf{K}} = 0 \Leftrightarrow \begin{cases} c_{\mathbf{k}-\mathbf{K}} = 0 \\ \text{ou} \\ E_{\mathbf{k}-\mathbf{K}}^0 = E \end{cases} \tag{7.69}$$

Deux cas se présentent alors. Soit il n'existe qu'un seul élément \mathbf{K} du réseau réciproque vérifiant la relation $E_{\mathbf{k}-\mathbf{K}}^0 = E$, c'est le cas non dégénéré, soit il en existe plusieurs auquel cas la solution est une combinaison linéaire de toutes les solutions. Dans ce cas, il n'y a aucune restriction sur les coefficients $c_{\mathbf{k}-\mathbf{K}}$.

Intéressons nous d'abord au cas non dégénéré. Pour \mathbf{k} donné, soit \mathbf{K}_1 un vecteur du réseau réciproque. On suppose que l'état d'énergie $E_{\mathbf{k}-\mathbf{K}_1}^0$ est loin des niveaux d'énergie dégénérés. On fait donc l'hypothèse suivante

$$\forall \mathbf{K} \in \mathcal{K}, \mathbf{K} \neq \mathbf{K}_1, |E_{\mathbf{k}-\mathbf{K}_1}^0 - E_{\mathbf{k}-\mathbf{K}}^0| \gg U \quad (7.70)$$

On cherche à savoir quel est l'effet du potentiel sur ce niveau d'énergie. On part pour cela de 7.68. Que peut-on dire des coefficients $c_{\mathbf{k}-\mathbf{K}}$? D'après cette équation, nous avons

$$c_{\mathbf{k}-\mathbf{K}} = \frac{V_{\mathbf{K}_1-\mathbf{K}} c_{\mathbf{k}-\mathbf{K}_1}}{E - E_{\mathbf{k}-\mathbf{K}}^0} + \sum_{\mathbf{K}' \neq \mathbf{K}_1} \frac{U_{\mathbf{K}'-\mathbf{K}} c_{\mathbf{k}-\mathbf{K}'}}{E - E_{\mathbf{k}-\mathbf{K}}^0} \quad (7.71)$$

Compte tenu de l'hypothèse sur le niveau d'énergie $E_{\mathbf{k}-\mathbf{K}_1}^0$, on s'attend à avoir

$$c_{\mathbf{k}-\mathbf{K}} = \frac{V_{\mathbf{K}_1-\mathbf{K}} c_{\mathbf{k}-\mathbf{K}_1}}{E - E_{\mathbf{k}-\mathbf{K}}^0} + O(V^2) \quad (7.72)$$

Grâce à cela et à l'équation 7.68, nous pouvons dire que le niveau d'énergie E perturbé diffère du niveau non perturbé $E_{\mathbf{k}-\mathbf{K}_1}^0$ par des termes d'ordre deux. On obtient alors directement

$$E = E_{\mathbf{k}-\mathbf{K}_1}^0 + \sum_{\mathbf{K}} \frac{|V_{\mathbf{K}-\mathbf{K}_1}|^2}{E_{\mathbf{k}-\mathbf{K}_1}^0 - E_{\mathbf{k}-\mathbf{K}}^0} + O(V^3) \quad (7.73)$$

Le point important à retenir est que pour des niveaux faiblement perturbés loin des niveaux dégénérés, le décalage en énergie est du second ordre en V . C'est la différence principale entre ces niveaux et les autres.

Pour des niveaux quasi-dégénérés, on suppose que les m énergies sont toutes du même ordre (éloignées d'au plus V les unes des autres). Mais on fait toujours l'hypothèse

$$\forall \mathbf{K} \in \mathcal{K}, \forall i \in [1, m], \mathbf{K} \neq \mathbf{K}_i, |E_{\mathbf{k}-\mathbf{K}_i}^0 - E_{\mathbf{k}-\mathbf{K}}^0| \gg U \quad (7.74)$$

En faisant le même type d'opérations que dans le cas non dégénéré, nous arrivons à l'expression généralisée

$$[E - E_{\mathbf{k}-\mathbf{K}_i}^0] c_{\mathbf{k}-\mathbf{K}_1} = \sum_{j=1}^m U_{\mathbf{K}_j-\mathbf{K}_i} c_{\mathbf{k}-\mathbf{K}_j} \quad (7.75)$$

Cette fois nous avons donc un système de m équations. Un des cas le plus important est celui où nous avons deux niveaux dégénérés. Dans ce cas, nous avons une ouverture d'un gap en limite de la zone de Brillouin dans le cas d'électrons qui auraient été dégénérés dans le modèle des électrons libres. La largeur du gap est alors de $2|V_{\mathbf{K}}|$. On peut voir la modification du modèle des électrons libres lorsque l'on prend en compte un faible potentiel dans les figures 7.13 et 7.14.

7.4.4 Modèle des liaisons fortes

Le modèle des liaisons fortes décrit un métal comme une collection d'atomes neutres interagissant faiblement entre eux. Sur chaque nœud du réseau nous avons un atome neutre. Tous ses électrons sont décrits par des fonctions d'onde solution du hamiltonien atomique. Cependant, on peut s'attendre à ce que les fonctions d'onde associées à différents atomes du réseau puissent se recouvrir. L'idée du modèle des liaisons fortes est de supposer ces recouvrements assez faibles pour que la description du métal comme un réseau d'atomes neutres aie toujours un sens, mais cependant non nuls pour que des corrections puissent avoir lieu. Ce modèle est utile pour décrire des électrons de couche d mais aussi les isolants.

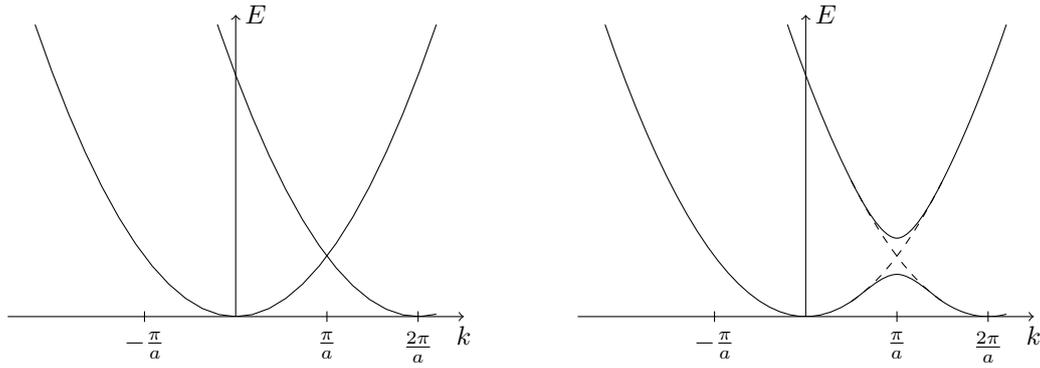


FIGURE 7.13 – Sur la figure de gauche est représenté le cas d’électrons libres. On voit qu’en limite de zone de Brillouin, on a des niveaux quasi-dégénérés pour deux paraboles distinctes. Lorsque l’on passe au modèle des électrons quasi-libres, cette dégénérescence est levée et on a ouverture d’un gap en limite de zone de Brillouin.

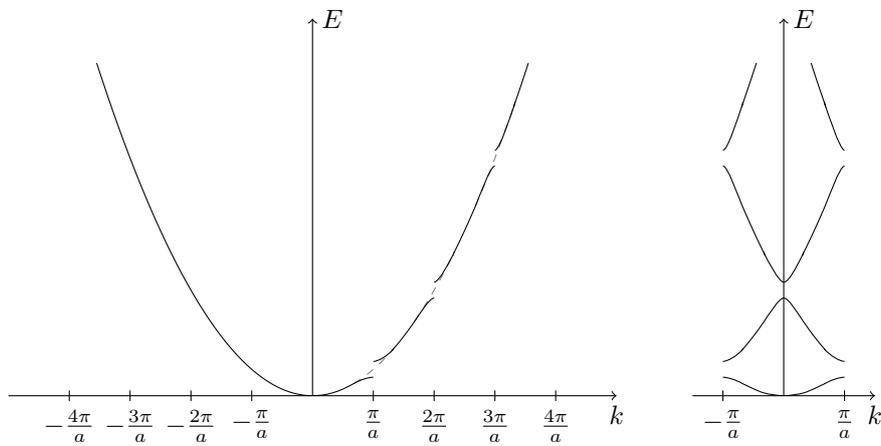


FIGURE 7.14 – Si l’on observe ce qui se passe en limite de zone de Brillouin à plus haute énergie, ou de façon équivalente ce qui se passe aux limites de zones suivantes, on se rend compte qu’on a ouverture de plusieurs gaps. En ramenant tout en première zone de Brillouin, on voit apparaître une structure en bandes pour l’énergie.

Approximation des liaisons fortes

Le hamiltonien général du système, constitué de N_n noyaux identiques et de N_e électrons s’écrit

$$H = \frac{1}{2M} \sum_{i=1}^{N_n} \mathbf{P}_i^2 + \frac{Z^2}{2} \sum_{i \neq j}^{N_n} V_c(\mathbf{R}_i - \mathbf{R}_j) + \frac{1}{2m_e} \sum_{i=1}^{N_e} \mathbf{p}_i^2 + \frac{1}{2} \sum_{i \neq j}^{N_e} V_c(\mathbf{r}_i - \mathbf{r}_j) + \sum_{i=1}^{N_n} \sum_{j=1}^{N_e} V_c(\mathbf{R}_i - \mathbf{r}_j) \quad (7.76)$$

Les trois derniers termes correspondent à la contribution électronique. En faisant une approximation de champ moyen, c’est-à-dire en considérant que chaque électron subit un potentiel moyen créé par chaque atome, le hamiltonien électronique s’écrit

$$H_e = \sum_{i=1}^{N_e} \left[\frac{1}{2m_e} \mathbf{p}_i^2 + \sum_{j=1}^{N_n} V_{at}(\mathbf{R}_j - \mathbf{r}_i) \right] \quad (7.77)$$

$$= \sum_{i=1}^{N_e} H_i \quad (7.78)$$

On doit donc résoudre le problème $H_i \psi_{k_i}(\mathbf{r}_i) = E_{k_i} \psi_{k_i}(\mathbf{r}_i)$. La solution globale s’écrit alors $\psi(\mathbf{r}_1, \dots, \mathbf{r}_{N_e}) = \psi_{k_1}(\mathbf{r}_1) \dots \psi_{k_{N_e}}(\mathbf{r}_{N_e})$ et l’énergie du système est simplement la somme des énergies $E = \sum_{i=1}^{N_e} E_{k_i}$. Mais, puisque les électrons sont des fermions, nous devons antisymétriser la fonction d’onde grâce à un déterminant de Slater. Finalement,

$$\psi(\mathbf{r}_1, \dots, \mathbf{r}_{N_e}) = \frac{1}{\sqrt{N_e!}} \begin{vmatrix} \psi_{k_1}(\mathbf{r}_1) & \cdots & \psi_{k_1}(\mathbf{r}_{N_e}) \\ \vdots & \ddots & \vdots \\ \psi_{k_{N_e}}(\mathbf{r}_1) & \cdots & \psi_{k_{N_e}}(\mathbf{r}_{N_e}) \end{vmatrix} \quad (7.79)$$

Le modèle des liaisons fortes suppose que les électrons sont fortement liés aux atomes qui, eux, interagissent très faiblement. Pour un électron d'un atome, nous avons l'équation

$$\left[\frac{\mathbf{p}^2}{2m} + V_{\text{at}}(\mathbf{r}) \right] \chi_n(\mathbf{r}) = E_n \chi_n(\mathbf{r}) \quad (7.80)$$

où $\chi_n(\mathbf{r})$ est une fonction d'onde électronique associée à l'énergie E_n . Toujours dans ce modèle, seul le dernier électron d'un atome est susceptible d'interagir avec les autres atomes. On note donc E_ν le dernier état occupé par les électrons d'un atome, $\chi_\nu(\mathbf{r})$ étant alors la fonction d'onde associée. Pour les atomes du cristal, on s'intéresse donc aux fonctions d'onde $\chi_\nu(\mathbf{r} - \mathbf{R})$ où \mathbf{R} est un vecteur du réseau de Bravais.

On supposera donc que les énergies propres du système total E_k sont proches de E_ν tout en étant éloignées des autres. On exprimera alors les fonctions d'onde ψ_k solution de l'équation de Schrödinger associé au hamiltonien H_k comme combinaison linéaire des orbitales atomiques

$$\psi_k(\mathbf{r}) = \frac{1}{\sqrt{N_e}} \sum_{l=1}^{N_e} a_{kl} \chi_\nu(\mathbf{r} - \mathbf{R}_l) \quad (7.81)$$

En notation de Dirac, on écrira $|\psi_k(\mathbf{r})\rangle = \frac{1}{\sqrt{N_e}} \sum_{l=1}^{N_e} a_{kl} |\mathbf{R}_l\rangle$. Finalement nous avons

$$\sum_{l=1}^{N_e} a_{kl} H_k |\mathbf{R}_l\rangle = \sum_{l=1}^{N_e} a_{kl} E_k |\mathbf{R}_l\rangle \quad (7.82)$$

Réécrivons le hamiltonien d'une autre façon

$$H_k = \underbrace{\frac{\mathbf{p}^2}{2m_e} + V_{\text{at}}(\mathbf{r} - \mathbf{R}_l)}_{\text{Hamiltonien de l'atome } l} + \underbrace{\sum_{i \neq l} V_{\text{at}}(\mathbf{r} - \mathbf{R}_i)}_{=V_l \text{ Potentiel des autres sites}} \quad (7.83)$$

Avec ceci, $H_k |\mathbf{R}_l\rangle = H_l |\mathbf{R}_l\rangle + V_l(\mathbf{r}) |\mathbf{R}_l\rangle = E_\nu |\mathbf{R}_l\rangle + V_l(\mathbf{r}) |\mathbf{R}_l\rangle$. D'où

$$\sum_{l=1}^{N_e} a_{kl} [E_\nu + V_l(\mathbf{r})] |\mathbf{R}_l\rangle = \sum_{l=1}^{N_e} a_{kl} E_k |\mathbf{R}_l\rangle \quad (7.84)$$

Dans l'hypothèse où les fonctions d'onde ont un recouvrement très faible, nous avons $\langle \mathbf{R}_k | \mathbf{R}_l \rangle = \delta_{kl}$, et on pose les *intégrales de transfert*

$$t_{nl} = -\langle \mathbf{R}_n | V_l(\mathbf{r}) | \mathbf{R}_l \rangle \quad (7.85)$$

Finalement, nous avons l'équation générale nous donnant les énergies du cristal dans le modèle des liaisons fortes

$$-\sum_{l=1}^{N_e} t_{nl} a_{kl} = (E_k - E_\nu) a_{kn} \quad (7.86)$$



Application

Considérons une chaîne linéaire monoatomique périodique de pas a à N_e sites. Nous avons donc $\psi_k(x + N_e a) = \psi_k(x)$ ce qui se traduit par $a_{kn} = a_{k(n+N_e)}$. Faisons l'approximation des premiers voisins : seuls t_0 et $t_{\pm 1}$ sont non nuls. Nous avons donc

$$-t_1 [a_{k(n+1)} + a_{k(n-1)}] = (E_k - E_\nu + t_0) a_{kn} \tag{7.87}$$

Grâce au théorème de Bloch, nous avons

$$-t_1 [e^{ika} + e^{-ika}] = (E_k - E_\nu + t_0) \Leftrightarrow E_k = E_\nu - t_0 - 2t_1 \cos(ka) \tag{7.88}$$

Nous voyons donc que les énergies accessibles pour les électrons de conduction sont contenues dans une bande appelée *bande de conduction*, qui, ici, est de largeur $4t_1$. On rappelle que dans le modèle des électrons libres, l'énergie était fonction du carré du vecteur d'onde k et n'était donc nullement bornée. Cependant, dans le cas particulier où $ka \gg 1$, nous avons $E_k = E'_0 + t_1 a^2 k^2$. Par analogie avec le cas libre, on pose m^* la *masse effective* de l'électron définie par

$$t_1 a^2 = \frac{\hbar^2}{2m^*} \Leftrightarrow m^* = \frac{\hbar^2}{2t_1 a^2} \tag{7.89}$$

Près des minima des bandes, les électrons peuvent être assimilés à des électrons libres simplement en modifiant leur masse par la masse effective. Cette masse inclut l'influence du potentiel sur l'électron. La figure 7.15 présente ce lien entre les différents modèles

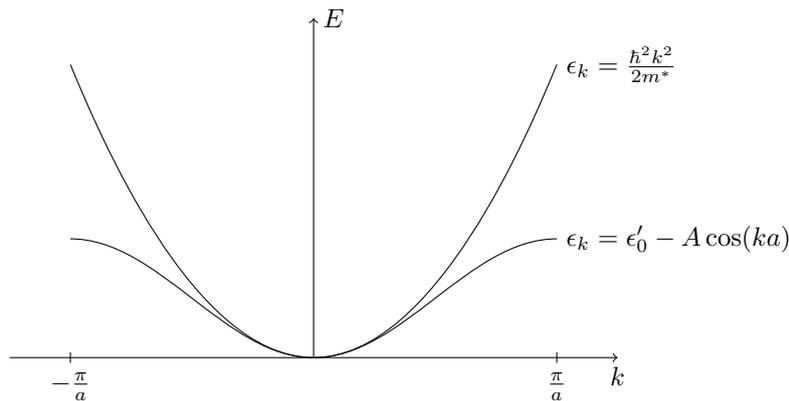


FIGURE 7.15 – La dépendance de l'énergie en k dans le cas du modèle des électrons libres (parabole) et dans celui des liaisons fortes (loi en cosinus). On voit qu'aux faibles énergies (c'est-à-dire aux grandes longueurs d'onde), les deux modèles donnent les mêmes résultats si l'on utilise la masse effective des électrons pour le modèle des électrons libres.

7.5 Sur la conduction dans les métaux : approche semi-classique

7.5.1 Modèle semi-classique

Lorsque nous avons étudié le modèle de Drude, deux questions sont restées en suspens : quelle était la vraie nature des collisions et quel était le mouvement entre ces collisions. Le modèle de Drude, qui supposait que les électrons subissaient des collisions avec les ions du réseau, n'est pas correct compte tenu des très grands libres parcours moyens. Dans un modèle quantique, les électrons de conduction sont décrits par un paquet d'ondes obéissant à l'équation de Schrödinger stationnaire en présence d'un potentiel périodique créé par le réseau. Dans ce cas, si l'électron a une vitesse non nulle, il la gardera indéfiniment (stationnaire). Ceci peut se comprendre grâce au caractère ondulatoire de l'électron : la fonction d'onde peut se propager sans atténuation dans un réseau périodique grâce aux interférences constructives des ondes diffusées.

Dit autrement, un électron étant décrit dans un réseau périodique comme une onde de Bloch, il devrait avoir une propagation non diffusive dans ce réseau. Cependant, il existe physiquement des mécanismes de diffusion



liés aux interactions électrons-phonons (et donc aux vibrations du réseau sous-jacent, pas uniquement à sa présence), électrons-électrons et électrons-impuretés. On ne cherchera pas ici à modéliser de manière complète ces interactions, mais on se contentera d'une approche effective en force visqueuse comme on le faisait dans le modèle de Drude.

L'idée du modèle semi-classique est de traiter le potentiel créé par les ions du réseau de manière quantique tandis que l'influence des champs extérieurs sera traitée de manière classique. Le but du modèle est de relier la structure électronique aux propriétés de transport. Un électron d'une bande n est caractérisé par trois paramètres, $E_n(\mathbf{k})$, \mathbf{k} et \mathbf{r} ⁴. On fera alors les hypothèses suivantes. Il n'y a pas de transition interbande, n est une constante du mouvement. La dynamique est donnée par

$$\mathbf{v}_n(\mathbf{k}) = \dot{\mathbf{r}} = \frac{1}{\hbar} \frac{\partial E_n(\mathbf{k})}{\partial \mathbf{k}} \quad (7.90a)$$

$$\dot{\mathbf{p}} = \hbar \dot{\mathbf{k}} = -e [\mathbf{E}(\mathbf{r}, t) + \mathbf{v}(\mathbf{r}, t) \times \mathbf{B}(\mathbf{r}, t)] \quad (7.90b)$$

L'équation sur la vitesse dit simplement que la vitesse de l'électron est égale à la vitesse de groupe du paquet d'ondes. On peut aussi définir des masses effectives de la façon suivante. Nous avons

$$\hbar \mathbf{k} = m \dot{\mathbf{v}} \quad (7.91)$$

On peut alors définir

$$m_{ij}^* = \hbar \frac{k_j}{v_i} = \hbar^2 \frac{1}{\frac{\partial^2 E(\mathbf{k})}{\partial k_i \partial k_j}} \quad (7.92)$$

On verra dans la suite que l'on peut définir cette matrice avec le signe opposé. Ce n'est cependant pas d'une importance capitale. Cette matrice est appelée *tenseur de masse effective*.

7.5.2 Bandes remplies

Les bandes d'énergie remplies sont celles qui se trouvent en-dessous du niveau de Fermi. Elles peuvent aussi être caractérisées par le fait que leur densité dans l'espace des vecteurs d'onde est égale à $\frac{d\mathbf{k}}{4\pi^3}$ puisque le facteur de Fermi est à peu près égal à l'unité pour ces bandes. Les équations semi-classiques impliquent qu'une bande pleine restera pleine même en présence de champs dépendant du temps. Ceci est conséquence directe du théorème de Liouville qui nous dit que le volume dans l'espace des phases est préservé par le flot semi-classique. En d'autres termes, la densité reste la même, d'où le résultat. La contribution au courant de ces bandes s'écrit alors

$$\mathbf{j} = -e \int \frac{1}{\hbar} \frac{\partial E_n(\mathbf{k})}{\partial \mathbf{k}} \frac{d\mathbf{k}}{4\pi^3} \quad (7.93)$$

L'intégrale d'une fonction périodique sur une cellule primitive étant nulle, les bandes pleines sont inertes du point de vue de la conduction.

D'un point de vue quantique, on peut comprendre facilement que les électrons d'une bande pleine ne peuvent participer à la conduction. Ceci est une conséquence du principe de Pauli. En effet, il est impossible d'augmenter la densité si un niveau contient le maximum d'électrons autorisé. De plus, si les transitions interbandes sont interdites, le nombre d'électrons d'une bande d'énergie ne peut être réduit. Les bandes pleines ne peuvent donc pas participer à la conduction.

Finalement, seuls les électrons appartenant à une bande partiellement remplie peuvent participer aux propriétés électroniques du solide. Cette conclusion justifie l'utilisation d'une densité d'électrons de conduction correspondant précisément à ces électrons. Dans le modèle de Drude, elle est souvent calculée à partir du nombre d'électrons de valence de chaque atome, qui appartiennent le plus souvent à une bande partiellement pleine.

⁴. Attention, les paramètres $E_n(\mathbf{k})$, \mathbf{k} et \mathbf{r} décrivent le même électron que $E_n(\mathbf{k})$, $\mathbf{k} + \mathbf{K}$ et \mathbf{r} où \mathbf{K} est un vecteur du réseau réciproque.



7.5.3 Les électrons et les trous

Un des succès du modèle semi-classique est d'expliquer pourquoi certaines propriétés ne peuvent être expliquées que si les porteurs de charges sont chargés positivement (signe du coefficient de Hall par exemple). La contribution au courant des électrons d'une bande partiellement remplie s'écrit

$$\mathbf{j} = -e \int_O \mathbf{v}(\mathbf{k}) \frac{d\mathbf{k}}{4\pi^3} \quad (7.94)$$

L'intégrale porte sur l'ensemble des niveaux occupés, noté ici O . Or, nous avons vu qu'une bande pleine ne contribue pas au courant donc

$$0 = e \int_O \mathbf{v}(\mathbf{k}) \frac{d\mathbf{k}}{4\pi^3} + e \int_E \mathbf{v}(\mathbf{k}) \frac{d\mathbf{k}}{4\pi^3} \quad (7.95)$$

où cette fois E désigne l'ensemble des niveaux de la bande non occupés. Le courant \mathbf{j} peut alors s'écrire

$$\mathbf{j} = e \int_E \mathbf{v}(\mathbf{k}) \frac{d\mathbf{k}}{4\pi^3} \quad (7.96)$$

Ainsi, le courant créé par des électrons occupant certains niveaux de la bande est équivalent au courant créé par des charges positives occupant les anciens niveaux vides tandis que les niveaux auparavant remplis par les électrons sont maintenant libres. Même si physiquement le courant est créé par les électrons, on peut voir les choses sous un autre angle. Ces particules fictives sont appelées des *trous*. Pour une bande d'énergie, on peut soit adopter le point de vue électronique soit le point de vue des trous. Il ne faut cependant pas mélanger les points de vue à l'intérieur d'une seule bande mais il est tout à fait légal de le faire pour différentes bandes.

Quelle est la dynamique des trous? En fait, l'évolution temporelle des niveaux inoccupés sous l'influence de champs est la même que s'ils étaient occupés par des électrons. Comment justifier le point de vue des trous? Pour cela, il faut connaître le lien entre l'accélération des électrons et le vecteur \mathbf{k} . Si les deux sont de sens opposé, les électrons se comporteront alors comme une particule de charge positive. Il s'avère que c'est souvent le cas pour un niveau inoccupé. En effet, à l'équilibre, et pas loin de l'équilibre, les niveaux non remplis d'une bande sont les niveaux d'énergie élevée. Si l'on note \mathbf{k}_0 le vecteur d'onde du maximum d'énergie, un développement autour du maximum donne

$$E(\mathbf{k}) = E(\mathbf{k}_0) - \frac{\partial^2 E(\mathbf{k}_0)}{\partial k_i \partial k_j} (\mathbf{k} - \mathbf{k}_0)_i (\mathbf{k} - \mathbf{k}_0)_j \quad (7.97)$$

Considérons le cas simple où le dernier terme est de plus symétrique, tel que

$$E(\mathbf{k}) = E(\mathbf{k}_0) - A(\mathbf{k} - \mathbf{k}_0)^2 \quad (7.98)$$

où A est une constante positive puisque l'on développe autour d'un maximum. La masse effective que nous avons définie au 7.92 vaut alors

$$m^* = -\frac{\hbar^2}{2A} \quad (7.99)$$

L'accélération, dérivée temporelle de la vitesse, est alors de sens opposé au vecteur d'onde \mathbf{k} . On peut aussi le voir à travers les équations semi-classique où la vitesse est directement reliée à la dérivée de l'énergie qui ici, autour du maximum, est négative et proportionnelle au vecteur d'onde. D'après l'équation de la dynamique 7.90a, l'électron se comporte comme s'il avait une masse négative. Il est en fait plus commode de penser que la dynamique est décrit par des particules de charge positive et de masse positive $|m^*|$. Cette masse est appelée *masse effective du trou*. Comme nous avons supposé que la dynamique d'un trou était la même que celle d'un électron présent dans le même niveau, nous avons bien répondu à la question posée.



7.6 Énergie de cohésion

L'énergie de cohésion est l'énergie nécessaire pour séparer un système condensé en ses constituants : atomes, ions ou molécules indépendants. On se limitera à une étude à température nulle. Cette énergie permet de déterminer la structure des cristaux mais aussi certaines propriétés thermodynamiques comme les constantes élastiques... L'énergie de point zéro est en général négligée bien que son influence soit parfois déterminante, comme, par exemple, pour les gaz nobles légers.

Les interactions interatomiques se hiérarchisent comme suit. Tout d'abord, nous avons les cristaux de Van der Waals dont la cohésion est assurée par les fluctuations dipolaires à courtes distances. Ensuite viennent les cristaux ioniques qui existent grâce à l'interaction électrostatique. Enfin, les cristaux covalents et métalliques sont ceux où les interactions sont les plus fortes mais aussi les plus compliquées.

7.6.1 Cristaux moléculaires

Les cristaux moléculaires ou cristaux de Van der Waals sont liés par une interaction dipolaire. Elle provient de la distorsion du nuage électronique qui crée un faible potentiel attractif de nature dipolaire. À très courte distance, la répulsion de l'ion rentre en jeu et domine l'interaction, qui devient répulsive. Le terme dipolaire est un terme en puissance six. La répulsion est en général représentée par un terme lui aussi en puissance, bien que cela ne soit pas un obligation. Le *potentiel de Lennard Jones* est le plus souvent utilisé et vaut

$$V(\mathbf{r}) = -4\epsilon \left(\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right) \quad (7.100)$$

La forme précise du potentiel n'est pas importante, il faut juste se rappeler la signification des différents termes. Par exemple, une version où le terme répulsif est une exponentielle existe : c'est le potentiel de Morse. Cela étant, l'énergie par particule du cristal u s'écrit alors

$$u = \frac{1}{2} \sum_{\mathbf{R} \in \mathcal{R}} V(\mathbf{R}) \quad (7.101)$$

Nous pouvons écrire cette énergie d'une manière plus simple en fonction de paramètres ne dépendant que de la structure géométrique du cristal

$$u = -2\epsilon \left(A_{12} \left(\frac{\sigma}{r} \right)^{12} - A_6 \left(\frac{\sigma}{r} \right)^6 \right) \quad (7.102)$$

où les constantes A_{12} et A_6 ne dépendent que de la géométrie cristalline. Nous pouvons alors déterminer certaines propriétés du solide. Par exemple, la distance entre plus proches voisins est donnée par

$$\frac{\partial u}{\partial r} = 0 \Leftrightarrow r_0 = \left(\frac{2A_{12}}{A_6} \right)^{1/6} \sigma \quad (7.103)$$

L'énergie de cohésion à l'équilibre est finalement

$$u(r_0) = -\frac{\epsilon A_6^2}{2A_{12}} \quad (7.104)$$

L'accord avec l'expérience est plutôt bon, même si la contribution de l'énergie de point zéro améliore les résultats. Nous pouvons de même déterminer le module de compressibilité B qui est donné à température nulle par

$$B = -V \frac{\partial P}{\partial V} = v \frac{\partial}{\partial v} \left(\frac{\partial u}{\partial v} \right) \quad (7.105)$$

7.6.2 Cristaux ioniques

Cette fois c'est l'interaction coulombienne qui maintient le réseau. Bien que la force de Van der Waals existe toujours, sa contribution peut être négligée. Cette contribution coulombienne statique est appelée énergie de Madelung. Pour un ion, la contribution à la cohésion est donnée par



$$U_i = \sum_{i \neq j} U_{ij}, \quad U_{ij} = \lambda e^{-\frac{r_{ij}}{\rho}} \pm \frac{q^2}{4\pi\epsilon_0 r_{ij}} \quad (7.106)$$

Le terme en exponentielle est un terme répulsif dont l'origine est fondée sur le principe de Pauli et est important pour les premiers voisins. En notant $r_{ij} = \rho_{ij}R$ où R est la distance interatomique, nous avons approximativement

$$U_i = \begin{cases} \lambda e^{-\frac{R}{\rho}} - \frac{q^2}{4\pi\epsilon_0 R} & \text{pour les premiers voisins} \\ \pm \frac{q^2}{4\pi\epsilon_0 \rho_{ij} R} & \text{pour les autres} \end{cases} \quad (7.107)$$

On considère un système avec N paires d'ions et Z premiers voisins. Finalement, l'énergie de cohésion vaut

$$U = NU_i = N \sum_{i \neq j} U_{ij} = N \left(\frac{Z}{N} e^{-\frac{R}{\rho}} - \frac{\alpha q^2}{4\pi\epsilon_0 R} \right) \quad (7.108)$$

où α est la constante de Madelung et vaut $\alpha = \sum_{i \neq j} \frac{\pm}{\rho_{ij}}$. Le cristal n'est stable que si α est positive. Pour un même type de charge, cette constante ne dépend que de la structure du cristal. En trois dimensions, le problème souvent rencontré est celui des séries divergentes. L'idée est alors de ne considérer que des sommes finies avec une charge de surface nulle. Ce type de sommes est appelé somme d'Ewald.

La plupart du temps, le terme répulsif exponentiel est changé par une contribution en puissance de la forme $\frac{\lambda}{r^n}$. Il est alors plus facile de déterminer les caractéristiques du réseau. La distance d'équilibre r_0 est donnée par

$$\frac{\partial U}{\partial R} = 0 \Leftrightarrow \lambda = \frac{\alpha q^2 R_0^{n-1}}{4\pi\epsilon_0 n} \quad (7.109)$$

Pour ce type de cristaux, il est en effet plus facile de déterminer la constante λ à partir d'une mesure expérimentale de la distance d'équilibre que de faire l'inverse. Plusieurs points peuvent être améliorés dans ce modèle. On peut d'abord bien prendre en compte la répulsion exponentielle, rajouter la contribution de Van der Waals ou encore compter l'énergie de point zéro. Ces améliorations n'altèrent cependant pas les conclusions, à savoir que la cohésion est principalement due à l'interaction de Coulomb entre les ions considérés comme des points fixes du réseau.

7.6.3 Cristaux covalents et métaux

La difficulté pour ce type de structure provient du fait qu'il faut à tout prix tenir compte de la structure en bandes. Précédemment cela n'était pas nécessaire puisque la structure électronique de ces atomes en réseau n'est pas très modifiée par rapport à celle des atomes isolés. Ce n'est plus le cas pour les métaux ou les cristaux covalents. Il faut absolument voir les électrons de conduction en interaction avec le réseau d'ions.

Pour obtenir la structure de bande, nous pouvons utiliser les différentes méthodes exposées dans les sections précédentes. Une méthode générale, qui de plus est l'une des plus utilisées actuellement est celle de la fonctionnelle de densité. On s'intéresse à la densité de probabilité électronique $n(\mathbf{r})$ qui vérifie

$$\lim_{\mathbf{r} \rightarrow \infty} n(\mathbf{r}) = 0 \quad (7.110)$$

$$\int n(\mathbf{r}) d\mathbf{r} = N \quad (7.111)$$

où N est le nombre d'électrons dans le système. Il existe un théorème qui permet d'affirmer que l'énergie est une fonctionnelle de la densité et que l'énergie de l'état fondamental est la valeur qui minimise cette fonctionnelle. La densité n_{eq} réalisant ce minimum correspond bien à la densité électronique du fondamental. La fonctionnelle d'énergie à la forme suivante

$$E[n(\mathbf{r})] = T[n(\mathbf{r})] + \int n(\mathbf{r}) \phi_N(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int n(\mathbf{r}) \phi_H(\mathbf{r}) d\mathbf{r} + E_{EC}[n(\mathbf{r})] \quad (7.112)$$

Le premier terme correspond à l'énergie cinétique. Le second correspond à l'énergie d'interaction élastique avec le réseau d'ions $\phi_N(\mathbf{r}) \sim_{\text{ions}} \frac{q_i}{4\pi\epsilon_0 |\mathbf{r} - (\mathbf{R})_i|}$ tandis que le troisième est l'énergie de Hartree, énergie issue de l'interaction électron-électron prise comme un champ moyen avec $\phi_H(\mathbf{r}) = \int \frac{n(\mathbf{r}')}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'$. Enfin, le dernier est un terme d'échange et de corrélations.

Chapitre 8

Transition de phases

Sommaire

8.1 Généralités sur les transitions de phases	195
8.2 Le modèle d'Ising	196
8.3 Théorie de Landau	201
8.4 Groupe de renormalisation	206

Le but de ce chapitre sera de traiter les changements de phases pouvant apparaître dans un système en se focalisant en particulier sur un modèle type de physique statistique, le modèle d'Ising. Ce modèle sera présenté au début du chapitre, et est un bon moyen d'étudier une transition de phase donnée, puisqu'il permet d'introduire des notions tels que le paramètre d'ordre. Nous verrons plusieurs moyens d'étudier les transitions de phases dans le modèle d'Ising, par l'approximation de champ moyen, puis par la théorie de Landau et enfin par les groupes de renormalisation.

Pour ce chapitre, les références utilisées sont [15], [4], [3] et [2].

Les parties traitant des gaz et liquides ont été traitées en physique statistique et ne seront pas refaites ici. Le point le plus important de ce chapitre est certainement la théorie de Landau. L'introduction peut paraître rugueuse au premier abord mais tente de résumer l'essentiel des notions nouvelles abordées dans l'étude moderne des transitions de phases.

FIGURE 8.1 – Opalescence critique : phénomène optique se produisant à la transition au point critique liquide/gaz (expérience réalisée avec la main experte d'Arnaud le Diffon).

8.1 Généralités sur les transitions de phases

En physique, il est assez commun de s'intéresser à des systèmes dont toutes les propriétés sont homogènes dans une certaine région de l'espace (par exemple, une casserole d'eau liquide est un système physique dans lequel la densité, la température ou la chaleur spécifique sont homogènes). Ce type de région où les propriétés d'un système sont essentiellement homogènes est ce que l'on appellera communément une *phase*. Il pourra selon les cas s'agir d'un état thermodynamique différent, par exemple les phases liquides, solides ou gazeuses d'un matériau quelconque, ou d'autres propriétés tels que des matériaux dans une phase ferromagnétique, des bosons condensés dans un condensat de Bose-Einstein, les phases supraconductrices de certains conducteurs, etc... De manière générale, plus que l'étude de ces phases, c'est l'étude des transformations du système qui nous intéresse, c'est-à-dire le passage d'une phase à une autre. Plusieurs observations expérimentales nous montrent que ces transformations (les *transitions de phase*) peuvent être de plusieurs types. En effet, certaines transitions impliquent une enthalpie de changement d'état, et le système absorbe ou émet une certaine quantité d'énergie lorsqu'il change d'état. Ces transitions sont appelées transitions de phase du *premier ordre*, et sont les plus communément rencontrées dans la vie de tous les jours (transitions de phase solide/liquide/vapeur dans les conditions habituelles par exemple). Mais toutes les transitions ne sont pas de ce type : les transitions de phase du *deuxième ordre*, autrement appelées phénomènes critiques, n'ont pas d'enthalpie associée au changement d'état. C'est notamment le cas de la transition liquide/vapeur au voisinage du point critique : un phénomène appelé opalescence critique a alors lieu, les phases liquides et gazeuses ne peuvent plus vraiment être distinguées, et la transition s'effectue de manière instantanée et homogène (contrairement au cas classique d'un liquide porté à ébullition). Cette transition bien particulière ainsi que le phénomène d'opalescence sont visibles dans la vidéo 8.1 située au début de ce chapitre. Cette classification historique souffre de quelques limites que l'on discutera.

L'origine physique d'une transition provient de la compétition entropie/ énergie, où l'on cherche à maximiser la première et minimiser la seconde. Pour pouvoir caractériser une transition, on utilise la notion de *paramètre d'ordre* qui est une variable prenant une valeur que l'on choisit nulle dans une phase et non nulle dans l'autre. Comme mentionné ci-dessus, on distingue différents types de transitions que l'on qualifie de premier ordre (parfois discontinue) et second ordre (ou ordre supérieur, continue...) que l'on relie à la régularité du paramètre d'ordre à la transition. Ce fut Landau qui le premier a introduit cette notion dans un cadre un peu plus restreint. En effet, il remarqua que la plupart des transitions du second ordre étaient accompagnées d'une *brisure spontanée de symétrie*. Cette idée, dont l'importance en physique dépasse le cadre de la physique statistique et est utilisée par exemple dans le Modèle Standard de la physique des particules, est en fait assez simple. On dit qu'il y a brisure spontanée de symétrie lorsque les symétries de l'état fondamental du système sont différentes (moindres) que les symétries de la dynamique. Bien que fructueuse, toutes les transitions du second ordre ne s'accompagnent pas toujours d'une brisure de symétrie (transition liquide/gaz au point critique par exemple) d'où la classification moderne donnée auparavant.

Dans toute la suite, on ne s'intéressera qu'à la physique des transitions du second ordre. Le prototype des transitions du premier ordre est la transition liquide/gaz de l'eau dont on connaît bien le diagramme des phases, la présence de chaleur latente... Le modèle d'Ising que l'on étudie dans la suite capture de nombreuses propriétés typiques des transitions du second ordre et permet d'illustrer simplement de nombreuses méthodes d'analyses. La physique des phénomènes critiques est remarquable sous plusieurs aspects. De nombreux systèmes physiques, dont la dynamique exacte diffère complètement, ont exactement les mêmes comportements à la transition. Expérimentalement, on observe que les grandeurs thermodynamiques, telle que la capacité thermique, ont un comportement algébrique en la température autour de la transition. Ce sont ces exposants appelés *exposants critiques* qui sont identiques pour de nombreux systèmes. On dit que l'ensemble des systèmes ayant les mêmes exposants critiques forme une *classe d'universalité*. Cette notion d'universalité est à la base des concepts de renormalisation et groupe de renormalisation que l'on abordera succinctement à la fin de ce chapitre. Ces méthodes permettent de construire à partir d'un modèle microscopique un modèle du comportement du système à la transition que l'on dénomme *théorie effective*. Parmi toutes les observables que l'on peut considérer, les fonctions de corrélations du paramètre d'ordre sont au premier plan de l'étude des phénomènes critiques. Leur comportement en fonction de la distance est caractérisé par *une longueur de corrélation*. A la transition, cette longueur caractéristique diverge, signe qu'il n'existe aucune taille typique, en particulier pour les fluctuations du paramètre d'ordre, au régime critique. On parle alors d'*invariance d'échelle* et cela est à l'origine de relations entre les exposants critiques appelées *relations d'échelles* qui elles aussi sont universelles. Ce sont ces fluctuations à toutes les échelles qui expliquent le phénomène d'opalescence critique : au point critique, les fluctuations de densité (densité qui est le paramètre d'ordre ici) engendrent des fluctuations énorme de l'indice optique du fluide qui le rendent alors opaque.

8.2 Le modèle d'Ising

8.2.1 Transition ferromagnétique : approche phénoménologique

Certaines substances ont la propriétés de conserver une aimantation à température pas trop élevée, même en l'absence de champ magnétique. C'est ce qui est appelé *aimantation spontanée*. Cette aimantation n'est observable qu'à une température plus faible qu'une *température critique*, T_C , aussi appelée température de Curie. Au-dessus de cette température, l'aimantation disparaît ; le corps devient *paramagnétique*. L'aimantation d'un tel corps en fonction de la température est tracée sur la figure 8.2.

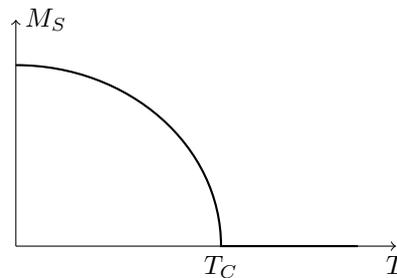


FIGURE 8.2 – Aimantation spontanée en fonction de la température. La transition ferromagnétique-paramagnétique se fait à la température critique T_C .

On peut comprendre schématiquement cette transition en regardant les spins des atomes constituant le matériau. À température nulle, du fait de leur interaction mutuelle, tous les spins sont alignés dans une même direction. À mesure que la température augmente, l'agitation thermique fait fluctuer l'alignement de chaque spin mais, globalement, une direction est toujours privilégiée. Cependant, il se trouve qu'à une température suffisante, la température de Curie, l'agitation est telle que tous les spins sont orientés aléatoirement, ce qui en moyenne donne une aimantation nulle. Ceci est visible sur la figure 8.3.

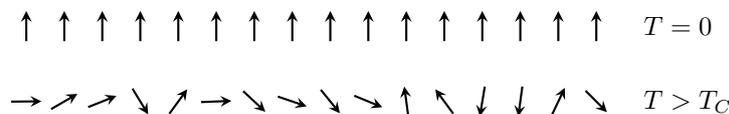


FIGURE 8.3 – Représentation de l'orientation des spins dans un matériau présentant une transition ferromagnétique-paramagnétique. Pour $T > T_C$, l'agitation thermique est telle qu'en moyenne l'aimantation est nulle.

Dans la phase désordonnée, aucune direction n'est privilégiée, alors que cela est le cas en-dessous de la température de Curie. Ce phénomène est appelée *brisure spontanée de symétrie*. Ici, l'aimantation, qui est nulle dans la phase désordonnée et différente de zéro à basse température, est appelé *paramètre d'ordre*.

La physique derrière la magnétisme dans la matière est en fait beaucoup plus riche et subtile. L'origine de l'ordre magnétique est en fait à chercher dans l'interaction de Coulomb entre les électrons et le principe de Pauli et non dans l'interaction magnétique qui est beaucoup trop faible. Ces deux aspects donnent lieu à une interaction d'échange qui peut être décrite de manière effective par le Hamiltonien de Heisenberg que l'on donnera dans la suite. Bien sûr, il faut encore un peu plus travailler pour comprendre le magnétisme des aimants que l'on rencontre dans la vie quotidienne (notion de domaine de Weiss, hystérésis...) mais ce n'est pas le sujet ici.

8.2.2 Description du modèle d'Ising

Il semble que le point le plus important dans la transition ferromagnétique soit l'interaction entre les spins du matériau, qui tend à les aligner dans une même direction. On considère donc N spins (par exemple des électrons) disposés sur un réseau (cubique pour simplifier) en dimension d . On suppose l'interaction à courte portée, c'est-à-dire que chaque électron interagit avec ses plus proches voisins. Sans champ magnétique extérieur, le hamiltonien que l'on peut imaginer remplissant ces prérequis est

$$H = -J \sum_{(i,j)} \sigma_i \cdot \sigma_j \quad (8.1)$$

où J est la *constante de couplage* et $\sigma_i = \boldsymbol{\sigma} \cdot \mathbf{n}_i$, avec \mathbf{n}_i l'orientation du spin i et $\boldsymbol{\sigma}$ le vecteur constitué des matrices de Pauli. La sommation est faite sur les plus proches voisins. Cet hamiltonien est celui du modèle de Heisenberg quantique. Le modèle de Heisenberg classique consiste à remplacer les matrices σ_i par de simples vecteurs; \mathbf{S}_i , représentant les spins.

Le modèle d'Ising consiste à faire une approximation supplémentaire : chaque vecteur \mathbf{S}_i est remplacé par un nombre S_i pouvant prendre deux valeurs : $+1$ ou -1 . En d'autres termes, on considère que les spins sont alignés dans une direction fixe et ont deux orientations possibles. L'hamiltonien correspondant est, en l'absence de champ extérieur,

$$H = -J \sum_{\langle i,j \rangle} S_i S_j \tag{8.2}$$

Pour pouvoir calculer les grandeurs thermodynamiques associées au système, il faut connaître la fonction de partition

$$Z = \sum_{\{S_i = \pm 1\}} e^{\beta J \sum_{\langle i,j \rangle} S_i S_j} \tag{8.3}$$

Cette somme contient 2^N termes. Toute la difficulté est de calculer cette fonction de partition.

Modèle d'Ising

Soit un réseau de dimension d . Un spin S_i est présent en chaque nœud du réseau. Tous les N spins sont alignés dans la même direction et ne peuvent prendre que deux orientations possibles $+1$ ou -1 . Chaque spin est en interaction avec ses plus proches voisins, cette interaction étant caractérisée par une constante de couplage J . En présence d'un champ extérieur B^a de même direction que les spins, le hamiltonien d'Ising est

$$H = -J \sum_{\langle i,j \rangle} S_i S_j - B\mu \sum_i S_i \tag{8.4}$$

La fonction de partition du modèle est finalement

$$Z = \sum_{\{S_i = \pm 1\}} e^{\beta (J \sum_{\langle i,j \rangle} S_i S_j + B\mu \sum_i S_i)} \tag{8.5}$$

a. On aurait pu prendre des constantes de couplage dépendant des spins en interaction J_{ij} , de même pour le champ magnétique B_i . Le modèle d'Ising que l'on considère est un modèle homogène et à courte portée.

8.2.3 Modèle à une dimension

Dans le cas à une dimension que nous allons considérer, les résultats attendus sont présentés sur la figure 8.4. Réécrivons d'abord la fonction de partition comme

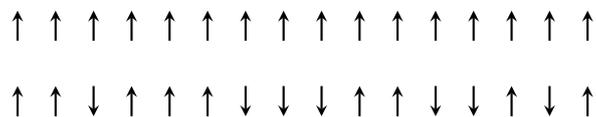


FIGURE 8.4 – Transition de phases attendue dans le modèle d'Ising unidimensionnel. Les spins ne peuvent prendre que deux valeurs : $+1$ et -1 .

$$\begin{aligned} Z &= \sum_{\{S_i = \pm 1\}} e^{\beta (J \sum_i S_i S_{i+1} + B\mu \sum_i S_i)} \\ &= \sum_{\{S_i = \pm 1\}} \prod_{i=1}^N e^{\beta (J S_i S_{i+1} + B\mu S_i)} \end{aligned} \tag{8.6}$$

Supposons d'abord le champ nul, $B = 0$. Une manière simple de résoudre est de constater que $S_i S_j = \pm 1$. En posant

$$\eta_i = \begin{cases} 1 & \text{si } S_i = S_{i+1} \\ -1 & \text{si } S_i = -S_{i+1} \end{cases} \quad (8.7)$$

la fonction de partition s'écrit simplement ¹

$$\begin{aligned} Z &= \sum_{\{\eta_i = \pm 1\}} \prod_{i=1}^N e^{\beta(J\eta_i)} \\ &= \prod_{i=1}^N \sum_{\eta_i = \pm 1} e^{\beta(J\eta_i)} \\ &= (2 \cosh(\beta J))^N \end{aligned} \quad (8.8)$$

Or, l'aimantation vaut

$$\begin{aligned} M &= \sum_{k=1}^N \langle S_k \rangle \\ &= \frac{1}{\beta \mu} \frac{\partial \ln(Z)}{\partial B} \\ &= 0 \end{aligned} \quad (8.9)$$

On voit donc que le modèle ne prédit pas d'aimantation spontanée en champ nul, puisque $M = 0$. Revenons au cas avec champ magnétique. Cette fois la fonction de partition est donnée par 8.6. On va résoudre le problème par la méthode de la matrice de transfert. On pose $t(S_i, S_{i+1}) = e^{\beta(JS_i S_{i+1} + B\mu S_i)}$. On peut voir $t(S_i, S_{i+1})$ comme l'élément d'une matrice, c'est-à-dire $t(S_i, S_{i+1}) = \langle S_i | T | S_{i+1} \rangle$ où T est appelé matrice de transfert. Comme nous avons $\sum_{S_i} |S_i\rangle \langle S_i| = \mathbb{1}$, la fonction de partition vaut

$$\begin{aligned} Z &= \sum_{\{S_i = \pm 1\}} \prod_{i=1}^N \langle S_i | T | S_{i+1} \rangle \\ &= \sum_{S_1 = \pm 1} \langle S_1 | T^N | S_1 \rangle \\ &= \text{Tr}(T^N) \end{aligned} \quad (8.10)$$

En déterminant les valeurs propres de cette matrice, on trouve la fonction de partition. Même en faisant tendre le champ vers zéro, le modèle d'Ising à une dimension ne prévoit pas d'aimantation spontanée ². Avec le modèle d'Ising en deux dimensions, on peut retrouver une aimantation spontanée. *Cependant, avec notre démarche, aucun modèle ne peut donner de transition de phase!*

8.2.4 La limite thermodynamique

Pour un système fini, la fonction de partition est une somme finie de fonctions analytiques de la température. Toutes les grandeurs thermodynamiques déterminées à partir de la fonction de partition, comme l'énergie libre, seront donc elles aussi analytiques. Or, une transition de phase correspond à une discontinuité dans une grandeur thermodynamique. Pour un système fini, aucune transition de phase n'est donc possible. Il faut donc analyser *la limite thermodynamique* pour trouver les transitions de phase. En effet, dans ce cas, rien ne garantit mathématiquement que la limite ait les mêmes propriétés de régularité que l'argument initial.

Pour pouvoir étudier une transition de phase, il est nécessaire en premier lieu de prendre la limite thermodynamique, $N \rightarrow \infty$.

1. On prend des conditions aux limites périodiques.

2. En fait, un système unidimensionnel ne peut pas présenter de transition de phase en l'absence d'interaction à longue portée.



8.2.5 Notion de paramètre d'ordre et classification

Une *phase* est une région de l'espace où toutes les propriétés physiques ou chimiques d'un matériau sont essentiellement homogènes. Une *transition de phase* est le passage d'une phase à une autre, ici la transition ferromagnétique-paramagnétique. Pour distinguer une transition, on utilise une grandeur appelée *paramètre d'ordre* qui, dans une des phases, est nulle alors qu'elle ne l'est pas dans l'autre (par exemple l'aimantation). Le plus souvent, l'apparition d'une valeur non nulle du paramètre d'ordre est le reflet d'une *brisure de symétrie*. La classification de Landau distingue deux types de transitions. On appelle une *transition du second ordre* une transition où le paramètre d'ordre est continu à la transition (mais pas nécessairement dérivable). On appelle *transition du premier ordre* une transition où le paramètre d'ordre n'est même pas continu.

Au cours de l'étude historique des transitions de phases, plusieurs classifications furent données. L'une distinguait la présence d'une chaleur latente (discontinuité de l'entropie) ou pas lors de la transition et était respectivement appelée de premier ou second ordre. Une limite à cette classification provient qu'il existe des transitions du premier ordre, c'est-à-dire clairement avec une discontinuité d'une grandeur thermodynamique, mais pas de chaleur latente (entropie continue). Une telle transition se rencontre pour les corps ferromagnétiques en présence d'un champ magnétique que l'on change de signe : l'entropie est continue mais la magnétisation est discontinue. Ehrenfest proposa une classification suivant la régularité d'une fonction thermodynamique et l'ordre de ces singularités. Ceci est encore une fois limité car la structure des singularités lors de certaines transitions est complexe. Landau proposa sa classification à partir des changements de symétrie et des brisures spontanées de symétrie. Malheureusement, il existe encore des transitions continues sans changeant de symétrie. Finalement, l'approche moderne exposée précédemment est un compromis entre toutes ces différentes idées.

8.2.6 Fonction de corrélation

La *fonction de corrélation* G_{ij} de deux spins S_i et S_j nous renseigne sur l'influence d'un spin sur un autre. Par définition, elle est égale à

$$G_{ij} = \langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle \quad (8.11)$$

Les moyennes se calculent de manière statistique

$$\begin{aligned} \langle S_i S_j \rangle &= \frac{1}{Z} \sum_{\{S_i\}} S_i S_j e^{-\beta H(\{S_i\})} \\ \langle S_i \rangle &= \frac{1}{Z} \sum_{\{S_i\}} S_i e^{-\beta H(\{S_i\})} \end{aligned} \quad (8.12)$$

que nous pouvons relier aux dérivées de l'énergie libre ou à la fonction de partition³ par

$$\langle S_i \rangle = \frac{1}{\beta} \frac{\partial \ln(Z)}{\partial B_i} = - \frac{\partial F}{\partial B_i} \quad (8.13)$$

La fonction de corrélation peut elle-même s'exprimer comme une dérivée de l'énergie libre par

$$G_{ij} = \frac{1}{\beta^2} \frac{\partial^2 \ln(Z)}{\partial B_j \partial B_i} = - \frac{\partial^2 F}{\partial B_j \partial B_i} \quad (8.14)$$

Dans le cas du modèle d'Ising à une dimension, on obtient

$$G_{ij} = e^{-\frac{r_{ij}}{\xi}} \quad (8.15)$$

où, si on note a le pas du réseau, $r_{ij} = a|i - j|$ et ξ est la *longueur de corrélation*. Ici

$$\xi = \frac{a}{|\ln \tanh \beta J|} \quad (8.16)$$

On constate que dans le cadre du modèle d'Ising à une dimension, la longueur de corrélation décroît quand la température augmente. Elle tend même vers zéro quand la température tend vers l'infini.

³. Par commodité, le facteur gyromagnétique μ est intégré à la valeur B ; B représente donc une énergie et non plus seulement le champ magnétique.

8.2.7 Approximation du champ moyen

La difficulté du modèle vient du terme d'interaction à deux corps. Dans le cas particulier d'une substance paramagnétique, c'est-à-dire $J = 0$, la fonction de partition est calculable facilement. On retrouve un résultat similaire aux précédents

$$Z = \sum_{\{S_i = \pm 1\}} \prod_{i=1}^N e^{\beta B S_i} = \prod_{i=1}^N \sum_{S_i = \pm 1} e^{\beta B S_i} = (2 \cosh(\beta B))^N \quad (8.17)$$

L'aimantation moyenne par site vaut $m = \langle S_i \rangle = \tanh(\beta B)$. Tous les spins étant équivalents, l'aimantation M vaut donc

$$M = \sum_{i=1}^N \langle S_i \rangle = N \tanh(\beta B) \quad (8.18)$$

L'*approximation du champ moyen* (Weiss 1907) consiste à considérer que chaque spin S_i ne ressent que la valeur moyenne $\langle S_j \rangle$ des autres spins du réseau. L'hamiltonien du champ moyen devient alors,

$$H_{\text{Champ moyen}} = \sum_i \left(-J \sum_{\langle j \rangle_i} \langle S_j \rangle - B \right) S_i = \sum_i (-2dJm - B) S_i \quad (8.19)$$

où d est la dimension du réseau, chaque spin ayant alors $2d$ voisins et $m = \langle S_i \rangle$ est l'aimantation moyenne par site. La relation avec le hamiltonien complet d'Ising

$$H = H_{\text{Champ moyen}} - J \sum_{\langle i, j \rangle} S_i (S_j - \langle S_j \rangle) \quad (8.20)$$

montre que l'approximation du champ moyen revient à négliger les fluctuations du spin autour de sa valeur moyenne. On constate que nous avons un hamiltonien similaire au hamiltonien paramagnétique. L'aimantation moyenne par site vérifie alors l'équation

$$m = \tanh(\beta(2dJm + B)) \quad (8.21)$$

Cette équation s'appelle *équation d'auto-cohérence*. Cette équation est une équation transcendante et ne peut être résolue que graphiquement. C'est ce qui est fait sur la figure 8.5. On voit alors que dans cette approximation, il existe une aimantation spontanée du milieu en champ nul : on a bien une transition de phase. La température critique donnée par le modèle est obtenue par $\beta 2dJ = 1$, et vaut

$$T_c = \frac{2dJ}{k_B} \quad (8.22)$$

L'approximation du champ moyen permet donc de prédire une transition de phase. Cependant, on remarque que l'apparition d'une transition ne dépend pas de la dimension du système étudié. Or, nous avons clairement vu que pour le modèle d'Ising unidimensionnel, la solution exacte montre qu'il n'existe pas de transition de phase : l'approximation du champ moyen échoue ici totalement ! Dans le cas bidimensionnel, la solution exacte d'Onsager montre bien l'existence d'une transition de phase. Cependant, la température critique exacte est plus faible que celle que nous avons trouvée. Cela peut se comprendre facilement en repensant que l'approximation néglige les fluctuations qui s'opposent à l'apparition d'une transition. En fait, plus la dimension est élevée, plus les résultats concordent avec la solution exacte.

Voyons maintenant ce que prédit le modèle. Pour cela, on se place au voisinage de la température critique à faible champ magnétique. Dans ce cas, $m \ll 1$ et nous pouvons utiliser le développement de \tanh^{-1} : $\tanh^{-1}(m) = m + \frac{1}{3}m^3 + O(m^5)$. Pour ce qui est de la dépendance de l'aimantation par site en fonction de la température au voisinage de la température critique, le modèle du champ moyen prédit

$$m + \frac{1}{3}m^3 = \beta 2dJm \quad (8.23)$$



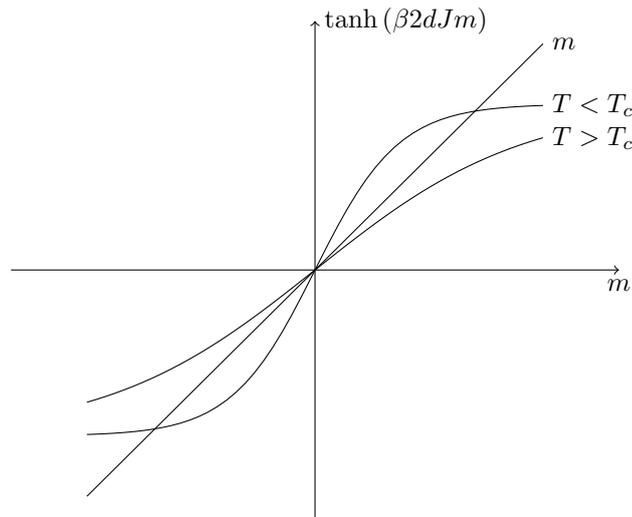


FIGURE 8.5 – Résolution graphique de l'équation d'auto-cohérence. A $T > T_c$, on a une seule valeur de m qui satisfait $m = \tanh(\beta 2dJm)$, $m = 0$. L'aimantation moyenne est donc nulle à haute température. Mais pour $T < T_c$, une seconde solution apparaît : c'est l'aimantation spontanée.

dont la solution est

$$m = \sqrt{3 \frac{(T_c - T)}{T_c}} \propto ((T_c - T))^{1/2} \quad (8.24)$$

On voit donc que l'aimantation suit une loi en puissance au voisinage de la température critique. Ce fait est en fait plus général : la chaleur spécifique, l'isotherme critique ou encore la susceptibilité suivent eux aussi une loi en puissance. Les exposants qui apparaissent sont appelés *exposants critiques*.

8.2.8 Exposants critiques

On constate expérimentalement que l'aimantation, la chaleur spécifique, la susceptibilité, l'isotherme critique ou la longueur de corrélation suivent tous une loi en puissance au voisinage de la température critique. On définit alors les *exposants critiques* α , β , γ , δ et ν par

$$C_v \propto |T - T_c|^{-\alpha} \quad (8.25a)$$

$$m \propto (T_c - T)^\beta \quad (8.25b)$$

$$\chi \propto |T - T_c|^{-\gamma} \quad (8.25c)$$

$$B \propto M^\delta \quad (8.25d)$$

$$\xi \propto |T - T_c|^{-\nu} \quad (8.25e)$$

Dans le cadre de l'approximation du champ moyen, nous avons $\alpha = 0$, $\beta = 1/2$, $\gamma = 1$, $\delta = 3$ et $\nu = 1/2$.

8.3 Théorie de Landau

Une transition de phase du second ordre est liée à un changement de symétrie. Pour caractériser les différentes phases, nous avons été amené à nous donner un paramètre d'ordre qui dans la plupart des cas est nul dans la phase désordonnée, qui est la phase possédant le plus de symétries, tandis qu'il est non nul dans la phase ordonnée qui est la phase possédant le moins de symétries. Dans la plupart des cas, mais pas tous, la phase désordonnée est celle qui existe aux températures les plus élevées. On rappelle que pour une transition de phase du second ordre, il y a continuité des grandeurs thermodynamiques, contrairement aux transitions du premier ordre.



On a vu précédemment pour le modèle d'Ising que, dans le cadre de l'approximation du champ moyen, le développement autour de la température d'équilibre de l'équation d'auto-cohérence donnait l'aimantation comme une solution d'un polynôme de degré trois. Cette expression, qui fait office d'équation d'état, peut être retrouvée de manière *ad hoc* en introduisant une énergie libre fonction du paramètre d'ordre η (qui est ici l'aimantation moyenne par site) $\Phi(T, P, \eta, \dots)$. La condition d'extrémalisation de cette énergie libre par rapport au paramètre d'ordre redonne l'équation d'état. On peut faire de même pour l'équation d'état de Van der Waals autour de la température de transition liquide-gaz.

L'idée de la théorie phénoménologique de Landau des transitions de phase est de postuler l'existence d'une énergie libre $\Phi(T, P, \eta, \dots)$, fonction du paramètre d'ordre η , développable en série autour du point de transition, et pour laquelle la condition d'extrémalisation par rapport au paramètre d'ordre redonne l'équation d'état.

8.3.1 Énergie libre de Landau

Soit un système subissant une transition de phase du second ordre. On fait l'hypothèse que cette transition est décrite par une énergie libre $\Phi(T, P, \eta, \dots)$ fonction du paramètre d'ordre η . On suppose développable en série l'énergie libre au voisinage de la transition.

Ces hypothèses ne tiennent pas compte du fait qu'en réalité, la transition de phase est caractérisée par une singularité de l'énergie libre. On suppose néanmoins que le développement est possible et nous en verrons une justification dans les sections suivantes. Cela étant, nous pouvons donc écrire au voisinage du point de transition

$$\Phi(T, P, \eta, \dots) = \Phi_0 + \alpha\eta + A\eta^2 + \beta\eta^3 + B\eta^4 + \dots \quad (8.26)$$

où les coefficients qui apparaissent sont fonctions de P et de T . Attention, β est ici un coefficient, et non la notation pour $\frac{1}{k_B T}$

Pour déterminer plus avant les coefficients, il faut discuter des symétries du système. En l'absence de champ, il n'y a aucune direction privilégiée, ce qui impose que tous les termes d'ordre impair soient nuls : nous devons avoir

$$\Phi(T, P, \eta) = \Phi(T, P, -\eta) \quad (8.27)$$

Concernant le coefficient $A(P, T)$, il doit être nul au point de transition. En effet, dans la phase désordonnée, nous avons à la fois $\eta = 0$ et η minimum de l'énergie libre donc dans cette phase $A(P, T) > 0$. En revanche, dans la phase ordonnée, nous avons cette fois $\eta \neq 0$. Pour qu'il corresponde en plus à un minimum de l'énergie libre, nous devons avoir dans cette phase $A(P, T) < 0$ mais aussi $B(P, T) > 0$ (sinon on aurait un maximum). C'est pourquoi à la transition $A_c(P, T) = 0$ où l'indice précise que nous sommes bien au point de transition. Cette condition nous donne de plus P et T l'un en fonction de l'autre : dans le plan (P, T) , la transition de phases est une courbe.⁴ Pour finir, dans la théorie de Landau, on suppose que le coefficient $A(P, T)$ est développable par rapport à la température. Au premier ordre, nous avons, compte tenu de la discussion précédente,

$$A(P, T) = a(P)(T - T_c) \quad (8.28)$$

où T_c désigne la température de transition ; tandis que l'on considère $B(P, T) = B(P, T_c) = B(P)$. On remarque que le signe de a fixe dans quel domaine de température existent les différentes phases. Généralement, la phase désordonnée existe à température élevée, nous supposons donc que $a(P) > 0$. Nous avons la forme finale de l'énergie libre de Landau

$$\Phi(T, P, \eta) = \Phi_0 + a(P)(T - T_c)\eta^2 + B(P)\eta^4, \quad B(P) > 0, \quad a(P) > 0 \quad (8.29)$$

Cette forme de l'énergie libre est visible sur la figure 8.6

8.3.2 Hamiltonien effectif de Ginzburg-Landau. Approximation de Landau

Nous pouvons voir que la théorie de Landau telle que nous l'avons formulée redonne les mêmes comportements que l'approximation de champ moyen, car elle a été construite sur la base de cette approximation. En fait, la

⁴ Si on avait le terme d'ordre trois non nul, nous aurions une autre condition à la transition $\beta_c(P, T) = 0$; les points de transition seraient alors des points isolés du plan (P, T) .

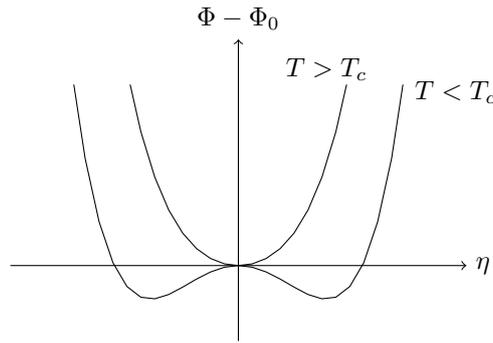


FIGURE 8.6 – Forme de l'énergie libre en fonction de la température. On voit qu'à $t < T_C$, on a apparition de deux nouveaux minima, et le système va aller dans un seul d'entre eux : on a une brisure spontanée de symétrie.

théorie de Landau elle aussi ne prend pas en compte les fluctuations du paramètre d'ordre. Ceci n'est pas très transparent dans la formulation donnée. De plus, on aimerait bien être capable de tenir compte des effets des fluctuations. Pour cela, nous allons reformuler et étendre la théorie de Landau. Dans ce qui précède, nous nous étions donné une énergie libre pour étudier la transition. Or, en mécanique statistique, l'énergie libre dérive de la fonction de partition qui, elle-même, est donnée par le hamiltonien du système. L'idée est donc de faire de même ici et de dériver la théorie de Landau par un procédé statistique. Nous devons donc trouver un hamiltonien⁵ effectif qui redonne,

sous une certaine approximation, l'énergie libre de Landau. De plus, grâce à cette formulation, nous pourrions enfin tenir compte des fluctuations.

Première approche

Considérons le cas d'un seul site dans le modèle d'Ising. Compte tenu de la forme de l'équation d'état, on considère le hamiltonien effectif suivant

$$H(\eta) = \frac{1}{2!}r_0\eta^2 + \frac{1}{4!}u_0\eta^4 \quad (8.30)$$

Dans la suite on rajoutera un terme $-\eta B$ lors de la présence d'un champ. La fonction de partition s'écrit alors

$$Z(\eta) = \int_{-\infty}^{+\infty} e^{-\beta[H(\eta)-B\eta]} d\eta \quad (8.31)$$

On omet toutes les constantes additives, puisqu'elles ne jouent aucun rôle dans le calcul des grandeurs thermodynamiques. Parfois, le facteur β est même omis et est intégré dans la définition de H puisque l'on ne s'intéresse qu'à une étude au voisinage du point critique. Pour les besoins de convergence, on doit avoir $u_0 > 0$.

Maintenant, on voit que l'intégrande possède soit un maximum soit deux suivant le signe de r_0 . De manière équivalente, ces maxima correspondent aux minima de $H(\eta) - B\eta$. Dans tous les cas, notons η_0 le minimum absolu.

L'approximation de Landau revient à considérer que la valeur de la fonction de partition est égale à la valeur maximale de l'intégrande,

$$Z(\eta) = \int_{-\infty}^{+\infty} e^{-\beta[H(\eta)-B\eta]} d\eta \approx e^{-\beta[H(\eta_0)-B\eta_0]} \quad (8.32)$$

Cette approximation est grossière puisqu'elle néglige la contribution du second minimum. Néanmoins, on la considère correcte dans la suite et on discutera son domaine de validité plus tard. On voit que cette approximation est bien du type champ moyen puisqu'on remplace η par sa valeur moyenne en négligeant les fluctuations. On en déduit l'énergie libre

$$F = -k_B T \ln Z = H(\eta_0) - B\eta_0 \quad (8.33)$$

⁵. Ce ne sera pas un hamiltonien au sens de la mécanique!

qui est bien l'énergie libre de Landau en présence d'un champ si on prend r_0 de la forme

$$r_0 = \bar{r}_0(T - T_c) \quad (8.34)$$

Formulation sur un réseau

On souhaite généraliser ces résultats sur un réseau à N sites, de pas a , en dimension d . Le paramètre d'ordre sera cette fois une fonction de la position $\eta(\mathbf{r})$. Comme pour le modèle d'Ising, nous devons maintenant considérer les interactions entre les différents sites. On considèrera une interaction entre premiers voisins de la forme

$$|\nabla\eta|^2 = \frac{1}{a^2} |\eta(\mathbf{r}_i) - \eta(\mathbf{r}_j)|^2 \quad (8.35)$$

qui représente un gradient discret. On postule alors une densité de hamiltonien que l'on écrit de la façon suivante

$$H[\eta] = a^d \sum_{i=1}^N \frac{c}{2} |\nabla\eta_i|^2 + \frac{1}{2!} r_0 \eta_i^2 + \frac{1}{4!} u_0 \eta_i^4 \quad (8.36)$$

où $\eta_i = \eta(\mathbf{r}_i)$ et c est une constante de "raideur". Cette fois H est une fonctionnelle puisqu'elle a pour argument la fonction η . Cet hamiltonien effectif étant donné, nous pouvons en déduire la fonction de partition en l'absence de champ comme

$$\begin{aligned} Z[\eta] &= \int_{\mathbb{R}^d} e^{-\beta H[\eta]} \prod_{j=1}^N d\eta_j \\ &= \int_{\mathbb{R}^d} \exp\left(-\beta a^d \sum_{i=1}^N \left[\frac{c}{2} |\nabla\eta_i|^2 + \frac{1}{2!} r_0 \eta_i^2 + \frac{1}{4!} u_0 \eta_i^4\right]\right) \prod_{j=1}^N d\eta_j \end{aligned} \quad (8.37)$$

Dans le cas de la présence d'un champ B , en notant $B_i = B(\mathbf{r}_i)$, la fonction de partition s'écrira alors

$$Z[\eta] = \int_{\mathbb{R}^d} e^{-\beta H[\eta] + \beta \sum_{k=1}^N B_k \eta_k} \prod_{j=1}^N d\eta_j \quad (8.38)$$

La fonction de partition est un produit de N intégrales identiques, les η_i étant tous indépendants. On peut alors utiliser l'approximation de Landau pour déterminer l'énergie libre du système. On remplace chaque intégrale par le maximum de son intégrande. Dans le cas d'un champ magnétique uniforme, toutes les aimantations moyennes par site sont égales et valent η_0 . L'énergie libre vaut finalement

$$F = -k_B T \ln Z = N [H(\eta_0) - B\eta_0] \quad (8.39)$$

Dans le cas d'un champ non uniforme, le maximum de l'intégrande est donné par

$$B_j = \left(\frac{\partial H}{\partial \eta_j}\right)_{\eta_j = \eta_{j_0}} \quad (8.40)$$

et l'énergie libre est donné par

$$F = H(\{\eta_{j_0}\}) - \sum_{j=1}^N B_j \eta_{j_0} \quad (8.41)$$



Formulation continue

La description sur un réseau est parfois lourde et on préfère passer à la limite continue. L'hamiltonien précédent, qui sommait toutes les contributions des sites, va donc naturellement devenir une intégrale. En ce qui concerne la fonction de partition, on doit sommer sur tous les états possibles que peut prendre le système. Précédemment, cela revenait à dire que, pour *chaque* site, nous devons sommer sur *toutes* les valeurs possibles que pouvait prendre le paramètre d'ordre η_i , d'où l'apparition naturelle d'intégrales multiples. Dans le cas continu qui nous intéresse ici, on doit donc sommer en *chaque* point de l'espace \mathbf{r} sur *toutes* les valeurs possible que peut prendre le champ η . En d'autres termes, nous devons effectuer une intégration sur toutes les champs possibles. On intègre sur un espace de fonctions. Nous avons affaire à une *intégrale fonctionnelle*, ou *intégrale de chemin* en mécanique quantique. Cette notion est expliquée dans l'annexe A.2.

En champ non nul, on postule le hamiltonien effectif de *Ginzburg-Landau* décrivant la transition

$$H[\eta] = \int_{\mathbb{R}^d} \left(\frac{c}{2} |\nabla\eta|^2 + \frac{1}{2!} r_0 \eta^2 + \frac{1}{4!} u_0 \eta^4 - B(\mathbf{r})\eta \right) d^d \mathbf{r} \quad (8.42)$$

La fonction de partition en champ non nul s'écrit donc ^a

$$Z[\eta] = \int \exp \left(-\beta \int_{\mathbb{R}^d} \left(\frac{c}{2} |\nabla\eta|^2 + \frac{1}{2!} r_0 \eta^2 + \frac{1}{4!} u_0 \eta^4 - B(\mathbf{r})\eta \right) d^d \mathbf{r} \right) \mathcal{D}\eta \quad (8.43)$$

^a. La notation $\mathcal{D}\eta$ désigne que l'on fait une intégrale fonctionnelle, sur tous les chemins possibles. Pour une introduction élémentaire à ce concept, consulter *La Bible*, le cours de Physique de Feynman, Livre 3, Chapitre 3, verset 1.

Si maintenant nous faisons l'approximation de Landau, la valeur de la fonction de partition est donnée en ne gardant que la fonction η qui minimise l'hamiltonien effectif. Cette fonction est obtenue en calculant la dérivée fonctionnelle (voir appendice A) de $H[\eta]$. Ceci amène aux équations d'Euler-Lagrange pour le paramètre d'ordre. Cela nous conduit à l'équation

$$-c\Delta\eta + r_0(T)\eta + \frac{1}{6}u_0\eta^3 = B(\mathbf{r}) \quad (8.44)$$

À partir de cette équation, nous pouvons retrouver tous les exposants critiques d'une théorie de champ moyen. Par exemple, en champ nul et pour un système homogène, on retrouve facilement que le paramètre d'ordre varie comme $(T - T_c)^{1/2}$. Si on note η_0 la solution à cette équation, nous aurons pour la fonction de partition

$$Z[\eta_0] = e^{-\beta H[\eta_0]} = \exp \left(-\beta \int_{\mathbb{R}^d} \left(\frac{c}{2} |\nabla\eta_0|^2 + \frac{1}{2!} r_0 \eta_0^2 + \frac{1}{4!} u_0 \eta_0^4 - B(\mathbf{r})\eta_0 \right) d^d \mathbf{r} \right) \quad (8.45)$$

Ce qui permet de déterminer l'énergie libre comme étant

$$F[\eta_0] = \int_{\mathbb{R}^d} \left(\frac{c}{2} |\nabla\eta_0|^2 + \frac{1}{2!} r_0 \eta_0^2 + \frac{1}{4!} u_0 \eta_0^4 - B(\mathbf{r})\eta_0 \right) d^d \mathbf{r} \quad (8.46)$$

8.3.3 Corrélation et critère de Ginzburg

Fonctions de corrélation

On généralise la notion de fonction de corrélation que nous avons introduite pour le modèle d'Ising.

On appelle *fonction de corrélation* $G(\mathbf{r}, \mathbf{r}')$ la quantité définie de la façon suivante

$$G(\mathbf{r}, \mathbf{r}') = \langle \eta(\mathbf{r})\eta(\mathbf{r}') \rangle - \langle \eta(\mathbf{r}) \rangle \langle \eta(\mathbf{r}') \rangle \quad (8.47)$$

Dans le cadre de l'approximation de Landau, nous pouvons montrer sans trop de difficultés que

$$G(\mathbf{r}, \mathbf{r}') = -\frac{1}{\beta} \frac{\delta^2 F[\eta_0]}{\delta B(\mathbf{r})\delta B(\mathbf{r}')} \quad (8.48)$$



La relation thermodynamique $\eta_0 = -\frac{\delta F[\eta_0]}{\delta B(\mathbf{r})}|_{\text{eq}}$ permet enfin de montrer la relation

$$G(\mathbf{r}, \mathbf{r}') = -\frac{1}{\beta} \frac{\delta \eta_0}{\delta B(\mathbf{r}')} \quad (8.49)$$

Pour pouvoir déterminer l'équation vérifiée par la fonction de corrélation, il suffit de dériver fonctionnellement l'équation 8.44 pour avoir

$$\left[-\Delta + \frac{1}{\xi^2} \right] G(\mathbf{r}, \mathbf{r}') = \frac{k_B T}{c} \delta(\mathbf{r} - \mathbf{r}') \quad (8.50)$$

où ξ est par définition la *longueur de corrélation* et vaut $\xi = \sqrt{\frac{c}{3u_0\eta_0^2 + r_0(T)}}$. On voit que la fonction de corrélation n'est rien d'autre qu'une fonction de Green. Pour un système macroscopique, elle ne dépend que de la différence $\mathbf{r} - \mathbf{r}'$. Nous pouvons résoudre cette équation à l'aide de la transformée de Fourier et du théorème des résidus. Dans le cas à trois dimensions, le résultat est le suivant

$$G(\mathbf{r} - \mathbf{r}') = \frac{k_B T}{4\pi c} \frac{e^{-\frac{r}{\xi}}}{r} \quad (8.51)$$

Au passage, nous avons retrouvé le comportement attendu pour la longueur de corrélation, à savoir qu'au voisinage du point critique elle est proportionnelle à $|T - T_c|^{-1/2}$. Si l'on fait le calcul pour une dimension quelconque, on constate que les résultats théoriques obtenus par l'approximation de Landau sont d'autant plus exacts que la dimension est élevée. Ceci se comprend intuitivement puisque, par cette approximation, nous avons négligé les fluctuations qui sont moins importantes lorsque la dimension est plus grande.

Critère de Ginzburg

On rappelle encore une fois que l'approximation de Landau néglige les fluctuations du paramètre d'ordre. Ceci n'est bien sûr qu'une approximation et elle n'est valable que sous conditions. Grâce à la notion de longueur de corrélation nous pouvons énoncer très simplement *le critère de Ginzburg*

$$G(\xi) \ll \eta_0^2 = \langle \eta \rangle^2 \quad (8.52)$$

Cette inégalité exprime le fait que, sur une longueur typique, les variations de la fonction de corrélation (qui mesure les fluctuations) sont négligeables devant le carré de la valeur moyenne.

8.4 Groupe de renormalisation

L'idée du groupe de renormalisation consiste à transformer un hamiltonien en un autre en réduisant le nombre de degrés de liberté, tout en conservant la fonction de partition. On passe en réalité d'une description microscopique à une description mésoscopique. Le but est, pour un système donné, de trouver un système équivalent qui, lui, peut être décrit par une théorie de Landau.

Le groupe de renormalisation est d'invention récente et n'est pas construit sur la base de théorèmes rigoureux. On ne cherchera pas dans la suite à développer de manière exhaustive cette méthode mais à décrire les concepts et les résultats intéressants à travers l'exemple du modèle d'Ising.

8.4.1 Modèle d'Ising unidimensionnel

Bien que le modèle d'Ising unidimensionnel puisse être résolu de manière exacte, nous allons lui appliquer les idées du groupe de renormalisation pour se familiariser avec cette méthode. Plaçons nous en champ nul. On rappelle que la fonction de partition s'écrit comme

$$Z = \sum_{\{S_i = \pm 1\}} \prod_{i=1}^N e^{\beta J(S_i S_{i+1} + B\mu S_i)} \quad (8.53)$$

L'idée est la suivante. Considérons trois spins voisins, disons S_1 , S_2 et S_3 . Si l'on s'intéresse au spin central S_2 , sa contribution à la fonction de partition apparaît dans le terme



$$\sum_{\{S_2=\pm 1\}} e^{K(S_1 S_2 + S_2 S_3)} = e^{K(S_1 + S_3)} + e^{-K(S_1 + S_3)} \quad (8.54)$$

où l'on a posé $K = \beta J$. Le but est alors de trouver un système de deux spins équivalent au trois autres. Pour cela, on conserve les spins S_1 et S_3 mais l'on modifie la constante de couplage de manière qu'elle tienne compte de l'influence de S_2 . On cherche donc à avoir une contribution à la fonction de partition de la forme

$$f(K) e^{K' S_1 S_3} \quad (8.55)$$

qui est bien la contribution de deux spins couplés. Pour que cette égalité ait un sens, elle doit être vérifiée quelles que soient les valeurs des spins S_1 et S_3 . Nous devons donc résoudre le système

$$\begin{cases} e^{2K} + e^{-2K} = f(K) e^{K'} \\ 2 = f(K) e^{-K'} \end{cases} \Leftrightarrow \begin{cases} f(K) = 2\sqrt{\cosh(2K)} \\ K' = \ln\left(\frac{f(K)}{2}\right) \end{cases} \quad (8.56)$$

Nous pouvons finalement appliquer ce procédé à tous les spins pairs. Ceci nous amène donc à considérer un système équivalent au système que l'on étudie mais avec un nombre inférieur de spins. Au prix d'une complexification de l'interaction entre les spins, nous en avons réduit le nombre. Pour finir, l'idée est d'à nouveau répéter ce procédé au nouveau système. On obtient alors une suite de systèmes tous équivalents ainsi qu'une suite de constantes de couplage. Dans le cas du modèle d'Ising unidimensionnel, la figure 8.7 représente graphiquement cette suite de constantes que l'on obtient.

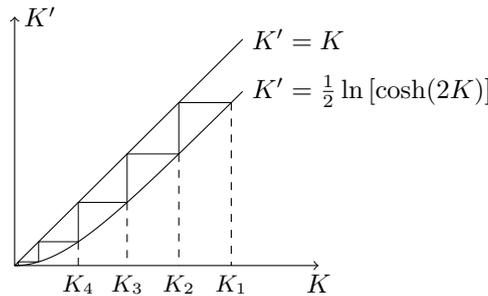


FIGURE 8.7 – Représentation des différentes constantes de couplage que l'on obtient par suite de l'application du procédé de renormalisation. On voit que cette suite tend vers zéro. Physiquement, cela veut dire que plus les spins sont éloignés moins ils interagissent entre eux. La longueur de corrélation reste finie.

L'interprétation physique de l'information que nous fournit le procédé du groupe de renormalisation est simple. À chaque itération du procédé, on regarde des spins qui, à l'origine, sont de plus en plus éloignés. La constante de couplage que l'on obtient à chaque fois nous renseigne sur l'interaction entre ces spins. En d'autres termes, on a accès à la corrélation entre des spins de plus en plus éloignés. Dans l'exemple du modèle d'Ising unidimensionnel, on voit que la corrélation diminue avec la distance : plus les spins sont éloignés, plus la constante de couplage est faible. De plus, on constate que la corrélation tend vers zéro quand la distance tend vers l'infini (à la limite thermodynamique) et cela est vrai pour n'importe quelle valeur de la température (pour n'importe quelle valeur de K). Ceci est caractéristique d'une absence de transition de phases car sinon, il existerait une température, la température critique, pour laquelle la longueur de corrélation divergerait à la limite thermodynamique (la suite des constantes de couplages tendrait vers l'infini). Donc, même s'il existe un fort couplage entre les plus proches voisins, il n'y en a aucun à une échelle mésoscopique.

8.4.2 Modèle d'Ising bidimensionnel

Nous pouvons faire le même type de raisonnement dans le cas bidimensionnel. Cependant, nous ne pouvons pas isoler uniquement trois spins. Pour pouvoir faire les choses correctement, nous devons sélectionner cinq spins. Si l'on se donne un spin que l'on nomme S_5 , nous devons considérer ses quatre plus proches voisins que l'on notera S_1, S_2, S_3 et S_4 , comme sur la figure 8.8

On cherche donc toujours un système équivalent mais avec un nombre de spins inférieur. Il aura non plus cinq mais quatre spins et l'influence du cinquième sera comptée dans les constantes de couplages. Nous aurons en

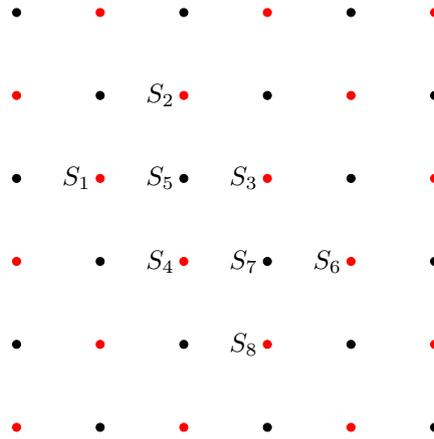


FIGURE 8.8 – L’ensemble des notations qui serviront lors de l’écriture des renormalisations du modèle d’Ising à deux dimensions. On peut considérer ce modèle d’Ising bidimensionnel comme contenant deux réseaux. Le but va être d’obtenir un système équivalent ne contenant plus que le réseau rouge, en réécrivant les couplages avec les spins du réseau noir.

effet besoin de trois constantes rendant compte du couplage entre plus proches voisins, les voisins suivants et les quatre spins ensemble. La contribution de S_5 à la fonction de partition est donnée par

$$\sum_{\{S_5=\pm 1\}} e^{K S_5(S_1+S_2+S_3+S_4)} \tag{8.57}$$

Nous cherchons donc à l’écrire sous la forme

$$\sum_{\{S_5=\pm 1\}} e^{K S_5(S_1+S_2+S_3+S_4)} = f(K) \underbrace{e^{\frac{K_1}{2}(S_1 S_2+S_2 S_3+S_3 S_4+S_4 S_1)}}_{\text{Plus proches voisins}} \underbrace{e^{K_2(S_1 S_3+S_2 S_4)}}_{\text{Plus proches voisins suivants}} e^{K_3 S_1 S_2 S_3 S_4} \tag{8.58}$$

Cette écriture doit être valable pour toutes les valeurs des quatre spins. Cela nous donne un système d’équations dont les solutions sont

$$f(K) = 2 \cosh^{1/8}(4K) \cosh^{1/2}(2K) \tag{8.59}$$

$$K_3 = \frac{1}{8} \ln(\cosh(4K)) - \frac{1}{2} \ln(\cosh(2K)) \tag{8.60}$$

$$K_2 = \frac{1}{8} \ln(\cosh(4K)) \tag{8.61}$$

$$K_1 = \frac{1}{4} \ln(\cosh(4K)) \tag{8.62}$$

Les interactions sont plus compliquées et n’aident pas à trouver une solution exacte. Cependant, si on oublie les termes d’interaction d’ordre supérieur, ceci pourrait changer la température critique mais pas les exposants critiques. Nous allons donc faire des approximations. Si on ignore les couplages K_2 et K_3 on retrouve un comportement similaire au cas unidimensionnel. En particulier, il n’existe pas de transition de phases. Cependant, on peut tenir compte du couplage entre les voisins suivants en changeant convenablement la constante K_1 caractérisant les interactions entre les plus proches voisins. En imposant que l’approximation reproduise le même comportement en champ moyen, et en remarquant que chaque interaction entre plus proches voisins apparaît deux fois dans la fonction de partition tandis que celle avec les voisins suivants n’apparaît qu’une seule fois, on introduit la constante

$$K' = K_1 + K_2 = \frac{3}{8} \ln(\cosh(4K)) \tag{8.63}$$

Cette fois le comportement du système est différent, comme cela est visible sur la figure 8.9. On voit qu’il existe un couplage K_c particulier, c’est-à-dire pour une température T_c particulière, qui sépare deux comportements différents. Lorsque $K < K_c$, ou $T > T_c$, à la limite thermodynamique, le couplage entre les spins tend vers

zéro. En d'autres termes, il n'y a pas de corrélation à grande distance. En revanche, pour $K > K_c$, ou $T < T_c$, le couplage à grande distance diverge. La longueur de corrélation est donc infinie et il y a toujours interaction entre des spins à grande distance. Ce phénomène est caractéristique d'une transition de phases qui a donc lieu à la température T_c .

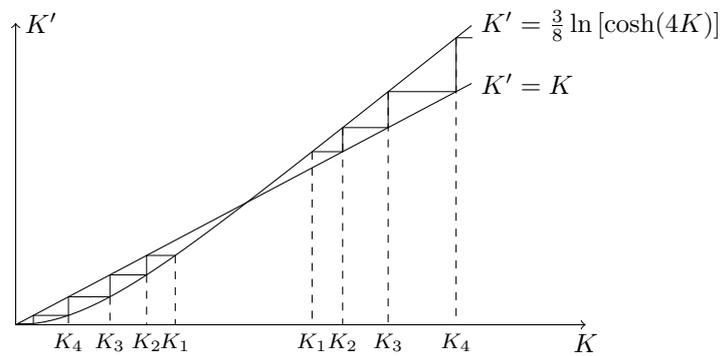


FIGURE 8.9 – Dans le cas bidimensionnel, on a deux cas qui apparaissent, car la courbe $K' = K_1 + K_2 = \frac{3}{8} \ln(\cosh(4K))$ croise la courbe $K' = K$. On appelle K_c le point de croisement. Alors, pour $K < K_c$, on retrouve le cas unidimensionnel, et la corrélation entre spins de plus en plus éloignés tend vers zéro. Mais, si la température est suffisamment faible, c'est à dire lorsque $K > K_c$, l'application itérée des renormalisations montre que le couplage devient de plus en plus important entre spins de plus en plus éloignés. Cela signe une transition de phases.

Annexe A

Dérivées et intégrales fonctionnelles¹

Sommaire

A.1 Dérivée fonctionnelle	210
A.2 Intégrale fonctionnelle	213

La dérivée et l'intégrale fonctionnelle sont des généralisations des notions classiques de dérivation et d'intégration pour le cas où les variables ne sont plus des éléments de \mathbb{R} ou de \mathbb{C} mais des éléments d'espaces fonctionnels comme $\mathcal{C}^\infty(\mathbb{R})$ par exemple. De telles applications sont en général appelées *fonctionnelles*. Dans ce qui suit, nous allons définir ces notions et donner les bases nécessaires pour les applications en physique.

A.1 Dérivée fonctionnelle

La dérivée fonctionnelle apparaît dans de nombreux domaines de la physique. À un niveau élémentaire, en mécanique analytique, l'application du Principe de Moindre Action conduit au calcul de dérivées fonctionnelles, même si cela n'est pas dit de manière explicite.

A.1.1 Définition

L'idée des dérivées fonctionnelles consiste à formellement mimer le calcul différentiel usuel en dimension infinie. Dans le cas usuel, on définit la différentielle de la manière suivante.

Soit $f \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$, fonction de n variables infiniment dérivable. Soit $x \in \mathbb{R}^n$. La différentielle de f en x , notée $(df)(x)$ ou encore df_x , est une application linéaire définie par

$$\forall h \in \mathbb{R}^n, f(x+h) = f(x) + (df)(x).h + O(h^2) \tag{A.1}$$

Si on explicite cette formule en fonction des dérivées partielles, nous avons

$$\forall h \in \mathbb{R}^n, f(x+h) = f(x) + \sum_{i=1}^n (\partial_i f)(x)h^i + O(h^2) \tag{A.2}$$

Sous certaines hypothèses, on peut avoir la formule de Taylor

$$\forall h \in \mathbb{R}^n, f(x+h) = f(x) + \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{(i_1 \dots i_k) \in (1, N)^k} (\partial_{i_1} \dots \partial_{i_k} f)(x)h^{i_1} \dots h^{i_k} \tag{A.3}$$

Le cas fonctionnel marche de la même façon mais en remplaçant i par la variable d'espace-temps $(x^\mu)_\mu$ et par les éventuels indices de composantes de champs.

1. Cet appendice est tiré du cours de Pascal Degiovanni, *Introduction à l'électrodynamique quantique*, de 1999.



Soit $S : \mathcal{F} \rightarrow \mathbb{K}$ une fonctionnelle, où \mathcal{F} est un espace fonctionnel et \mathbb{K} représente le corps des réels ou des complexes. Soient $\phi, \eta \in \mathcal{F}$. La dérivée fonctionnelle $\frac{\delta S}{\delta \phi(x)}$ de S est définie par

$$S[\phi + \eta] = S[\phi] + \int \frac{\delta S}{\delta \phi(x)} \eta(x) dx + O(\eta^2) \quad (\text{A.4})$$

Bien sûr, comme $[\phi] \rightarrow \left(\frac{\delta S}{\delta \phi(x)}\right)[\phi]$ est encore une fonctionnelle de ϕ , on peut définir des dérivées fonctionnelles multiples $\frac{\delta^n S}{\delta \phi(x_1) \cdots \delta \phi(x_n)}$. L'analogie de la formule de Taylor devient

$$S[\phi + \eta] = S[\phi] + \sum_{n=1}^{\infty} \frac{1}{n!} \int \frac{\delta^n S}{\delta \phi(x_1) \cdots \delta \phi(x_n)} \eta(x_1) \cdots \eta(x_n) dx_1 \cdots dx_n \quad (\text{A.5})$$

A.1.2 Quelques exemples

Un premier calcul

Soit

$$S[\phi] = \frac{1}{2} \int \phi(x) F(x, y) \phi(y) d^d x d^d y \quad (\text{A.6})$$

Calculons la dérivée fonctionnelle première de cette fonctionnelle. Pour cela, revenons à la définition. Nous avons

$$S[\phi + \eta] = \frac{1}{2} \int (\phi(x) + \eta(x)) F(x, y) (\phi(y) + \eta(y)) d^d x d^d y \quad (\text{A.7})$$

$$\simeq S[\phi] + \frac{1}{2} \int (\phi(x) F(x, y) \eta(y) + \eta(x) F(x, y) \phi(y)) d^d x d^d y + O(\eta^2) \quad (\text{A.8})$$

Ceci permet de conclure que

$$\frac{\delta S}{\delta \phi(x)} = \frac{1}{2} \int (\phi(y) F(y, x) + F(x, y) \phi(y)) d^d y \quad (\text{A.9})$$

Le calcul de la dérivée fonctionnelle seconde donnerait

$$\frac{\delta^2 S}{\delta \phi(x) \delta \phi(y)} = \frac{1}{2} (\phi(y) F(y, x) + F(x, y) \phi(y)) \quad (\text{A.10})$$

Une relation importante

Calculons la dérivée fonctionnelle de la fonctionnelle identité, c'est à dire de l'application suivante

$$\begin{aligned} \mathcal{F} &\rightarrow \mathbb{K} \\ \phi &\mapsto \phi(x) \end{aligned} \quad (\text{A.11})$$

On note sa dérivée fonctionnelle comme $\frac{\delta \phi(x)}{\delta \phi(y)}$. Nous avons donc

$$\phi(x) + \eta(x) = \phi(x) + \int \delta(x - y) \eta(y) d^d y \quad (\text{A.12})$$

Ceci nous donne donc directement

$$\frac{\delta \phi(x)}{\delta \phi(y)} = \delta(x - y) \quad (\text{A.13})$$



Une identité utile en mécanique statistique

Imaginons observer un système statistique qui fait intervenir un champ $\phi(x)$. Alors, si on désigne par $\langle \dots \rangle$ les valeurs moyennes statistiques, il est intéressant de considérer la fonctionnelle suivante

$$K[J] = \langle \exp \left(\int J(x)\phi(x)d^d x \right) \rangle \quad (\text{A.14})$$

Le sens de la moyenne sera donné dans les sections qui suivent. On peut tout de même faire un parallèle entre cette formule et la fonction de partition. En développant l'exponentielle, nous avons

$$K[J] = \langle 1 \rangle + \sum_{n=1}^{\infty} \frac{1}{n!} \int \langle \phi(x_1) \cdots \phi(x_n) \rangle \left(\prod_{k=1}^n J(x_k) d^d x_k \right) \quad (\text{A.15})$$

En clair, cette fonctionnelle contient toutes les corrélations $\langle \phi(x_1) \cdots \phi(x_n) \rangle$. En rapprochant cette formule de la formule de Taylor, nous avons la relation

$$\left(\frac{\delta^n K[J]}{\delta J(x_1) \cdots \delta J(x_n)} \right)_{J=0} = \langle \phi(x_1) \cdots \phi(x_n) \rangle \quad (\text{A.16})$$

A.1.3 Résumé

Définition

Soit $S : \mathcal{F} \rightarrow \mathbb{K}$ une fonctionnelle, où \mathcal{F} est un espace fonctionnel et \mathbb{K} représente le corps des réels ou des complexes. Soient $\phi, \eta \in \mathcal{F}$. La dérivée fonctionnelle $\frac{\delta S}{\delta \phi(x)}$ de S est définie par

$$S[\phi + \eta] = S[\phi] + \int \frac{\delta S}{\delta \phi(x)} \eta(x) dx + O(\eta^2) \quad (\text{A.17})$$

Formule de Leibnitz

$$\frac{\delta (A[\phi] B[\phi])}{\delta \phi(x)} = A[\phi] \frac{\delta B[\phi]}{\delta \phi(x)} + \frac{\delta A[\phi]}{\delta \phi(x)} B[\phi] \quad (\text{A.18})$$

Formule de Schwartz

En général, on n'hésite pas à intervertir les dérivées :

$$\frac{\delta^2}{\delta \phi(x) \delta \phi(y)} = \frac{\delta^2}{\delta \phi(y) \delta \phi(x)} \quad (\text{A.19})$$

Dérivée de l'identité

$$\frac{\delta \phi(x)}{\delta \phi(y)} = \delta(x - y) \quad (\text{A.20})$$

Formule de Taylor

$$S[\phi + \eta] = S[\phi] + \sum_{n=1}^{\infty} \frac{1}{n!} \int \frac{\delta^n S}{\delta \phi(x_1) \cdots \delta \phi(x_n)} \eta(x_1) \cdots \eta(x_n) dx_1 \cdots dx_n \quad (\text{A.21})$$

Ces formules sont en gros tout ce que l'on a besoin de savoir en pratique pour utiliser les dérivées fonctionnelles. La seule chose à se rappeler est que tout marche comme pour les dérivées ordinaires. Une fois que l'on sait cela, il n'y a plus de difficultés.



A.2 Intégrale fonctionnelle

Il existe plusieurs méthodes pour introduire l'intégrale fonctionnelle. Historiquement, c'est Feynman qui en a eu l'idée en 1948 pour reformuler la mécanique quantique traditionnelle en termes d'intégrales sur des trajectoires. L'introduction la plus élémentaire se trouve dans *La Bible*, le cours de Physique de Feynman, Livre 3, Chapitre 3, verset 1.



Annexe B

Théorie des groupes¹

Sommaire

B.1 Groupes et exemples	214
B.2 Action de groupes	215
B.3 Représentation de groupe	216
B.4 Groupes de Lie, algèbres de Lie et représentations	216

Le concept de groupe est étroitement lié à la notion de symétrie. Un groupe de symétrie d'un objet géométrique est un ensemble de transformations géométriques laissant invariant l'objet. Une des propriétés de ce groupe est que l'application de plusieurs symétries redonne une symétrie. Une autre est que l'inverse d'une symétrie existe et est une symétrie. Ces propriétés sont à la base de la définition des groupes. Historiquement, la définition moderne des groupes fut formulée par Évariste Galois alors qu'il étudiait le problème de la résolubilité par radicaux des équations polynomiales.

Ce complément est un bref exposé de certaines notions de théorie des groupes qui sont utilisées en physique. Ces quelques pages ne sont pas suffisantes pour appréhender complètement la notion de groupe, encore moins celles de groupe et d'algèbre de Lie. On essaiera tout de même de donner un aperçu de ces objets pour au moins comprendre ce qui est dit quand ils sont utilisés.

B.1 Groupes et exemples

B.1.1 Définitions

Un groupe G est un ensemble non vide muni d'une loi $\star : G \times G \rightarrow G$ tel que

$$\begin{cases} \forall (x, y, z) \in G^3, & (x \star y) \star z = x \star (y \star z) & \text{Associativité de la loi} \\ \exists e \in G, \forall x \in G & x \star e = e \star x = x & \text{Existence d'un élément neutre} \\ \forall x \in G, \exists x^{-1} \in G & x^{-1} \star x = x \star x^{-1} = e & \text{Existence d'un inverse} \end{cases} \quad (\text{B.1})$$

La notion de groupe est centrale en physique puisque c'est elle qui permet de discuter de manière convenable des symétries comme par exemple le changement d'observateurs. Avant de rentrer dans les détails, voyons quelques exemples d'importance.

1. Cet appendice est tiré du cours de Pascal Degiovanni, *Introduction à l'électrodynamique quantique*, de 1999.

B.1.2 Un exemple de groupe

Le groupe des permutations (\mathfrak{S}_n, \circ) est le groupe des bijections de l'ensemble $\{1, \dots, n\}$. C'est un groupe fini d'ordre (de cardinal) $n!$.

Un *cycle* σ noté $(i_1 \cdots i_r)$ est une permutation telle que

$$\sigma(i_k) = i_{k+1} \quad \text{si } k \in \llbracket 1, r-1 \rrbracket \quad (\text{B.2a})$$

$$\sigma(i_r) = i_1 \quad (\text{B.2b})$$

$$\sigma(i_k) = i_k \quad \text{sinon} \quad (\text{B.2c})$$

On dit que σ est un cycle de longueur r . On peut montrer qu'alors $\sigma^r = \mathbb{1}$. En particulier, un cycle de longueur deux est une *transposition*. On définit le *support* d'un cycle $(i_1 \cdots i_r)$ comme étant l'ensemble $\{i_1 \cdots i_r\}$. Deux cycles sont disjoints si leurs supports le sont. Dans ce cas, les deux cycles commutent. Un théorème stipule que toute permutation de \mathfrak{S}_n peut être décomposée en un produit de cycles disjoints et que cette décomposition est unique à l'ordre des facteurs près. En particulier, un cycle $(i_1 \cdots i_r)$ est le produit de $r-1$ transpositions,

$$(i_1 \cdots i_r) = (i_1 i_2)(i_2 i_3) \cdots (i_{r-1} i_r) \quad (\text{B.3})$$

De là, on en déduit le corollaire suivant : toute permutation de \mathfrak{S}_n est produit d'au plus $n-1$ transpositions.

Un morphisme important que nous allons définir provient de la signature d'une permutation. Par définition, on appelle *signature* d'une permutation $\sigma \in \mathfrak{S}_n$, notée $\epsilon(\sigma)$, la quantité

$$\epsilon(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(i) - \sigma(j)}{i - j} \quad (\text{B.4})$$

La signature d'une transposition est égale à 1. On peut montrer que nous avons les propriétés suivantes

$$\epsilon(\sigma) = \pm 1 \quad (\text{B.5a})$$

$$\forall \sigma, \rho \in \mathfrak{S}_n, \epsilon(\rho\sigma) = \epsilon(\rho)\epsilon(\sigma) \quad (\text{B.5b})$$

Ceci permet d'affirmer, en notant n le nombre de transpositions dont est composé une permutation σ , que $\epsilon(\sigma) = (-1)^n$. En particulier, la signature d'un cycle de longueur r vaut $(-1)^{r-1}$. Finalement, l'application

$$\epsilon : \begin{array}{l} \mathfrak{S}_n \rightarrow \{-1; 1\} \\ \sigma \mapsto \epsilon(\sigma) \end{array} \quad (\text{B.6})$$

est un morphisme de groupe dont le noyau, sous-groupe de \mathfrak{S}_n , est appelé *le groupe alterné*. C'est l'ensemble de toutes les permutations de signature égale à l'unité. Son ordre est égal à $\frac{n!}{2}$.

B.2 Action de groupes

Soit G un groupe et E un espace vectoriel de dimension n .

On dit que G agit sur l'ensemble E lorsque l'on dispose de $\cdot : G \times E \rightarrow E^a$

$$\forall (g, g') \in G^2, \forall x \in E \quad g \cdot (g' \cdot x) = (g \star g') \cdot x \quad (\text{B.7a})$$

$$\forall x \in E, e \cdot x = x \quad (\text{B.7b})$$

a. Attention, en général, on utilise les mêmes notations pour la loi de groupe et l'action. Il faut donc bien réfléchir au sens de chaque opération.

En fait, l'action de G sur E est décrite par un morphisme de G dans le groupe des permutations de l'ensemble E , \mathfrak{S}_n . Par exemple, en relativité restreinte, tous les observateurs sont reliés entre eux par un groupe, le groupe de Poincaré, qui est le groupe des isométries affines de l'espace de Minkowski. En fait, on peut voir G comme fournissant des moyens d'indexer l'ensemble des observateurs. Pour cela, on en choisit un (c'est là qu'est l'arbitraire) θ_e et on définit, pour tout $g \in G$, $\theta_g = g \cdot \theta_e$. Il ne faut pas perdre de vue qu'en changeant le référentiel de "référence", on change l'indexation. Toutes les indexations doivent être mises sur le même plan si on veut éviter de retrouver la notion de référentiel absolu.

B.3 Représentation de groupe

Continuons en utilisant les changements de référentiels. En fait, dans tout le traitement des symétries, l'index du référentiel peut être vu comme l'élément d'un groupe. Au niveau de l'espace-temps, comme au niveau des champs, on associe à chaque changement de référentiel une transformation

$T_{\theta\theta'} : X_\theta \rightarrow X_{\theta'}$. Si $\theta' = g \cdot \theta$ avec $g \in G$, on a donc $T_\theta(g) : X_\theta \rightarrow X_{g \cdot \theta}$ et bien sûr on s'attend à avoir la relation de composition

$$T_\theta(g'g) = T_{g \cdot \theta}(g') \circ T_\theta(g) \quad (\text{B.8})$$

En général, on identifie les espaces X_θ à un seul espace X ². Dans ce cas, on a $T(g) : X \rightarrow X$ et l'équation ci-dessus devient

$$T(g'g) = T(g') \circ T(g) \quad (\text{B.9})$$

Cela signifie que l'on dispose d'un morphisme de groupe qui va de G dans le groupe des automorphismes de l'espace X considéré. Ceci porte le nom de représentation linéaire de G . C'est un concept absolument central en physique. Formellement, nous avons donc

Une représentation linéaire du groupe G est un morphisme $\rho : G \rightarrow \text{Aut}(X)$ où X est un espace vectoriel de dimension finie. C'est une action où ρ_g est linéaire.

L'intérêt est de pouvoir représenter un groupe abstrait par des objets plus habituels comme des matrices.

B.4 Groupes de Lie, algèbres de Lie et représentations

B.4.1 Groupe de Lie : définition

C'est la notion pertinente pour les familles continues d'observateurs : c'est le cas pour les translations, les rotations, les transformations de Lorentz ou les symétries que l'on rencontre en physique des particules. Même les transformations de jauge rentrent dans ce cadre mais en dimension infinie.

Un groupe de Lie est un groupe qui est aussi une variété différentielle avec une loi compatible avec la structure différentielle. C'est un groupe "différentiable".

Un groupe de Lie G est un groupe muni d'une structure de variété différentielle (réelle ou complexe) telle que la loi de groupe ainsi que l'application inverse soient différentiables (ou holomorphes).

B.4.2 Algèbre et groupe de Lie

Dans ce cas, on imagine bien qu'il soit intéressant d'étudier un tel groupe au voisinage de son élément neutre. Sans être précis, cela revient à faire un développement des éléments du groupe au premier ordre

$$g \simeq \mathbb{1} + h + O(h^2) \quad (\text{B.10})$$

La commutativité du groupe est codée dans les propriétés de h . Nous avons pour $g_1, g_2 \in G$, en utilisant l'identité $(\mathbb{1} + h)^{-1} = \mathbb{1} - h$,

$$g_1 g_2 g_2^{-1} g_1^{-1} \simeq (\mathbb{1} + h_1 + O(h^2))(\mathbb{1} + h_2 + O(h^2))(\mathbb{1} - h_2 + O(h^2))(\mathbb{1} - h_1 + O(h^2)) \quad (\text{B.11})$$

$$\simeq \mathbb{1} + [h_1, h_2] + O(h^2) \quad (\text{B.12})$$

De manière plus précise, ce qui correspond à ce développement est ce que l'on appelle *l'algèbre de Lie* associée au groupe de Lie G . Formellement,

². Le fait que l'on identifie les espaces X_θ est simplement un choix commode. On pourrait ne pas le faire mais cela ne serait pas très pratique.

L'algèbre de Lie \mathfrak{g} associée au groupe de Lie G est par définition l'espace tangent à G en l'identité e .

$$\mathfrak{g} = (TG)_e \quad (\text{B.13})$$

On peut montrer que \mathfrak{g} est muni d'une loi $[\cdot, \cdot]$ que l'on appelle *crochet de Lie* et qui, pour le groupe des matrices, correspond au commutateur. Cette loi vérifie les propriétés suivantes :

$$\begin{aligned} \forall (x, y) \in \mathfrak{g}^2, & \quad [x, y] = -[y, x] \\ \forall (x, y, z) \in \mathfrak{g}^3, & \quad [x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \\ \forall x \in \mathfrak{g}, & \quad [x, \cdot] \in \mathcal{L}(\mathfrak{g}) \end{aligned} \quad (\text{B.14})$$

Cet espace est donc bien une algèbre, en particulier un espace vectoriel. Lorsque l'algèbre de Lie est de dimension finie, les éléments formant une base de cet espace vectoriel sont appelés *générateurs* de l'algèbre de Lie.

D'un point de vue indépendant des groupes, une algèbre de Lie est un espace vectoriel muni d'un crochet de Lie (défini par les propriétés ci-dessus). Compte tenu de cette définition, il faudrait montrer en toute rigueur que l'espace tangent en l'identité d'un groupe de Lie est bien une algèbre de Lie. Cependant, nous n'irons pas jusque-là et nous admettrons la cohérence des définitions.

Comme G est un groupe, on peut translater ses éléments. Nous avons l'application

$$L_g : \begin{array}{ccc} G & \rightarrow & G \\ h & \mapsto & gh \end{array} \quad (\text{B.15})$$

qui est une application différentiable dont la différentielle

$$(dL_g)(e) : \begin{array}{ccc} (TG)_e & \rightarrow & (TG)_g \\ x & \mapsto & g \cdot x \end{array} \quad (\text{B.16})$$

transporte $(TG)_e$ sur $(TG)_g$ de manière \mathcal{C}^∞ . On voit que l'on dispose d'une manière commode de transporter les espaces tangents. On dit que le groupe de Lie est parallélisable³

B.4.3 L'application exponentielle

Soit $x \in \mathfrak{g}$. On peut alors essayer de résoudre l'équation différentielle (avec $g : \mathbb{R} \rightarrow G$)

$$\frac{dg(t)}{dt} = g(t) \cdot x \quad (\text{B.17})$$

A priori, $\left(\frac{dg(t)}{dt}\right) \in TG_{g(t)}$ et $g(t) \cdot x \in TG_{g(t)}$ aussi. Cette équation a donc un sens. Par le théorème de Cauchy, on montre qu'elle admet une solution unique avec la condition aux limites $g(0) = e$ et on dit que cette solution est

$$g(t) = e^{tx} \quad (\text{B.18})$$

Dans le cas matriciel, cela correspond bien au cas que l'on connaît mais le concept est plus général. L'exponentielle est une application de l'algèbre de Lie dans un groupe de Lie qui est définie dans le cas de n'importe quel groupe. C'est même elle qui permet de montrer que, étant donné une algèbre de Lie, il existe un groupe de Lie dont l'espace tangent en l'identité est cette même algèbre. L'algèbre de Lie permet de "cartographier" naturellement un groupe, au moins de manière locale. En clair, la notion d'algèbre de Lie englobe la notion physique de "transformation infinitésimale".

3. Il existe des variétés différentielles parallélisables qui ne sont pas des groupes de Lie.

B.4.4 Groupes matriciels

Groupe spécial unitaire

$$SU(n) = \{M \in \mathcal{M}_n(\mathbb{C}) / M^\dagger M = \mathbb{1}, \det(M) = 1\} \quad (\text{B.19})$$

L'algèbre de Lie $\mathfrak{su}(n)$ est isomorphe à l'ensemble des matrices de $\mathcal{M}_n(\mathbb{C})$ qui sont anti-hermitiennes et de trace nulle muni du commutateur des matrices

$$\mathfrak{su}(n) = \{M \in \mathcal{M}_n(\mathbb{C}) / M^\dagger = -M, \text{tr}(M) = 0\}. \quad (\text{B.20})$$

Étudions le cas particulier de $SU(2)$. Un élément de ce groupe s'écrit de la manière suivante

$$\begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix} \text{ avec } |a|^2 + |b|^2 = 1 \quad (\text{B.21})$$

L'algèbre de Lie est $\mathfrak{su}(2)$ formée des matrices suivantes

$$\begin{pmatrix} i\alpha & -\bar{\beta} \\ \beta & -i\alpha \end{pmatrix} \text{ avec } \alpha \in \mathbb{R}, \beta \in \mathbb{C} \quad (\text{B.22})$$

et elle admet pour base $(i\sigma^x, -i\sigma^y, i\sigma^z)$ avec

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (\text{B.23})$$

La base de l'algèbre de Lie vérifie les mêmes relations de commutation que celle des matrices de Pauli, à savoir

$$[\sigma^\alpha, \sigma^\beta] = 2\epsilon^{\alpha\beta\gamma} \sigma^\gamma \quad (\text{B.24})$$

Cette relation définit la structure de l'algèbre de Lie $\mathfrak{su}(2)$. On note que $\dim_{\mathbb{R}}(\mathfrak{su}(2)) = 3$. On peut procéder de la même façon pour $SU(3)$ mais c'est plus fastidieux. Remarquons juste que la condition d'anti-hermité dit que la dimension de l'algèbre de Lie $\mathfrak{su}(3)$ est $9 - 1 = 8$ (le -1 provient de $\text{tr}(H) = 0$). On retrouve le nombre total de gluons différents, ce qui n'est pas un hasard...

Groupe spécial orthogonal

$$SO(n) = \{M \in \mathcal{M}_n(\mathbb{R}) / {}^t M M = \mathbb{1}, \det(M) = 1\} \quad (\text{B.25})$$

L'algèbre de Lie $\mathfrak{so}(n)$ est isomorphe à l'ensemble des matrices de $\mathcal{M}_n(\mathbb{R})$ antisymétriques muni de la relation de commutation matricielle.

$$\mathfrak{so}(n) = \{M \in \mathcal{M}_n(\mathbb{R}) / {}^t M = -M\}. \quad (\text{B.26})$$

Pour $\mathfrak{so}(3)$, cela donne une dimension trois. On a une base donnée par

$$t^x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad t^y = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad t^z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (\text{B.27})$$

Nous avons par exemple $e^{i\theta t^x}$ qui correspond à la rotation d'angle θ autour de l'axe x . De manière similaire au cas complexe, nous avons les relations de commutations suivantes

$$[it^\alpha, it^\beta] = \epsilon^{\alpha\beta\gamma} it^\gamma. \quad (\text{B.28})$$

Ceci montre immédiatement que les algèbres de Lie $\mathfrak{so}(3)$ et $\mathfrak{su}(2)$ sont isomorphes. Mais $SU(2)$ et $SO(3)$ ne le sont pas du tout. Cet exemple montre que nous pouvons avoir $G_1 \neq G_2$ et $\mathfrak{g}_1 \cong \mathfrak{g}_2$.

B.4.5 Action des groupes de Lie et des algèbres de Lie

La notion de représentation d'un groupe de Lie est importante pour l'étude des changements d'observateurs. C'est un cas particulier de morphisme de groupe de Lie. Soit $f : G \rightarrow H$ un tel morphisme; on peut le différencier et on a alors $(df)(g) : (TG)_g \rightarrow (TG)_{f(g)}$. En particulier $(df)(e) : (TG)_e \rightarrow (TG)_e$. On montre alors, puisque f est un morphisme de groupe, que $(df)(e)$ est un morphisme d'algèbre de Lie

$$(df)(e) ([x, y]) = [(df)(e) \cdot x, (df)(e) \cdot y] \quad (\text{B.29})$$

En clair, dès que l'on a une représentation du groupe, on a une représentation de l'algèbre. Attention, la réciproque est fautive (exemple du spin 1/2).



Bibliographie

- [1] <http://www.cea.fr/recherche-fondamentale/les-faces-cachees-du-noyau/avis-de-tempete-sur-la-vallee>.
- [2] D. Lederer B. Roullet B. Diu, C. Guthmann. *Physique Statistique*. Hermann, 1995.
- [3] D. Lederer B. Roullet B. Diu, C. Guthmann. *Thermodynamique*. Hermann, 2007.
- [4] M. Le Bellac. *Des phénomènes critiques aux champs de jauge*. CNRS Éditions, 2002.
- [5] M. Le Bellac. *Physique Quantique*. CNRS Éditions, 2003.
- [6] A. Einstein. *La théorie de la relativité restreinte et générale*. Dunod, 2004.
- [7] R. Feynman. *The Feynman Lectures on Physics, Vol. 3*. Pearson- Addison Wesley, 2006.
- [8] R. P. Feynman. Space-time approach to quantum electrodynamics. *Physical Review*, 76(6) :769–789, September 1949.
- [9] R. P. Feynman. The theory of positrons. *Physical Review*, 76(6) :749–759, September 1949.
- [10] R. P. Feynman. Mathematical formulation of the quantum theory of electromagnetic interaction. *Physical Review*, 80(3) :440–457, November 1950.
- [11] E.ourgoulhon. *Relativité restreinte*. CNRS Éditions, 2010.
- [12] A. Hey I. Aitchinson. *Gauge Theories in Particle Physics*. Adam Hilger, 2006.
- [13] J. Jackson. *Électrodynamique classique*. Dunod, 2001.
- [14] E. Lifshitz L. Landau. *Quantum Mechanics*. Butterworth-Heinemann, 1977.
- [15] E. Lifshitz L. Landau. *Physique Statistique - 4ème Édition*. Ellipses, 1994.
- [16] A. Messiah. *Mécanique quantique, tome 2*. Dunod, 1995.
- [17] D. Mermin N. Ashcroft. *Solid State Physics*. Brooks/Cole, 1976.
- [18] W. Williams. *Nuclear and Particle Physics*. Clarendon Press, 1991.

Postface

ムサシ: なんだかんだと聞かれましたら
コジロウ: 答えてあげるのが世の情け
ムサシ: 世界の破壊を防ぐため
コジロウ: 世界の平和を守るため
ムサシ: 愛と真実の悪をつらぬく
コジロウ: ラブリー・チャーミーな敵役
ムサシ: ムサシ
コジロウ: コジロウ
ムサシ: 銀河をかけるロケット団の二人には.....
コジロウ: ホワイトホール、白い明日が待ってるぜ
ニャース: にゃんてにゃ

Putain, c'est fini!

Ce sont les dernières lignes de ce document et, de façon amusante, ce sont aussi les premières lignes que j'écris. Les premières lignes lisibles directement, du moins. Je suis le gribouilleur de schémas du groupe.

C'est le genre de situation qui vous prend complètement au dépourvu. Un jour, quelqu'un vous dit « je vais écrire un petit résumé » et le lendemain vous êtes en train d'illustrer un document de plus d'une centaine de pages. Évidemment, quand vous vous rendez compte de ce que vous faites, il est trop tard. Vous voilà prisonnier des entrailles de \TeX , de \PGF et de \TikZ . Au cœur d'une nuit sans lune, avec pour seul espoir la chaleur des processeurs maltraités, vous vous retrouvez à faire des trucs et des machins de plus en plus fous... ou de plus en plus stupides.

Je dois dire que je n'aurais pu progresser à la fois aussi vite et aussi bien dans cette quête pour perdre la raison sans le soutien d'Alexandre ou de Clément. Le premier s'est lancé corps et âme dans ce projet aberrant. D'une manière un peu conne certes, car on ne débarque pas avec un court résumé d'une septantaine de pages en ayant l'air très malin. Le second a su pérenniser cette névrose collective et grâce à lui, nous avons su passer le point de non retour sans heurts, sans nous en rendre compte. Pour cela, je les remercie chaleureusement tous les deux. Travailler en équipe était un vrai plaisir, et j'espère que le résultat sera à la hauteur de nos déraisons.

Voilà, c'est la dernière ligne de ce document et c'est aussi la dernière ligne que j'écris... pour l'instant.