

4P066 : Symétries en Physique

Jean-Bernard Zuber

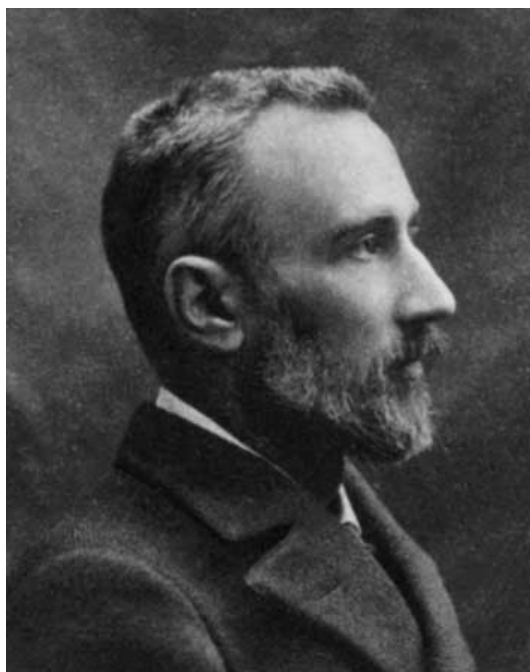


FIGURE 1 –

Pierre Curie (1859-1906) qui donne à notre Université (la moitié de) son nom, est un grand physicien. Il est en particulier à l'origine de nombreuses idées sur le rôle des symétries en Physique, comme on verra dans ces pages ...

Table des matières

1	Symétries géométriques. Des molécules aux cristaux	1
1.1	Généralités sur les symétries	1
1.1.1	Transformations de l'espace et symétries	1
1.1.2	Un petit album d'images...	2
1.1.3	Isométries et déplacements	7
1.2	Réflexion-miroir et inversion/parité	8
1.2.1	Isométries impropres	8
1.2.2	Réflexion-miroir	9
1.3	Rotations et réflexions de $\mathbb{R}^2$ et de $\mathbb{R}^3$	10
1.3.1	Groupes $O(d)$ et $SO(d)$	10
1.3.2	Sous-groupes finis des groupes de rotation à $d = 2$ et $d = 3$	11
1.3.3	Groupes d'invariance des polygones et des polyèdres	11
1.3.4	Groupes ponctuels	13
1.4	Symétries des molécules	14
1.5	Symétries d'un cristal	15
1.5.1	Réseaux de Bravais	16
1.5.2	Cristallographie à deux dimensions	18
1.5.3	Cristallographie à trois dimensions	22
1.5.4	Autres ordres translationnels ou orientationnels	29
1.6	Exercices	30
2	Transformations des grandeurs physiques	33
2.1	Rotations	34
2.1.1	Vecteurs et scalaires	34
2.1.2	Tenseurs. Exemples physiques	36
2.2	Réflexions et Inversion/Parité	38
2.2.1	Scalaires et pseudoscalaires, vecteurs et pseudovecteurs	38
2.2.2	Covariance des grandeurs et invariance des lois de la physique	40
2.3	Principe de Curie	41
2.3.1	Applications physiques d'arguments de symétrie	41
2.3.2	Principe de Curie	43
2.3.3	Autres illustrations du principe de Curie	43
2.4	Translations	44

2.4.1	Diffraction cristalline, loi de von Laue	44
2.4.2	Zones de Brillouin	46
2.5	Modes propres de vibration d'une molécule	47
2.6	Exercice	48
3	Brisure spontanée de symétrie	49
3.1	Brisure spontanée de symétrie discrète	49
3.1.1	Un modèle mécanique	49
3.1.2	Ferromagnétisme	50
3.1.3	Modèle d'Ising	51
3.1.4	Un paradoxe ?	54
3.2	Brisure spontanée de symétrie continue	54
3.3	Autres exemples de symétries brisées spontanément	55
3.4	Exercice : Champ moyen	57
4	Groupes continus. Générateurs infinitésimaux. Lois de conservation	61
4.1	Transformations continues, générateurs infinitésimaux	61
4.1.1	Groupes des translations $\mathbb{R}$, $\mathbb{R}^d$	61
4.1.2	Groupe des phases $U(1)$	62
4.1.3	Groupe de rotation à deux dimensions $SO(2)$	63
4.1.4	Groupe de rotation à trois dimensions $SO(3)$	64
4.1.5	Groupe de Lie, algèbre de Lie, et leur dimension	65
4.2	Symétries continues et lois de conservation	68
4.2.1	Formalisme lagrangien et formalisme hamiltonien	68
4.2.2	Invariances et lois de conservation	70
4.3	Invariance de jauge de l'électrodynamique	73
4.4	Exercices en théorie classique des champs	74
5	Représentations des groupes	77
5.1	Définition et propriétés générales des représentations	77
5.1.1	Définitions de base	77
5.1.2	Représentations équivalentes. Caractères	78
5.1.3	Représentations réductibles et irréductibles	80
5.1.4	Représentations unitaires	82
5.1.5	Lemme de Schur	83
5.2	Représentations des groupes finis	84
5.2.1	Orthogonalité et complétude des représentations	85
5.2.2	Conséquences	87
5.2.3	Un groupe continu : $U(1)$	89
5.3	Produits directs de groupes ou de représentations ; décomposition de Clebsch-Gordan	90
5.3.1	Produits directs de groupes et leurs représentations	90
5.3.2	Produit tensoriel de représentations	91

5.4	Décomposition des représentations d'un groupe sur celles d'un sous-groupe . . .	93
5.5	Applications physiques	93
5.5.1	Modes de vibration des molécules	93
5.5.2	Vibration des cristaux	95
5.6	Exercices	95
6	Symétries en Mécanique Quantique	99
6.1	Théorème de Wigner	99
6.1.1	Théorème de Wigner.	99
6.1.2	Invariance d'un système quantique	100
6.2	Symétries discrètes en physique quantique : parité, renversement du temps, . . .	103
6.2.1	Transformation d'un état quantique sous l'effet de la parité	103
6.2.2	Brisure de la symétrie de parité en physique des particules	106
6.2.3	Renversement du temps	106
6.2.4	Brisure de la symétrie de renversement du temps en physique des particules	107
6.3	Symétries de translation, théorème de Bloch	108
6.4	Application à la physique moléculaire	109
6.5	Symétries de la physique des particules	111
7	Rotations à trois dimensions. Le groupe $SO(3)$	113
7.1	Rotations de $\mathbb{R}^3$, le groupe $SO(3)$, ses générateurs infinitésimaux	113
7.1.1	Le groupe $SO(3)$, groupe à trois paramètres	113
7.1.2	Générateurs infinitésimaux de $SO(3)$	115
7.1.3	Algèbre de Lie $so(3)$	117
7.2	Représentations de $so(3)$ et de $SO(3)$	117
7.2.1	Représentations de l'algèbre $so(3)$	118
7.2.2	Représentation de spin j du groupe $SO(3)$	120
7.3	Coordonnées sphériques. Harmoniques sphériques.	123
7.4	Produit tensoriel de représentations	126
7.4.1	Produit direct de représentations et l'“addition de moments angulaires” .	126
7.4.2	Coefficients de Clebsch-Gordan, symboles $3-j$	128
7.4.3	Théorème de Wigner–Eckart	130
7.5	Quelques applications physiques	132
7.5.1	Moments multipolaires	132
7.5.2	Etats propres de moment angulaire en Mécanique Quantique	133
7.5.3	L'isospin	133
7.6	2 Problèmes	135
7.6.1	Molécule de fullerène	135
7.6.2	Champ cristallin	138
8	Invariance relativiste. Les groupes de la Relativité	141
8.1	Principes de la Relativité restreinte. Conséquences physiques	141
8.1.1	Référentiels, synchronisation des horloges et simultanéité.	141

8.1.2	Postulats de la Relativité	142
8.1.3	Transformations de Lorentz.	142
8.1.4	Conséquences des transformations de Lorentz. Cinématique relativiste . .	145
8.2	Espace de Minkowski. Groupe de Lorentz $O(3,1)$	148
8.2.1	Métrie de l'espace de Minkowski. Groupes de Lorentz et de Poincaré .	148
8.2.2	Vecteurs et tenseurs contravariants et covariants.	149
8.2.3	Géométrie de l'espace de Minkowski, temps propre, lignes d'univers. . . .	150
8.3	Formalisme lagrangien et relativité. Cinématique relativiste	152
8.3.1	Particule libre relativiste	152
8.3.2	Cinématique relativiste : collisions et désintégration de particules	154

Bibliographie

- [1] C. Aslangul, Cours de magistère ENS-P6, *Physique de la Matière Condensée*
- [2] C. Aslangul, *Mécanique Quantique*, tome 2, de Broeck, 2009
- [3] J.-P. Blaizot et J.-C. Tolédano, *Symétries et physique microscopique*, Ellipses
- [4] É. Brézin, R. Mosseri et J.-C. Tolédano, *Symétries, symétries brisées*, Cours à l'École Polytechnique
- [5] J. Hladik, *La théorie des groupes en physique et chimie quantiques*, Masson 1995.
- [6] C. Kittel, *Introduction to Solid State Physics*, John Wiley & Sons (7th edition 1995)
- [7] L. Landau et E. Lifshitz, *Théorie de l'élasticité*, éditions Mir.
- [8] L. Landau et E. Lifshitz, *Électrodynamique des milieux continus*, éditions Mir.
- [9] L. Landau et E. Lifshitz, *Mécanique quantique*, éditions Mir.
- [10] L. Landau et E. Lifshitz, *Mécanique statistique*, éditions Mir.
- [11] W. Ludwig and C. Falter, *Symmetries in Physics*, Springer.
- [12] A. Messiah, *Mécanique quantique*, Dunod.
- [13] S. Sternberg, *Group theory and physics*, Cambridge University Press.
- [14] M. Tinkham, *Group Theory and Quantum Mechanics*, McGraw-Hill
- [15] L. Valentin, *Physique subatomique : noyaux et particules*, Hermann
- [16] S. Weinberg, *The Quantum Theory of Fields*, Cambridge Univ. Press

Parmi ces ouvrages, certains sont écrits dans un esprit assez proche de celui du présent cours, en particulier, [3, 4, 5]. On trouvera beaucoup de détails supplémentaires dans [9, 10, 14, 15] et de nombreux exemples et applications dans les autres.

Chapitre 1

Symétries géométriques. Des molécules aux cristaux

1.1 Généralités sur les symétries

1.1.1 Transformations de l'espace et symétries

Tout au long de ce cours on va s'intéresser à des *transformations* de l'espace.

Par “espace”, on entend l'espace physique, espace euclidien à d dimensions, avec $d = 3$ dans la situation la plus courante, mais $d = 1$ et $d = 2$ pouvant aussi jouer un rôle (chaînes polymériques unidimensionnelles, substrats et autres composés bidimensionnels, etc). Par la suite, ce contexte sera élargi à d'autres situations, telle $d = 4$ avec l'espace-temps de Minkovski (qui n'est plus euclidien), ou à des espaces plus abstraits, espaces de symétrie interne d'isospin ou de jauge, etc.

Par “transformation”, on entend une transformation ponctuelle inversible, qui associe à chaque point de l'espace considéré un point et un seul du même espace, et telle qu'à partir de tout point image, on puisse revenir de façon unique au point d'origine par une autre transformation du même type : cette condition d'inversibilité exclut par exemple des transformations comme la projection sur un sous-espace.

D'un point de vue mathématique, ces transformations forment un *groupe*. On peut en effet les composer : selon la convention usuelle, la composition des transformations t_1 puis t_2 effectuées dans cet ordre est notée $t_2 \circ t_1$, ou $t_2.t_1$ ou simplement t_2t_1 . Elle associe à tout point x de l'espace le point transformé $x \mapsto x' = t_2.t_1(x) = t_2(t_1(x))$. Cette opération de composition est bien associative (c'est-à-dire $t_3.(t_2.t_1) = (t_3.t_2).t_1$, pourquoi?), elle a un élément neutre (l'opération triviale $x \mapsto x$) et toute transformation admet un inverse (notre hypothèse d'inversibilité).

Par ailleurs, nous allons nous intéresser à des *symétries*, c'est-à-dire des transformations agissant sur un objet géométrique ou un système physique et en préservant des propriétés géométriques (par exemple les dimensions, ou les angles, ou le volume...), ou des propriétés physiques, structure interne, dynamique, etc. On distinguera donc le concept de symétrie de celui de transformation. Une transformation agit sur un système physique, ses coordonnées et ses autres degrés de liberté, mais n'est pas nécessairement une symétrie telle qu'on vient de la

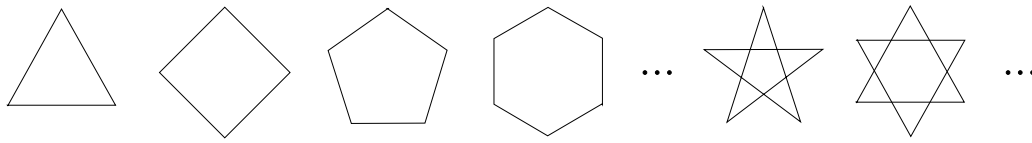


FIGURE 1.1 – Les premiers polygones réguliers convexes ou non-convexes.

définir. Pour nous toute symétrie est donc liée à une certaine *invariance* de l'objet considéré.

Tout comme les transformations, les symétries d'un objet géométrique ou d'un système physique forment un groupe : la composition de deux symétries est encore une symétrie etc. On parle donc du groupe de symétrie (ou d'invariance) de l'objet considéré.

1.1.2 Un petit album d'images...

L'observation de symétries est commune dans la vie courante comme dans le monde mathématique ou physique.

Elles se manifestent en géométrie, bien sûr ; en physique, depuis la mécanique (lois de Kepler, toupie, ...), à la minéralogie et la cristallographie et aux symétries plus cachées du monde atomique ou subatomique ; en botanique et en zoologie ; et dans les constructions humaines, architecturales, artistiques ou technologiques...

Quelques premiers exemples sont donnés, en géométrie, par les polygones réguliers, en commençant par le triangle équilatéral, Fig. 1.1 ; et par les polyèdres réguliers, qui en sont l'analogue à 3 dimensions. Les polyèdres réguliers sont des solides *convexes* dont toutes les arêtes ont même longueur et tous les sommets et toutes les faces sont de même nature. Ils ont été classifiés par l'école de Platon, d'où le nom de *solides platoniciens*. Ce sont les cinq solides de la figure 1.2, le tétraèdre, le cube et l'octaèdre, et le dodécaèdre et l'icosaèdre. Les quatre derniers forment des paires duales : les milieux des faces du cube sont les sommets d'un octaèdre régulier et vice versa, et de même pour le dodécaèdre et l'icosaèdre.

Le fait que la liste des polyèdres réguliers s'arrête là est une conséquence simple de la caractéristique d'Euler : les nombres V , A et F de sommets, d'arêtes et de faces satisfont $V - A + F = 2$ pour tout polyèdre, régulier ou pas. Mais dans un polyèdre régulier, si v est la valence (nombre de voisins) de chaque sommet et v' celle des faces (nombre de faces adjacentes à chaque face), on a $2A = vV = v'F$. Donc en reportant dans la relation d'Euler, $(\frac{2}{v} - 1 + \frac{2}{v'})A = 2$, et on se convainc rapidement que les seules valeurs possibles de (v, v') sont (3,3) tétraèdre ; (3,4) cube ou (4,3) octaèdre ; (3,5) dodécaèdre et (5,3) icosaèdre. Toute valeur supérieure de v ou v' conduit à un membre de gauche de la relation nul ou négatif.

Dans le monde cristallin, de très belles structures observées reflètent une symétrie microscopique sous-jacente.

- Ainsi la glace présente une grande variété de cristaux de symétrie hexagonale, Fig. 1.3 ;
- le quartz, qui chimiquement est de la silice SiO_2 , cristallise en de très beaux cristaux avec une symétrie d'ordre 6, Fig. 1.4, (mais en fait de nombreuses autres formes cristallines du quartz existent dans la nature, selon la température ou la pression de formation) ;
- on trouve beaucoup d'autres exemples, tel ce cristal d'alun de chrome, un octaèdre parfait de taille imposante (162 mm de diagonale), Fig. 1.4 ;

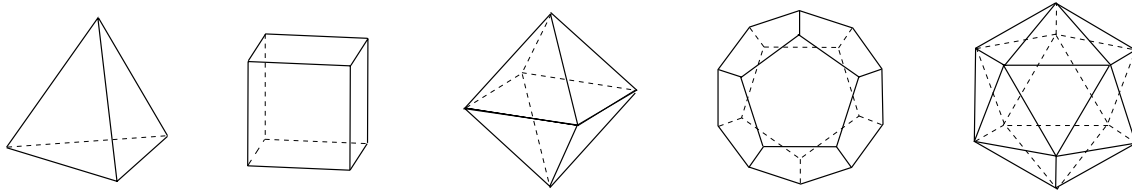


FIGURE 1.2 – Les cinq polyèdres réguliers convexes ou “solides platoniciens” : tétraèdre, cube, octaèdre, dodécaèdre et icosaèdre (cf http://fr.wikipedia.org/wiki/Solide_de_Platon pour plus de détails).



FIGURE 1.3 – Cristaux de glace. Quelles sont leurs symétries ? Source Google Cristaux de glace

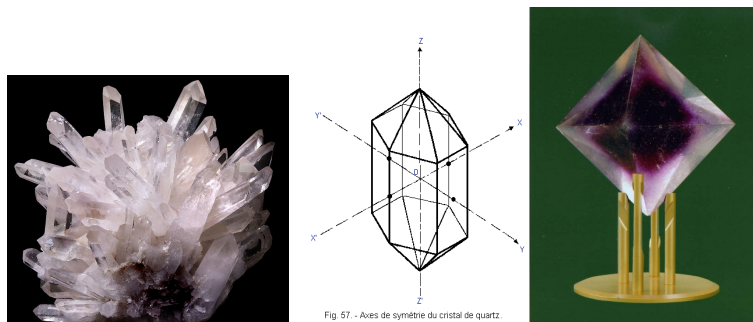


FIGURE 1.4 – Cristaux de quartz, avec au centre, le schéma d'un cristal idéal, avec ses axes de symétrie. À droite, cristal d'alun de chrome, Sources http://pagesperso-orange.fr/daniel.robert9/Digit/Digit_6TS2.html ; Source <http://www.dacg.nl/fotogalerij.html>, [R. Wagenaar, Europhysicsnews 40/1 2009 p 05]

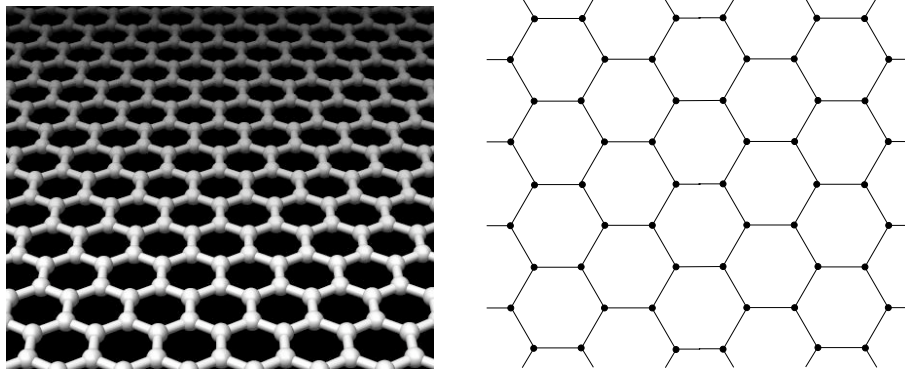


FIGURE 1.5 – Graphène Source Wikipedia

– le graphène, un matériau récemment découvert, formé d’un cristal de carbone bidimensionnel (d’une épaisseur d’un seul atome), qui exhibe une remarquable structure hexagonale “en nid d’abeilles” et des propriétés physiques –transport électronique en particulier– tout à fait inattendues, Fig. 1.5.

En physique moléculaire, des molécules simples (ammoniac NH_3 , pyramide à base équilatérale ; eau H_2O , quelles symétries ?), voir Fig 1.18, ou plus complexes telle le fullerène, Fig. 1.6.

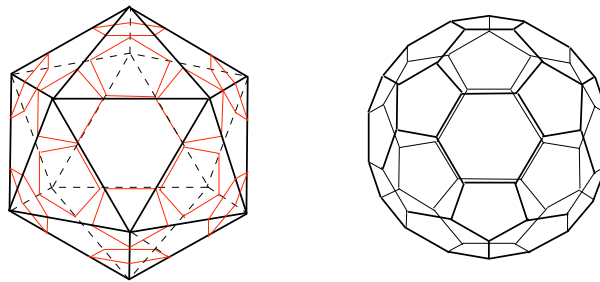


FIGURE 1.6 – Molécule C_{60} de fullerène, obtenue à partir d’un icosaèdre en tronquant (traits rouges) des calottes autour de chacun des 12 sommets et en plaçant aux 60 nouveaux nœuds des atomes de carbone.

En botanique et en zoologie, la Nature semble n’être jamais à court d’idées de symétries originales, cf Fig. 1.7.

Les mammifères ont en général une symétrie de réflexion d’ordre 2, au moins globalement, (Fig. 1.8), sur laquelle nous reviendrons plus longuement au § 1.2.

En matière architecturale et artistique, les exemples abondent aussi : Fig. 1.9, Fig. 1.10, mais aussi dans le domaine des objets manufacturés par l’homme ; ainsi les ailes des moulins à vent et autres éoliennes, les hélices ou turbines telles celles de la Fig. 1.11 se doivent d’avoir une symétrie par rotation (discrète) autour de leur axe aussi parfaite que possible pour éviter du “ballant”, cause de vibrations et d’usure, mais en revanche ne sont pas symétriques par réflexion, faute de quoi elles ne tourneraient pas !

C’est un bon exercice d’identifier toutes les symétries de chacun des exemples présentés sur



FIGURE 1.7 – carambole, sanguinaire du Canada, rudbeckie, marguerite; étoiles de mer, oursin, ruche d'abeilles, Sources shortcourses.com, <http://britton.disted.camosun.bc.ca/fibslide/jbfibslide.htm>, <http://antimon-santo.a.n.pic.centerblog.net/4yff7p39.jpg>

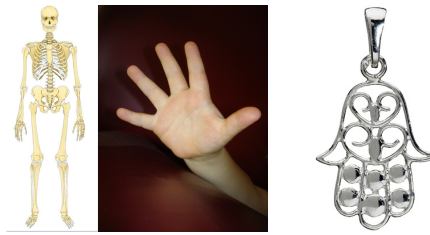


FIGURE 1.8 – Le squelette humain présente une symétrie par réflexion dans un plan, ce qui n'est évidemment pas le cas de la main, sauf quand elle est stylisée ...Source <http://www.pendentifs-argent.fr/images/pendentif-main-fatma.jpg>



FIGURE 1.9 – Palais de Versailles. Tour Eiffel. Dôme de l'église de Mostar, Malte. Sources <http://rpeyre.free.fr/versailles/05.htm>, M. Talon



FIGURE 1.10 – “grecques”, mosaïque d’Aquileia (Italie), azulejo de l’Alcazar, Séville ; gravure d’Escher

Source <http://mcescher.frloup.com/affichediapo.php?cat=6>



FIGURE 1.11 – éolienne, hélice de bateau ; turbine. Sources https://www.educ.usherb.ca/05667735/mes_voyages.htm,

<http://www.nauticexpo.fr/fabricant-bateau/helice-420.html> et <http://www.barrettsturbineengine.com/storage/turbine.jpg>

les figures précédentes. Quel est l'*ordre* du groupe, c'est-à-dire le nombre total de toutes ces symétries, y compris l'identité, (ne pas oublier les réflexions!).

Les symétries qu'on vient de considérer ici sont des *isométries* de l'espace à deux ou trois dimensions, c'est-à-dire des opérations préservant les longueurs et les angles (peut-être au signe près, voir ci-dessous). Nous rencontrerons plus tard des symétries plus générales, soit des déplacements dans des espaces de dimension plus élevée (telles les transformations relativistes de l'espace-temps), soit des transformations de nature différente, dilatations ou changements d'échelle, importantes dans l'étude des phénomènes critiques ou le comportement asymptotique de haute énergie en physique des particules, ou d'autres plus abstraites encore, comme les rotations d'isospin ou les transformations de jauge.

Ce premier chapitre va se concentrer sur les transformations géométriques de base que sont les déplacements.

1.1.3 Isométries et déplacements

On appelle *isométrie* une transformation géométrique de l'espace euclidien $\mathbb{R}^d$ qui préserve les longueurs. (C'est donc une symétrie de la géométrie euclidienne : tout théorème vrai pour une figure l'est aussi pour ses transformées par toute isométrie). Par les règles de la trigonométrie, une isométrie préserve donc aussi les angles au signe près. On appelle *déplacement* une isométrie qui préserve l'orientation (de l'espace ou d'un repère), donc les angles orientés. On démontre qu'à 1 dimension, les isométries sont les translations et les réflexions par rapport à un point de la droite. Seules les translations sont des déplacements. En dimension 2 ou 3, peuvent aussi intervenir des rotations (d'angle θ , autour d'un point à $d = 2$, autour d'un axe $\vec{u}$ à $d = 3$). On démontre que toute isométrie est le "produit" (ou la composition) de trois types de transformations : les translations, les rotations et les symétries par rapport à une droite (à deux dimensions) ou à un plan (à trois dimensions).

Composition des isométries

Il est clair d'après ce qui précède que la composition de ces isométries ou déplacements produit d'autres isométries ou déplacements : les isométries dans une dimension donnée forment un *groupe*, au sens mathématique du mot, et il en est de même des déplacements.

Les translations et les rotations forment indépendamment des sous-groupes d'isométries. La composée de deux translations de vecteurs $\vec{T}_1$ et $\vec{T}_2$ est la translation de vecteur $\vec{T} = \vec{T}_1 + \vec{T}_2$; la composée de deux rotations de même centre (à $d = 2$) ou d'axes concourants (à $d = 3$) est une autre rotation. La composée de deux réflexions par rapport à deux plans P et P' se coupant le long d'une droite D et faisant un angle α est une isométrie respectant les orientations et laissant D invariante, c'est donc une rotation d'axe D dont on calcule aisément l'angle $= 2\alpha$, voir fig. 1.12. Que se passe-t-il si les deux plans sont parallèles ?

On étudie en géométrie la composition des autres types d'isométries, mais nous ne poursuivons pas cette discussion ici.

Dans ce premier chapitre, nous allons nous restreindre la plupart du temps à des transforma-

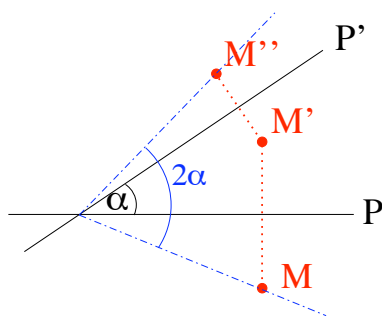


FIGURE 1.12 – Le produit de deux réflexions par rapport à deux axes P et P' du plan faisant un angle α est une rotation d'angle 2α . La même conclusion est vraie à $d = 3$ pour le produit de deux réflexions dans des plans P et P' se coupant selon une droite D avec un angle α ; la figure peut alors être considérée comme la vue dans un plan orthogonal à la droite D .

tions de translation et de rotation *discrètes*, c'est-à-dire telles que pour chaque transformation il n'en existe pas d'autre arbitrairement proche. Mieux encore, les rotations seront en nombre fini. Pour un groupe fini, c'est un exercice souvent utile et toujours instructif de construire la table de composition des opérations du groupe, voir TD.

1.2 Réflexion-miroir et inversion/parité

1.2.1 Isométries impropres

Dans ce paragraphe, nous nous intéressons au groupe de transformations le plus simple qui soit, le groupe à deux éléments engendré par une réflexion, par rapport à un point, une droite ou un plan, respectivement à une, deux ou trois dimensions. Cette réflexion est désignée en général par *réflexion-miroir*.

En dimension impaire, on peut aussi considérer l'opération de symétrie par rapport à un point, disons l'origine O , appelée *inversion* (d'espace) par les cristallographes (à ne pas confondre avec l'inversion en géométrie!¹) et *parité* en physique atomique ou des particules, et définie par $\overrightarrow{OM'} = -\overrightarrow{OM}$. Elle a pour effet de changer le signe de toutes les coordonnées dans un repère attaché à cette origine O , $\overrightarrow{OM} = (x, y, z) \mapsto \overrightarrow{OM'} = (-x, -y, -z)$. En dimension paire, l'inversion/parité est une rotation, cf. le cas $d = 2$. En dimension 1, elle s'identifie à la réflexion (miroir) par rapport à l'origine. En dimension 3, sa composition avec la réflexion par rapport au plan Oxy ($z \mapsto -z$) donne la rotation d'angle π autour de l'axe des z .

Toutes ces opérations, réflexions-miroir ou inversion/parité en dimension impaire, ont en commun une propriété géométrique importante : celle de renverser l'orientation d'un repère orienté, ou le signe d'un angle dans un plan orienté. On parle d'isométries "impropres".

Chacune de ces transformations impropres peut être considérée seule ou composée avec d'autres transformations, propres (rotations) ou impropres. C'est le cas quand on cherche à déterminer le groupe de symétrie de rotation et réflexion d'un objet donné. On verra des

1. qui est la transformation $\vec{x} \mapsto k \frac{\vec{x}}{x^2}$ mais que nous n'utiliserons pas dans ce cours.

exemples au § 1.3.3 ci-dessous.

1.2.2 Réflexion-miroir

En physique, en chimie ou dans les sciences de la vie, ces opérations de réflexions-miroir jouent un rôle important. Mais aussi dans la vie courante ! Une main (Fig. 1.8), une vis, un tire-bouchon, une clé en général ... sont des objets *chiraux*², c'est-à-dire non superposables à leur image-miroir. Comme on l'a dit, une éolienne, une turbine ne tourneraient pas si elles n'étaient pas chirales, Fig. 1.11. En chimie, une molécule est en général chirale. Des molécules simples comme l'eau, l'ammoniac, le benzène, sont *achirales*, c'est-à-dire superposables à leur image-miroir. Des molécules plus complexes telles les protéines sont le plus souvent chirales. Dans le monde minéral, elles apparaissent généralement sous forme de mélange *racémique*, c'est-à-dire avec des populations égales de chaque chiralité.

La question se pose aussi en mathématiques, en théorie des nœuds, par exemple : le "nœud de trèfle" ci-contre est-il ou non chirale, c'est-à-dire identique (à déformation près) à son image miroir ? Qu'en est-il de l'"entrelacs" (nœud à trois composantes) de la mosaïque de la figure 1.10 ?



Le concept de chiralité est également important en physique. La chiralité se manifeste ainsi dans les cristaux, par exemple le quartz, (chimiquement, on l'a vu, de la silice SiO_2), qui vient en deux variétés dans lesquelles les tétraèdres (Si entouré de 4 O) forment des hélices de chiralités opposées. On y reviendra au paragraphe suivant.

Un phénomène important où apparaît la chiralité est celui de lumière polarisée circulairement.

On rappelle que dans un faisceau de lumière monochromatique de vecteur d'onde $\vec{k}$ et de fréquence ω , le champ électrique est donné par (la partie réelle de) la fonction de la position $\vec{x}$ et du temps t

$$\vec{E}(\vec{x}, t) = A_1 \vec{e}_1 \exp(i(\vec{k} \cdot \vec{x} - \omega t + \alpha_1)) + A_2 \vec{e}_2 \exp(i(\vec{k} \cdot \vec{x} - \omega t + \alpha_2))$$

où $(\vec{e}_1, \vec{e}_2, \vec{k})$ forment un trièdre orthonormé et les amplitudes A_1 et A_2 sont choisies réelles positives. À un point $\vec{x}$ donné, l'extrémité de $\vec{E}$ décrit en fonction du temps une ellipse en général (polarisation elliptique), un segment de droite (polarisation droite) si A_1 ou A_2 s'annule, et un cercle (polarisation circulaire) si $A_1 = A_2$ et $|\alpha_1 - \alpha_2| = \pi/2$. Dans ce dernier cas, la polarisation est dite gauche ou droite selon que $\alpha_1 - \alpha_2 = \pm\pi/2$.

L'image-miroir d'un faisceau de polarisation gauche est un faisceau de polarisation droite.

C'est d'ailleurs en observant le "pouvoir rotatoire" de certains cristaux ou molécules qu'on détermine leur caractère chiral ou non. Un faisceau lumineux polarisé *linéairement* voit son plan de polarisation tourner vers la droite (quand on observe le rayon) à la traversée d'une substance "dextrogyre", vers la gauche pour une substance "lévogyre", voir fig 1.13. C'est l'effet observé d'abord par Arago (sur le quartz) et Biot (sur un corps en solution) au début du XIX^{ème} siècle, et relié à la structure microscopique chirale de la substance selon une hypothèse de Herschel.

Dans un travail fameux, Louis Pasteur (en 1848, à l'âge de 25 ans)³ résout un paradoxe : on avait découvert que l'acide tartrique $\text{HOOC-CHOH-CHOH-COOH}$ (ou son sel, le paratartrate de sodium et d'ammonium)

2. du grec *kheir*, main

3. "Aussi bien ce sont mes premières joies de chercheur. Celles qu'inspire la science n'ont pas moins de poésie que les autres." (L. Pasteur)

existait sous deux formes de même formule chimique, l'une, sous-produit de la vinification, venant des dépôts dans les tonneaux de vin, est dextrogyre ; l'autre forme, de synthèse, nommée acide paratartrique, n'a pas de pouvoir rotatoire. Pasteur montre que l'acide paratartrique est racémique, c'est-à-dire qu'il vient en deux espèces de chiralités opposées ; après cristallisation et tri manuel méticuleux des cristaux, il vérifie le pouvoir rotatoire opposé de chacune des deux chiralités.

On constate que le saccharose (sucre extrait de la betterave) est dextrogyre, tandis que le fructose, présent dans les fruits, est lévogyre.

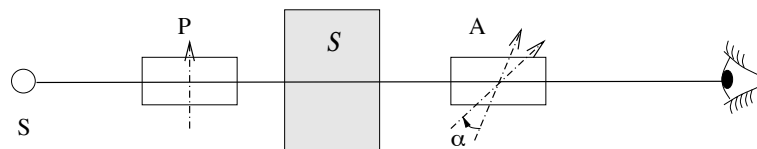


FIGURE 1.13 – Pouvoir rotatoire d'une substance : un faisceau polarisé linéairement par un polariseur P traverse ensuite une substance "optiquement active" $\mathcal{S}$: on constate que l'analyseur A doit être tourné d'un angle α pour retrouver l'extinction du rayon lumineux.

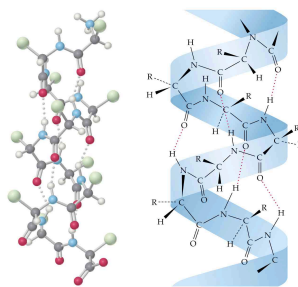


FIGURE 1.14 – Kératine, Source http://wps.prenhall.com/wps/media/objects/602/616516/Chapter_24.html

Dans le monde vivant, la chiralité est en effet très répandue. L'ADN, les protéines, les acides aminés, ... sont des molécules chirales. Par exemple, la kératine, protéine constitutive des ongles, cheveux et autres cornes a une structure hélicoïdale, voir fig 1.14. Le goût et d'autres activités biologiques dépendent de façon cruciale de la chiralité de la molécule.

L'origine de ces propriétés chirales du vivant demeure pour une large part mal comprise. Ainsi, pourquoi une seule chiralité, et toujours la même, (droite), apparaît-elle dans tous les ARN de toutes les espèces ? Pourquoi les acides aminés des protéines sont-ils gauches ? et les sucres présents dans l'ADN des organismes vivants droits ?

1.3 Rotations et réflexions de $\mathbb{R}^2$ et de $\mathbb{R}^3$

1.3.1 Groupes $O(d)$ et $SO(d)$

Les isométries de $\mathbb{R}^d$ qui "fixent un point" (c'est-à-dire laissent un point invariant) sont, on l'a vu, les rotations (autour de ce point) et réflexions (par rapport à ce point ou à une droite ou un plan le contenant) et leurs produits. Si on choisit le point fixe comme origine de l'espace, ces transformations peuvent être considérées comme applications linéaires de l'espace

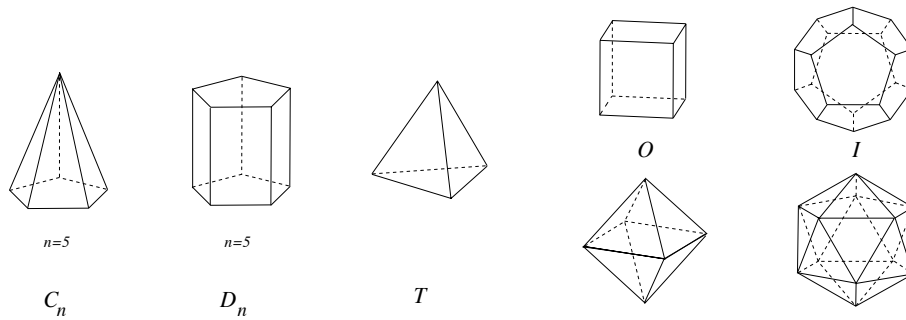


FIGURE 1.15 – Solides dont les groupes de symétrie de rotation décrivent tous les sous-groupes finis de $\text{SO}(3)$. Les groupes C_n et D_n sont illustrés ici par le cas $n = 5$.

$\mathbb{R}^d$ préservant la norme $|\vec{r}|^2$ de tout vecteur $\vec{r}$, et sont donc décrites par des matrices orthogonales O , $O.O^T = \mathbb{1}$.

Rappelons l'argument : si $\vec{r}$ est représenté par un vecteur colonne X , la transformation linéaire $X \mapsto X' = OX$ est une isométrie si $|\vec{r}'|^2 = X'^T \cdot X' = X^T O^T O X = X^T \cdot X = |\vec{r}|^2$, donc $O^T \cdot O = \mathbb{1}$ ou de façon équivalente, $O.O^T = \mathbb{1}$. En prenant le déterminant de cette relation, $\det O.O^T = (\det O)^2 = \det \mathbb{1} = 1$, on voit qu'une telle matrice orthogonale a un déterminant égal à ± 1 . Par ailleurs, l'orientation d'un repère de d vecteurs $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_d$ est définie par le signe du déterminant $\det(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_d)$. Par une transformation $\vec{a}_i \mapsto \vec{a}'_i$, $i = 1, \dots, d$, par la matrice O , $\det(\vec{a}'_1, \vec{a}'_2, \dots, \vec{a}'_d) = \det O \det(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_d)$. L'orientation du repère change si et seulement si $\det O = -1$.

Ces matrices O forment un groupe de matrices noté $O(d)$. Ce groupe orthogonal a comme sous-groupe les matrices orthogonales de déterminant 1, noté $\text{SO}(d)$ (groupe spécial orthogonal). Seules les matrices de $\text{SO}(d)$ décrivent des rotations. Les matrices de déterminant -1 décrivent des “transformations impropres”, qui sont d'une façon générale le produit d'un nombre impair de réflexions.

1.3.2 Sous-groupes finis des groupes de rotation à $d = 2$ et $d = 3$

Il est utile de connaître les sous-groupes finis du groupe des rotations à $d = 2$ et $d = 3$. À deux dimensions, toute rotation d'angle $2\pi/n$ autour de l'origine engendre un sous-groupe d'ordre n . C'est le *groupe cyclique* C_n , commutatif (ou “abélien”). La situation à trois dimensions, où le groupe $\text{SO}(3)$ est non commutatif (ou “non abélien”) comme tous les amateurs du Rubik Cube le savent !, est beaucoup plus intéressante. On démontre [13] que tous les sous-groupes finis du groupe $\text{SO}(3)$ figurent dans la liste suivante : groupes cycliques C_n de rotation d'angle $2\pi/n$ autour d'un axe $\vec{u}$, groupes “diédraux” D_n , d'ordre $2n$, obtenus en adjoignant aux n rotations de C_n une rotation de π autour d'un axe orthogonal à $\vec{u}$; et les trois groupes de symétrie de rotation des solides platoniciens (voir § suivant et exercice en TD). On peut donc résumer la situation en disant que tous les groupes finis de $\text{SO}(3)$ apparaissent comme groupe de symétrie de rotation d'un des solides de la figure 1.15.

1.3.3 Groupes d'invariance des polygones et des polyèdres

Il est instructif et utile pour la suite de consacrer un peu de temps à l'étude des groupes d'invariance des figures géométriques simples, les polygones et les polyèdres réguliers.

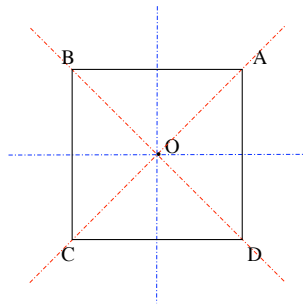


FIGURE 1.16 – Symétries de rotation et de réflexion du carré.

Exemple : groupe du carré (voir TD). On désigne ainsi le groupe d'invariance (d'isométrie) du carré. Ce groupe comporte des rotations d'angle multiple de $\pi/2$ autour du centre O du carré : le sous-groupe de rotation est donc d'ordre 4. Mais il contient aussi des réflexions par rapport aux deux médiatrices des paires de côtés opposés, et celles par rapport aux deux diagonales. Noter que l'inversion par rapport à son centre n'est autre que la rotation de π autour de son centre (à deux dimensions, l'inversion respecte l'orientation). Le groupe complet du carré est donc d'ordre 8.

Les invariances de la figure se traduisent par une permutation de ses sommets. Pour calculer la composition de deux transformations et construire la table de multiplication, il peut être plus commode de composer les deux permutations correspondantes. Ainsi par exemple, la réflexion par rapport à la diagonale AC donne la permutation (ADCB) de la configuration initiale (ABCD). La rotation de $-\pi/2$ donne de même (BCDA), donc en les composant, on trouve la permutation (BADC) qui correspond à la réflexion par rapport à la médiatrice de AB. Le produit effectué dans l'ordre inverse conduit à la réflexion par rapport à l'autre médiatrice.

On peut répéter cette discussion pour les autres polygones réguliers, cf fig 1.1. Mais plus intéressant, on peut aussi l'étendre aux polyèdres réguliers de la fig 1.2. Leur apparition fréquente dans la suite, surtout celle du tétraèdre et du cube, justifie qu'on étudie leur groupe d'invariance, voir aussi exercice en TD.

Un tétraèdre régulier (voir figure 1.17(a)) est invariant par les rotations de $\pm 2\pi/3$ autour de chacun des 4 axes joignant un sommet au centre de gravité de la face opposée ; et aussi par les rotations de π (demi-tour) autour de chacun des 3 axes joignant les milieux des paires de côtés opposés (AB et CD par exemple). Toute rotation de cet ensemble est *engendrée* par deux d'entre elles, une de chaque type, comme on s'en convainc en construisant la table de multiplication. (Selon l'observation faite plus haut pour le carré, il est utile de regarder ces transformations comme des permutations des 4 sommets.) On désigne par T ce *groupe du tétraèdre*, groupe des rotations le laissant invariant. Il est donc d'ordre $2 \times 4 + 3 + 1 = 12$.

Pour faire bien apparaître les symétries impropres (réflexions et autres) du tétraèdre, il est utile de le voir comme inscrit dans un cube (voir figure 1.17(b)). Le tétraèdre est invariant par réflexion dans l'un des 6 plans bissecteurs passant par une de ses arêtes et le milieu de l'arête opposée. Mais il est aussi invariant sous la combinaison d'une inversion par rapport au centre du cube suivie d'une rotation de $\pm\pi/2$ autour d'un des 3 axes perçant le cube aux centres d'une paire de faces opposées. Le groupe complet d'isométrie du tétraèdre noté T_d est donc d'ordre 24, qui est le nombre de permutations des 4 sommets : on a bien ainsi "épuisé" toutes

les symétries du tétraèdre.

Pour le cube, (ou ce qui revient au même, pour l'octaèdre), une discussion du même genre conduit à un groupe de rotation O d'ordre 24, fait de $4 \times 2 = 8$ rotations de $\pm 2\pi/3$ autour des 4 diagonales, de $3 \times 3 = 9$ rotations de $\pi/2, \pi, 3\pi/2$ autour des axes joignant les centres des paires de faces opposées, et des 6 rotations de π autour des axes joignant les milieux de paires d'arêtes opposées, plus l'identité bien sûr.

Comme dans tout objet géométrique fini ayant un centre de symétrie (centre d'inversion), toute isométrie du cube est soit une rotation (y compris l'identité, rotation d'angle nul), soit le produit d'une telle rotation par l'inversion. En effet considérons une isométrie : ou bien elle préserve l'orientation et c'est donc une rotation, ou bien elle ne la préserve pas, mais alors son produit par l'inversion la préserve et on est ramené au cas précédent. On dit donc que le groupe complet noté O_h du cube est le produit de O par le groupe $\mathbb{Z}_2$ engendré par l'inversion, et $O_h = O \times \mathbb{Z}_2$ est d'ordre 48.

Enfin le groupe I de rotations de l'icosaèdre est d'ordre 60 : il comporte 6×4 rotations d'angle multiple de $2\pi/5$ autour d'un des 6 axes joignant des paires de sommets opposés ; 10×2 rotations de $\pm 2\pi/3$ autour d'un axe joignant les milieux des faces triangulaires opposées ; et 15 rotations de π autour de l'un des axes joignant les milieux des paires d'arêtes opposées ; plus l'identité. Voir Fig. 1.17. Le groupe I_h d'isométries, propres ou impropres, selon l'argument donné pour le cube est de la forme $I \times \mathbb{Z}_2$, il a l'ordre 120.

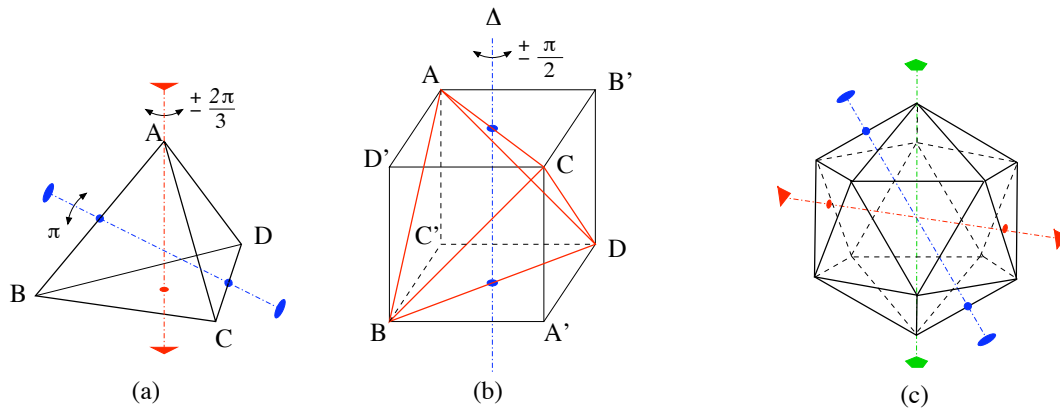


FIGURE 1.17 – Symétries propres et impropres du tétraèdre : (a) rotations d'angle $\pm 2\pi/3$ et d'angle π ; (b) le tétraèdre est invariant par l'inversion par rapport au centre O du cube suivie d'une rotation par exemple de $\pi/2$ autour de l'axe Δ qui amène B' sur A , D' sur C , etc ; et aussi par la réflexion par rapport au plan passant par CD et $D'C'$. (c) Symétries de rotation de l'icosaèdre.

1.3.4 Groupes ponctuels

Ce qui vient d'être dit sur les groupes d'invariance des polyèdres réguliers est le début de la classification des *groupes ponctuels* à 3 dimensions, c'est-à-dire de tous les groupes finis engendrés par des rotations, réflexions ou l'inversion par rapport à un point. Ce sont donc les sous-groupes finis du groupe $O(3)$. Le mot "ponctuel" signifie que chacun de ces groupes laisse invariant un point : ce point est le centre d'inversion éventuel, et les axes de rotation ou les plans de réflexion passent par ce point. La terminologie la plus simple pour dénoter ces

groupes est celle de Schönflies⁴ utilisée en spectroscopie ([9], § 93) : C_n pour le groupe cyclique engendré par la rotation de $2\pi/n$ autour d'un axe qu'on peut imaginer vertical ; C_{nh} si un miroir perpendiculaire à l'axe est aussi présent ; C_{nv} , si le miroir est parallèle à l'axe ; S_{2n} pour le groupe engendré par une "rotation-réflexion" d'ordre $2n$ (ce qui signifie un produit d'une rotation d'angle $2\pi/2n$ par une inversion par rapport à un point de l'axe)⁵. D_n groupe diédral engendré par C_n et les rotations de π autour d'un axe orthogonal, D_{nh} a en outre un miroir perpendiculaire à l'axe et D_{nd} a, outre les générateurs de D_n , un miroir vertical faisant un angle de π/n avec l'axe de rotation horizontal ; T et O sont les groupes de rotation du tétraèdre et de l'octaèdre ou cube ; T_d et O_h en sont les groupes complets d'invariance (rotations, inversions et réflexions), cf. § 1.3.3, et $T_h = T \times \mathbb{Z}_2$ combine les rotations de T avec une inversion ; enfin I et I_h sont les groupes d'invariance de l'icosaèdre rencontrés plus haut. Pour un exposé mathématique précis de la classification des sous-groupes finis de $SO(3)$ et $O(3)$, voir [13].

1.4 Symétries des molécules

Les premiers objets physiques dont nous pouvons analyser les propriétés de symétrie sont des molécules. Nous nous bornerons pour le moment à en décrire les symétries géométriques, avant de discuter aux chapitres 5 et 6 les implications sur leurs spectres d'excitation etc.

Selon la conformation spatiale de la molécule, le groupe de symétrie est évidemment très différent, voir Fig. 1.18.

Pour des molécules linéaires (en particulier diatomiques, telles H_2 , HCl , ... mais pas nécessairement, par exemple l'acétylène C_2H_2), cette symétrie inclut une rotation d'angle quelconque autour de l'axe Δ de la molécule, et la réflexion dans tout plan contenant cet axe de la molécule. Avec les notations de Schönflies (cf § 1.3.4), on parlera du groupe de rotation C_∞ et du groupe complet $C_{\infty v}$. Mais la molécule peut avoir une symétrie supplémentaire par rotation de π autour d'un axe orthogonal à Δ et l'inversion par rapport à son centre de gravité : c'est le cas des molécules diatomiques composées de deux atomes ou de deux motifs d'atomes identiques, O_2 , H_2 , etc, C_2H_2 , et le groupe est alors noté D_{nh} .

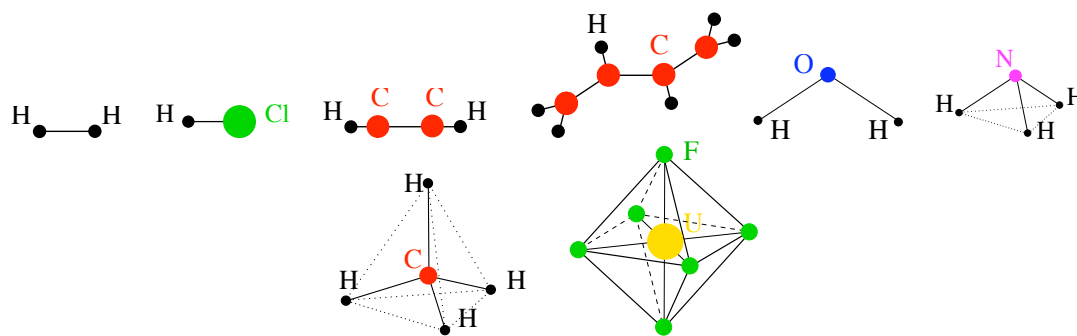


FIGURE 1.18 – Molécules H_2 , HCl , acétylène C_2H_2 , butadiène C_4H_6 , H_2O , NH_3 , CH_4 , hexafluorure d'uranium UF_6

4. Arthur Moritz Schönflies (ou Schoenflies) (1853-1928), mathématicien allemand, auteur d'importants travaux en cristallographie.

5. Pourquoi se limite-t-on à des S d'ordre pair ? Examiner par exemple ce qu'est le groupe engendré par une rotation-réflexion d'ordre 3.

La molécule d'eau n'est pas linéaire : il y a un angle de 104° environ entre les deux liaisons O-H. Le groupe de symétrie est le produit du groupe $\mathbb{Z}_2$ de réflexion dans le plan de la molécule par le groupe $\mathbb{Z}_2$ de réflexion dans le plan médiateur.

Exemples d'autres molécules planes : benzène, ozone, éthylène $\text{H}_2\text{C}=\text{CH}_2$, etc, ou non planes : éthane $\text{H}_3\text{C}-\text{CH}_3$, etc.

La molécule d'ammoniac NH_3 a la forme d'une pyramide à base triangulaire (tétraèdre irrégulier), avec un angle $\widehat{\text{H}\text{N}\text{H}} = 107,8^\circ$ et une distance interatomique $d(\text{H}-\text{N}) = 1,02 \text{ \AA}$. Le groupe de symétrie de rotation est C_3 , cf. § 1.3.2. S'y ajoutent les réflexions, formant le groupe C_{3v} , voir § 1.3.4.

Le méthane CH_4 a la configuration d'un tétraèdre régulier, avec le C au centre de symétrie, et $d(\text{H}-\text{C}) = 1,08 \text{ \AA}$. Son groupe de symétrie est le groupe T_d étudié plus haut (cf. 1.3.3). On rencontre aussi des configurations plus compliquées, par exemple octaédrique (hexafluorure d'uranium UF_6) ou icosaédrique, avec le fullerène C_{60} , voir fig 1.6. Le groupe de symétrie de rotation du fullerène est donc I , (cf § précédent), le groupe complet, rotations et réflexions, est I_h .

1.5 Symétries d'un cristal

Un cristal est un arrangement régulier, *périodique*, de constituants –atomes, ions, molécules. Dans l'idéal, le cristal s'étend à l'infini dans toutes les directions. De façon plus réaliste, on suppose seulement que ses dimensions sont très grandes devant les dimensions microscopiques entre constituants.

Pour évaluer ces dimensions microscopiques, on suppose que les constituants sont essentiellement au contact. Pour des cristaux atomiques ou ioniques, les distances sont donc de l'ordre de quelques $\text{\AA}$. En comparaison, un calcul simple conduit à une distance moyenne entre atomes dans un gaz de l'ordre de $a = (22,4 \cdot 10^{-3} / 6 \cdot 10^{23})^{1/3} \text{ m} \approx 30 \text{ \AA}$ (pourquoi ? rétablir l'argument). Dans un cristal biologique, les distances entre les macromolécules constituantes sont évidemment plus élevées.

Les cristaux les plus intéressants pour le physicien sont tridimensionnels, mais le progrès des techniques –épitaxie et autres– de fabrication de “couches minces” ainsi que leur plus grande simplicité théorique amènent à s'intéresser aussi aux objets bidimensionnels. Rien n'interdit d'étudier également la situation à une dimension (cf exercices en TD) ou en dimension 4 et plus.

Les cristaux en tant qu'objets géométriques ont des propriétés de symétrie manifestes. La plus évidente est celle d'invariance par les translations (les périodes) dont on a supposé l'existence. Cela va nous amener à examiner des *réseaux* et à en étudier les différents types de symétries supplémentaires, rotations, réflexions, etc les laissant invariants ou laissant invariante la structure cristalline dans son ensemble. Ce faisant, on va être conduit à une *classification* complète des différents types de structures cristallines. L'intérêt pour le physicien de ce travail de classification est que des substances cristallisant selon la même structure ont des propriétés physiques communes, obéissent aux mêmes règles de sélection etc, cf Chapitre 2.

Par définition, un cristal jouit de propriétés d'invariance par les translations ou périodes d'un réseau sous-jacent (cf § suivant), dont les nœuds peuvent être ou non localisés sur des atomes

du cristal. L'ensemble du cristal peut donc être reconstruit à partir d'un motif élémentaire, ou *maille*, sur lequel on fait agir les translations. Chaque maille peut elle-même contenir un ou plusieurs atomes. Les symétries du cristal vont donc résulter de celles du réseau sous-jacent et de celles de la maille élémentaire.

La discussion qui va suivre va s'organiser en plusieurs étapes :

- classification des *groupes ponctuels*, d'isométries qui laissent fixe un point du réseau ;
- classification des réseaux périodiques (définis plus bas) et de leurs types de symétries ;
- classification des *groupes d'espace*, qui incluent les translations.

Dimension d	Groupes ponctuels	Réseaux de Bravais	Systèmes réticulaires	Groupes d'espace
$d = 1$	2	1	1	2
$d = 2$	10	5	4	17
$d = 3$	32	14	7	230

Nombre de classes de réseaux de Bravais et de groupes de symétries cristallines, selon la dimension d d'espace

Les *systèmes réticulaires* (on dit aussi parfois système cristallin) rassemblent les cristaux (les groupes d'espace) dont les réseaux de Bravais ont le même groupe ponctuel.

1.5.1 Réseaux de Bravais

Définition

On appelle *réseau* en mathématiques, et *réseau de Bravais*⁶ en cristallographie, un ensemble infini discret $\mathcal{R}$ de points de l'espace $\mathbb{R}^d$ obtenus comme combinaisons linéaires entières de d vecteurs de base linéairement indépendants $\vec{a}_1, \dots, \vec{a}_d$ appelés *périodes*

$$\mathcal{R} : \quad \vec{t} = \sum_{i=1}^d n_i \vec{a}_i \quad n_i \in \mathbb{Z} . \quad (1.1)$$

On appelle *maille élémentaire*⁷ le solide géométrique, parallélogramme à $d = 2$, parallélépipède à $d = 3$, (et “parallélotope” en dimension quelconque), sous-tendu par les vecteurs $\vec{a}_1, \dots, \vec{a}_d$, c'est-à-dire l'ensemble des points M tels que $\overrightarrow{OM} = \sum_{i=1}^d x_i \vec{a}_i$ avec $0 \leq x_i \leq 1$.

Le choix de l'origine du réseau, dans un cristal, est arbitraire. Il est commode, mais pas obligatoire, de le prendre sur un atome. Le choix des périodes $\vec{a}_i$ et donc celui de la maille n'est pas unique non plus. Si on effectue un changement de base $\vec{a}'_i = \sum_{j=1}^d \vec{a}_j S_{ji}$ (et donc des composantes $n'_i = \sum_{j=1}^d S_{ij} n_j$), il faut que les S_{ij} soient entiers (de $\mathbb{Z}$) et (pour que la matrice inverse soit elle aussi à coefficients entiers) que $\det S = 1$ (matrice “unimodulaire”). Montrer que la condition $\det S = 1$ implique que le volume de la maille élémentaire est indépendant de son choix. Exemple à $d = 2$, cf figure 1.19.

Montrer pourquoi le “réseau en nid d'abeilles” tel celui de la figure 1.5 n'est pas un réseau de Bravais mais plutôt une superposition de deux réseaux triangulaires.

6. Auguste Bravais, 1811–1863, a classé vers 1850 les réseaux dans $\mathbb{R}^3$ selon leurs groupes de symétrie en 7 “systèmes”, et plus finement en 14 “classes”, comme on va l'expliquer.

7. Les mathématiciens parlent de “domaine fondamental”.

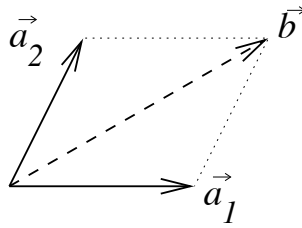


FIGURE 1.19 – Les réseaux à $d = 2$ définis par $\vec{a}_1$ et $\vec{a}_2$, ou par $\vec{a}_1$ et $\vec{b} = \vec{a}_1 + \vec{a}_2$, sont identiques, pourquoi ?

Symétries de rotation d'un réseau

Supposons qu'un réseau est invariant par une rotation d'angle θ défini modulo 2π radians = 360° . Pour un réseau bidimensionnel, il s'agit d'une rotation dans le plan, c'est-à-dire autour d'un point de ce plan (ou si on préfère autour d'un axe orthogonal au plan) ; à trois dimensions, la rotation est autour d'une direction $\vec{u}$ de l'espace. Comparons la description de cette rotation dans deux repères de coordonnées. Dans un repère orthonormé choisi dans le plan de rotation, (éventuellement complété à $d = 3$ par le vecteur unitaire $\vec{u}$), la matrice de rotation s'écrit

$$O = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \text{ respectivement } O = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Dans une base de périodes du réseau, la matrice O' n'est *a priori* plus orthogonale mais elle doit n'avoir que des coefficients *entiers* puisque la rotation transforme les périodes en périodes. Les matrices O et O' sont semblables, $O = XO'X^{-1}$ (avec X la matrice de changement de base), et doivent donc avoir la même trace. Cela impose que $2 \cos \theta$ est entier (positif, nul ou négatif) de valeur absolue ≤ 2 . Il ne peut donc prendre que les 5 valeurs $0, \pm 1, \pm 2$, ce qui restreint les angles de rotation possibles :

$$2 \cos \theta \in \mathbb{Z} \implies |\theta| = 2\pi/n, n = 1, 2, 3, 4, 6, \text{ soit } \theta = 180^\circ, \pm 120^\circ, \pm 90^\circ, \pm 60^\circ \text{ (et } 0^\circ \text{ bien sûr)}.$$

Noter qu'en particulier un axe de rotation d'ordre 5, c'est-à-dire $\theta = 2\pi/5$, est interdit (d'où la stupéfaction des physiciens Shechtman (prix Nobel de chimie 2011), Blech, Gratias et Cahn en 1984 à la vue d'images de diffraction de rayons X présentant une symétrie d'ordre 5, cf ci-dessous § 1.5.4).

Cet argument s'applique donc aux réseaux à 2 et 3 dimensions, tandis qu'en dimension supérieure, d'autres valeurs des angles de rotation deviennent possibles.

Il faut maintenant voir quelles réflexions sont compatibles avec les rotations et translations d'un réseau. On va procéder en plusieurs étapes, en déterminant d'abord les *groupes ponctuels*, c'est-à-dire fixant un point du réseau, donc excluant les translations ; puis en faisant rentrer les translations dans la discussion et en examinant quels réseaux de Bravais sont possibles ; et enfin en identifiant quelles décorations de chaque maille élémentaire sont possibles par un "motif" d'atomes, et quels *groupes (de symétrie) d'espace* en résultent.

Nous allons décrire assez en détail la situation à $d = 2$ dimensions et nous contenter de l'esquisser à $d = 3$.

1.5.2 Cristallographie à deux dimensions

Groupes ponctuels à $d = 2$

Les cinq types de rotations possibles qu'on a déterminés plus haut peuvent se combiner avec une ou des réflexion(s). On obtient donc 10 groupes ponctuels à $d = 2$ dimensions. La notation traditionnelle en cristallographie est la suivante :

$$1, 1m, 2, 2mm, 3, 3m, 4, 4mm, 6, 6mm.$$

Le chiffre désigne l'ordre de la rotation, tandis que le m signale l'existence d'une réflexion-miroir, et mm celle de deux miroirs indépendants (non reliés par l'action d'une rotation). Ces dix groupes sont représentés sur la figure 1.20 par l'*orbite* d'un point "générique" (=quelconque), c'est-à-dire par l'ensemble de ses images par l'action du groupe. Si le groupe G est d'ordre (=nombre d'éléments) $|G|$, une orbite générique compte donc $|G|$ points.

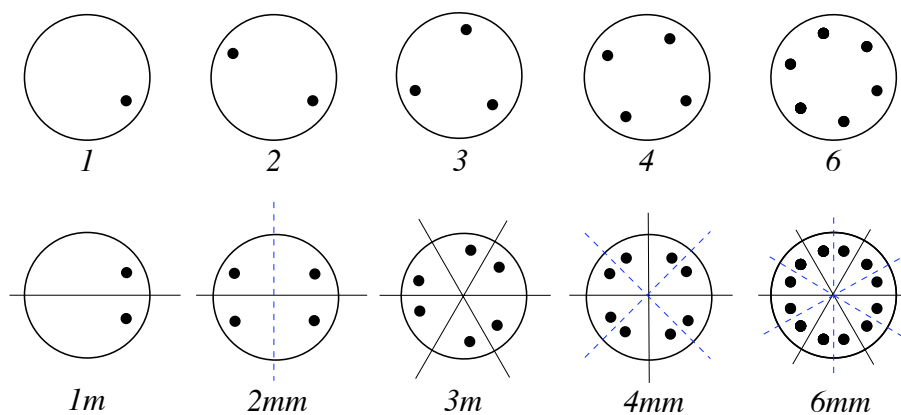


FIGURE 1.20 – Les 10 groupes ponctuels à $d = 2$ dimensions, symbolisés par l'“orbite”, c'est-à-dire l'ensemble des images par l'action du groupe, d'un point générique. Sur la deuxième ligne, on a tracé les axes de réflexion et leurs images par les rotations soit en trait plein, soit en ligne brisée.

Types de réseaux de Bravais à $d = 2$

Examinons quelles rotations (autour d'un point du réseau) et réflexions par rapport à un axe passant par un point du réseau sont possibles, pour différentes configurations de périodes de base $\vec{a}_1$ et $\vec{a}_2$. Il est clair que tout réseau de Bravais à $d = 2$ est invariant par les rotations de π autour de l'origine : les périodes $\vec{a}_1$ et $\vec{a}_2$ sont transformées en leurs opposées, qui appartiennent toujours au réseau. C'est la seule symétrie du réseau “générique”, dont la maille Δ est un parallélogramme arbitraire (voir fig. 1.21(a)). D'autre part une rotation d'ordre 4, donc d'angle $\pi/2$, n'est possible que si $\vec{a}_1$ et $\vec{a}_2$ sont orthogonaux et de même longueur. Dans ce cas, Δ est un carré, le réseau est le réseau carré, invariant aussi sous les deux réflexions par rapport aux directions de $\vec{a}_1$ et $\vec{a}_2$, et aussi par rapport à celles de $\vec{a}_1 \pm \vec{a}_2$ (cf étude du “groupe du carré” au § 1.3.3 et fig 1.21(b)). De même une rotation d'ordre 3 ou 6 (angles de $2\pi/3$ ou $\pi/3$) n'est possible que si $\vec{a}_1$ et $\vec{a}_2$ font un angle de $\pm\pi/3$ ou $\pm 2\pi/3$. Dans les deux cas, le réseau qui en résulte est le réseau triangulaire (dit aussi “hexagonal”) (fig 1.21(c)), invariant aussi par les

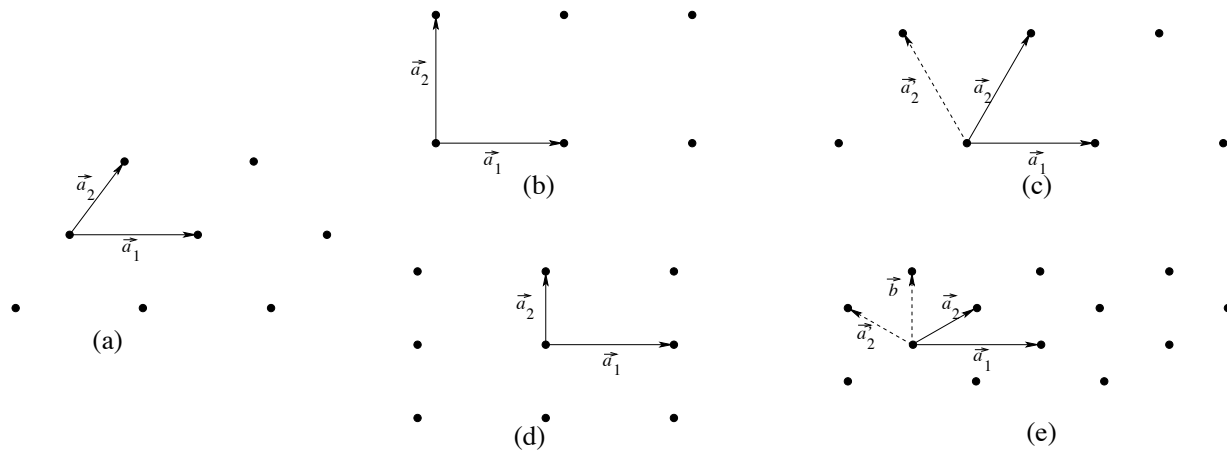


FIGURE 1.21 – Les 5 classes de réseaux de Bravais à 2 dimensions. Successivement, (a) réseau *monoclinique* : $\vec{a}_1$ et $\vec{a}_2$ arbitraires, maille Δ = parallélogramme générique ; (b) réseau carré (ou *tétragonal*) : $|\vec{a}_1| = |\vec{a}_2|$ et $\vec{a}_1 \perp \vec{a}_2$, Δ = carré ; (c) réseau hexagonal : $|\vec{a}_1| = |\vec{a}_2|$, Δ = losange d'angles $\pi/3$, $2\pi/3$; (d) réseau *orthorhombique* (ou rectangulaire) : $\vec{a}_1 \perp \vec{a}_2$, Δ = rectangle ; (e) “réseau” rectangulaire centré $\vec{a}_1, \vec{a}_2 = \frac{1}{2}\vec{a}_1^2$, Δ = losange générique.

réflexions par rapport aux trois directions portées par $\vec{a}_1$, $\vec{a}_2$ et $\vec{a}_2'$. Examinons enfin le cas de réseaux qui n'ont pas d'invariance par rotation (autre que la rotation de π) mais seulement par réflexion. On vérifie que les deux seules possibilités sont celles d'un réseau rectangulaire, $\vec{a}_1 \perp \vec{a}_2$ mais $|\vec{a}_1| \neq |\vec{a}_2|$ (fig 1.21(d)) et d'un réseau engendré par $\vec{a}_1$ et $\vec{a}_2$, tels que $\vec{a}_1 \cdot \vec{a}_2 = \frac{1}{2}\vec{a}_1^2$ (fig 1.21(e)). Ce dernier cas peut être vu comme un réseau de domaine fondamental losange porté par les vecteurs $\vec{a}_2$ et $\vec{a}_2 - \vec{a}_1$, mais on préfère souvent le décrire comme un “réseau rectangulaire centré” de périodes $\vec{a}_1$ et $\vec{b} = 2\vec{a}_2 - \vec{a}_1$; noter que dans cette dernière description, ce “réseau” rectangulaire centré n'est pas un réseau de Bravais, (pourquoi ?), mais l'avantage est que les symétries apparaissent plus clairement.

On appelle *réseau primitif* un réseau dont la maille élémentaire ne contient qu'un nœud. Ainsi le réseau rectangulaire centré ci-dessus n'est pas primitif. On ne considère un réseau non primitif que quand sa symétrie est plus évidente que celle du réseau primitif sous-jacent. Pourquoi ne pas considérer aussi un réseau carré centré ? Montrer qu'un tel réseau est en fait un réseau carré, de maille $a/\sqrt{2}$ si a est la maille initiale.

Cela achève l'énumération des 5 types de réseaux de Bravais à 2 dimensions.

Remarques et exercices.

1. Noter que les deux derniers cas (fig 1.21(d) et (e)) reproduisent les réseaux carré (b) et hexagonal (c) comme cas particuliers (pour quelles valeurs des modules et des angles des deux périodes ?)
2. Noter que les groupes de symétrie de ces 5 types de réseaux ne constituent que 4 des 10 groupes trouvés plus haut. Lesquels et pourquoi ?

Groupes d'espace à $d = 2$

Pour terminer la discussion des réseaux cristallins à $d = 2$, il nous faut encore examiner de combien de façons indépendantes les réseaux de Bravais que nous venons de classer peuvent

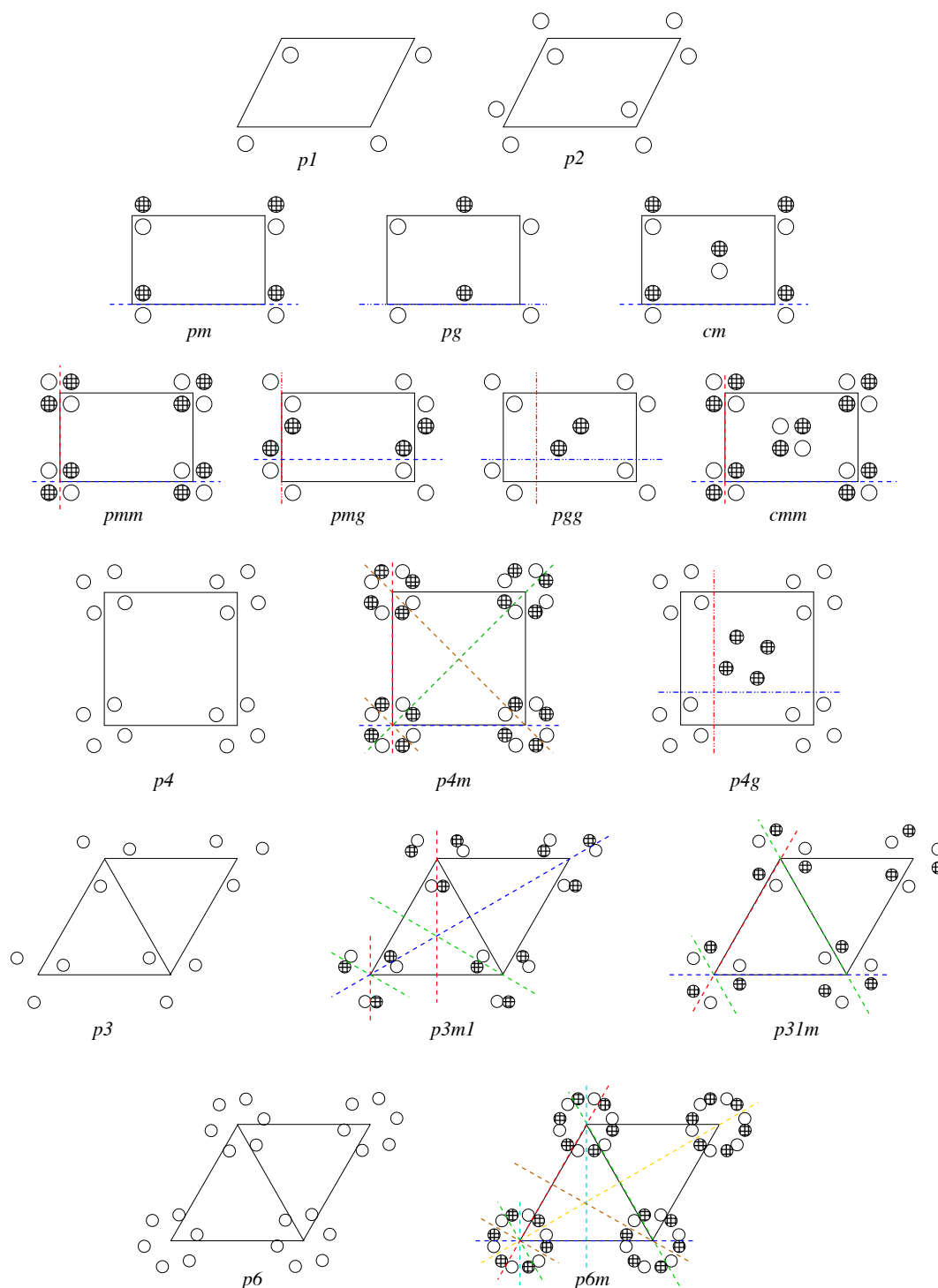


FIGURE 1.22 – Les 17 types de réseaux cristallins à $d = 2$ dimensions. On a dessiné une maille élémentaire ainsi que, sur quelques mailles adjacentes, le motif arbitraire représenté par un cercle ouvert ou son image miroir représentée par un cercle quadrillé. Les axes de réflexion ont été indiqués en ligne brisée, ceux de glissement en trait mixte. Noter les deux réseaux $p3m1$ et $p31m$ qui ont le même réseau de Bravais comme base, mais deux types de symétrie distincts qu'on identifiera. Noter aussi que tous les axes de réflexion et/ou glissement n'ont pas été dessinés. Exercice : identifier des axes de glissement dans cm ou de réflexion dans $p4g$.

être décorés par des motifs atomiques et quelle est la symétrie finale du cristal. Les symétries du réseau ainsi décoré peuvent être différentes de celles du réseau “nu”, elles peuvent impliquer non seulement des translations, rotations et réflexions mais aussi des *glissements*, obtenus par composition d’une translation par une demi-période et d’une réflexion par rapport à l’axe de translation. Les 17 groupes d’espace qu’on obtient au final se reflètent dans les 17 types de *pavages* périodiques que les artistes du monde entier (de l’Égypte antique au monde arabe à la Chine et à Hawaï) ont découvert bien longtemps avant les cristallographes (Fedorov, 1891)⁸. Voir http://en.wikipedia.org/wiki/Wallpaper_group pour une très jolie collection de pavages, appelés dans ce contexte “wallpapers” (papiers peints) par les anglo-saxons. La figure 1.22 montre ces 17 types de réseaux à deux dimensions.

La notation est conventionnelle et due à Hermann et Mauguin⁹. Elle comporte une lettre p ou c suivie de 3 symboles, dont chacun est un chiffre 1,2,4,6 ou la lettre m ou g . La première lettre p ou c indique si la maille est primitive ou centrée. L’entier n qui suit indique le plus grand ordre de symétrie de rotation présente : 1 si aucune symétrie de rotation, n si une rotation de $2\pi/n$ est une symétrie. Le symbole suivant est un m ou un g s’il existe un axe perpendiculaire à un axe de translation du réseau qui est un axe de réflexion-miroir (m) ou un axe de glissement (g), c’est un 1 s’il n’y a pas de tel axe. Le quatrième symbole précise l’existence d’un tel axe parallèle ou faisant un angle de π/n avec l’axe de translation précédent. On utilise généralement une notation abrégée (deuxième colonne de la Table 1 et figure 1.22). Voir l’article de Wikipedia juste mentionné pour plus de détails.

Chacun de ces types de réseau cristallin est caractérisé par un groupe, voir Table 1. En répondant aux questions successives

- le cristal est-il chiral, c’est-à-dire non superposable (même après un éventuel déplacement) à son image miroir ?
- quel est l’ordre n de la plus grande symétrie de rotation ?
- comment sont réalisées les symétries impropres, réflexion ou glissement ?
- le réseau est-il “primitif” (c’est-à-dire de Bravais) ou centré ?

on peut déterminer de façon non ambiguë le type cristallin, c’est-à-dire le groupe d’espace de ce cristal.

Exercice : auquel de ces 17 types de réseaux appartient chacun des deux pavages de la figure 1.10 ? Voir aussi TD.

Réseaux à une dimension

Il est instructif de regarder aussi la situation à *une* dimension. Dans ce cas, il n’existe qu’un seul réseau de Bravais et deux groupes ponctuels (selon que l’on considère ou non une symétrie de réflexion par rapport à un point). *Stricto sensu* il n’y a que deux groupes d’espace, mais il y a 7 façons de décorer le réseau uni-dimensionnel avec des motifs, pour constituer ce qu’on appelle une frise. Certains auteurs parlent donc de 7 groupes d’espace à une dimension. Ces 7 groupes sont illustrés par les 7 types de frise, cf TD.

8. Evgraf Stepanovich Fedorov (1853–1919), cristallographe russe, auteur avec A. Schönflies de la classification des groupes d’espace.

9. Carl Hermann (1898–1961), cristallographe allemand, et Charles Victor Mauguin (1878–1958), cristallographe français, sont les pères du système international de notation des systèmes cristallins.

Nom complet	Nom abrégé	Groupe ponctuel	Chiral ?	Rotation $2\pi/n$	# axes réflexion
<i>p111</i>	<i>p1</i>	1	oui	$n = 1$	0
<i>p211</i>	<i>p2</i>	2	oui	$n = 2$	0
<i>p1m1</i>	<i>pm</i>	1 <i>m</i>	non	$n = 1$	1
<i>p1g1</i>	<i>pg</i>	1 <i>m</i>	non	$n = 1$	0
<i>c1m1</i>	<i>cm</i>	1 <i>m</i>	non	$n = 1$	1
<i>p2mm</i>	<i>pmm</i>	2 <i>mm</i>	non	$n = 2$	2
<i>p2mg</i>	<i>pmg</i>	2 <i>mm</i>	non	$n = 2$	1
<i>p2gg</i>	<i>pgg</i>	2 <i>mm</i>	non	$n = 2$	0
<i>c2mm</i>	<i>cmm</i>	2 <i>mm</i>	non	$n = 2$	2
<i>p411</i>	<i>p4</i>	4	oui	$n = 4$	0
<i>p4mm</i>	<i>p4m</i>	4 <i>mm</i>	non	$n = 4$	4
<i>p4gm</i>	<i>p4g</i>	4 <i>mm</i>	non	$n = 4$	2
<i>p311</i>	<i>p3</i>	3	oui	$n = 3$	0
<i>p3m1</i>	<i>p3m1</i>	3 <i>m</i>	non	$n = 3$	3
<i>p31m</i>	<i>p31m</i>	3 <i>m</i>	non	$n = 3$	3
<i>p611</i>	<i>p6</i>	6	oui	$n = 6$	0
<i>p6mm</i>	<i>p6m</i>	6 <i>mm</i>	non	$n = 6$	6

Table 1 : Les 17 groupes d'espace, ou types de réseaux cristallins, à $d = 2$.

1.5.3 Cristallographie à trois dimensions

Passant de 2 à 3 dimensions, on procède de manière analogue, l'analyse étant évidemment beaucoup plus longue et les cas plus nombreux.

On trouve ainsi successivement 32 groupes ponctuels¹⁰, 14 types de réseaux de Bravais possibles et 230 groupes d'espace. On en trouvera la description détaillée dans les ouvrages de cristallographie¹¹. Nous nous contentons de donner ici la table des 14 réseaux de Bravais, cf fig. 1.23. En fait, plutôt que des réseaux de Bravais, dont par définition chaque maille ne contient de nœuds qu'à ses coins, on préfère souvent représenter des réseaux "non primitifs" où la maille peut en contenir plusieurs (comme le réseau rectangulaire centré au § précédent, fig. 1.21), afin de mieux faire apparaître les symétries du réseau.

On distingue les réseaux primitifs, qui ne possèdent qu'un nœud par maille (P), les réseaux centrés où un nœud supplémentaire est placé au centre de la maille (I), les réseaux où une face est centrée (C) et ceux où toutes les faces sont centrées (F). Les notations sont traditionnelles.

Exercice, dans chaque cas non primitif, identifier la sous-maille primitive. Inversement, pourquoi le dernier cas est-il qualifié d'hexagonal ?

10. Bien noter que les groupes ponctuels qui apparaissent ici sont des groupes ponctuels cristallographiques, compatibles avec l'existence d'un réseau, au contraire de ceux évoqués au § 1.3.4. Par exemple, le groupe de l'icosaèdre qui apparaissait dans la liste du § 1.3.4 est interdit comme groupe cristallographique par la discussion du § 1.5.1.

11. par exemple [6], voir aussi le site <http://it.iucr.org> des Tables Internationales de Cristallographie

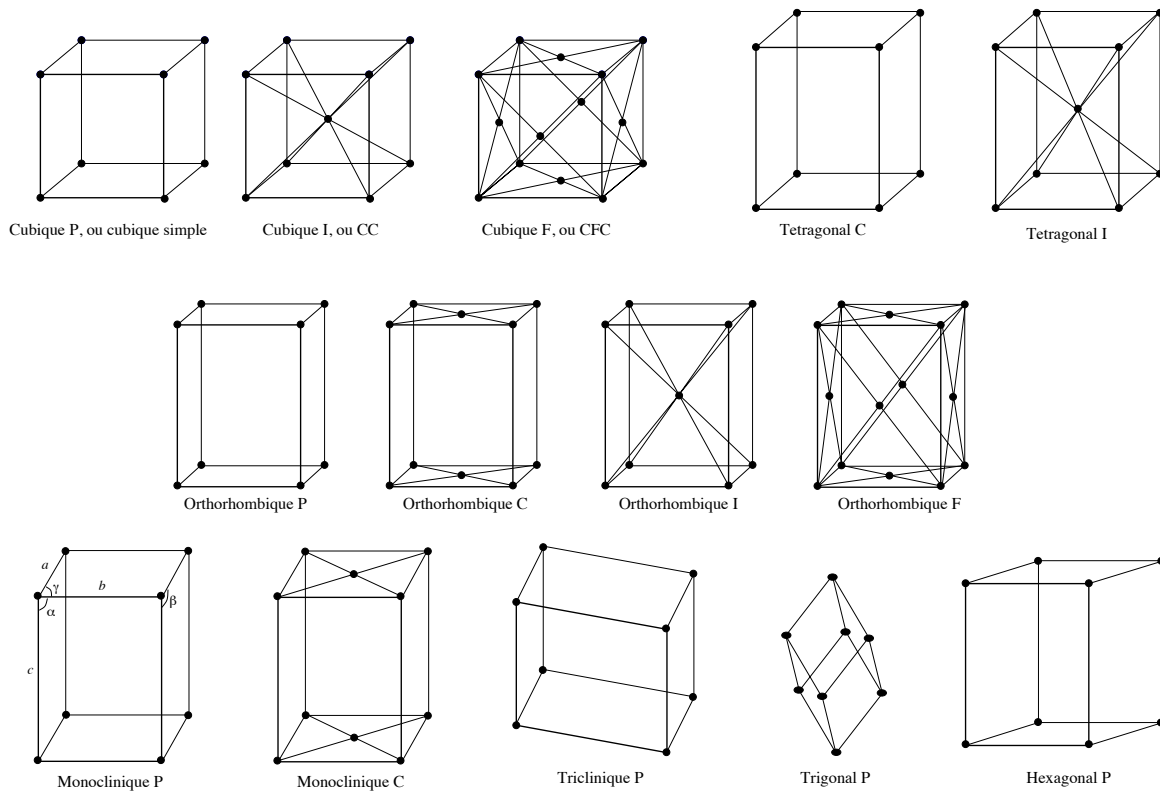
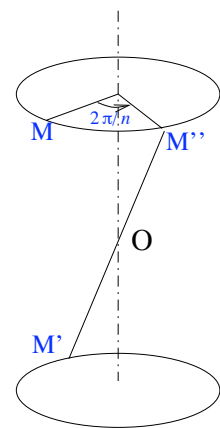


FIGURE 1.23 – Les 14 types de réseaux de Bravais à 3 dimensions. On a successivement les trois réseaux cubiques $a = b = c$, $\alpha = \beta = \gamma = \frac{\pi}{2}$; les deux réseaux tétragonaux $a = b \neq c$, $\alpha = \beta = \gamma = \frac{\pi}{2}$; les quatre réseaux orthorhombiques $a \neq b \neq c$, $\alpha = \beta = \gamma = \frac{\pi}{2}$; les deux réseaux monocliniques $a \neq b \neq c$, $\alpha = \beta = \frac{\pi}{2}$ et γ quelconque et le réseau triclinique $a \neq b \neq c$, α , β et γ quelconques; le réseau trigonal $a = b = c$, α , β et γ quelconques et le réseau hexagonal $a = b \neq c$, $\alpha = \beta = \frac{\pi}{2}$ et $\gamma = \frac{\pi}{3}$. Question : pourquoi n'y a-t-il pas de réseau monoclinique avec un nœud au centre de la maille ?

Les groupes ponctuels sont répertoriés soit dans la notation de Schönflies, (cf. § 1.3.4), soit dans la notation cristallographique de Hermann-Mauguin, comparable à celle utilisée plus haut à $d = 2$, avec une chaîne de symboles qui code précisément toutes les opérations du groupe : le premier chiffre n désigne un axe de rotation d'ordre n , avec $n = 1, 2, 3, 4, 6$; $\bar{n}$ un axe de rotation-inversion d'ordre n (c'est-à-dire la composée d'une rotation de $2\pi/n$ par une inversion par rapport à un point sur l'axe de rotation, cf ci-contre); $\frac{n}{m}$ dénote un plan de réflexion-miroir orthogonal à un axe d'ordre n ; $n2$, un axe d'ordre 2 orthogonal à celui d'ordre n ; nm , un plan de réflexion-miroir parallèle à l'axe d'ordre n ; $\bar{n}2$ et $\bar{n}m$, idem pour un axe de rotation-inversion; $\frac{n}{m}m$, un axe Δ d'ordre n avec un plan de réflexion-miroir orthogonal à Δ et n plans de réflexion-miroir parallèles à Δ . Ainsi le groupe de rotation du cube est noté O ou 432 , le groupe complet (rotations et réflexions) du cube est O_h ou $\frac{4}{m}\bar{3}\frac{2}{m}$ (exercice : le vérifier). La liste des 32 groupes ponctuels est donnée dans la table 2.



Système réticulaire	Cellule élémentaire	Groupe G , nom de Schönflies	Nom de Hermann–Mauguin	$ G $ ordre de G
<i>Triclinique</i>	$a \neq b \neq c$ $\alpha \neq \beta \neq \gamma$	C_1	1	1
		S_2	$\bar{1}$	2
<i>Monoclinique</i>	$a \neq b \neq c$ $\alpha = \beta = \pi/2 \neq \gamma$	C_{1h}	m	2
		C_2	2	2
		C_{2h}	$\frac{2}{m}$	4
<i>Orthorhombique</i>	$a \neq b \neq c$ $\alpha = \beta = \gamma = \pi/2$	C_{2v}	$mm2$	4
		D_2	222	4
		D_{2h}	$\frac{2}{m} \frac{2}{m} \frac{2}{m}$	8
<i>Tétragonal</i>	$a = b \neq c$ $\alpha = \beta = \gamma = \pi/2$	C_4	4	4
		S_4	$\bar{4}$	4
		C_{4h}	$\frac{4}{m}$	8
		D_4	422	8
		C_{4v}	$4mm$	8
		D_{2d}	$\bar{4}2m$	8
		D_{4h}	$\frac{4}{m} \frac{2}{m} \frac{2}{m}$	16
<i>Trigonal</i>	$a = b = c$ $\alpha = \beta = \gamma < 2\pi/3, \neq \pi/2$	C_3	3	3
		S_6	$\bar{3}$	6
		C_{3v}	$3m$	6
		D_3	32	6
		D_{3d}	$\bar{3} \frac{2}{m}$	12
<i>Hexagonal</i>	$a = b \neq c$ $\alpha = \beta = \pi/2, \gamma = 2\pi/3$	C_6	6	6
		C_{3h}	$\bar{6}$	6
		C_{6h}	$\frac{6}{m}$	12
		D_6	622	12
		C_{6v}	$6mm$	12
		D_{3h}	$\bar{6}m2$	12
		D_{6h}	$\frac{6}{m} \frac{2}{m} \frac{2}{m}$	24
<i>Cubique</i>	$a = b = c$ $\alpha = \beta = \gamma = \pi/2$	T	23	12
		T_h	$\frac{2}{m} \bar{3}$	24
		T_d	$\bar{4}3m$	24
		O	432	24
		O_h	$\frac{4}{m} \bar{3} \frac{2}{m}$	48

Table 2 : Les 32 groupes ponctuels à $d = 3$ et les 7 systèmes réticulaires auxquels ils s'appliquent.

Des exemples que l'on rencontre souvent en physique sont ceux des trois réseaux cubiques, usuellement désignés en français par les acronymes C, CC et CFC (cubique simple, cubique centré, cubique à faces centrées) et en anglais par C, BCC (body centered) et FCC, et aussi celui du réseau hexagonal, voir ci-dessous.

Les groupes d'espace peuvent aussi impliquer non seulement des glissements comme à $d = 2$

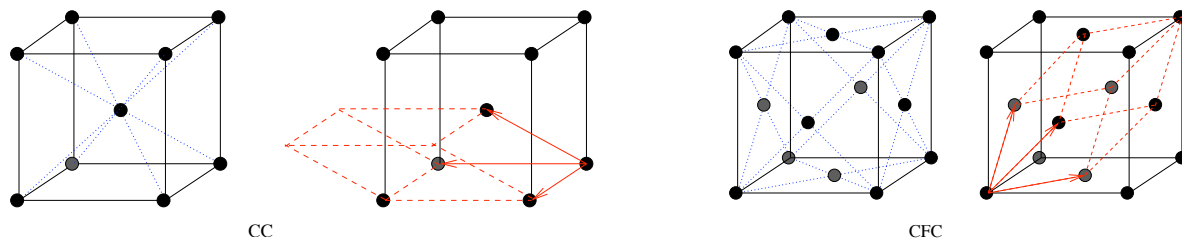


FIGURE 1.24 – Réseaux cubique centré et cubique à faces centrées. On a dessiné un choix de maille primitive en lignes (tirets) rouges. La symétrie cubique n'est pas évidente sur cette maille primitive !

(avec maintenant un *plan* de glissement) mais aussi des “vissages”, qui sont des mouvements hélicoïdaux, composés d'une rotation d'axe $\vec{u}$ et d'une translation parallèle à $\vec{u}$...

Les 230 types de groupes cristallins ont tous été observés expérimentalement à quelques exceptions près.

Quelques exemples de corps cristallisant dans les systèmes C, CFC, CC, hexagonal etc.

Expérimentalement on constate que le cuivre cristallise dans un réseau CFC, de même que Al, Ni, Cu, Ag, Au et le fer $\text{Fe}\gamma$ (une des phases du fer). Le fer $\text{Fe}\alpha$ ainsi que Ba, Cr, V, Mo, W ont des cristaux CC. Les corps Mg, Ca, Sr, Ti, Co adoptent une structure hexagonale compacte, voir Fig. 1.25.

Il est instructif de considérer et de comparer les cristaux de diamant et de blende de zinc (sulfure de zinc ZnS). L'un et l'autre sont basés sur un réseau cubique à faces centrées, dans lequel en plus des atomes (de C ou de Zn, respectivement) aux 8 coins du cube et aux centres des 6 faces, quatre atomes supplémentaires (de C ou de S, resp.) occupent les milieux de 4 des segments joignant le centre du cube aux coins et forment avec les atomes voisins des coins ou des faces un petit tétraèdre, voir Fig. 1.26. Chaque atome de C a quatre voisins C, chaque atome de Zn de la blende a 4 voisins S et vice-versa (Exercice : le vérifier sur la figure). Mais ces atomes voisins étant identiques dans le diamant et différents (Zn et S) dans la blende, la symétrie ponctuelle d’“inversion” par rapport au milieu des segments joignant les atomes voisins est une symétrie du cristal de diamant, mais pas de la blende, puisqu'elle échange les Zn et les S, voir Fig. 1.26. Les groupes d'espace du diamant et de la blende (ou “sphalérite”) sont donc distincts (et répondent aux doux noms de $Fd\bar{3}m$ et $F\bar{4}3m$, respectivement).

Autre forme cristallisée de carbone pur, le graphite est très différent du diamant, Fig. 1.27. Il est composé de feuillets hexagonaux non compacts, ayant la structure du graphène, (Fig. 1.5), avec une distance de plan à plan de l'ordre de 0,335 nm. Dans chaque feuillet, les atomes de carbone sont fortement liés par des liaisons covalentes de type sp^2 , tandis que les liaisons entre les feuillets sont faibles, de type π , ce qui explique les propriétés très anisotropes, mécaniques (clivage et faible dureté, cf la mine des crayons!) ou de transport (conductibilité électrique), très différentes dans les feuillets et entre feuillets.

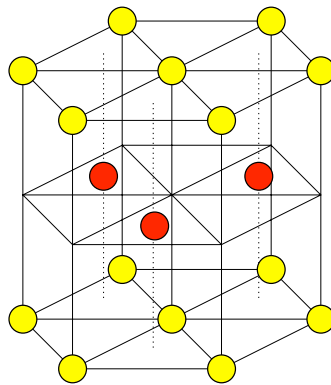


FIGURE 1.25 – Le réseau hexagonal compact. Les atomes de la couche intermédiaire sont à la verticale des centres de gravité de la moitié des triangles des couches voisines.

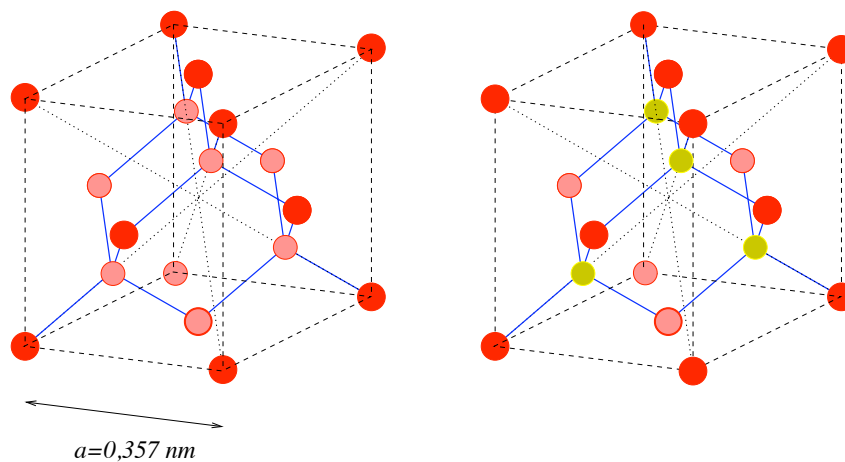


FIGURE 1.26 – Les réseaux cristallins du diamant et de la blende ZnS. En tirets noirs, le réseau cubique sous-jacent, en traits continus (bleus) les segments entre atomes voisins formant un tétraèdre. En pointillé les diagonales du cube, se croisant au centre 0. Vérifier que chaque cube contient autant d'atomes de Zn que de S.

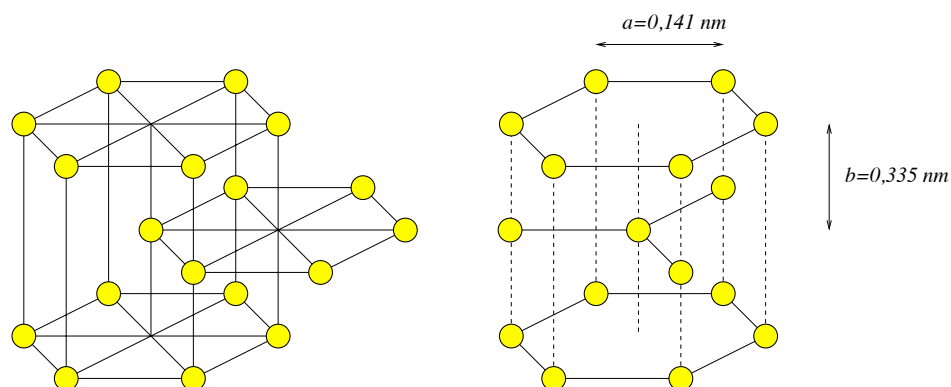


FIGURE 1.27 – Deux représentations de la maille du graphite (pas à l'échelle !)

Le chlorure de sodium cristallise dans une structure cubique, voir Fig. 1.28. Les cations Na^+ forment un sous-réseau cubique à faces centrées dans lequel les cations Cl^- viennent s'intercaler, formant un deuxième réseau cubique à faces centrées, décalé par rapport au premier de la moitié du côté de sa maille. Chaque ion Cl est entouré de 6 ions Na formant un octaèdre régulier autour de lui et vice-versa, et on dit que les Cl^- occupent les “sites octaédriques” de la maille des Na , voir figure. Dans cette description, les ions Na^+ et Cl^- sont interchangeables.

La wurtzite, une autre forme cristalline du sulfure de zinc, forme deux empilements hexagonaux compacts enchevêtrés, voir Fig. 1.29.

Le quartz α cristallise dans le système cristallin trigonal, avec un axe d'ordre 3. Comme mentionné plus haut, il vient en deux espèces, gauche et droite, selon la façon dont les tétraèdres SiO_4 qui le composent s'enroulent autour de cet axe en hélices gauches ou droites, cf Fig. 1.30. C'est là l'origine des propriétés chirales observées, pouvoir rotatoire, etc.

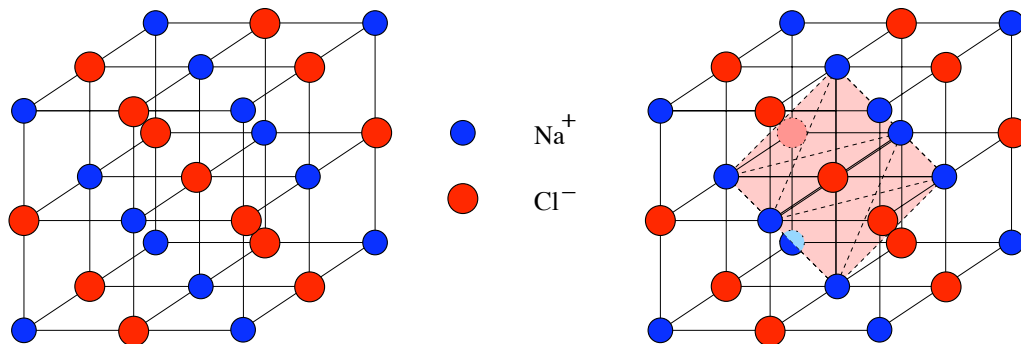


FIGURE 1.28 – Le cristal de NaCl . Chaque ion de Cl a pour plus proches voisins 6 ions Na qui forment un octaèdre centré sur lui, (ici en tirets), et vice versa.

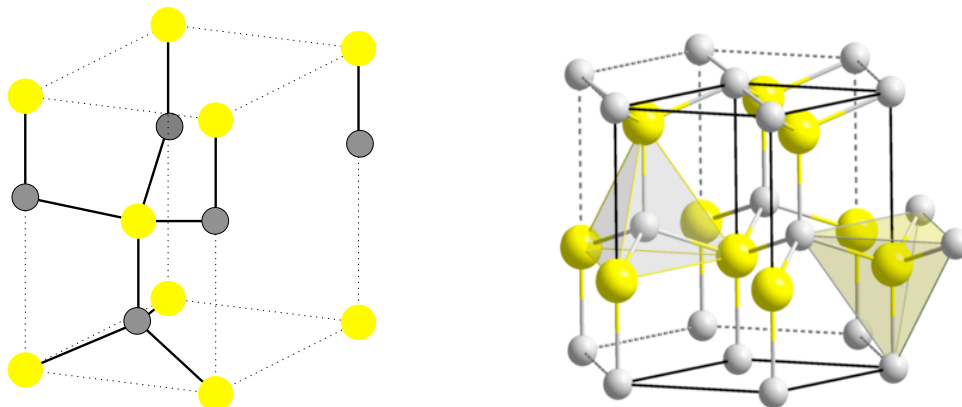


FIGURE 1.29 – Deux représentations du cristal de wurtzite, formé de deux réseaux hexagonaux de Zn et de S qui s'enchevêtrent. Chaque Zn est au centre d'un tétraèdre de S , et vice versa. Source Wikipedia

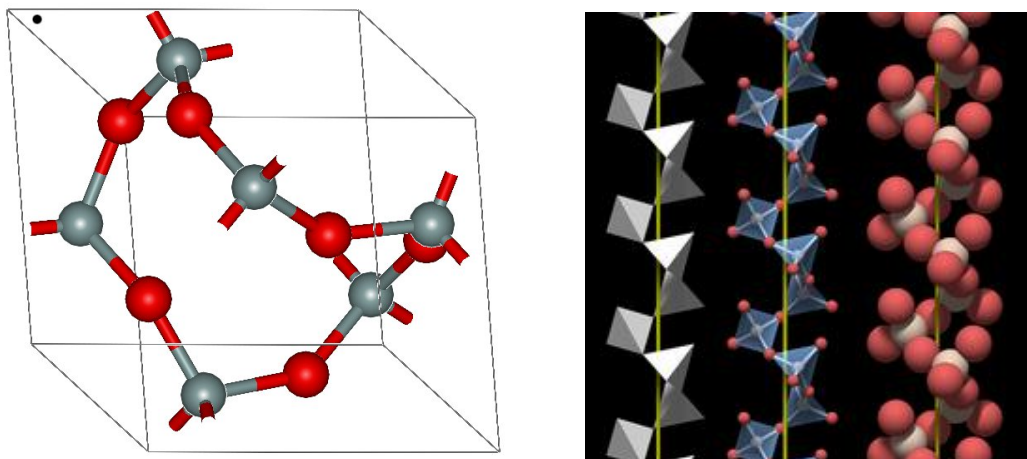


FIGURE 1.30 – Différentes représentations de la structure cristalline du quartz α (Sources : <http://en.wikipedia.org/wiki/Quartz>, http://www.quartzpage.de/gen_struct.html).

Densités d'atomes. Compacité des empilements

Combien y-a-t-il d'atomes de carbone par maille cubique du diamant ? Il convient de compter chaque atome en lui affectant un “poids” qui tient compte du nombre de mailles cubiques auxquelles il appartient. Ainsi chacun des 8 atomes des 8 coins est commun aux huit mailles qui touchent ce coin et a un poids $\frac{1}{8}$. Chacun des 6 atomes placés aux centres des faces des cubes est partagé par deux faces et a donc a poids $\frac{1}{2}$. Enfin les 4 atomes placés à l'intérieur viennent avec le poids 1. Au total, chaque maille cubique du diamant (ou de la blende ZnS, ou du carbure de Si) possède $8 \times \frac{1}{8} + 6 \times \frac{1}{2} + 4 = 8$ atomes, tandis qu'une maille cubique à faces centrées ordinaire en a 4.

De la même façon, on calcule qu'une maille cubique centrée avec un atome par nœud du réseau compte $8 \times \frac{1}{8} + 1 = 2$ atomes par maille.

Si on assimile chaque atome (ou ion) d'un cristal à une petite sphère dont le rayon est connu, on peut calculer la *compacité* τ (ou taux de remplissage) définie comme le rapport du volume occupé par ces sphères au volume total. Ainsi pour le chlorure de césium, à maille cubique centrée de côté $a = 411$ pm, étant donnés le rayon de l'ion Cs^+ , $r_+ = 174$ pm et celui de Cl^- , $r_- = 181$ pm, montrer que la compacité est $\tau = \frac{4\pi}{3}(r_+^3 + r_-^3)/a^3 = 0,676$.

Calculer ainsi la compacité du chlorure de sodium NaCl : $\tau = 0.644$ à partir de la distance interatomique $d_{\text{Na}^+ - \text{Cl}^-} = 2.82$ Å, $r_{\text{Na}^+} = 0.99$ Å, $r_{\text{Cl}^-} = 1.81$ Å.

Imaginons que les atomes soient des petites sphères jointives. On parle alors d'empilement compact de sphères. Dans l'empilement compact CFC, avec une maille de côté a , il y a 4 sphères, deux sphères voisines ont par exemple les coordonnées $(0,0,0)$ et $(\frac{1}{2}a, \frac{1}{2}a, 0)$, leur distance est $a/\sqrt{2}$, le rayon de chacune est $r = a/2\sqrt{2}$ et son volume $v = \frac{4}{3}\pi r^3$, donc la compacité $\tau = 4v/a^3 = \frac{\pi}{3\sqrt{2}} = 0.74$.

Calculer de même la compacité d'un empilement hexagonal compact (fig. 1.25), ou celui d'un empilement CC (cubique centré).

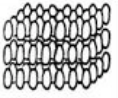
Cristallin	Colonnaire	Smectique	Nématique		Isotrope
C	Col	S _B S _C S _A	N _D	N	I
ordre	ordre d'orientation et de position		ordre d'orientation		désordre
					
	molécules plutôt symétriques avec 2 chaînes grasses		molécules discotiques	<chem>N#Cc1ccc(cc1)-c2ccc(cc2)C#N</chem> Ex : cyanobiphényles	

FIGURE 1.31 – Différents types de cristaux liquides ou mésophases Source Wikipedia

Un problème important pour le mathématicien ou le physicien comme dans la vie pratique est celui des empilements les plus compacts possibles. Comment disposer dans l'espace $\mathbb{R}^3$ des sphères de même rayon pour obtenir la plus grande densité? Comme tous les marchands de fruits le savent, la réponse est donnée par des dispositions en couches successives suivant un réseau triangulaire, chaque nouvelle couche venant se loger dans les creux de la précédente. Montrer que pour la 3ème couche, il y a deux choix possibles inéquivalents, dont l'un correspond à un réseau CFC. Et l'autre? L'un et l'autre remplissent une fraction $\frac{\pi}{3\sqrt{2}} \simeq 0.74048$ de l'espace. Ce résultat conjecturé par Kepler il y a 4 siècles (1611) n'a été démontré rigoureusement qu'il y a quelques années (T. Hales 1998)! S'il est assez aisé à démontrer en supposant l'empilement régulier, donc obéissant à l'une des symétries cristallographiques, il est infiniment plus complexe si on ne suppose *a priori* aucune symétrie.

Quand on connaît les masses des atomes constituants, ce même genre de considérations permet de calculer la masse volumique (ou "densité") d'un solide cristallisé. Voir l'exemple du diamant et du graphite dans les Exercices.

1.5.4 Autres ordres translationnels ou orientationnels

Nous nous contenterons de quelques mots sur des objets physiques présentant d'autres types d'ordre translationnel ou orientationnel :

- les cristaux liquides
- les quasi-cristaux
- les nanotubes de carbone.

Cristaux liquides : il s'agit de molécules de forme allongée ou aplatie en galette, et donc enclines à adopter un ordre orientationnel, et éventuellement un ordre translationnel partiel, intermédiaire entre les ordres solides et liquides. On distingue les

- nématiques : molécules oblongues parallèles (ordre orientationnel uniquement)
- smectiques : molécules dans des couches parallèles équidistantes, ordre translationnel partiel, mais pas d'ordre au sein des couches ("plan liquide orienté")
- cholestériques : plans liquides orientés, les orientations de plan à plan forment une hélice

mais il existe aussi d'autres phases, colonnaires, etc, voir Figure 1.31.

Quasi-cristaux : Découverts au milieu des années 80, les quasi-cristaux ont dans un premier temps paru paradoxaux aux physiciens et cristallographes : des images de diffraction par rayons X (cf Chapitre 2) faisaient apparaître une symétrie d'ordre 5 ou 10, (Fig. 1.32(a)), semblant indiquer une structure cristalline ayant cette même symétrie. Cela était donc en contradiction avec le résultat du § 1.5.1 qui interdit une symétrie d'ordre autre que 1,2,3,4 ou 6 dans un cristal. On a compris alors que la symétrie des images de diffraction était compatible

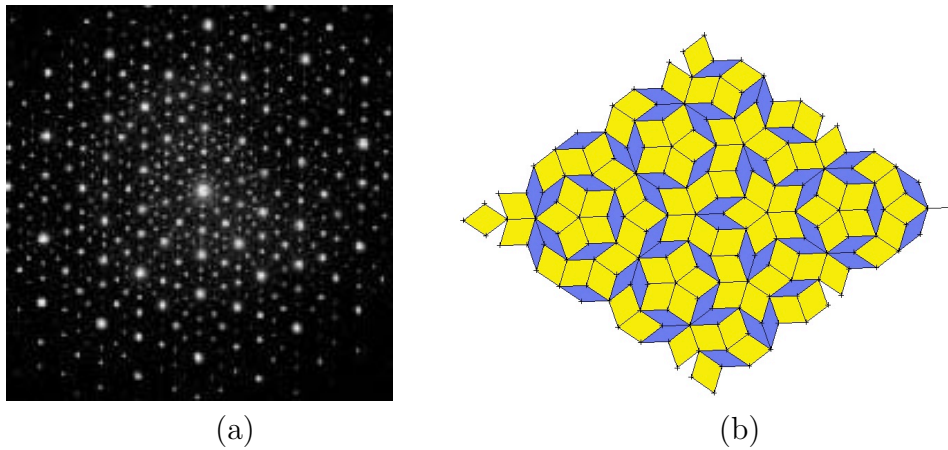


FIGURE 1.32 – (a) : une image de diffraction de rayons X avec une symétrie d’ordre 10 ; (b) un pavage de Penrose : noter la symétrie *locale* d’ordre 5 mais l’absence de symétrie de translation (Source Wikipedia).

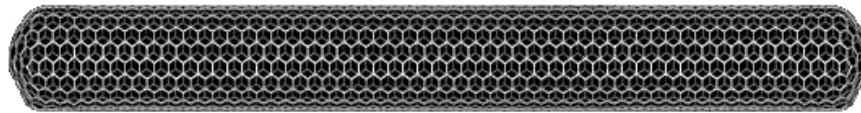


FIGURE 1.33 – Un nanotube de carbone (Source Wikipedia).

avec une symétrie plus faible que celle d’un cristal véritable, dans une structure dite *quasipériodique*, où une symétrie *locale* de rotation d’ordre 5 (ou 8, ou 10) peut exister sans ordre translationnel. Les atomes du corps ne sont pas disposés aux nœuds d’un réseau de Bravais. Un bon exemple de cette situation est donné par les fameux *pavages de Penrose*, voir Fig. 1.32(b) pour un exemple.

Nanotubes de carbone : Les nanotubes de carbone sont des objets très étudiés dans les nanotechnologies contemporaines et très intéressants pour leurs propriétés mécaniques (“résilience”, dureté) et optiques, et pour leurs conductibilités thermique et électrique très élevées. On peut se les représenter comme constitués d’un feuillet de graphène enroulé sur lui-même, et éventuellement refermé à ses extrémités. Un tel nanotube a des dimensions typiques de 1 à 10 nanomètres en diamètre et de plusieurs micromètres en longueur, donc un rapport de 100 à 1000 entre ses deux dimensions. Dans l’approximation d’un cylindre infini, un tel nanotube a comme symétries des rotations discrètes d’angle $2\pi/n$ si n est le nombre d’hexagones homologues dans une section transversale, des translations discrètes et aussi des déplacements hélicoïdaux.

L’article de Wikipedia http://fr.wikipedia.org/wiki/Nanotube_de_carbone donne une bonne introduction au sujet.

1.6 Exercices

1. Pavages et cristaux 2d

Pour chacune des gravures d’Escher de la Fig. 1.34, dans lesquelles on ne tiendra pas compte des couleurs, identifier les symétries et le “groupe d’espace” (ici à $d = 2$) selon l’algorithme



FIGURE 1.34 – 4 gravures d’Escher, Source : <http://www.mcescher.com/>

expliqué à la Table 1.

2. Systèmes cristallins à 3d

Examiner les 14 réseaux de la figure 1.23 et déterminer leur groupe ponctuel de symétrie. Montrer que seuls 7 des 32 groupes ponctuels possibles apparaissent. Comparer avec les données de la Table 2 et commenter.

3. Pérovskite

Les pérovskites (du nom du minéralogiste russe L.A. Perowski (1792-1856)) désignent à l’origine une classe de cristaux de titanates, de calcium, de baryum, de strontium, etc. Voir figure 1.35, où un ion Ti^{4+} est au centre de la maille cubique, des Ca^{2+} aux coins et O^{2-} aux centres des faces. Vérifier la neutralité du cristal. Donner la formule chimique de ce corps. Montrer qu’on peut aussi représenter ce cristal comme une juxtaposition d’octaèdres se joignant par les O aux sommets, avec un Ti au centre de chaque octaèdre et un Ca ou un Ba au milieu des “cuboctaèdres” intercalaires. L’appellation pérovskite a été étendue à des silicates XSiO_3 , etc, qui cristallisent dans la même structure.

4. Densité du diamant et du graphite

En s’appuyant sur les dessins des mailles du diamant et du graphite, Fig. 1.26 et 1.27, et sur les dimensions interatomiques qui y figurent,

- calculer le nombre d’atomes par maille ;
- calculer le volume de la maille ;
- en prenant $m_C = 12\text{g}/\mathcal{N}$ pour masse de l’atome de C, $\mathcal{N}$ le nombre d’Avogadro, calculer la masse volumique (“densité”) du diamant et celle du graphite. Les valeurs expérimentales

sont respectivement de 3.5 à 3.53 g/cm³ et de 2.09 à 2.23 g/cm³ selon les variétés.
Conclure.

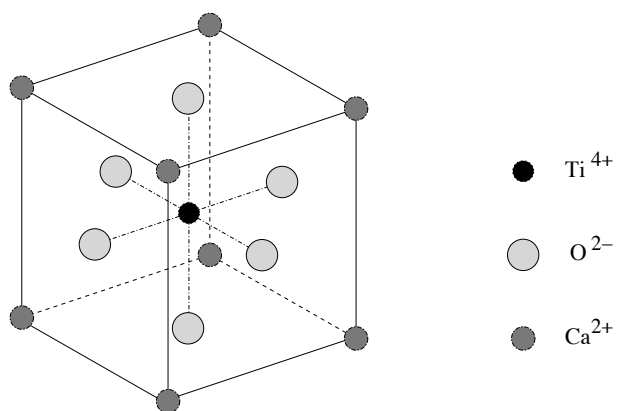


FIGURE 1.35 – Pérovskite

Chapitre 2

Transformations des grandeurs physiques

Dans ce chapitre, nous allons étudier quelques premières conséquences physiques, essentiellement d'ordre qualitatif, des symétries que nous avons étudiées au chapitre précédent. Pour cela, on va commencer par décrire comment se transforment diverses quantités physiques, sous l'effet de rotations ou de réflexions ou de l'inversion/parité : cette loi de transformation est dictée par la nature géométrique, scalaire, vectorielle, tensorielle etc de la quantité considérée. On examinera ensuite quelle est la conséquence d'une hypothèse de symétrie (ou d'invariance) du système sur ces quantités. Le principe de Curie, qui lie symétries des causes et symétries des effets, apparaîtra naturellement à la lueur de nombreux exemples. Enfin on reviendra sur le rôle des translations dans les systèmes cristallins.

C'est sans doute le bon endroit pour clarifier un point de principe. Chaque fois qu'on parle de transformations d'espace (et éventuellement de symétries, cf le 1er § du Chap. 1), agissant sur les coordonnées d'un système selon une loi inversible $x \mapsto x' = f(x)$, deux interprétations sont possibles, cf fig 2.1 :

1. le point de vue *actif* : la transformation agit effectivement sur le système considéré $\mathcal{S}$ et le transforme en un autre système $\mathcal{S}'$ décrit dans le même repère par les coordonnées $x' = f(x)$;
2. le point de vue *passif* : le système n'est pas affecté par la transformation, mais le repère de coordonnées change, avec les nouvelles coordonnées x' au lieu des anciennes x .

Il doit être clair que ces deux points de vue sont équivalents et qu'on est totalement libre de choisir l'un ou l'autre selon le problème, ou selon sa propre intuition !

Noter toutefois que si on veut que les coordonnées du système $\mathcal{S}'$ dans le repère xyz (point de vue actif) soient les mêmes que celles de $\mathcal{S}$ dans le nouveau repère $x'y'z'$ du point de vue passif, il faut avoir opéré avec la transformation $\mathcal{T}^{-1}$ dans le point de vue passif si on a opéré avec $\mathcal{T}$ dans l'actif. Cela est bien apparent sur la figure 2.1 où on a basculé dans des sens opposés le corps $\mathcal{S}'$ et le trièdre $Oxy'z'$.

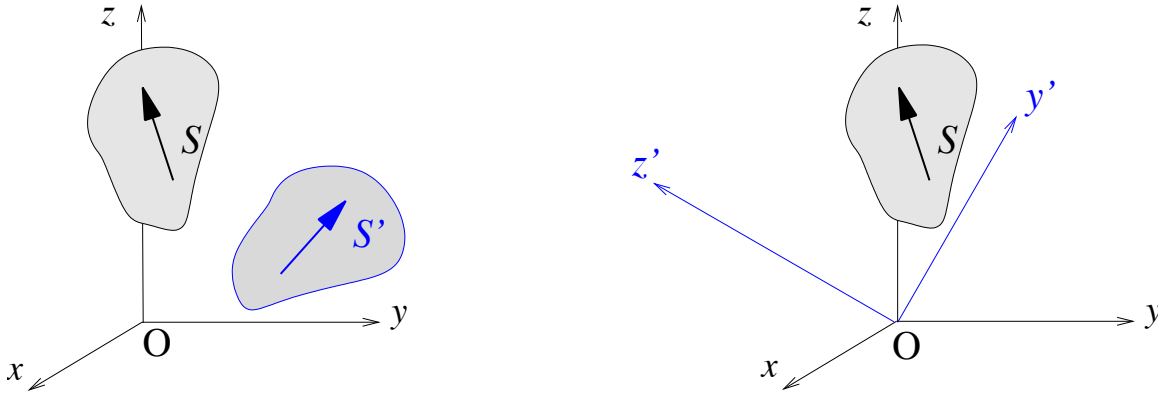


FIGURE 2.1 – Les points de vue actif et passif dans une transformation.

2.1 Rotations

2.1.1 Vecteurs et scalaires

On est familier avec la notion de vecteur, disons de l'espace $\mathbb{R}^3$. Sous l'effet d'une rotation du groupe $SO(3)$, décrite par une matrice orthogonale O , le vecteur $\vec{V}$ de coordonnées V^i dans un repère orthonormé $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$, $\vec{V} = \sum_{i=1}^3 V^i \vec{e}_i$, se transforme selon

$$\vec{V} \mapsto \vec{V}' = O\vec{V} = \sum_{i=1}^3 V^i O\vec{e}_i = \sum_{i,j=1}^3 V^i \vec{e}_j O_{ji}$$

donc

$$V'^i = \sum_{j=1}^3 O_{ij} V^j . \quad (2.1)$$

Rappelons pourquoi on écrit le transformé par O du vecteur $\vec{e}_i$ comme $O\vec{e}_i = \sum_{j=1}^3 \vec{e}_j O_{ji}$. De cette façon, si on compose O puis O' , $(O' \circ O)\vec{e}_i = O'(\sum_{j=1}^3 \vec{e}_j O_{ji}) = \sum_{j=1}^3 O_{ji}(O'\vec{e}_j) = \sum_{j,k=1}^3 O_{ji} \vec{e}_k O'_{kj} = \sum_{k=1}^3 \vec{e}_k (O'.O)_{ki}$, avec le produit matriciel $O'.O$ dans cet ordre. À la composition $O' \circ O$ correspond la matrice $O'.O$.

Inversement toute quantité se transformant selon la loi (2.1) sera appelé vecteur (pour les rotations). En revanche, une quantité invariante par rotation est dite *scalaire* (pour les rotations). C'est bien sûr le cas du produit scalaire de deux vecteurs, ou du carré de la norme d'un vecteur, par exemple.

Remarque. Nous notons ici les indices des composantes en position supérieure, et celle des vecteurs de base en position inférieure. Il faudrait pour être cohérent noter les éléments de matrice O^i_j . Dans le cas présent du groupe orthogonal, ces conventions ne prêtent pas à conséquence et on mettra indifféremment en position inférieure ou supérieure les indices des vecteurs, tenseurs et matrices. Il en va autrement pour d'autres types de transformations, telles les transformations du groupe de Lorentz en Relativité restreinte, ou les changements de coordonnées en Relativité générale.

Exercice : sur l'exemple de transformation d'un vecteur, comparer les points de vue actif et passif et vérifier l'assertion faite ci-dessus que les rotations active et passive doivent être inverses l'une de l'autre pour que le vecteur soit représenté finalement par les mêmes coordonnées.

D'un point de vue de physicien, des exemples de fonctions scalaires sont fournis par des densités (de masse, de charge, etc), ou par la température, la pression, l'indice de réfraction etc. Les quantités notées vectoriellement, telles la position $\vec{x}$ d'un mobile ou ses dérivées vitesse $\vec{v} = \dot{\vec{x}}$ ou accélération $\vec{a} = \ddot{\vec{x}} = \dot{\vec{v}}$, et donc aussi l'impulsion alias quantité de mouvement $\vec{p}$ sont des vecteurs au sens précédent. En physique, on rencontre aussi des *champs*, qui sont des grandeurs définies en tout point $\vec{x}$ de l'espace, tels les champs électrique et magnétique, mais aussi des champs de pression, de température, etc, qui peuvent varier de point à point. Un champ $T(\vec{x})$, scalaire pour les rotations, est tel que

$$T(\vec{x}) \mapsto T'(\vec{x}') = T(\vec{x}) , \quad (2.2)$$

ou encore $T'(\vec{x}) = T(O^{-1}\vec{x})$, cf fig 2.1 : par exemple, la nouvelle température T' au point $\vec{x}'$ du système transformé est égale à celle du système initial au point initial, (ou encore, dans le point de vue passif, la nouvelle valeur de la température dans les nouvelles coordonnées est la même que l'ancienne ...). Un champ de vecteurs $\vec{A}(\vec{x})$ est tel que sous l'effet des rotations

$$\vec{A} \mapsto \vec{A}'(\vec{x}') = O\vec{A}(\vec{x}) \quad (2.3)$$

ou encore $A^i(\vec{x}') = O_{ij}A^j(\vec{x})$.

Il en est bien ainsi du champ électrique $\vec{E}$, et (Exercice) on se convaincra de la compatibilité de la formule donnant son expression en termes d'une densité de charges scalaire ρ , $\vec{E}(\vec{x}) = \frac{1}{4\pi\epsilon_0} \int d^3y \rho(\vec{y}) \frac{(\vec{x}-\vec{y})}{|\vec{x}-\vec{y}|^3}$ avec les transformations (2.2) et (2.3) de ρ et du champ $\vec{E}$, respectivement. Il en est de même du champ magnétique $\vec{B}$ (penser à la loi de Biot et Savart $\vec{B}(\vec{x}) = \frac{\mu_0}{4\pi} \oint I \frac{d\vec{y} \wedge (\vec{x}-\vec{y})}{|\vec{x}-\vec{y}|^3}$; ou encore à la loi de Laplace, $d\vec{F} = \vec{j} \wedge \vec{B} dV$ donnant la force s'exerçant sur un volume élémentaire dV traversé par un courant de densité $\vec{j}$ et soumis à un champ magnétique $\vec{B}$, etc).

On peut aussi s'intéresser aux opérateurs différentiels gradient, rotationnel etc. Le gradient $\vec{\nabla}$ d'une fonction scalaire f se transforme comme un vecteur. La preuve en est presque élémentaire, mais un peu de soin est requis dans les manipulations d'indices

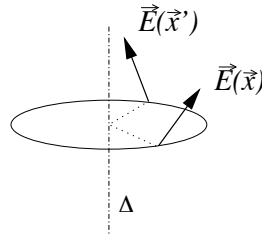
$$\frac{\partial}{\partial x^i} \mapsto \frac{\partial}{\partial x'^i} = \sum_j \frac{\partial x^j}{\partial x'^i} \frac{\partial}{\partial x^j} = \sum_j O_{ij} \frac{\partial}{\partial x^j} \quad (2.4)$$

la dernière équation provenant de l'inversion de $x'^i = \sum_j O_{ij}x^j$ en $x^j = \sum_i O_{ij}x'^i$ (par l'orthogonalité de O) donc $\partial x^j / \partial x'^i = O_{ij}$. Cela justifie au passage l'appellation de "champ de gradient". Le rotationnel d'un champ de vecteurs est un vecteur pour les rotations, sa divergence est un scalaire.

L'électromagnétisme introduit aussi d'autres champs, les potentiels scalaire $V(\vec{x})$ et vecteur $\vec{A}$, dont "dérivent" les champs $\vec{E}$ et $\vec{B}$

$$\vec{E} = -\vec{\nabla}V - \frac{\partial \vec{A}}{\partial t} , \quad \vec{B} = \text{rot} \vec{A}. \quad (2.5)$$

Vérifier que ces deux champs V et $\vec{A}$ sont effectivement respectivement scalaire et vectoriel, sous l'effet des rotations. On pourra soit utiliser les lois de transformations des gradients et

FIGURE 2.2 – Champ à symétrie cylindrique, autour d'un axe Δ .

rotationnels, soit se rappeler les formules qui fournissent les expressions de V et $\vec{A}$ en termes des distributions de charge (scalaire) et de courant (vectorielle).

Les équations de Maxwell

$$\operatorname{div} \vec{E} = \frac{\rho}{\epsilon_0} \quad \operatorname{div} \vec{B} = 0 \quad (2.6)$$

$$\overrightarrow{\operatorname{rot}} \vec{B} - \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t} = \mu_0 \vec{j} \quad \overrightarrow{\operatorname{rot}} \vec{E} + \frac{\partial \vec{B}}{\partial t} = 0 \quad (2.7)$$

impliquent bien des objets de même nature aux membres de gauche et de droite, scalaires pour la première ligne, vecteurs à la seconde. On dira qu'elles sont invariantes par les transformations de $\operatorname{SO}(3)$, voir plus bas.

◦ Invariance d'un champ.

Nous venons d'examiner les transformations de champs scalaires ou vectoriels sous les rotations. Supposons maintenant que le système qu'ils décrivent est *invariant* par ces rotations. On dit qu'un champ V (scalaire ou vectoriel) est invariant si

$$V'(\vec{x}) = V(\vec{x}), \quad (2.8)$$

autrement dit si les *fonctions* V et V' sont identiques. Comparer à (2.2-2.3). Exemples : un champ électrique est invariant pas les rotations autour d'un axe Δ si $\vec{E}'(\vec{x}) = \vec{E}(\vec{x})$, soit en utilisant (2.3), $\vec{E}'(\vec{x}') = \vec{E}(\vec{x}') = \vec{E}(O\vec{x}) = O\vec{E}(\vec{x})$: le champ $\vec{E}$ au point transformé $\vec{x}' = O\vec{x}$ est le transformé du champ au point $\vec{x}$, ce qui correspond bien à notre intuition d'un champ à symétrie cylindrique, voir Fig. 2.2. Autre exemple, une distribution de charge ou de masse est invariante par rotation autour de l'origine si $\rho'(\vec{x}) = \rho(\vec{x})$, donc $\rho(\vec{x}) = \rho(O\vec{x})$. (Il est alors naturel d'utiliser des coordonnées sphériques r, θ, ϕ et $\rho(\vec{x}) = \rho(r)$ ne dépend que de la coordonnée radiale r .)

2.1.2 Tenseurs. Exemples physiques

Après cette discussion des scalaires et vecteurs, un mot sur les *tenseurs* avant de refermer ce paragraphe.

On appelle tenseur de rang n un objet X décrit (à nouveau dans une base orthonormée) par des composantes à n indices, $X^{i_1 i_2 \dots i_n}$, (les indices i prenant 3 valeurs possibles, ici à $d = 3$), et

se transformant (sous l'effet des rotations) selon

$$X \mapsto X' = \{X'^{i_1 i_2 \dots i_n}\} \quad X'^{i_1 i_2 \dots i_n} = \sum_{j_1, \dots, j_n=1}^3 O_{i_1 j_1} O_{i_2 j_2} \dots O_{i_n j_n} X^{j_1 j_2 \dots j_n} \quad (2.9)$$

Autrement dit, $X^{i_1 i_2 \dots i_n}$ se transforme comme le produit $x^{i_1} x^{i_2} \dots x^{i_n}$ de n composantes d'un vecteur $\vec{x}$ (ou encore, X se transforme comme le *produit tensoriel* $\vec{x} \otimes \vec{x} \otimes \dots \otimes \vec{x}$).

NB. Dans le cas présent du groupe orthogonal, il n'y a pas lieu de distinguer entre les tenseurs "contravariants" (à indices en position supérieure) et tenseurs "covariants" (position inférieure).

Noter qu'en "contractant" un tenseur X de rang n avec un tenseur Y de rang $p < n$, c'est-à-dire en construisant

$$Z^{i_1 \dots i_{n-p}} = \sum_{i_{n-p+1}, \dots, i_n} X^{i_1 \dots i_{n-p} i_{n-p+1} \dots i_n} Y^{i_{n-p+1} \dots i_n}$$

on fabrique un nouveau tenseur de rang $n - p$: vérifier que c'est bien un tenseur en utilisant la définition du groupe orthogonal. Une contraction partielle sur $p' < p$ indices conduit à un tenseur de rang $n + p - 2p'$.

Deux exemples de tenseurs souvent rencontrés :

- le tenseur de rang 2 de Kronecker δ_{ij} . (On rappelle $\delta_{ij} = 1$ si $i = j$, $= 0$ sinon.) Sa contraction avec deux vecteurs $\vec{A}$ et $\vec{B}$ (deux tenseurs de rang 1) donne un tenseur de rang 0, un scalaire, $\sum_{ij} \delta_{ij} A^i B^j$ qui n'est autre que le produit scalaire de $\vec{A}$ et $\vec{B}$. Cela montre que le tenseur δ_{ij} est un tenseur *invariant* : $\delta'_{ij} = \delta_{ij}$ avec les notations de (2.9). C'est (à un facteur près) le **seul** tenseur symétrique de rang 2 avec cette propriété (pourquoi?) ;
- le tenseur ε_{ijk} complètement antisymétrique, $\varepsilon_{ijk} = \varepsilon_{kij} = -\varepsilon_{jik}$, $\varepsilon_{123} = 1$, qui sert à fabriquer le produit vectoriel

$$(\vec{A} \wedge \vec{B})^i = \sum_{j,k} \varepsilon_{ijk} A^j B^k .$$

Pour trois vecteurs le "produit mixte" $(\vec{A}, \vec{B}, \vec{C}) = \vec{A} \cdot (\vec{B} \wedge \vec{C}) = \sum_{j,k} \varepsilon_{ijk} A^i B^j C^k$, c'est-à-dire le déterminant de la matrice ayant pour colonnes les composantes de ces trois vecteurs, est le volume orienté du parallélépipède construit sur $\vec{A}$, $\vec{B}$ et $\vec{C}$. Le tenseur ε_{ijk} est lui aussi invariant par l'action des rotations. (En effet par (2.9), $\varepsilon'_{ijk} = \sum_{i',j',k'} O_{ii'} O_{jj'} O_{kk'} \varepsilon_{i'j'k'} = \det O \varepsilon_{ijk} = \varepsilon_{ijk}$, puisque $\det O = 1$, cqfd.)

Des exemples de quantités tensorielles apparaissant en physique sont fournis par les coefficients de proportionnalité (pouvant varier de point à point) reliant deux quantités vectorielles ou tensorielles,

- en mécanique, le tenseur d'inertie $I_{ij} = \int d^3x \rho(\vec{x}) (\delta_{ij} x^2 - x^i x^j)$, où ρ est la densité de masse, qui relie le moment angulaire au vecteur vitesse angulaire $\vec{J} = I \cdot \vec{\omega}$ (voir TD) ;
- en théorie de l'élasticité, pour un corps soumis à une déformation $\vec{u}(\vec{x}) = \vec{x}' - \vec{x}$ en chaque point, on introduit d'une part le tenseur des déformations ε , $\varepsilon_{ij} := \frac{\partial u^i}{\partial x^j} + \frac{\partial u^j}{\partial x^i}$; et d'autre part le tenseur des contraintes σ qui décrit la distribution des forces selon les directions : σ_{ij} est égal à la i -ème composante de la force exercée sur le corps, s'appliquant sur l'élément de surface orthogonale à la direction $\hat{n}$, $F_i = \sigma_{ij} \hat{n}_j$, voir fig 2.3 ; ce sont tous

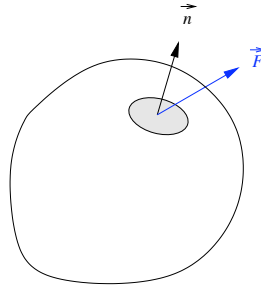


FIGURE 2.3 – Le tenseur des contraintes.

deux des tenseurs symétriques (évident pour ε , moins évident pour σ , [7]) et ils sont reliés par la loi de Hooke

$$\sigma = \underline{H} \cdot \varepsilon$$

impliquant le tenseur d'élasticité $\underline{H}$ de rang 4 ;

- en physique des solides, le tenseur piézoélectrique, (voir ci-dessous) ;
- en théorie de la diffusion, le tenseur de diffusion $\underline{D}$ de rang 2, avec la loi de Fick-Fourier $\vec{j} = -\underline{D} \cdot \vec{\nabla} \phi$ reliant le flux de chaleur ou de particules au gradient de la cause ϕ (température, concentration). Rappelons que, combinée à une relation de conservation du courant $\vec{j}$ de la forme $\partial_t \phi = -c \vec{\nabla} \cdot \vec{j}$, cette loi de Fick-Fourier mène à une équation de diffusion (ou équation de la chaleur) $\partial_t \phi = \kappa \vec{\nabla} \cdot \underline{D} \cdot \vec{\nabla} \phi$ avec au second membre un opérateur laplacien anisotrope ;
- en électromagnétisme, les tenseurs de conductivité $\underline{\sigma} : \vec{j} = \underline{\sigma} \cdot \vec{E}$, ou de polarisabilité électrique $\underline{\alpha}$ qui relie le moment dipolaire induit $\vec{p}$ au champ qui le produit $\vec{p} = \underline{\alpha} \cdot \vec{E}$, etc, etc. Voir TD et [7, 8].

Dans chacun de ces exemples, le cas “isotrope” est celui où le tenseur a une symétrie maximale. Par exemple, si le tenseur de conductibilité électrique $\underline{\sigma}$ est isotrope, c'est-à-dire $\sigma_{ij} = \sigma \delta_{ij}$, on retrouve une loi familière, équivalente à la loi d'Ohm. Cette propriété d'isotropie de tenseurs symétriques de rang 2 est une conséquence de l'invariance du système considéré par les rotations de $SO(3)$: en effet, selon une remarque faite plus haut, un tenseur symétrique invariant (par rotation) de rang 2 est nécessairement proportionnel au tenseur identité δ_{ij} . Mais une symétrie plus faible, par un sous-groupe de $SO(3)$, peut suffire à assurer l'isotropie. Voir des exemples en TD.

2.2 Réflexions et Inversion/Parité

2.2.1 Scalaires et pseudoscalaires, vecteurs et pseudovecteurs

La discussion du paragraphe précédent a porté sur les transformations sous l'effet des rotations. Examinons maintenant ce qui se passe sous l'action des isométries “impropres”, celles

qui renversent les orientations, les réflexions-miroirs, ou l'inversion/parité (nous supposons dans tout ce chapitre que nous sommes en dimension 3). Autrement dit, nous étendons au groupe $O(3)$ la discussion précédente qui ne traitait que de $SO(3)$.

Supposons que sous l'effet de l'inversion/parité, $\vec{x} \mapsto \vec{x}' = -\vec{x}$, une quantité scalaire f pour les rotations reste proportionnelle à elle-même¹

$$f \mapsto f' \quad \text{avec} \quad f'(\vec{x}') = f(\vec{x}) \quad \text{donc} \quad f'(\vec{x}) = f(-\vec{x}) = \lambda f(\vec{x}) . \quad (2.10)$$

Comme une réflexion-miroir à trois dimensions est le produit de l'inversion/parité par une rotation de π d'axe orthogonal au plan de réflexion, et que nous avons supposé la fonction f scalaire pour les rotations, la même formule (2.10) s'applique aussi à toute réflexion, et donc à toute isométrie impropre. Comme la transformation considérée (réflexion ou inversion/parité) est "involutive", c'est-à-dire de carré 1, en répétant deux fois la transformation, on revient à la quantité initiale et on doit donc avoir $\lambda^2 = 1$, ce qui laisse deux possibilités $\lambda = \pm 1$. Le premier cas, $\lambda = 1$, définit les fonctions scalaires, le deuxième $\lambda = -1$ les fonctions pseudo-scalaires. Le même raisonnement s'applique aux objets de nature vectorielle (pour les rotations), ... à cela près que certains signes sont renversés ! En effet sous l'effet de la parité, toutes les composantes d'un vecteur comme la position changent de signe, tandis que celles d'un produit vectoriel ou d'un rotationnel ne changent pas. On dira donc que la position $\vec{x}$, la vitesse $\vec{v}$, le champ électrique, l'impulsion, le gradient $\vec{\nabla} f$ d'une fonction scalaire sont des vecteurs (par rotation et parité), tandis que le moment angulaire $\vec{J} = \vec{x} \wedge \vec{p}$, le champ magnétique $\vec{B}$, etc sont des *pseudovecteurs*. Le rotationnel $\vec{\text{rot}} \vec{A}$ de tout vecteur $\vec{A}$ est un pseudovecteur, et vice versa. Quelle est la nature du "vecteur de Poynting", $\vec{E} \wedge \vec{B}$ (qui à un facteur μ_0 près décrit le flux d'énergie électromagnétique) ? Vecteur ou pseudo-vecteur ?

Nous pouvons maintenant fabriquer des *pseudoscalaires* par produit scalaire d'un vecteur et d'un pseudovecteur : c'est le cas de $\vec{E} \cdot \vec{B}$, une quantité qui apparaît en électromagnétisme ; ou du produit scalaire $\vec{J} \cdot \vec{V}$ du moment angulaire avec un vecteur, etc. C'est aussi le cas du "produit mixte" de trois vecteurs, c'est-à-dire du déterminant de leurs composantes. Géométriquement, cela mesure le volume *orienté* du parallélépipède construit sur ces 3 vecteurs. Ce volume est donc un pseudoscalaire, il change de signe par parité, comme on l'a vu au chapitre 1, § 1.3.1, en discutant de l'orientation d'un repère.

On peut aussi combiner une rotation de π autour d'un axe Δ et l'inversion/parité par rapport à un point de Δ pour étudier comment se transforment différentes quantités sous l'action d'une réflexion-miroir par rapport à un plan perpendiculaire à Δ . Montrer ainsi qu'un champ magnétique $\vec{B}$ porté par Δ et donc orthogonal à ce plan est invariant, un champ $\vec{B}$ dans le plan (donc orthogonal à Δ) change de signe, tandis qu'un champ électrique a une image réfléchie dans le plan. Inversement quelles symétries sont compatibles avec la présence d'un champ électrique uniforme $\vec{E}$, ou avec celle d'un champ magnétique $\vec{B}$? Montrer que toute rotation d'axe $\vec{E}$ et toute réflexion-miroir dans un plan contenant $\vec{E}$ laisse $\vec{E}$ invariant ; et que

1. Toute fonction $f(\vec{x})$ peut s'écrire comme $\frac{1}{2}((f(\vec{x}) + f(-\vec{x})) + (f(\vec{x}) - f(-\vec{x})))$, c'est-à-dire comme la somme d'une fonction paire et d'une impaire, et nous supposons ici qu'un seul des deux termes est non nul. Nous reformulerons plus tard cette hypothèse en disant que nous considérons ici des fonctions se transformant comme une représentation irréductible du groupe à deux éléments engendré par la parité : elle est ou bien paire, ou bien impaire.

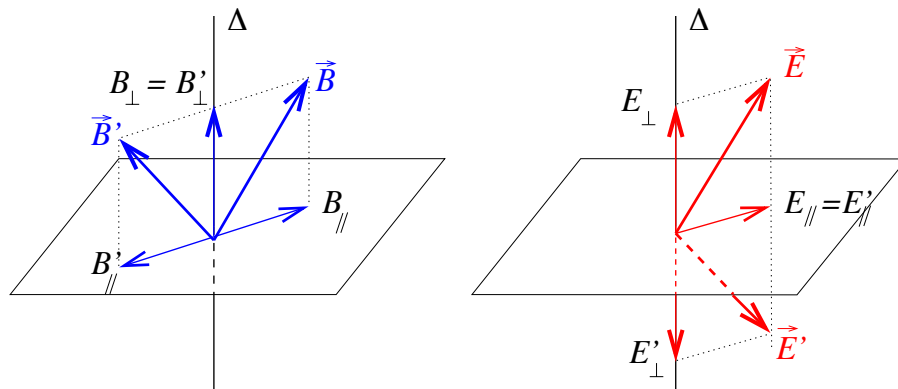


FIGURE 2.4 – Réflexion dans un plan d'un champ magnétique et d'un champ électrique.

toute rotation d'axe $\vec{B}$ ou toute réflexion-miroir dans un plan *orthogonal* à $\vec{B}$ le laisse invariant, voir Fig. 2.4. Ces observations seront utiles dans le paragraphe suivant.

2.2.2 Covariance des grandeurs et invariance des lois de la physique

Toutes les lois de la Physique que nous connaissons sont invariantes par l'action des rotations. Dans le point de vue passif, cela signifie que les dites lois ne dépendent pas de l'orientation du repère de coordonnées, une chose heureuse pour l'expérimentation ! Dans le point de vue actif, cela veut dire qu'elles s'appliquent de la même façon à un système qu'au système obtenu par rotation. Cela implique aussi que toute équation du mouvement (par exemple $\vec{J} = \vec{M}$, moment des forces appliquées, en mécanique) ou de contrainte (par exemple $\text{div} \vec{E} = \rho$ en électricité) reliant différentes quantités physiques implique au membre de gauche et au membre de droite des objets de nature tensorielle identique, se transformant donc de la même façon par rotation. Cela est apparent sur les lois de la mécanique, sur celles de l'électromagnétisme, y compris les lois de Maxwell, sur les lois de Hooke, de Fick qu'on vient de rappeler, etc. Cela est dû au caractère *covariant* des grandeurs considérées. Ce mot "covariant" couvre les différentes situations rencontrées ci-dessus, scalaire, vecteur, tenseur de rang approprié...

Il est naturel de se demander si cette invariance des lois de la physique s'étend à d'autres transformations que les rotations

- translations d'espace ou de temps (la réponse est oui, sinon les lois de la physique dépendraient de l'endroit ou du moment de l'expérience!);
- transformations de Galilée mêlant transformations d'espace et de temps, dans lesquelles les deux repères de coordonnées diffèrent non seulement par une combinaison des transformations précédentes (rotation $\mathcal{R}$ et translation d'espace $\vec{a}$ et de temps δ) mais aussi par une translation de vitesse constante $\vec{v}$: $\vec{x}' = \mathcal{R}\vec{x} + \vec{v}t + \vec{a}$, $t' = t + \delta$ (repères *inertiels*); l'invariance de Galilée, vraie pour les lois de la mécanique, ne l'est plus pour celles de l'électromagnétisme : on doit lui substituer l'invariance relativiste...
- ou encore réflexions et inversion/parité. Pour ces deux dernières, les choses sont plus subtiles, nous y reviendrons (chapitre 6).

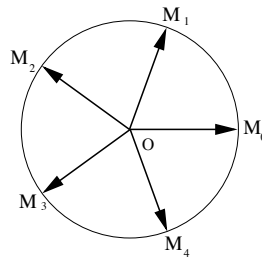


FIGURE 2.5 – La somme $\vec{S}$ des vecteurs OM_i est nulle par symétrie de rotation de $2\pi/5$. Si elle ne l'était pas, $\vec{S}$ serait un vecteur invariant par ces rotations, ce qui est absurde.

2.3 Principe de Curie

2.3.1 Applications physiques d'arguments de symétrie

Nous venons d'étudier comment les grandeurs physiques se transforment sous l'effet des transformations ponctuelles les plus couramment rencontrées, rotations, réflexions, inversion/parité. Supposons maintenant que le système considéré est *invariant* sous certaines de ces transformations. Bien souvent un simple argument de symétrie permet d'éviter des calculs explicites.

C'est le cas en géométrie où on démontre ainsi que la somme des rayons vecteurs du centre d'un polygone régulier à ses n sommets s'annule, cf. fig 2.5 pour le cas d'un pentagone, ce qui résulte aussi de l'identité $\sum_{p=0}^{n-1} e^{2\pi ip/n} = 0$; ou encore que pour deux points A et A' symétriques par rapport à un plan passant par O, la somme $\vec{OA} + \vec{OA'}$ appartient au plan, etc.

Mais c'est aussi le cas dans de nombreux problèmes de physique. Bornons nous à quelques exemples.

- Une charge électrique initialement au repos et soumise à un champ électrique et un champ magnétique orthogonaux acquiert un mouvement dans le plan orthogonal à $\vec{B}$. En effet une réflexion dans ce plan laisse $\vec{E}$ et $\vec{B}$ invariants, (voir remarque à la fin du § 2.2), c'est donc une symétrie du problème. On retrouve bien sûr ce résultat élémentaire par application de la loi de Laplace $\vec{F} = q(\vec{v} \wedge \vec{B} + \vec{E})$, qui montre que l'accélération $\vec{a}$ de la charge est contenue dans le plan, et donc aussi la vitesse et la position si elles l'étaient à l'instant initial.

- Considérons une molécule à symétrie axiale. Elle peut porter un moment électrique dipolaire. Mais la présence d'une symétrie d'inversion de la molécule par rapport à son centre assure que le moment dipolaire est nul, puisqu'il change de signe par cette inversion.

- Considérons le tenseur d'inertie d'un solide I_{ij} : c'est un tenseur de rang 2 pour le groupe $O(3)$, puisqu'il relie deux pseudovecteurs, $\vec{J}$ et $\vec{\omega}$. Si on a invariance du solide par le groupe $O(3)$ tout entier, (cas d'un corps à symétrie sphérique), I_{ij} doit être proportionnel au tenseur δ_{ij} qui est le seul tenseur invariant de rang 2. Ce tenseur est donc alors *isotrope*. Mais l'existence d'une symétrie par un sous-groupe de $O(3)$, comme un des groupes ponctuels étudiés au chapitre 1, peut suffire à contraindre la forme de I_{ij} , voir TD. Un raisonnement similaire s'applique au tenseur diélectrique, qui relie dans un milieu polarisable le vecteur "déplacement électrique" $\vec{D}$

à l'intensité du champ $\vec{E}$: $\vec{D} = \underline{\varepsilon} \cdot \vec{E}$, cf TD.

Inversement certains phénomènes physiques ne peuvent se manifester dans une situation (un matériau, tel un cristal, ou une configuration de champs extérieurs, etc) qui présente “trop” de symétrie. L'existence d'un tenseur $\underline{\gamma}$ de rang 3 non nul dans un corps cristallin n'est compatible qu'avec certains types de symétrie. Par exemple, sous l'effet de l'inversion/parité, un tel tenseur change de signe. Si le corps (par exemple un cristal) est invariant par inversion, ce tenseur s'annule. Un exemple va être fourni par les effets piézoélectrique et pyroélectrique.

Piézoélectricité. Pyroélectricité. Ferroélectricité

On appelle *piézoélectricité* la propriété de certains cristaux d'acquies une polarisation sous l'effet de déformations mécaniques, telles des tractions ou compressions. Ce phénomène, découvert par Jacques et Pierre Curie en 1880, a été exploité dans les sonars et dans les têtes de lecture des tourne-disques. On définit le *tenseur piézoélectrique*, qui relie le champ de déplacement électrique $\vec{D}$ induit au tenseur $\underline{\varepsilon}$ de déformation du solide (cf § 2.1.2) : $\vec{D} = \underline{\gamma} \cdot \underline{\varepsilon}$, ou $D_i = \sum_{jk} \gamma_{i,jk} \varepsilon_{jk}$. Selon l'argument qu'on vient de donner, seuls les cristaux ne possédant pas de centre de symétrie peuvent être sujets au phénomène de piézoélectricité. Seules 21 classes cristallines sur les 32 remplissent ce critère, et 20 présentent effectivement le phénomène de piézoélectricité.

Parmi ces classes piézoélectriques, 10 possèdent une polarisation électrique spontanée et sont dites polaires. Le phénomène est un analogue électrique du ferromagnétisme, où une aimantation spontanée peut apparaître (cf Chap 3). Au contraire de cette dernière, toutefois, la polarisation est assez rapidement neutralisée par l'accumulation de charges (issues de l'environnement ou de l'intérieur du cristal). La présence du phénomène de polarisation n'est souvent observable que par le biais de sa variation sous l'effet d'un échauffement, et on parle de pyroélectricité.

La *pyroélectricité* (du grec *pyr*, feu) est la propriété d'un matériau dans lequel un changement de température entraîne une variation temporaire de polarisation électrique. L'effet est dû aux déplacements des positions des atomes causés par la variation de température. L'effet pyroélectrique ne doit pas être confondu avec l'effet thermoélectrique, dans lequel un gradient de température donne naissance à une tension permanente.

Si la polarisation permanente peut être renversée par l'application d'un champ électrique extérieur, (comme on le fait avec l'aimantation permanente d'un ferromagnétique par application d'un champ magnétique), le corps est dit *ferroélectrique*.

Ces ferroélectriques font partie de la classe des *ferroïques*², qui présentent des propriétés électriques, magnétiques ou élastiques couplées. Une propriété étudiée actuellement dans ces matériaux “actifs” est celle qui couple phénomènes électriques et magnétiques, avec une polarisation électrique $\vec{P}$ causée par l'application d'un champ magnétique $\vec{H}$, et vice versa une aimantation $\vec{M}$ causée par un champ électrique $\vec{E}$. On définit les tenseurs de susceptibilité χ^{EM} et χ^{ME}

$$P_i = \chi_{ij}^{EM} H_j \quad M_i = \chi_{ij}^{ME} E_j$$

et on montre que $\chi_{ij}^{EM} = \chi_{ji}^{ME}$. À nouveau, seules certaines classes cristallines sont susceptibles de manifester ces phénomènes. . .

Exemple : la tourmaline, un silicate de métaux qui peuvent être l'aluminium, le fer, le magnésium, le sodium, le lithium, le calcium, le potassium, etc, cristallise dans le système cristallin trigonal à réseau rhomboédrique, en baguettes ou aiguilles allongées à section triangulaire et faces courbes. Elle présente l'effet piézoélectrique, et également des propriétés pyroélectriques : quand elle est chauffée, les deux extrémités des cristaux accumulent des charges électriques opposées. Le quartz aussi manifeste l'effet piézoélectrique mais pas la pyroélectricité.

2. Le préfixe “ferro” a une origine historique et se réfère à la présence de fer dans les premiers corps étudiés présentant ces propriétés. Les corps ferroïques modernes ne contiennent plus nécessairement de fer, tels les céramiques PZT, $\text{PbZr}_{1-x}\text{Ti}_x\text{O}_3$, très utilisées dans l'industrie pour leurs propriétés piézoélectriques et ferroélectriques. La structure cristalline des PZT est de type pérovskite, voir exercice du Chap. 1.

2.3.2 Principe de Curie

Tous les exemples que nous venons de considérer conduisent au

Principe de Curie (1894)

Lorsque certaines causes produisent certains effets, les éléments de symétrie des causes doivent se retrouver dans les effets produits.

*Lorsque certains effets révèlent une certaine dissymétrie, cette dissymétrie doit se retrouver dans les causes qui lui ont donné naissance.*³

Autrement dit, les symétries des “causes” (conformation du système, ...) se répercutent dans celles des “effets” (conséquences observées). Mais les symétries de ces effets peuvent être plus grandes que celles des causes. Par exemple, considérons le cas d’un cristal de symétrie cubique. Tout tenseur de rang 2 (piézoélectrique, de conductivité ou autre) doit être invariant par les opérations du groupe ponctuel de symétrie (donc un sous-groupe de $O(d)$), ce qui lui impose d’être isotrope, donc finalement invariant par le groupe de rotation tout entier. Mais les symétries des effets, selon ce principe, ne peuvent être moindres que celles des causes. Réciproquement, (deuxième partie du principe), un effet (phénomène) impliquant la brisure d’une symétrie ne peut résulter de certaines causes si cette symétrie est présente (et non brisée) dans les causes.

Une illustration de cette réciproque a été fournie par un argument de Curie lui-même sur l’impossibilité de synthèse chirale : il est impossible qu’un champ magnétique uniforme seul permette de synthétiser des molécules chirales, comme l’avait tenté Pasteur dans ses jeunes années. En effet la présence d’un tel champ laisse une symétrie de réflexion par rapport au plan orthogonal, comme on l’a vu (Fig. 2.4), ce qui rend impossible tout phénomène brisant cette symétrie⁴.

Noter l’importance de la “réciproque” (la deuxième partie) du principe de Curie : selon le propos de P. Curie lui-même, *C’est la dissymétrie qui crée le phénomène*. Un exemple en a été donné plus haut avec la piézoélectricité dans un cristal, dont l’existence impose l’*absence* de symétrie par inversion.

2.3.3 Autres illustrations du principe de Curie

Dans son article de 1894, Curie classe les groupes de symétrie possibles, engendrés par des rotations discrètes ou continues, des réflexions-miroir et/ou l’inversion/parité, compatibles avec l’existence d’un axe d’isotropie Δ , qu’on peut toujours considérer comme vertical. Ce sont donc tous des sous-groupes du groupe $O(3)$ des isométries de l’espace.

- (a) le groupe de symétrie $D_{\infty h}$ du cylindre fini ou infini d’axe Δ , engendré par les rotations d’angle quelconque autour de l’axe Δ , la rotation de π autour d’un axe orthogonal à Δ passant par le milieu O du cylindre, les réflexions dans tout plan passant par Δ ou dans un plan orthogonal en O à Δ , l’inversion par rapport à O . C’est la symétrie maximale d’un corps doté d’un axe optique (une lentille par exemple) ;

3. Pierre Curie, *Sur la symétrie dans les phénomènes physiques, symétrie d’un champ électrique et d’un champ magnétique*, Journ. de Physique **3** 393-415 (1894).

Rappelons qu’avant ses travaux sur la radioactivité avec Marie Sklodowska (et son mariage avec elle!), Pierre Curie (1859-1906) a à son actif des travaux nombreux sur le ferromagnétisme, les symétries en physique, la piézoélectricité etc. Il a partagé le prix Nobel de physique 1903 avec Henri Becquerel et Marie Curie.

4. On pourrait imaginer que la présence d’un champ électrique *et* d’un champ magnétique, donc l’existence d’un pseudo-scalaire $\vec{E} \cdot \vec{B}$, permet de tourner cette objection et autorise cette synthèse. Comme l’a montré à son tour P.-G. de Gennes, cela est impossible à l’équilibre [4].

inversement tout corps doté d'un axe optique a un groupe de symétrie qui est un sous-groupe du groupe du cylindre ;

- (b) symétrie D_∞ d'un cylindre torsadé (comme une tige filetée) : rotations d'axe Δ , rotation de π autour de l'axe orthogonal en O à Δ , (mais plus de symétrie “impropre”, réflexion-miroir ou inversion/parité) ; c'est la symétrie d'un corps doté de pouvoir rotatoire. Un tel corps est chiral, n'étant pas identique à son image miroir ; par exemple un cristal de quartz, dont le groupe de symétrie est un sous-groupe du groupe considéré.
- (c) symétrie $C_{\infty v}$ d'un tronc de cône : rotations d'axe Δ et réflexions dans tout plan passant par Δ . (On a perdu la symétrie haut-bas, par rotation de π ou par réflexion dans un plan orthogonal à Δ). C'est la symétrie de tout problème physique en présence d'un champ vertical (parallèle à Δ), champ de pesanteur, champ électrique, etc, ayant la symétrie du cylindre. Par exemple, le champ électrique créé par deux plateaux circulaires parallèles d'axe Δ et portant une distribution de charges isotrope. Cette symétrie doit se retrouver dans les “effets”, en particulier dans la dynamique d'un corps massif, resp. chargé, initialement au repos, placé dans ce champ.
- (d) symétrie $C_{\infty h}$ d'un cylindre tournant : rotations d'axe Δ , réflexions dans tout plan orthogonal à Δ , inversion de centre sur Δ ; rappelons en effet que la rotation est décrite par un pseudovecteur $\vec{\omega}$, laissé invariant par l'inversion. C'est aussi la symétrie d'un champ magnétique uniforme, ou d'un champ magnétique doté d'une symétrie de révolution autour d'un axe Δ , tel celui créé par une boucle de courant circulaire centrée sur et orthogonale à Δ : la symétrie de la cause (courant) –rotation et réflexion dans le plan– se retrouve dans l'effet, le champ magnétique.
- (e) enfin symétrie C_∞ consistant en les seules rotations autour d'un axe Δ , sans réflexion ni inversion. Ce groupe est un sous-groupe commun aux cas (c) et (d). Donner une situation physique où on rencontre ce groupe de symétrie.

Exercice Montrer que les seules symétries possibles d'une molécule diatomique sont $D_{\infty h}$ et $C_{\infty v}$.

2.4 Translations

Nous examinons ici brièvement les premières conséquences physiques de l'existence de symétries de translation dans un système cristallin. On va voir apparaître le rôle important joué par le réseau réciproque.

2.4.1 Diffraction cristalline, loi de von Laue

Considérons un réseau cristallin $\mathcal{R}$ engendré par les vecteurs $\vec{a}_1, \vec{a}_2, \vec{a}_3$, non coplanaires, cf (1.1). Supposons qu'on dirige un faisceau monochromatique (de rayons X, disons, mais les électrons ou neutrons sont aussi utilisés) de vecteur d'onde $\vec{k}$ sur le réseau et qu'on en observe la diffusion élastique dans la direction de vecteur d'onde $\vec{k}'$, $|\vec{k}| = |\vec{k}'| = 2\pi/\lambda$. On sait que la diffraction va donner lieu à des phénomènes d'interférences, les pics étant donnés par la condition de von Laue⁵

$$\vec{a}_i \cdot (\vec{k}' - \vec{k}) = 2\pi n_i \in \mathbb{Z}, \quad i = 1, 2, 3 \quad (2.11)$$

pour chacun des vecteurs de base du réseau, où n_i doit être un entier, voir Fig. 2.6.

L'interprétation géométrique de cette condition est que le vecteur $\Delta\vec{k} = \vec{k}' - \vec{k}$ doit appartenir

5. Max von Laue (1879-1960), prix Nobel de physique 1914, eut le premier l'idée d'irradier un cristal par un faisceau des rayons X découverts peu avant par Röntgen, prouvant ainsi sa structure périodique.

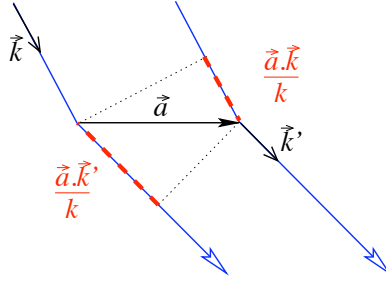


FIGURE 2.6 – Deux rayons diffractés par deux sites séparés par un vecteur $\vec{a}$ du réseau interfèrent constructivement si la différence des chemins optiques (en tirets) est un multiple de la longueur d’onde, $\vec{a} \cdot (\vec{k}' - \vec{k})/k = n\lambda$, d’où la condition (2.11).

au *réseau réciproque* $\tilde{\mathcal{R}}$, engendré par les vecteurs $\vec{b}_1, \vec{b}_2, \vec{b}_3$ satisfaisant

$$\vec{a}_i \cdot \vec{b}_j = 2\pi \delta_{ij} \quad (2.12)$$

(le facteur 2π n’est pas conventionnel, mais fort utile pour les applications physiques), voir TD.

Plus précisément, l’onde incidente au point $\vec{x}$ a pour amplitude $Ae^{i\vec{k} \cdot \vec{x}}$, et l’onde diffractée par $\vec{x}$ et observée en $\vec{y}$ a pour amplitude $Ae^{i\vec{k} \cdot \vec{x}} e^{i\vec{k}' \cdot (\vec{y} - \vec{x})}$. Si la diffusion s’effectue sur des points de densité $\rho(\vec{x})$, l’amplitude totale est proportionnelle à

$$\int d^3x \rho(\vec{x}) e^{i\vec{k} \cdot \vec{x}} e^{i\vec{k}' \cdot (\vec{y} - \vec{x})} = e^{i\vec{k}' \cdot \vec{y}} \tilde{\rho}(\vec{k}' - \vec{k})$$

où $\tilde{\rho}$ est la transformée de Fourier de la densité ρ , et l’intensité mesurée est proportionnelle à $|\tilde{\rho}(\vec{k}' - \vec{k})|^2$.

Pour un cristal idéal de taille infinie, en prenant pour ρ une somme de fonctions delta sur les sites du réseau $\mathcal{R}$, $\rho(\vec{x}) = \sum_{\vec{a} \in \mathcal{R}} \delta(\vec{x} - \vec{a})$, on montre (exercice 3.2 du TD02) que $\tilde{\rho}(\vec{k})$ est une somme de fonctions delta sur les sites du réseau réciproque, $\tilde{\rho}(\vec{k}) = \sum_{\vec{b} \in \tilde{\mathcal{R}}} \delta(\vec{k} - \vec{b})$. L’effet de la taille finie du cristal est d’“adoucir” les pics infinis de la fonction delta.

Indices de Miller

Dans un réseau $\mathcal{R}$ engendré par les vecteurs $\vec{a}_1, \vec{a}_2, \vec{a}_3$, considérons un plan, dit *plan réticulaire*, P passant par trois points du réseau : il en contient donc une infinité, et il constitue un sous-réseau à deux dimensions, engendré par deux vecteurs $\vec{a}'_1$ et $\vec{a}'_2$ de $\mathcal{R}$. On peut toujours choisir comme nouvelle base de $\mathcal{R}$ $\vec{a}'_1$ et $\vec{a}'_2$, complétés par un troisième vecteur $\vec{a}'_3$. Considérons maintenant le réseau réciproque et sa base $\vec{b}'_j$ duale des $\vec{a}'_i$. En vertu de (2.12) le vecteur $\vec{b}'_3$ est normal à $\vec{a}'_1$ et $\vec{a}'_2$, donc au plan réticulaire P . C’est le plus petit vecteur de $\tilde{\mathcal{R}}$ ayant cette propriété (le vérifier). Sa longueur est

$$\|\vec{b}'_3\| = \frac{2\pi}{d}$$

où d est la distance entre P et le plan réticulaire parallèle à P le plus proche.

On appelle *indices de Miller* (h, k, l) du plan P les coordonnées de $\vec{b}'_3$ dans la base initiale du réseau réciproque $\tilde{\mathcal{R}}$

$$\vec{b}'_3 = h\vec{b}_1 + k\vec{b}_2 + l\vec{b}_3.$$

Le vecteur $\vec{b}'_3$ étant le plus petit vecteur du réseau réciproque orthogonal à P , ses composantes (h, k, l) sont nécessairement des entiers premiers entre eux.

Exemple, dans le réseau cubique, avec un repère orthogonal $\vec{a}_i$ et normé par $\vec{a}_i \cdot \vec{a}_j = a^2 \delta_{ij}$, on peut prendre $\vec{b}_i = 2\pi \vec{a}_i / a^2$. Le plan engendré par $\vec{a}_1$ et $\vec{a}_3$ a $\vec{b}'_2 = \vec{b}_2$, et a donc pour indices de Miller $(0, 1, 0)$; le plan passant par l'extrémité de $\vec{a}_1$ et $\vec{a}_2$ et parallèle à $\vec{a}_3$ a $\vec{b}'_3 = 2\pi(\vec{a}_1 + \vec{a}_2)/a$, donc $(h, k, l) = (1, 1, 0)$, et $d = a/\sqrt{2}$; le plan passant par l'extrémité de $\vec{a}_1$, $\vec{a}_2$ et $\vec{a}_3$ a $(h, k, l) = (1, 1, 1)$, etc.

Ces indices de Miller sont très utilisés en cristallographie et en physique du solide pour identifier les plans réticulaires.

2.4.2 Zones de Brillouin

Anticipant un peu sur la discussion des symétries en mécanique quantique (chapitre 6), considérons le cas d'une équation de Schrödinger satisfaite par la fonction d'onde d'un électron dans un potentiel périodique créé par un cristal

$$H\psi(\vec{r}) = E\psi(\vec{r}) \quad (2.13)$$

Le théorème de Bloch, que nous démontrerons plus bas (chap. 6), affirme⁶ que toute fonction propre de (2.13) est de la forme

$$\psi(\vec{r}) = e^{i\vec{k} \cdot \vec{r}} u_{\vec{k}}(\vec{r}) \quad (2.14)$$

où $u_{\vec{k}}(\vec{r})$ est périodique sous l'action des périodes du réseau (1.1)

$$u_{\vec{k}}(\vec{r} + \vec{t}) = u_{\vec{k}}(\vec{r}) \quad \forall \vec{t} \in \mathcal{R}$$

et donc

$$\psi(\vec{r} + \vec{t}) = e^{i\vec{k} \cdot \vec{t}} \psi(\vec{r}). \quad (2.15)$$

Ce sont les fonctions d'onde de Bloch, cf le cours de Physique de matière condensée.

Il est clair que le vecteur $\vec{k}$ qui apparaît dans (2.15) est défini modulo l'addition de tout vecteur du réseau réciproque $\tilde{\mathcal{R}}$. Pour cette raison, on choisit de le restreindre à un "domaine fondamental" (une maille élémentaire) de ce réseau réciproque, appelé *1ère zone de Brillouin*. La forme de cette 1ère zone de Brillouin dépend du réseau considéré, et il est d'usage de la centrer sur un site du réseau $\tilde{\mathcal{R}}$ (plutôt que de considérer un parallélépipède construit sur les 3 vecteurs $\vec{b}_i$). La construction de cette 1ère zone de Brillouin peut se faire en utilisant les cellules de Wigner-Seitz (appelées aussi cellules de Voronoi en mathématiques), qui sont définies comme suit. Autour de chaque nœud du réseau réciproque, on définit le domaine des points de l'espace plus proches de ce nœud que de tout autre nœud. Cette cellule est donc un polyèdre dont les plans sont des plans médiateurs entre paires de nœuds voisins.

Voir figure 2.7 pour les réseaux réciproques des 5 réseaux de Bravais à $d = 2$ de la figure 1.21 et leur 1ère zone de Brillouin.

Voir aussi [14] chap. 8, et l'exercice ci-dessous.

6. Dans l'étude des systèmes dynamiques, ce théorème, appliqué à des phénomènes périodiques en *temps*, porte le nom de théorème de Floquet.

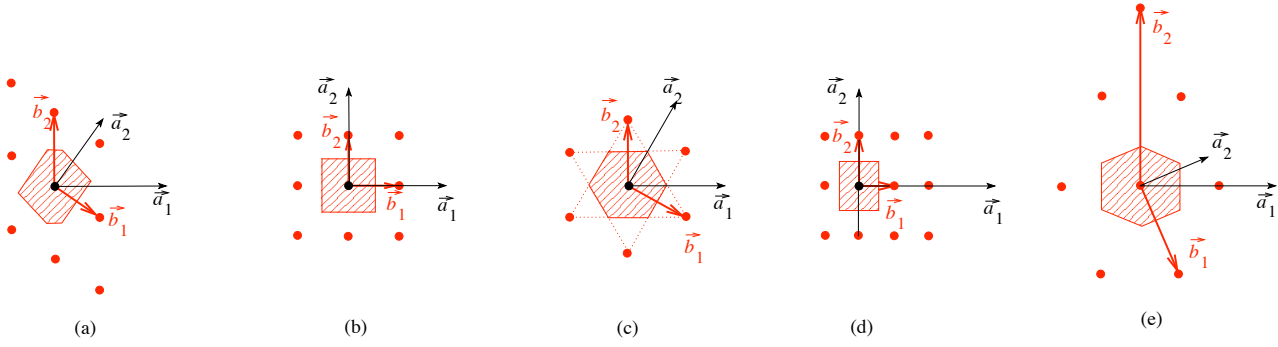


FIGURE 2.7 – Réseaux réciproques des 5 réseaux de Bravais à $d = 2$ de la figure 1.21 et leur 1ère zone de Brillouin

2.5 Modes propres de vibration d'une molécule

On rappelle (cf le cours de physique moléculaire) que dans le calcul des niveaux d'énergie d'une molécule, interviennent trois effets : les interactions des électrons dans le champ créé par une configuration fixe des noyaux ; les mouvements de vibration de ces noyaux autour de leur position d'équilibre ; et le mouvement de rotation de la molécule. De façon générique, les énergies en jeu diffèrent largement

$$\Delta E_{\text{élect}} \gg \Delta E_{\text{vibr}} \gg \Delta E_{\text{rot}}$$

ce qui autorise un traitement indépendant des différents effets. (Voir [2] chap 28, p 1255 ff pour une discussion de ces différentes échelles d'énergie). Dans ce paragraphe, nous abordons la question des niveaux d'énergie de vibration d'une molécule, sur lesquels nous reviendrons par la suite.

Problème : étant donnée une molécule dotée d'une certaine symétrie, exploiter cette symétrie pour construire les modes propres de vibration. Si la molécule est faite de n atomes, on note ξ_i , ($i = 1, \dots, N'$), les $N' = 3n$ coordonnées décrivant l'écart des n noyaux à leur position d'équilibre. Dans une approximation harmonique, l'hamiltonien (classique ici, quantique plus tard) de la molécule est de la forme

$$H = \sum_{i=1}^{N'} \frac{1}{2} m_i \dot{\xi}_i^2 + \frac{1}{2} \sum_{1 \leq i, j \leq N'} \left. \frac{\partial V}{\partial \xi_i \partial \xi_j} \right|_0 \xi_i \xi_j, \quad (2.16)$$

ou encore avec les variables $q_i = \sqrt{m_i} \xi_i$

$$H = \sum_{i=1}^{N'} \frac{1}{2} \dot{q}_i^2 + \frac{1}{2} \sum_{1 \leq i, j \leq N'} \left. \frac{\partial V}{\partial q_i \partial q_j} \right|_0 q_i q_j, \quad (2.17)$$

On cherche à diagonaliser simultanément les deux termes d'énergie cinétique et d'énergie potentielle, c'est-à-dire à trouver des coordonnées Q_i , fonctions linéaires des q_j , telles que

$$H = \frac{1}{2} \sum_{1 \leq i \leq N} (\dot{Q}_i^2 + \omega_i^2 Q_i^2) \quad (2.18)$$

où ω_i est la *fréquence propre* du i -ème *mode normal* Q_i . Le problème admet une solution en vertu du théorème de diagonalisation de toute matrice symétrique A (ici $A_{ij} = \left. \frac{\partial V}{\partial q_i \partial q_j} \right|_0$) par une matrice orthogonale O , $A = ODO^T$, $D = \text{diag}(\omega_i^2)$. Le changement de variables cherché est donc $Q_i = \sum_j O_{ji} q_j$, qui n'affecte pas la forme quadratique de l'énergie cinétique (O orthogonale) et diagonalise bien l'énergie potentielle.

Noter que l'invariance de l'hamiltonien par une translation arbitraire globale ou une rotation arbitraire globale de la molécule implique l'existence de "modes nuls". Ces modes, qui ne nous intéressent pas ici et seront donc éliminés, sont en nombre égal à 5 ou 6, selon que la molécule est linéaire ou non. Il y a en effet 3 modes de translation globale de la molécule, et 3 de rotation globale pour une molécule non linéaire; mais si la molécule a un axe de symétrie Δ , seules les 2 rotations indépendantes d'axe différent de Δ agissent effectivement sur la molécule. Une molécule possède donc $N = 3n - 5$ (cas linéaire) ou $N = 3n - 6$ degrés de liberté de vibration.

Dans la recherche de ces N modes propres et dans leur interprétation, on est aidé par l'utilisation des symétries de la molécule, et plus précisément par l'étude des représentations de son groupe de symétrie. Nous y reviendrons au chapitre 5.

2.6 Exercice

Cellules de Wigner–Seitz–Voronoi

Étude des cellules de Wigner-Seitz-Voronoi à 2 dimensions

- Soient $\vec{a}_1, \vec{a}_2$ deux vecteurs de base de $\mathcal{R}$, $\vec{b}_1, \vec{b}_2$ la base duale dans $\tilde{\mathcal{R}}$. Montrer que pour tout vecteur $\vec{b}$ de $\tilde{\mathcal{R}}$ il existe un vecteur $\vec{a}$ de $\mathcal{R}$ qui lui est orthogonal.
- Montrer en utilisant le point (a.) que les faces des cellules de WSV dans $\tilde{\mathcal{R}}$ sont parallèles à des vecteurs de $\mathcal{R}$.

Chapitre 3

Brisure spontanée de symétrie

Nous avons vu qu’une symétrie peut être présente dans un système physique et brisée par la modification de son environnement –branchement d’un champ extérieur, déformation du système, etc. Une tout autre façon de briser une symétrie est par le mécanisme de brisure spontanée, auquel ce chapitre est consacré. Ce mécanisme apparaît dans de très nombreuses situations physiques. Nous l’illustrerons plus particulièrement par l’étude de la transition ferromagnétique–paramagnétique.

3.1 Brisure spontanée de symétrie discrète

3.1.1 Un modèle mécanique

Considérons d’abord un système mécanique très simple, celui d’un point matériel à une dimension soumis à un potentiel pair $V(x) = V(-x)$, par exemple

$$V(x) = \alpha x^4 + \beta x^2 \quad (3.1)$$

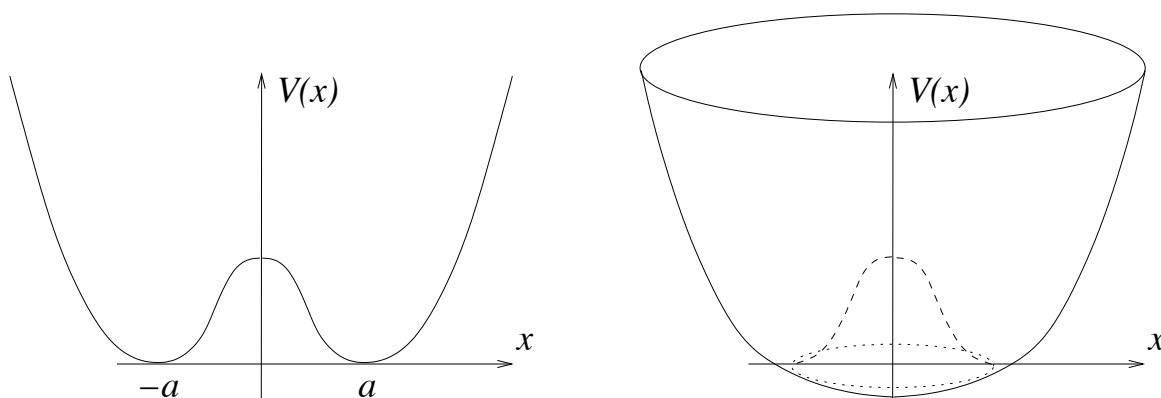


FIGURE 3.1 – Double puits de potentiel. Le potentiel en “chapeau mexicain” : en pointillés, le “cercle de gorge”.

avec des coefficients α et β constants, $\alpha > 0$ et β pour le moment quelconque. L'équation du mouvement est $\ddot{x} = -V'(x)$, soit ici

$$\ddot{x} + (4\alpha x^2 + 2\beta)x = 0 \quad (3.2)$$

et le système est clairement invariant pas la réflexion $x \mapsto -x$, on parlera de symétrie $\mathbb{Z}_2$. Cette réflexion doit s'appliquer aussi aux conditions initiales, et l'assertion correcte et complète consiste donc à dire :

L'unique solution $x(t)$ de l'équation du mouvement satisfaisant des conditions initiales $x(0) = x_0$, $\dot{x}(0) = v_0$ admet une image par réflexion, $\tilde{x}(t) = -x(t)$, elle aussi solution de l'équation du mouvement et satisfaisant aux conditions initiales $\tilde{x}(0) = -x_0$, $\dot{\tilde{x}}(0) = -v_0$.

La symétrie $\mathbb{Z}_2$ est brisée *explicitement* si on ajoute un terme par exemple linéaire au potentiel, $V_\gamma(x) = V(x) + \gamma x$; le potentiel V_γ n'est plus pair, et la propriété qu'on vient d'énoncer n'est plus vraie.

Revenons au cas symétrique, avec le potentiel pair (3.1). Les positions d'équilibre possibles sont à chercher parmi les extrema du potentiel, et donc parmi les solutions de $V'(x_c) = (4\alpha x_c^2 + 2\beta)x_c = 0$. Si $\beta > 0$ (et $\alpha > 0$ comme précédemment), $x_c = 0$ est la seule solution. En revanche si $\beta < 0$, on a trois extrema, soit $x_c = 0$ ou $\pm a$, avec $a^2 = -\beta/2\alpha$, mais notre intuition nous dit –et une analyse de stabilité nous le confirme–, le point $x_c = 0$ n'est pas stable¹. Si on choisit $x(0) = x_0 = 0$, $v_0 = 0$ comme condition initiale, la perturbation la plus infime va précipiter le point soit vers la gauche soit vers la droite, dans un mouvement oscillant dans un des deux puits de potentiel. On parle alors de symétrie $\mathbb{Z}_2$ “brisée spontanément”, le mot *spontané* se référant au fait qu'il suffit d'une perturbation infinitésimale pour la déclencher, soit par le choix d'une condition initiale $x_0 \neq 0$ arbitrairement petite, soit par l'addition d'un terme γx dans le potentiel, brisant la symétrie, si petit qu'il soit. Elle se distingue ainsi d'une brisure explicite de la symétrie, telle celle mentionnée plus haut pour un terme γx fini.

Ce cas très simple de brisure spontanée d'une symétrie, ici $\mathbb{Z}_2$ à une dimension, est juste l'illustration d'un mécanisme qui gouverne de très nombreux phénomènes physiques, depuis les transitions de phase de la matière condensée ou des systèmes ferromagnétiques aux théories modernes de physique des particules et de cosmologie.

3.1.2 Ferromagnétisme

Rappelons l'observation de base. Au dessus d'une température T_c (“c” pour critique ou pour Curie), un système ferromagnétique est dans une *phase paramagnétique*, avec une aimantation nulle en champ extérieur nul. En dessous de cette température, il présente une aimantation “spontanée” M_{sp} non nulle même en champ nul, il est dans une *phase ferromagnétique*. Plus précisément, à $T < T_c$ fixée, si on part d'une situation où le matériau est soumis à un champ

1. Rappelons qu'une *linéarisation* de la solution au voisinage de l'une de ces solutions, $x = x_c + \xi$, avec un développement de l'équation (3.2) au premier ordre en ξ supposé petit, mène à une équation d'oscillateur harmonique $\ddot{\xi} + \omega^2 \xi = 0$, avec une fréquence carrée ω^2 *négative* pour $x_c = 0$ et positive pour $x_c = \pm a$: on a donc des petites oscillations dans ce dernier cas, et un ξ croissant dans le premier cas, ce qui signale l'instabilité de la solution $x_c = 0$.

magnétique $\vec{H}$ suffisamment fort, les moments magnétiques des atomes constitutants ont tendance à s'aligner avec $\vec{H}$ et il en résulte une aimantation $\vec{M}$ parallèle à $\vec{H}$. Si on fait décroître puis changer de signe $\vec{H}$, l'aimantation décroît mais on constate qu'à $\vec{H} = 0$, l'aimantation ne s'est pas encore annulée mais prend (en module) la valeur M_{sp} et qu'il faut aller à des valeurs de $\vec{H}$ opposées à la direction initiale et suffisamment fortes pour la voir s'annuler puis s'aligner à nouveau avec $\vec{H}$: c'est le phénomène d'hystérésis, cf Fig.3.2(a).

L'aimantation spontanée décroît avec la température et s'annule à T_c , Fig. 3.2(b). Il y a donc une transition de phase en $T = T_c$ entre une phase ordonnée, d'aimantation non nulle, et une phase non ordonnée, symétrique (aimantation nulle). La transition est une transition continue ("du deuxième ordre"), l'aimantation M_{sp} s'annulant à $T = T_c$. L'apparition d'une direction privilégiée dans laquelle pointe l'aimantation et la non-analyticité de l'aimantation en $T = T_c$ (non nulle en deçà, identiquement nulle au delà), sont des phénomènes très remarquables qui exigent une explication de la part du physicien... Une analyse plus poussée des propriétés du système révèle que d'autres grandeurs physiques ont également un comportement non analytique, voir plus bas. On parle de *phénomène critique* pour décrire le comportement du système au voisinage d'une telle transition.

Le phénomène de ferromagnétisme peut aussi se décrire comme la brisure spontanée d'une symétrie d'orientation (isotropie) : le matériau, un système de moments magnétiques, ne présente pas a priori de direction privilégiée en champ nul. Néanmoins à basse température et en champ nul, l'aimantation spontanée semble choisir une telle direction. Voir plus bas (§ 3.1.4) une discussion de l'apparente contradiction avec le principe de Curie.

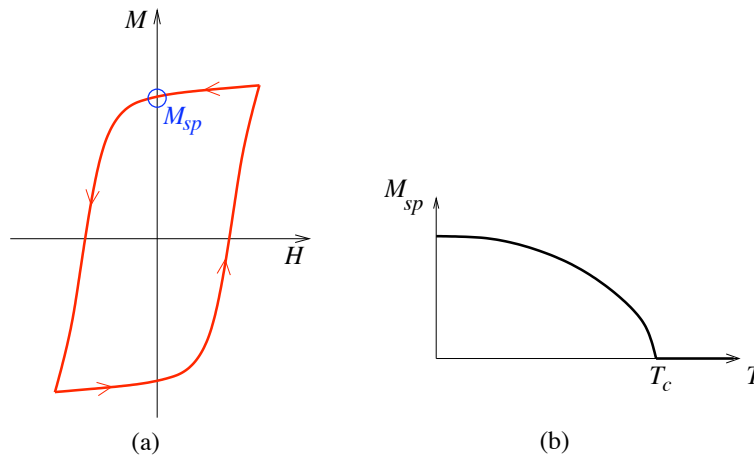


FIGURE 3.2 – (a) Cycle d'hystérésis d'un système ferromagnétique; (b) Aimantation "spontanée" en champ nul

3.1.3 Modèle d'Ising

Il s'agit d'un modèle très simplifié de ferromagnétisme. Les moments magnétiques, ou par abus de langage les "spins", sont classiques et ne prennent que deux valeurs ± 1 . Le modèle est défini sur un réseau à d dimensions. Le cas $d = 1$ constitue un exercice classique de mécanique

statistique mais ne nous intéressera pas ici, ne présentant pas de phénomène de brisure spontanée de symétrie. Examinons plutôt le cas $d = 2$: on considère un réseau carré en chaque site i duquel on attache un spin $\sigma_i = \pm 1$. L'énergie d'une configuration de spins sur le réseau est donnée par

$$E(\{\sigma\}) = -J \sum_{(i,j)} \sigma_i \cdot \sigma_j \quad (3.3)$$

où on somme sur toutes les paires (i, j) de sites voisins sur le réseau. Le paramètre J est appelé énergie d'échange, car il mesure (à un facteur 2 près) la différence entre les énergies de paires de spins alignés et antialignés. Le poids de Boltzmann attaché à une telle configuration est donc

$$w(\{\sigma\}) = \exp -\beta E(\{\sigma\}) ,$$

$\beta = \frac{1}{kT}$, et on est supposé calculer la fonction de partition $Z = \sum_{\{\sigma\}} w(\{\sigma\})$ et les valeurs moyennes des grandeurs thermodynamiques $A(\{\sigma\})$ fonctions des spins σ , soit $\langle A \rangle = Z^{-1} \sum_{\{\sigma\}} A(\{\sigma\}) w(\{\sigma\})$, etc. Par exemple, l'aimantation moyenne doit se calculer par $M = \lim_{N \rightarrow \infty} \langle \frac{1}{N} \sum_{i=1}^N \sigma_i \rangle$ où on se restreint d'abord à un nombre N fini de spins puis on prend la limite $N \rightarrow \infty$ ²; ou encore, en invoquant un argument d'homogénéité, $M = \langle \sigma \rangle$ pour un spin unique du réseau infini. Il convient de bien comprendre que le phénomène de transition de phase avec la non-analyticité qui lui est attachée ne peut se produire qu'à la limite $N \rightarrow \infty$. Pour tout N grand mais fini, la fonction de partition et les valeurs moyennes sont des sommes de fonctions analytiques en β , donc ne manifestent pas de non-analyticité.

Si $J > 0$, le système est ferromagnétique : à basse température $\beta \gg 1$, le poids de Boltzmann favorise les paires de spins voisins alignés, c'est-à-dire ici dans le modèle d'Ising, valant tous deux $+1$ ou -1 . Mais cela ne suffit pas pour conclure que le modèle a une aimantation spontanée non nulle : il faut en effet examiner le poids (entropie – énergie) des autres configurations, par exemple des îlots de spins retournés.

À l'évidence, le modèle présente une symétrie $\mathbb{Z}_2$: le retournement de *tous* les spins $\sigma_i \mapsto -\sigma_i$ ne change pas l'énergie $E(\{\sigma\})$ (3.3) mais change $M \mapsto -M$. On pourrait donc s'attendre à ce que l'aimantation s'annule identiquement (à toute température). Mais cette conclusion est elle aussi incorrecte : des calculs *a priori* de comparaison énergie-entropie de configurations (Peierls)³ d'une part, un calcul explicite rendu possible par le caractère exactement soluble de ce modèle de l'autre, permettent d'affirmer que le modèle présente le phénomène de brisure spontanée de symétrie en dimension $d \geq 2$: à basse température, $T < T_c$ (ou $\beta > \beta_c$), la symétrie est brisée, l'aimantation $M \neq 0$ et le système présente une *phase ordonnée*, où les spins pointent majoritairement dans la direction de M . (En revanche à $d = 1$, les fluctuations thermiques annulent l'aimantation à toute température $T > 0$ et le système est toujours dans sa phase symétrique désordonnée, voir TD.) L'aimantation est appelée "paramètre d'ordre". En fait, deux phases d'aimantation de signes opposés sont possibles à $T < T_c$ et ce sont les conditions initiales (par exemple un champ magnétique infinitésimal H) qui déterminent quelle

2. À N fini, il faudrait préciser les conditions aux bords, puis s'assurer que quand $N \rightarrow \infty$, leur contribution est négligeable. . .

3. Rudolf Peierls, *Proc. Cambridge Phil. Soc.* **32** 477 (1936) ; Robert B. Griffiths, *Phys. Rev.* **136**, A437–A439 (1964)

phase le système choisit. Cela est apparent sur la figure 3.3 où dans le plan (T, H) , les deux phases sont séparées par une ligne de transition à $H = 0$. À haute température, $T \geq T_c$, à l'inverse, l'aimantation est nulle, la symétrie est restaurée, et le système est dans une phase désordonnée. À $T = T_c$ le modèle subit une transition de phase du deuxième ordre.

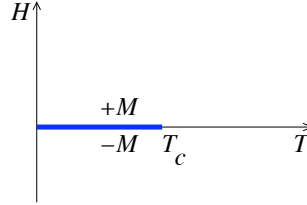


FIGURE 3.3 – Les deux phases de basse température du modèle d’Ising

Le calcul exact de l’aimantation spontanée est possible dans le modèle d’Ising bidimensionnel [Onsager et Yang⁴]. On trouve

$$M = \begin{cases} \left(1 - (\sinh(2J\beta))^{-\frac{1}{4}}\right)^{\frac{1}{8}} & \text{si } T \leq T_c \\ 0 & \text{si } T \geq T_c \end{cases} \quad (3.4)$$

qui s’annule à la température critique définie par $\sinh(2J\beta_c) = 1$ et qui y présente une non-analyticité, $M \sim (T_c - T)^{\frac{1}{8}}$.

Plus généralement, l’étude du comportement singulier des différentes grandeurs thermodynamiques dans des systèmes critiques quelconques, énergie libre, aimantation, etc, a été l’objet de beaucoup de travail et de progrès au cours des 40 dernières années⁵. On cherche en particulier à déterminer les exposants critiques qui décrivent l’annulation ou le caractère singulier des quantités thermodynamiques au voisinage de $T = T_c$. Ainsi la chaleur spécifique, l’aimantation ou la susceptibilité magnétique $\chi = \partial M / \partial H|_{H=0}$ ont des comportements critiques obéissant à des lois de puissance

$$c \sim |T - T_c|^{-\alpha} \quad M \sim (T_c - T)^\beta, \quad T \leq T_c \quad \chi \sim |T - T_c|^{-\gamma}, \quad \text{etc} \quad (3.5)$$

avec des *exposants critiques* $\alpha, \beta, \gamma, \dots$ (notations traditionnelles) qu’il s’agit de déterminer⁶. Pour le modèle d’Ising à deux dimensions, $\alpha = 0$ (singularité logarithmique), $\beta = \frac{1}{8}$, $\gamma = \frac{7}{4}$.

De façon remarquable, la clé du problème se trouve dans l’existence d’une nouvelle symétrie qui apparaît au point critique, une *invariance par dilatation* des échelles. Cette invariance provient du fait qu’à $T = T_c$, une longueur fondamentale du système, sa *longueur de corrélation* ξ , qui mesure la distance sur laquelle les spins s’influencent, diverge : $\xi \sim |T - T_c|^{-\nu}$ avec $\nu > 0$. L’outil théorique pour traiter cette invariance de dilatation s’appelle le groupe de renormalisation, il est maintenant d’usage très général en physique théorique, depuis l’étude du comportement critique en physique statistique ou physique de l’état condensé, à l’étude du comportement asymptotique à très haute énergie en physique des particules, ou à celle du comportement turbulent de flots hydrodynamiques.

4. Lars Onsager, (1903–1976), physico-chimiste norvégien, prix Nobel de chimie 1968, auteur de travaux remarquables en chimie et en physique, en particulier la première solution exacte du modèle d’Ising bidimensionnel.

Chen Ning Yang, (1922–), physicien sino-américain, prix Nobel de physique 1957 avec T.D. Lee pour leur découverte de la violation de la parité, voir chap. 6 ; auteur de travaux fondamentaux en physique statistique, physique des particules et théorie des champs.

5. tout particulièrement par Kenneth Wilson, (1936–2013), prix Nobel 1982, auteur de travaux fondamentaux sur les transitions de phase et le *groupe de renormalisation*, voir plus bas

6. Attention de ne pas confondre cet exposant β avec $\beta = 1/kT$

3.1.4 Un paradoxe ?

Ce phénomène de brisure spontanée de symétrie semble à première vue contredire le principe de Curie du paragraphe précédent. La symétrie des causes (interaction entre les spins d’Ising symétrique par l’action du groupe $\mathbb{Z}_2$) ne se retrouve pas dans celle des effets (aimantation non nulle brisant cette symétrie $\mathbb{Z}_2$). Ou à l’inverse, l’apparition d’une aimantation spontanée implique une brisure de l’isotropie de l’espace qui n’est pas présente dans les causes (la physique des moments magnétiques microscopiques). Il faut y regarder de plus près, et bien comprendre le rôle crucial des conditions initiales dans le déclenchement de la brisure spontanée. Or puisque ces conditions initiales brisent la symétrie –explicitement même si ce n’est qu’infinitésimalement–, nous sortons des conditions d’application du principe de Curie. L’orientation de l’aimantation spontanée dépend en fait de l’histoire passée du matériau, typiquement du champ magnétique auquel il a été soumis. Cela est l’analogue des conditions initiales différant de façon infinitésimale de 0, ou du terme de brisure γx infinitésimal, dans l’exemple mécanique examiné au paragraphe 3.1.1.

3.2 Brisure spontanée de symétrie continue

Considérons maintenant le cas d’une symétrie continue, tout d’abord avec le problème mécanique du § précédent, cf fig. 3.1(b), où le système vit maintenant à deux dimensions, avec un potentiel invariant par les rotations du groupe $O(2)$. Le même mécanisme est à l’œuvre, avec la même conséquence : dans les circonstances illustrées sur la figure, le “vrai” état stable de la théorie (en mécanique quantique, nous dirions le vrai “fondamental”) n’est pas à $x = 0$, mais quelque part dans le “cercle de gorge” du minimum du potentiel, et c’est le choix d’une condition initiale, infinitésimale ou non mais non symétrique, qui va le déterminer.

On conçoit aussi au vu de la figure qu’il soit très aisé –cela ne coûtera pas d’énergie potentielle– de passer d’un tel état fondamental à un état voisin, lui aussi au fond de la gorge. Quand le même phénomène se produit dans un contexte de physique statistique ou quantique, cela se manifeste par l’existence d’excitations d’énergie nulle dans la théorie. Dans un contexte de physique des particules, ces excitations correspondent à des particules de masse nulle. On a donné le nom de *bosons de Nambu–Goldstone* aux excitations ou particules correspondantes. Ce résultat présenté ici de façon très heuristique peut être justifié par des arguments précis et cela nous permet donc d’affirmer : la brisure spontanée d’une symétrie continue s’accompagne de l’apparition de bosons de Nambu–Goldstone ayant un spectre d’énergie allant jusqu’à 0. Une telle excitation sera qualifiée de “sans gap” dans le langage de la physique de la matière condensée, ou de masse nulle dans celui de la physique des particules.

Donnons un exemple, tiré une fois encore de la physique du ferromagnétisme. Plutôt que le modèle d’Ising précédent, intéressons-nous au modèle dit de Heisenberg classique, un autre modèle de mécanique statistique classique. Pour rester simples, considérons-le à deux dimensions : il est défini sur un réseau carré, et les variables dynamiques sont des vecteurs unitaires à n dimensions $\vec{S}_i \in \mathbb{R}^n$, $|\vec{S}_i| = 1$, attachés à chaque site i du réseau. L’énergie d’une configuration

de tels “spins classiques” est

$$E = -J \sum_{(i,j)} \vec{S}_i \cdot \vec{S}_j \quad (3.6)$$

avec les mêmes notations que dans (3.3). À nouveau, le cas ferromagnétique correspond à $J > 0$, et on s’intéresse à calculer fonction de partition et énergie libre, aimantation, etc. Par exemple, l’aimantation doit se calculer par $\vec{M} = \langle \vec{S} \rangle$ dans le système infini.

Ce modèle présente une symétrie de rotation $O(n)$: l’expression (3.6) de l’énergie est évidemment invariante par l’action d’un même élément $O \in O(n)$ sur *tous* les $\vec{S}_i$, $\vec{S}_i \mapsto O\vec{S}_i$. Noter qu’il s’agit là d’une *symétrie interne*, agissant dans l’espace des vecteurs spins, distinct a priori de l’espace physique (ici à $d = 2$ dimensions).

Noter encore que le cas $n = 1$, donc de vecteurs unitaires à une dimension, $S_i = \pm 1$, se réduit au modèle d’Ising. Mais alors que le modèle d’Ising présente une symétrie discrète par le groupe $\mathbb{Z}_2$, la symétrie du modèle de Heisenberg, pour $n \geq 2$, est continue.

Cette différence est importante en ce qui concerne l’existence d’une transition de phase avec symétrie brisée spontanément. On démontre qu’au contraire du modèle d’Ising, le modèle de Heisenberg (pour $n \geq 2$) ne présente pas de transition en dimension $d = 2$ (cf TD). En revanche le même modèle à $d = 3$ dimensions, (défini sur un réseau cubique, disons), présente une transition, avec brisure spontanée de la symétrie continue $O(n)$ et apparition d’un paramètre d’ordre, l’aimantation, dans la phase de basse température. Et selon l’argument heuristique donné ci-dessus, il existe dans cette phase ordonnée de basse température des excitations de basse énergie, appelées dans ce contexte “ondes de spins” ou “magnons”.

3.3 Autres exemples de symétries brisées spontanément

Les applications du mécanisme de brisure spontanée ne se limitent pas aux systèmes ferromagnétiques. En fait, dans de très nombreux phénomènes, on y fait appel :

- Landau a suggéré que la cristallisation pouvait s’interpréter comme brisure spontanée des symétries de translation et de rotation.
- Dans les transitions de phase des cristaux liquides (cf supra, § 1.5.4), Pierre-Gilles de Gennes a montré que la transition de phase liquide isotrope-nématique est associée à un paramètre d’ordre : l’orientation moyenne des molécules, voir ci-dessous.
- La théorie moderne de physique des particules appelée “Modèle Standard” est fondée sur deux idées-clés : une symétrie de jauge (voir chapitre suivant), et la brisure spontanée de cette symétrie. Il en résulte non pas l’apparition de particules de masse nulle comme expliqué plus haut, mais au contraire dans ce cas, l’acquisition d’une masse par des particules *a priori* de masse nulle ! C’est le mécanisme de Brout–Englert–Higgs⁷, dont l’accélérateur LHC du CERN vient de confirmer tout récemment la validité dans ce contexte.
- Un mécanisme du même genre est à l’œuvre dans la transition conducteur-superconducteur ou dans la transition superfluide-fluide normal de l’hélium : dans la phase de basse température, un condensat (d’électrons dans un cas, d’atomes d’He⁴ dans l’autre) se forme, et la “pseudo-fonction d’onde” ψ de ce condensat, qui satisfait une généralisation non-linéaire de l’équation de Schrödinger, sert de paramètre d’ordre, avec $\langle \psi \rangle \neq 0$ signalant la brisure de l’invariance $\psi \rightarrow e^{i\alpha}\psi$.
- Le mécanisme de brisure spontanée est aussi évoqué en cosmologie, pour expliquer la phase inflationnaire de l’Univers et l’apparition de défauts topologiques entre les différents “vides” de la théorie...

7. Robert Brout (1928–2011). François Englert (1932–) et Peter Higgs, (1929–), prix Nobel 2013

Approximation de Ginzburg–Landau. Cristaux liquides nématiques

L'approximation de Ginzburg–Landau est une méthode très puissante et universelle. Elle consiste à identifier une grandeur macroscopique $\Phi(\vec{x})$, le paramètre d'ordre, variant lentement dans l'espace et décrivant les degrés de liberté les plus importants d'un système proche d'une transition de phase. Dans un ferromagnétique, ce sera l'aimantation locale, dans un cristal liquide nématique, une orientation etc. On écrit alors une énergie libre "effective", qui est une fonctionnelle de ce champ Φ , qu'on limite à un nombre fini de termes de degré bas en Φ et ses dérivées $\nabla\Phi$. Par exemple pour un ferromagnétique de type Ising

$$\mathcal{F}[\Phi] = \int d^d x \left(\frac{1}{2} (\vec{\nabla}\Phi)^2 + \frac{r}{2} \Phi^2 + g\Phi^4 \right)$$

où on a supposé que pour des raisons de symétrie par $\Phi \leftrightarrow -\Phi$, seuls des termes de degré pair apparaissent. On discute alors les extrema de cette énergie libre effective, supposés "uniformes" (Φ indépendant de $\vec{x}$), et la discussion est très semblable à ce qu'on a fait au § 3.1 ou dans l'approximation du champ moyen (cf § suivant), avec des situations très différentes selon le signe du paramètre r qui joue le rôle d'écart à la température critique. Comme on l'a mentionné plus haut, le même genre d'approximation peut s'appliquer à d'autres systèmes, tels qu'un supraconducteur ou un superfluide, ou la cristallisation (Landau 1937), ou encore la transition de phase des cristaux liquides nématiques. Rappelons qu'une phase nématique est faite de molécules essentiellement parallèles, manifestant donc un ordre orientationnel, cf § 1.5.4. On pourrait penser décrire l'orientation des molécules par un vecteur unitaire $\hat{n}$, mais l'identification des directions de $\hat{n}$ et $-\hat{n}$ conduit à utiliser un tenseur de rang 2 : la modélisation par de Gennes⁸ consiste à introduire un paramètre d'ordre Q qui est un tenseur 3×3 symétrique

et de trace nulle. En le diagonalisant dans une base orthogonale, $Q = \begin{pmatrix} -\frac{1}{2}(q + \eta) & 0 & 0 \\ 0 & -\frac{1}{2}(q - \eta) & 0 \\ 0 & 0 & q \end{pmatrix}$. Le cas

$\eta = 0$ correspond au cas "uniaxe" où on a symétrie de rotation autour d'un axe. On écrit alors une fonctionnelle de Ginzburg–Landau de la forme

$$\mathcal{F}[Q] = \int d^3 x \left(\frac{1}{2} \text{tr} (\vec{\nabla}Q)^2 + \frac{a}{2} (T - T_*) \text{tr} Q^2 + \frac{b}{3} \text{tr} Q^3 + \frac{c}{4} (\text{tr} Q^2)^2 \right).$$

On cherche les configurations uniformes minimisant $\mathcal{F}[Q]$, donc minimisant le potentiel $\frac{a}{2}(T - T_*) \text{tr} Q^2 + \frac{b}{3} \text{tr} Q^3 + \frac{c}{4} (\text{tr} Q^2)^2$. Dans le cas uniaxe, ($\eta = 0$), ce potentiel s'écrit

$$V(q) = \frac{3a}{4} (T - T_*) q^2 + \frac{b}{4} q^3 + \frac{9c}{32} q^4$$

et la figure 3.4 montre la forme typique de cette fonction pour différentes valeurs de T . Si $b < 0$, il existe une valeur T_c telle que pour $T > T_c$, le minimum de V est en $q = 0$, tandis que pour $T = T_c$, un nouveau minimum apparaît pour $q \neq 0$. Il y a donc en $T = T_c$ une *transition du premier ordre*, avec discontinuité du paramètre d'ordre q , entre la phase ordonnée nématique et la phase désordonnée isotrope.

Question : pourquoi n'écrit-on pas aussi un terme $\text{tr} Q^4$ dans $\mathcal{F}[Q]$? Indication : se rappeler que toute matrice satisfait sa propre équation caractéristique et montrer que pour une matrice 3×3 de trace nulle, un tel terme n'est pas indépendant des précédents.

Pour en savoir plus, voir P. G. de Gennes et J. Prost, *The Physics of Liquid Crystals*, Oxford U.P.

8. Pierre Gilles de Gennes, (1932-2007), a reçu le prix Nobel de physique 1991 pour ses travaux sur les cristaux liquides et les polymères.

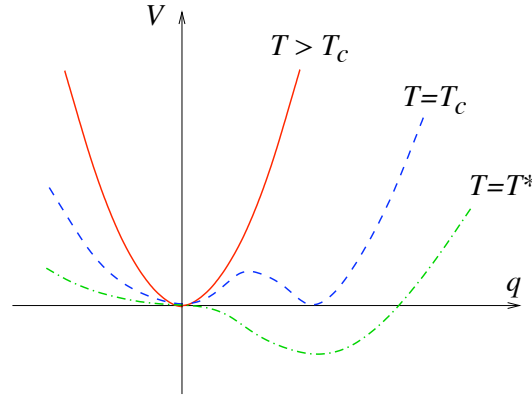


FIGURE 3.4 – Potentiel de Landau-Ginzburg-de Gennes pour des cristaux liquides nématiques

3.4 Exercice : Champ moyen

L'approximation du champ moyen peut s'obtenir de multiples façons. Illustrons la sur le cas du modèle d'Ising sur un réseau hypercubique en dimension d fait de N sites (avec $N \rightarrow \infty$)

$$Z_{\text{Ising}} = \sum_{\{\sigma=\pm 1\}} \exp(\beta J \sum_{(i,j)} \sigma_i \sigma_j - \beta h \sum_i \sigma_i)$$

où on suppose les spins couplés aussi à un champ magnétique extérieur h .

1. Rappeler l'approximation *auto-cohérente* du champ moyen (cf. cours de Mécanique Statistique). On écrit

$$Z_{\text{Ising}} = \sum_{\{\sigma=\pm 1\}} \exp(\beta J \underbrace{\sum_{(i,j)} \sigma_i \sigma_j}_{S(\{\sigma\})} - \beta h \sum_i \sigma_i)$$

L'approximation consiste à remplacer les interactions entre spins par leur interaction avec un champ effectif local H_i , $S(\{\sigma\}) \approx \beta \sum_i \sigma_i H_i$, tel que $\beta H_i := \frac{\partial S}{\partial \sigma_i} = \beta (J \sum_{j:i} \sigma_j - h) \approx 2d\beta J M - \beta h$ et $M := \langle \sigma_i \rangle = \tanh H$, en approximant H_i et $\langle \sigma_i \rangle$ par des valeurs uniformes H et M . En déduire l'équation du champ moyen

$$2\beta J d \tanh \beta H = \beta (H + h). \quad (3.7)$$

En champ extérieur nul, $h = 0$, on a à résoudre l'équation

$$\tanh \beta H = \frac{H}{2Jd}.$$

Discuter ses solutions en fonction de la valeur de $2\beta J d$ par rapport à 1. Comparer avec l'équation $V'(x_c) = 0$ du § 3.1.1. Conclure sur l'apparition ou non d'une aimantation non nulle, c'est-à-dire d'une éventuelle brisure spontanée de symétrie.

2. On va montrer que l'approximation du champ moyen fournit une borne inférieure à $\log Z$.

i) Rappeler pourquoi la propriété de convexité de la fonction exponentielle se traduit par l'inégalité

$$\langle e^X \rangle \geq e^{\langle X \rangle}, \quad (3.8)$$

pour toute moyenne $\langle \cdot \rangle$ à poids positifs.

ii) Montrer qu'on peut récrire Z_{Ising} comme

$$Z_{\text{Ising}} = \sum_{\{\sigma=\pm 1\}} e^{\beta H \sum_i \sigma_i} \exp(\beta J \sum_{(i,j)} \sigma_i \sigma_j - \beta(H+h) \sum_i \sigma_i) \quad (3.9)$$

$$= \langle \exp(\beta J \sum_{(i,j)} \sigma_i \sigma_j - \beta(H+h) \sum_i \sigma_i) \rangle_H Z_H \quad (3.10)$$

où H est un champ arbitraire, $Z_H = \sum_{\{\sigma=\pm 1\}} \prod_i e^{\beta H \sum_i \sigma_i} = (2 \cosh \beta H)^N$ et $\langle X \rangle_H$ désigne la moyenne sur les $\sigma_i = \pm 1$ avec le poids factorisé $\prod_i e^{\beta H \sum_i \sigma_i}$: $\langle X \rangle_H = \frac{1}{Z_H} \sum_{\{\sigma=\pm 1\}} \prod_i e^{\beta H \sum_i \sigma_i} X$.

(iii) Montrer que $M := \langle \sigma \rangle_H = \tanh \beta H$.

(iv) Appliquer alors (3.8) pour obtenir

$$Z_{\text{Ising}} \geq \exp(\langle \beta J \sum_{(i,j)} \sigma_i \sigma_j - \beta(H+h) \sum_i \sigma_i \rangle_H) Z_H = \exp\{N(\beta J d M^2 - \beta(H+h)M)\} (2 \cosh \beta H)^N$$

soit $\frac{1}{N} \log Z_{\text{Ising}} \geq (\beta J d M^2 - \beta(H+h)M) + \log(2 \cosh \beta H)$, et cela pour tout H .

(v) Montrer que le maximum du membre de droite de cette inégalité est atteint pour H satisfaisant (3.7), l'équation familière du champ moyen obtenue plus haut. Le champ moyen fournit donc une borne inférieure à l'énergie libre (au signe près, selon les conventions) $\log Z_{\text{Ising}}$.

3. Voici une troisième méthode qui permet le calcul systématique de corrections au champ moyen. On remplace d'abord la somme sur les variables discrètes $\sigma = \pm 1$ par une intégrale continue, qu'on traitera par la méthode du col, à justifier a posteriori.

i) Montrer qu'on peut écrire pour toute fonction $f(\cdot)$ d'un spin

$$\sum_{\sigma=\pm 1} f(\sigma) = \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} \frac{d\alpha}{2\pi i} \sum_{\sigma=\pm 1} e^{(\sigma-u)\alpha} f(u) = \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} \frac{d\alpha}{2\pi i} e^{\omega(\alpha)-u\alpha} f(u)$$

où $\omega(\alpha) := \log \sum_{\sigma=\pm 1} e^{\sigma\alpha} = \log(2 \cosh \alpha)$.

Appliquer cela à la fonction de partition du modèle d'Ising sur un réseau hypercubique à d dimensions (en champ h nul)

$$Z_{\text{Ising}} = \prod_i \int_{-\infty}^{\infty} du_i \int_{-\infty}^{\infty} \frac{d\alpha_i}{2\pi i} e^{\sum_i \omega(\alpha_i) - u_i \alpha_i + \beta J \sum_{(i,j)} u_i u_j}.$$

(ii) Écrire l'équation de stationnarité de l'argument de l'exponentielle (méthode du col), soit

$$u_i = \tanh \alpha_i, \quad \beta J \sum_{j:i} u_j = \alpha_i$$

dont on cherche des solutions uniformes, indépendantes du site i . Montrer qu'on retrouve bien les équations de la méthode précédente. La méthode du col utilisée est justifiée pour $d \rightarrow \infty$, (et α d'ordre d), puisqu'alors les trois termes de l'exponentielle sont grands et d'ordre d . L'avantage

de cette approche est qu'on a maintenant une manière systématique de calculer des corrections, en développant au voisinage du col et en intégrant en présence du terme gaussien. Les corrections successives forment un développement en puissances inverses de d .

Retenons que la méthode du champ moyen est justifiée pour d grand. En fait, on démontre dans l'approche du groupe de renormalisation que les exposants critiques calculés dans l'approximation du champ moyen sont exacts pour $d > 4$ (pour le modèle d'Ising). La dimension 4 est appelée *dimension critique supérieure*.

Chapitre 4

Groupes continus. Générateurs infinitésimaux. Lois de conservation

On va s'intéresser maintenant plus particulièrement à des groupes de transformations dotés de propriétés de continuité et de différentiabilité dans les paramètres de la transformation. On va en particulier étudier des transformations infinitésimalement proches de l'identité. Cela va conduire à la notion importante de générateur infinitésimal.

D'un point de vue mathématique, ces générateurs satisfont des relations algébriques (relations de commutation), essentiellement caractéristiques du groupe considéré, et forment une *algèbre de Lie*. Pour le physicien, ces relations sont très familières dans le cadre de la mécanique quantique, cf les générateurs des rotations, égaux à un facteur $\hbar$ près aux composantes du moment angulaire, et leurs relations de commutation.

Par ailleurs, chaque symétrie dynamique par l'action d'un groupe continu se traduit par l'existence de *relations de conservation*. Là encore, il s'agit d'un résultat généralisant des choses bien connues : conservation de l'impulsion, respectivement du moment angulaire, associée à l'invariance par translation d'espace, resp. par rotation.

4.1 Transformations continues, générateurs infinitésimaux

4.1.1 Groupes des translations $\mathbb{R}$, $\mathbb{R}^d$

Considérons le plus simple des groupes continus, celui des translations à une dimension : $t(a) : x \mapsto x' = x + a$, avec x et a des variables réelles $x, a \in \mathbb{R}$. Pour toute fonction réelle $f(x)$, définissons la nouvelle fonction $f' = t(a)f$ par $f'(x') = f(x)$ ou de façon équivalente, $f'(x) = f(x - a)$. La loi de composition de ces transformations (loi de groupe) est

$$t(b) \circ t(a) = t(a + b) . \quad (4.1)$$

La loi est commutative, le groupe est abélien. Ce groupe est en fait $\mathbb{R}$ avec la loi de groupe fournie par l'addition des réels.

Si a est infinitésimal, on peut écrire $f'(x) \approx f(x) - a \frac{d}{dx} f(x)$ ou encore

$$f'(x) = \left(1 - a \frac{d}{dx}\right) f(x)$$

soit encore, en définissant la variation de la fonction

$$\Delta f(x) = f'(x) - f(x) = -a \left(\frac{d}{dx}\right) f(x) .$$

La variation infinitésimale de la fonction est linéaire dans le paramètre a (on travaille au 1er ordre!) et dans la fonction f , et donnée par l'action du *générateur infinitésimal* $T_a = -aT = -a \left(\frac{d}{dx}\right)$.

Noter que la loi de groupe (4.1) et sa propriété de commutativité se traduisent au niveau infinitésimal par la commutativité des générateurs

$$[T_a, T_b] = T_a T_b - T_b T_a = 0 .$$

Les transformations continues (et différentiables) du groupe impliquent donc l'existence des générateurs infinitésimaux avec des propriétés qui reflètent la loi de groupe. Inversement une fois connu le générateur infinitésimal $T = -\frac{d}{dx}$, il est possible de reconstruire la transformation finie par *action exponentielle*

$$\begin{aligned} \exp(-aT)f(x) &= (1 - aT + \frac{1}{2}a^2T^2 + \frac{-a^3}{3!}T^3 + \dots)f(x) \\ &= \left(1 - a \left(\frac{d}{dx}\right) + \frac{a^2}{2} \left(\frac{d}{dx}\right)^2 + \frac{(-a)^3}{3!} \left(\frac{d}{dx}\right)^3 + \dots\right) f(x) \\ &= f(x - a) \end{aligned}$$

qui n'est autre que la série de Taylor, que nous supposons convergente (f analytique réelle).

Ces considérations s'étendent sans difficulté à des translations $t(\vec{a})$ dans l'espace $\mathbb{R}^d$ et à des fonctions f des d variables (coordonnées) $x_1, \dots, x_d$. La loi de groupe est toujours additive et donc commutative $t(\vec{a}) \circ t(\vec{b}) = t(\vec{a} + \vec{b}) = t(\vec{b}) \circ t(\vec{a})$, le groupe n'est autre que $\mathbb{R}^d$ avec son addition. L'opérateur infinitésimal est maintenant l'opérateur différentiel $T_{\vec{a}} = -\vec{a} \cdot \vec{T}$ avec $\vec{T}$ le vecteur gradient $\vec{T} = \vec{\nabla} = \left(\frac{d}{dx_1}, \dots, \frac{d}{dx_d}\right)$, et la formule de Taylor à d variables s'applique.

4.1.2 Groupe des phases U(1)

Considérons le groupe multiplicatif des nombres complexes de module 1, $\zeta_\alpha = e^{i\alpha}$. Le nombre réel α est défini modulo 2π . La loi de groupe est

$$e^{i\alpha} e^{i\beta} = e^{i(\alpha+\beta)} . \quad (4.2)$$

Il s'agit donc encore d'un groupe abélien. Sa dénomination mathématique est U(1) (groupe unitaire à 1 dimension). On s'intéresse maintenant à l'action de ce groupe sur les nombres complexes non nuls

$$z \mapsto z' = e^{i\alpha} z$$

et sur les puissances entières (positives ou négatives) de z

$$z^p \mapsto z'^p = e^{ip\alpha} z^p .$$

Ces puissances entières z^p sont invariantes par l'action de $\zeta_{2\pi} = e^{i(2\pi)}$. (Noter qu'en revanche, pour une puissance r non entière de z , l'action de $e^{i(2\pi)}$ n'est pas l'identité.) Inversement, les théorèmes sur les séries de Fourier nous disent que toute fonction (suffisamment régulière) f de z telle que $f(z) = f(e^{2i\pi}z)$ peut se développer sur les puissances z^p , p entier positif ou négatif, $p \in \mathbb{Z}$. Pour étudier l'action du groupe sur ces fonctions f , il suffit donc de l'étudier sur les z^p .

À nouveau, on peut considérer des transformations infinitésimales, avec $|\alpha| \ll 1$,

$$z \mapsto z' \approx (1 + i\alpha)z \quad z^p \mapsto z'^p = (1 + ip\alpha)z^p = (1 + T_\alpha)z^p$$

Le générateur infinitésimal s'écrit $T_\alpha = -i\alpha z \frac{d}{dz}$. On a encore commutation, $[T_\alpha, T_\beta] = T_\alpha T_\beta - T_\beta T_\alpha = 0$, reflétant la commutativité de la loi de groupe, et l'exponentielle $\exp T_\alpha$ agissant sur $f(z)$ réalise une transformation finie $\exp T_\alpha f(z) = f(e^{-i\alpha}z)$. (Effectuer le changement de variable $z \mapsto iw = \log z$ et constater que T_α est le générateur infinitésimal de la translation de w , comme ci-dessus au 4.1.1.)

4.1.3 Groupe de rotation à deux dimensions $SO(2)$

Considérons maintenant le groupe des rotations à deux dimensions, noté $SO(2)$. L'élément générique $R(\alpha)$ dépend d'un paramètre continu, l'angle α de rotation défini modulo 2π . La rotation $R(\alpha)$ agit sur les vecteurs $\vec{x}$ du plan par

$$\vec{x} \mapsto \vec{x}' = R(\alpha)\vec{x}$$

où la notation doit être comprise comme l'action d'un opérateur linéaire sur le vecteur $\vec{x}$, ou encore, dans un repère orthogonal, comme l'action d'une matrice orthogonale sur les composantes x_1, x_2 de $\vec{x}$

$$\vec{x}' = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (4.3)$$

(la matrice est orthogonale par définition de $SO(2)$, ou encore, géométriquement, parce que les transformations considérées préservent la norme $\|\vec{x}'\| = \|\vec{x}\|$, or $\vec{x}' \cdot \vec{x}' = x'^T (R^T(\alpha) \cdot R(\alpha)) x = \vec{x} \cdot \vec{x}$ donc $R^T(\alpha) \cdot R(\alpha) = \mathbb{1}$, cf Chap. 1 et 2). Ces rotations se composent comme on sait bien

$$R(\alpha)R(\beta) = R(\alpha + \beta) = R(\beta)R(\alpha) \quad (4.4)$$

de façon à nouveau additive et commutative. En fait la loi est la même que celle (4.2) du groupe $U(1)$ du § précédent. En effet il s'agit de deux descriptions équivalentes de la même géométrie du plan, la première à l'aide d'une variable complexe z , la seconde à l'aide de vecteurs à deux composantes $\vec{x}$.

Dans ce dernier langage, les générateurs infinitésimaux sont des opérateurs, ou des matrices 2×2 . En effet, prenant α infinitésimal et développant au premier ordre, $\cos \alpha \approx 1$, $\sin \alpha \approx \alpha$, donc

$$\alpha \ll 1 \quad R(\alpha) = \mathbb{1} - i\alpha J = \mathbb{1} - i\alpha \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{avec} \quad J = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (4.5)$$

où l'apparition du i peut paraître curieuse dans ce cadre réel, mais est utile si on veut considérer le générateur J comme hermitien (plutôt qu'antisymétrique réel donc antihermitien).

Considérons maintenant la loi de groupe (4.4) pour $\beta = d\alpha$ infinitésimal, en faisant usage de (4.5) :

$$R(\alpha + d\alpha) = R(d\alpha)R(\alpha) = (\mathbb{1} - id\alpha J)R(\alpha) = R(\alpha) - id\alpha JR(\alpha)$$

donc $dR(\alpha) = -id\alpha JR(\alpha)$. On obtient donc l'équation différentielle

$$\frac{d}{d\alpha} R(\alpha) = -iJR(\alpha) \quad (4.6)$$

où J est, rappelons le, une matrice fixe, indépendante de α . Cette équation différentielle du premier ordre à coefficients constants et homogène, complétée par la condition initiale $R(0) = \mathbb{1}$, se résout immédiatement en

$$R(\alpha) = \exp -i\alpha J, \quad (4.7)$$

avec l'exponentielle d'une matrice définie par son développement en série (toujours convergent)

$$\exp A = \sum_{n=0}^{\infty} \frac{A^n}{n!}. \quad (4.8)$$

Exercice. Pour $J = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, vérifier que l'exponentiation de $-i\alpha J$ reproduit bien la matrice de $R(\alpha)$ donnée plus haut. (Indication : calculer les puissances successives $J^2, J^3, \dots, J^n$ et construire $\exp i\alpha J$ par son développement en série (4.8).)

De la même façon qu'on a dans le sous-paragraphe précédent appliqué les transformations du groupe $U(1)$ à des puissances quelconques (entières) de z , on peut aussi étudier l'application des transformations à des tenseurs à p indices, se transformant donc comme le tenseur de composantes $(x_{i_1}, \dots, x_{i_p})$. Cela fait apparaître de nouveaux générateurs infinitésimaux, qui sont des matrices $2^p \times 2^p$ agissant sur les composantes des tenseurs. Et à nouveau, l'action pour une transformation finie s'obtient par exponentiation de la transformation infinitésimale. Au chapitre suivant, on parlera de *représentation* du groupe $SO(2)$ pour cette action sur les tenseurs et on analysera les choses de façon plus systématique.

4.1.4 Groupe de rotation à trois dimensions $SO(3)$

Considérons enfin le groupe des rotations à trois dimensions, noté $SO(3)$, en nous bornant à un mot, avant l'étude plus approfondie du Chapitre 7.

Comme on le sait bien après le cours de Physique Quantique, les rotations infinitésimales du groupe $SO(3)$ font apparaître trois générateurs infinitésimaux indépendants, notés J_i , $i = 1, 2, 3$ et dotés de relations de commutation non triviales

$$[J_i, J_j] = i \sum_k \epsilon_{ijk} J_k . \quad (4.9)$$

L'exponentiation permettant de reconstruire la transformation finie à partir de ces générateurs infinitésimaux est plus délicate à mettre en œuvre, à cause de ces relations de commutation, mais elle peut encore se démontrer, voir chapitre 7.

4.1.5 Groupe de Lie, algèbre de Lie, et leur dimension

Groupes de Lie

Les groupes $\mathbb{R}^d$, $U(1)$, $SO(2)$ et $SO(3)$ que nous venons d'évoquer sont des exemples de groupes dont les éléments dépendent de façon différentiable de leurs paramètres (ici, des translations ou des angles de rotation). On appelle *groupes de Lie* de tels groupes. D'autres exemples fréquemment rencontrés par le physicien sont les *groupes linéaires* $GL(n, \mathbb{C})$ et $GL(n, \mathbb{R})$ des matrices $n \times n$ inversibles à coefficients complexes ou réels, et leurs sous-groupes respectifs $U(n)$, $SU(n)$ et $O(n)$, $SO(n)$.

Rappelons la définition des groupes orthogonaux $O(n)$, $SO(n)$ (cf § 1.3.1). Le groupe $O(n)$ est le groupe des matrices $n \times n$ orthogonales réelles, c'est-à-dire satisfaisant $O.O^T = I$ ou de façon équivalente $O^T.O = I$. Le groupe $SO(n)$ est le sous-groupe du groupe $O(n)$ constitué des matrices orthogonales de déterminant 1

$$SO(n) = \{O : O.O^T = I, \quad \det O = 1\} .$$

Le *groupe unitaire* $U(n)$ est le groupe de matrices $n \times n$ complexes et unitaires, $U.U^\dagger = I$ ou de façon équivalente $U^\dagger.U = I$. (On rappelle que $X^\dagger := (X^T)^*$.) Le groupe $SU(n)$ est le sous-groupe du groupe $U(n)$ constitué des matrices unitaires de déterminant 1

$$SU(n) = \{U : U.U^\dagger = I, \quad \det U = 1\} .$$

Transformations infinitésimales, algèbres de Lie

Les transformations d'un groupe de Lie étant différentiables, on peut considérer un élément infinitésimalement proche de l'identité et le développer au voisinage de l'identité. Par exemple, pour une matrice de $O(n)$ infinitésimalement proche de l'identité, écrivons $O = I + X$, avec X infinitésimal : la relation d'orthogonalité s'écrit alors

$$O.O^T = (I + X).(I + X^T) = I$$

donc au premier ordre en X ,

$$o(n) : \quad X + X^T = 0 , \quad (4.10)$$

qui exprime que X est une matrice réelle antisymétrique. On notera par $\mathfrak{o}(n)$ l'ensemble de ces matrices $n \times n$ antisymétriques. De la même façon, vérifier qu'une matrice unitaire infinitésimalement proche de l'identité est de la forme $U = I + X$ avec X complexe antihermitique, c'est-à-dire satisfaisant

$$\mathfrak{u}(n) : \quad X + X^\dagger = 0 . \quad (4.11)$$

On note $\mathfrak{u}(n)$ l'ensemble de ces matrices $n \times n$ antihermitiques. Passant alors de $U(n)$ à $SU(n)$ et à sa version infinitésimale, il faut imposer aussi que le déterminant de $U = I + X$ est 1. On sait que toute matrice $(I + X)$ peut être écrite (dans une base adéquate) sous forme triangulaire supérieure, avec des éléments diagonaux notés $1 + x_i$, et donc $\det(I + X) = \prod_i (1 + x_i)$. Pour X infinitésimale, les x_i sont aussi infinitésimaux, et on peut écrire

$$\begin{aligned} \det(I + X) &= \prod_i (1 + x_i) = 1 + \sum_i x_i + O(x^2) = 1 + \operatorname{tr} X + O(X^2) \\ &= 1 \quad \text{si et seulement si } \operatorname{tr} X = 0. \end{aligned} \quad (4.12)$$

Au premier ordre, la condition sur les matrices X est donc d'être antihermitiques et de trace nulle. Cela définit un ensemble de matrices que nous noterons $\mathfrak{su}(n)$.

$$\mathfrak{su}(n) : \quad X + X^\dagger = 0 \quad \text{et} \quad \operatorname{tr} X = 0. \quad (4.13)$$

Noter que, contrairement au cas unitaire, pour des matrices orthogonales, la condition de déterminant 1 n'apporte pas de nouvelle contrainte à (4.10). Cela est dû au fait que pour toute matrice orthogonale réelle, $O.O^T = I \implies (\det O)^2 = 1$, donc $\det O = \pm 1$. Pour une matrice infinitésimalement proche de l'identité, la condition $\det O = +1$ est automatiquement satisfaite par continuité. Autrement dit, les ensembles $\mathfrak{o}(n)$ et $\mathfrak{so}(n)$ sont identiques.

On note que chacun de ces ensembles $\mathfrak{o}(n)$, $\mathfrak{u}(n)$ ou $\mathfrak{su}(n)$, (4.10), (4.11) ou (4.13), forme un espace vectoriel réel : toute combinaison linéaire à coefficients réels de deux matrices satisfaisant une de ces conditions la satisfait aussi. Plus remarquable, le *commutateur* $[X, Y] = X.Y - Y.X$ de deux matrices X et Y d'un de ces ensembles est aussi dans l'ensemble. Par exemple pour $\mathfrak{o}(n)$,

$$X, Y \in \mathfrak{o}(n) \quad [X, Y]^T = (X.Y - Y.X)^T = Y^T.X^T - X^T.Y^T = Y.X - X.Y = -[X, Y] \Rightarrow [X, Y] \in \mathfrak{o}(n)$$

Vérifier que cela est bien satisfait aussi dans $\mathfrak{u}(n)$ ou $\mathfrak{su}(n)$. On appelle *algèbre de Lie* un espace vectoriel doté de cette propriété d'être stable par le crochet de commutation.

Cette structure est générale : pour tout groupe de Lie G , les éléments proches de l'identité forment une algèbre de Lie, notée $\mathfrak{g}$.

Dimension d'une algèbre de Lie

Il est aisé de compter le nombre de paramètres réels dont dépend une matrice satisfaisant (4.10), (4.11) ou (4.13). Par définition, ce nombre est appelé la *dimension* de l'algèbre de Lie ou du groupe de Lie correspondant.

Ainsi une matrice réelle $n \times n$ dépend de n^2 paramètres, mais les conditions (4.10) imposent $n + (n - 1) + (n - 2) + \dots + 1 = n(n + 1)/2$ conditions indépendantes $X_{ij} = -X_{ji}$, $i \leq j$. Il

demeure donc $n(n-1)/2$ éléments indépendants, qui est la dimension de l'algèbre de Lie $\mathfrak{o}(n)$ ($= \mathfrak{so}(n)$) ou des groupes de Lie $O(n)$ et $SO(n)$.

Exercice : vérifier que les dimensions des algèbres de Lie précédentes sont données par le tableau suivant

Groupe de Lie G	Algèbre de Lie $\mathfrak{g}$	Dimension d
$O(n)$	$\mathfrak{o}(n)$	$\frac{n(n-1)}{2}$
$SO(n)$	$\mathfrak{so}(n)$	$\frac{n(n-1)}{2}$
$U(n)$	$\mathfrak{u}(n)$	n^2
$SU(n)$	$\mathfrak{su}(n)$	$n^2 - 1$
$\mathbb{R}^d$	$\mathbb{R}^d$	d

Puisque les algèbres de Lie sont des espaces vectoriels, on peut y choisir une base. On appelle *système de générateurs infinitésimaux* du groupe de Lie G toute base de son algèbre de Lie. La dimension d'une algèbre de Lie (ou du groupe de Lie correspondant) est donc le nombre maximal de générateurs infinitésimaux indépendants. Par exemple, dans le cas de $SO(3)$, un système de générateurs infinitésimaux, c'est-à-dire une base de matrices antisymétriques 3×3 est donnée par

$$T_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad T_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad T_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (4.14)$$

Noter que ces formules peuvent s'exprimer de façon compacte sous la forme

$$(T_k)_{ij} = -\epsilon_{ijk} \quad (4.15)$$

à l'aide du tenseur antisymétrique ϵ_{ijk} (cf Chap 2) et que les relations de commutation de ces générateurs sont

$$[T_i, T_j] = \epsilon_{ijk} T_k, \quad (4.16)$$

à nouveau en termes de ce même tenseur.

Groupes compacts

Rappelons qu'en mathématiques, un sous-ensemble de $\mathbb{R}^d$ qui est fermé et borné est dit *compact*. Au contraire, des intervalles de la droite réelle comme $]a, b]$ ou $[a, \infty)$ ne sont pas compacts. On dira d'un groupe de Lie de dimension d qu'il est compact si l'ensemble de ses paramètres varie dans un sous-ensemble compact de $\mathbb{R}^d$. Nous admettrons que $O(n)$, $SO(n)$, $U(n)$, $SU(n)$ sont compacts, (on le verra au chap. 7 pour $SO(3)$), et il est clair que $\mathbb{R}^d$ ne l'est pas, n'étant pas borné.

Exercice : groupe de Lorentz. C'est l'ensemble des changements de coordonnées de la Relativité Restreinte, qui préservent la forme quadratique $ds^2 = c^2 dt^2 - d\vec{x}^2$. Quelle est sa dimension, quels sont ses générateurs infinitésimaux, est-il compact ?

La morale de ce paragraphe 4.1, qu'on va reprendre et développer aux chapitres suivants, est que, dans un groupe continu comme $U(1)$, $SO(2)$ ou $SO(3)$, et en fait de façon générale dans tout groupe de Lie comme on vient de les définir, une étude *locale* du groupe, au voisinage de l'identité, donc une étude des générateurs infinitésimaux, suffit pour reconstruire essentiellement toute la structure du groupe.

Plus précisément, on reconstruit tout élément d'un groupe compact connexe par exponentiation d'un élément de l'algèbre.

4.2 Symétries continues et lois de conservation

On va maintenant supposer qu'un groupe de transformations continues, comme ceux étudiés au § précédent, est un groupe de symétrie d'un système physique. En d'autres termes, la dynamique est invariante sous l'action du groupe. Des exemples sont fournis par des systèmes invariants par translation ou par rotation.

Selon le formalisme utilisé, hamiltonien ou lagrangien, les choses se présentent un peu différemment, mais la conclusion est la même : une invariance par un groupe continu se traduit par l'existence de quantités conservées, comme on va le voir.

4.2.1 Formalisme lagrangien et formalisme hamiltonien

Rappelons d'abord la formulation lagrangienne de la dynamique d'un système entre des temps t_1 et t_2 , qui applique un principe variationnel à une *fonctionnelle d'action* dépendant de la trajectoire $q(t)$

$$S[q(\cdot); t_1, t_2] = \int_{t_1}^{t_2} dt L(q(t), \dot{q}(t)) . \quad (4.17)$$

Ici q désigne l'ensemble des coordonnées des n degrés de liberté, $q = (q_1, q_2, \dots, q_n)$; L est une fonction, dite fonction de Lagrange, de $q(t)$ et $\dot{q}(t)$ évalués à une certaine valeur de t . Le principe variationnel consiste à dire que dans le mouvement entre les points q_1 au temps t_1 et q_2 au temps t_2 , l'action S est stationnaire par rapport à des variations infinitésimales de la trajectoire $q(t)$ sujettes aux conditions aux limites que $q(t_1) = q_1$ et $q(t_2) = q_2$ sont fixés. On écrit donc $q(t) \mapsto q(t) + \delta q(t)$, avec $\dot{q}(t) \mapsto \dot{q}(t) + \dot{\delta} q(t)$, et $\delta q(t_1) = \delta q(t_2) = 0$, et on calcule la variation au premier ordre de $S : S \mapsto S + \delta S = S + \sum_i \int_{t_1}^{t_2} dt \left(\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \dot{\delta} q_i \right)$, soit, après intégration par parties (rendue possible sans termes de bord grâce à l'annulation de δq aux extrémités de l'intervalle (t_1, t_2))

$$\delta S = \sum_i \int_{t_1}^{t_2} dt \delta q_i(t) \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) .$$

Cela ne peut être vrai indépendamment de la fonction $\delta q_i(t)$ que si la parenthèse s'annule pour tout $t \in (t_1, t_2)$, ce qui constitue les équations d'Euler-Lagrange

$$\frac{\partial L}{\partial q_i(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i(t)} = 0 \quad \forall i = 1, \dots, n . \quad (4.18)$$

On définit alors les variables d'impulsion p_i comme variables conjuguées des q_i en ce sens que

$$p_i = \frac{\partial L}{\partial \dot{q}_i}(q, \dot{q}) , \quad (4.19)$$

et l'équation d'Euler-Lagrange (4.18) nous dit que

$$\dot{p}_i = \frac{\partial L}{\partial q_i}(q, \dot{q}) . \quad (4.20)$$

Le hamiltonien se construit alors par une “transformée de Legendre” de L par rapport aux variables $\dot{q}$, ce qui veut dire que c'est la fonction de q et p donnée par

$$H(p, q) = \sum_i p_i \dot{q}_i(p, q) - L(q, \dot{q}(p, q)) \quad (4.21)$$

où la fonction $\dot{q}(p, q)$ est obtenue en résolvant (4.19). La fonctionnelle $H(p, q)$ mesure l'énergie du système. En différentiant (4.21), on a $dH = \sum_i \left(dp_i \dot{q}_i - \frac{\partial L}{\partial q_i} dq_i + \left(p_i - \frac{\partial L}{\partial \dot{q}_i} \right) d\dot{q}_i \right)$. Le dernier terme entre parenthèses s'annule en vertu de (4.19), et on lit alors les deux équations de Hamilton

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad (4.22)$$

qui réexpriment donc pour la première la relation (4.19) entre p et $\dot{q}$ et pour la seconde une équation dynamique, l'équation du mouvement équivalente à (4.18). On peut encore les récrire à l'aide des crochets de Poisson

$$\dot{q}_i = \{H, q_i\} \quad \dot{p}_i = \{H, p_i\} \quad (4.23)$$

où on rappelle que le crochet de Poisson de deux fonctions dans l'espace de phases est défini par

$$\{f(p, q), g(p, q)\} := \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} . \quad (4.24)$$

Exemple : Pour une particule de coordonnée $q = \vec{x}$, soumise à un potentiel $V(\vec{x})$, le lagrangien est simplement $L = \frac{1}{2} m \dot{\vec{x}}^2 - V(\vec{x})$ ce qui conduit à l'équation du mouvement (Euler-Lagrange)

$$m \ddot{\vec{x}} = -\vec{\nabla} V(\vec{x}) =: \vec{F} . \quad (4.25)$$

L'impulsion est $\vec{p} = m \dot{\vec{x}}$, l'hamiltonien $H = \frac{1}{2} \frac{\vec{p}^2}{m} + V(\vec{x})$ et les équations de Hamilton sont simplement $\dot{\vec{x}} = \frac{\vec{p}}{m}$, $\dot{\vec{p}} = -\vec{\nabla} V(\vec{x})$. Pour une particule libre, le membre de droite s'annule, et l'impulsion est conservée $d\vec{p}/dt = 0$.

Au chapitre 8, nous verrons comment ce principe variationnel d'action peut se généraliser dans un cadre relativiste, où le temps et les coordonnées d'espace jouent des rôles analogues.

4.2.2 Invariances et lois de conservation

En formalisme lagrangien, appelons symétrie une transformation des coordonnées laissant invariante la fonction de Lagrange $L(q(t))$. Pour une transformation infinitésimale, $q_i(t) \mapsto q_i(t) + \delta q_i(t)$, la condition d'invariance est

$$\delta L(t) = \sum_i \left(\delta q_i(t) \frac{\partial L}{\partial q_i(t)} + \delta \dot{q}_i(t) \frac{\partial L}{\partial \dot{q}_i(t)} \right) = 0$$

soit en utilisant les équations (4.19, 4.20),

$$\sum_i (\delta q_i(t) \dot{p}_i(t) + \delta \dot{q}_i(t) p_i(t)) = 0 \quad (4.26)$$

qui exprime que la quantité

$$\sum_i \delta q_i(t) p_i(t) = \text{const.} \quad (4.27)$$

est conservée dans le temps. On vient donc de montrer qu'une invariance dynamique par des transformations continues se traduit par une loi de conservation. C'est le premier exemple d'une situation très générale, s'étendant à des systèmes à nombre infini de degrés de liberté (théories de champs) et à la physique quantique et/ou relativiste. (On donne alors le nom de théorème de Noether à ce résultat important, cf Exercice A à la fin de ce chapitre.)

Examinons quelques exemples.

Invariance par translation d'espace

Supposons que les $N = 3n$ degrés de liberté q soient les coordonnées dans $\mathbb{R}^3$ de n particules, $\vec{r}_a$, $a = 1, \dots, n$. La fonction de Lagrange est de la forme usuelle en mécanique, $L = T - V$, avec T l'énergie cinétique et V (attention au signe !) le potentiel dont dérivent les forces. Supposons que ce potentiel ne dépend que des différences $\vec{r}_a - \vec{r}_b$ et donc qu'il est invariant par translation simultanée de tous les $\vec{r}_a : \vec{r}_a \mapsto \vec{r}_a + \delta \vec{a}$ avec $\delta \vec{a}$ indépendant du temps et de l'indice a . La loi de conservation précédente (4.27) se réduit à celle de $\delta \vec{a} \cdot \sum_a \vec{p}_a$ quel que soit $\delta \vec{a}$, et donc à celle de

$$\vec{P} = \sum_{a=1}^n \vec{p}_a \quad (4.28)$$

c'est-à-dire de l'impulsion totale. L'invariance par translation s'est donc traduite par la **loi de conservation de l'impulsion totale**.

Deux remarques

Bien noter que l'hypothèse faite d'invariance par translation d'espace exclut entre autres la présence dans la fonction de Lagrange de champs extérieurs non uniformes.

On sait que le formalisme lagrangien en électrodynamique fait apparaître non pas les champs $\vec{E}$ et $\vec{B}$ mais leurs potentiels V et $\vec{A}$. Donc inversement, la loi de conservation nous apprend : nul espoir d'accélérer une particule en lui appliquant un potentiel électrique ou magnétique uniforme. Cela n'est pas tellement pour nous surprendre...

Invariance par rotation

Avec les mêmes notations que précédemment, supposons que le potentiel $V(\vec{r}_a)$ est invariant par rotation : il ne dépend donc que des invariants $r_a = \|\vec{r}_a\|$ et $\vec{r}_a \cdot \vec{r}_b$, et peut-être aussi des pseudoscalaires $\det(\vec{r}_a, \vec{r}_b, \vec{r}_c)$. Une rotation infinitésimale s'écrivant

$$\vec{r}_a \mapsto \vec{r}_a + \delta\vec{\omega} \wedge \vec{r}_a$$

la quantité conservée (4.27) se récrit

$$\sum_a \delta\vec{\omega} \wedge \vec{r}_a \cdot \vec{p}_a = \delta\vec{\omega} \cdot \sum_{a=1}^n \vec{r}_a \wedge \vec{p}_a \quad (4.29)$$

et donc, puisque cela est vrai pour tout $\delta\vec{\omega}$, la loi de conservation est celle du moment cinétique

$$\vec{L} = \sum_{a=1}^n \vec{r}_a \wedge \vec{p}_a . \quad (4.30)$$

L'invariance par rotation s'est donc traduite par la **loi de conservation du moment cinétique total** ("orbital" à ce niveau classique : pas de spin!).

Invariance par translation du temps

Terminons cette revue des cas classiques par une situation légèrement différente de celles étudiées ci-dessus. On suppose le potentiel indépendant du temps (sans dépendance explicite en t , sa seule dépendance étant par l'intermédiaire des fonctions $q_i(t)$). Examinons l'effet sur la fonction de Lagrange d'une translation du temps

$$q_i(t) \mapsto q'_i(t) = q_i(t + \delta t) . \quad (4.31)$$

Pour une transformation infinitésimale (δt infinitésimal), $q_i(t) \mapsto q_i(t) + \delta q_i(t)$, $\delta q_i = \dot{q}_i(t) \delta t$. Cette fois nous n'écrivons pas que la fonction de Lagrange est invariante mais que sa variation, c'est-à-dire δt fois sa dérivée par rapport au temps, est donnée par

$$\delta L(t) = \delta t \frac{d}{dt} L(t) = \sum_i \left(\delta t \dot{q}_i(t) \frac{\partial L}{\partial q_i(t)} + \delta t \ddot{q}_i(t) \frac{\partial L}{\partial \dot{q}_i(t)} \right) = \delta t \sum_i (\dot{q}_i(t) \dot{p}_i(t) + \ddot{q}_i(t) p_i(t))$$

ou encore que

$$\frac{d}{dt} \left(\sum_i \dot{q}_i p_i - L \right) = 0 . \quad (4.32)$$

On reconnaît la conservation de l'hamiltonien, c'est-à-dire de l'énergie totale.

Récapitulons ce que nous avons trouvé à ce stade

Invariance	Loi de conservation	Quantité conservée
Translation d'espace	Impulsion	$\vec{P}$
Rotation d'espace	Moment angulaire	$\vec{L}$
Translation dans le temps	Énergie	$H(p, q)$

mais on verra d'autres exemples dans la suite (conservation de l'isospin et autres "symétries internes", de charges ...). D'une façon générale, l'existence d'une invariance sous l'action d'un groupe continu avec d générateurs infinitésimaux indépendants, donc de dimension d , se traduit par la conservation de d quantités indépendantes. Ainsi $d = 3$ pour les translations et pour les rotations de $\mathbb{R}^3$ avec les 3 composantes de $\vec{P}$ et de $\vec{L}$ respectivement, et $d = 1$ pour l'invariance dans le temps.

Formalisme hamiltonien

Pour finir, mentionnons rapidement comment les choses s'expriment en formalisme hamiltonien. On se rappelle que pour toute fonction $F(p, q)$ sur l'espace des phases sans dépendance explicite par rapport au temps t , l'évolution temporelle est dictée par l'équation

$$\begin{aligned}
 \dot{F}(p, q) &= \sum_i \left(\frac{\partial F}{\partial p_i} \dot{p}_i + \frac{\partial F}{\partial q_i} \dot{q}_i \right) \\
 &= \sum_i \left(-\frac{\partial F}{\partial p_i} \frac{\partial H}{\partial q_i} + \frac{\partial F}{\partial q_i} \frac{\partial H}{\partial p_i} \right) \\
 &= \{H, F\}
 \end{aligned} \tag{4.33}$$

où on a utilisé les équations (4.23) et la définition du crochet de Poisson (4.24). (En cas de dépendance explicite de F par rapport à t , il faudrait ajouter un terme supplémentaire $\partial F / \partial t$ au membre de droite.) Au vu de (4.33) on dit que l'hamiltonien est le "générateur au sens du crochet de Poisson" de l'évolution temporelle.

Plus généralement, considérons une transformation des variables de l'espace des phases de la forme $\delta q_i = \{f(p, q), q_i\} = \frac{\partial f}{\partial p_i}$, $\delta p_i = \{f(p, q), p_i\} = -\frac{\partial f}{\partial q_i}$ où f est une fonction indépendante du temps. Pour la fonction $H(p, q)$ (comme pour toute autre fonction sur l'espace de phases), la variation est donnée par le crochet de Poisson avec f

$$\delta H(p, q) = \sum_i \left(\frac{\partial H}{\partial p_i} \delta p_i + \frac{\partial H}{\partial q_i} \delta q_i \right) = \{f(p, q), H(p, q)\} . \tag{4.34}$$

La fonction $f(p, q)$ est le générateur de la transformation au sens du crochet de Poisson, avec la terminologie précédente. Une telle transformation est une invariance si $\delta H(p, q) = 0$ et on a alors $\dot{f} = \{H, f\} = 0$, selon (4.33). La fonction $f(p, q)$ est donc une quantité conservée.

Exemple des translations : $f(p, q) = \delta a \sum p_i$ de telle sorte que $\delta q_i = \{f, q_i\} = \delta a$, $\delta p_i = 0$. Traiter de même l'exemple des rotations. Quel est le générateur de la translation de temps ?

On retrouvera toutes ces notions en mécanique quantique, où le crochet de Poisson $\{f, g\}$ devient (à un facteur $-i\hbar$ près) le commutateur des observables $\hat{f}$ et $\hat{g}$: $[\hat{f}, \hat{g}] = -i\hbar \{f, g\}$.

4.3 Invariance de jauge de l'électrodynamique

Nous terminerons cette discussion générale des lois de conservation en discutant rapidement un cas d'invariance un peu différent mais fort intéressant, celui de l'*invariance de jauge* en électrodynamique. En électromagnétisme classique, cela se réfère à l'indépendance des quantités physiques, vecteurs champs électrique $\vec{E}$ et magnétique $\vec{B}$, par rapport à la transformation des potentiels dont ils dérivent, cf (2.5)

$$V \mapsto V + \frac{\partial f(\vec{x}, t)}{\partial t} \quad \vec{A} \mapsto \vec{A} - \vec{\nabla} f(\vec{x}, t) . \quad (4.35)$$

En effet, $\vec{E} = -\vec{\nabla}V - \frac{\partial \vec{A}}{\partial t}$ et $\vec{B} = \text{rot } \vec{A}$ sont bien invariants sous la transformation (4.35). Mais la transformation prend tout son sens dans un cadre plus général, couplant le champ électromagnétique à des charges, soit en électrodynamique classique, soit en physique quantique. Nous rappelant la description en mécanique quantique d'une particule de charge q couplée à un champ électrique ou magnétique, considérons l'équation de Schrödinger

$$i\hbar \frac{\partial \psi(\vec{x}, t)}{\partial t} = \left[-\frac{\hbar^2}{2m} \nabla^2 + \frac{iq\hbar}{m} \vec{A} \cdot \vec{\nabla} + \frac{iq\hbar}{2m} \text{div } \vec{A} + \frac{q^2}{2m} \vec{A}^2 + qV \right] \psi(\vec{x}, t) \quad (4.36)$$

ou encore, en utilisant l'identité $\vec{A} \cdot \vec{\nabla} + \vec{\nabla} \cdot \vec{A} = 2\vec{A} \cdot \vec{\nabla} + \text{div } \vec{A}$

$$i\hbar \frac{\partial \psi}{\partial t} = \frac{(\vec{p} - q\vec{A})^2}{2m} \psi + qV(\vec{x}, t) \psi \quad (4.37)$$

où $\vec{p} = -i\hbar \vec{\nabla}$. Noter que le champ électromagnétique n'y apparaît que par l'intermédiaire de ses potentiels scalaire V et vecteur $\vec{A}$. On peut remettre cette équation sous la forme

$$(i\hbar \frac{\partial}{\partial t} - qV) \psi = \frac{(-i\hbar \vec{\nabla} - q\vec{A})^2}{2m} \psi \quad (4.38)$$

sur laquelle l'invariance sous la transformation de jauge étendue à ψ

$$V \mapsto V + \frac{\partial f(\vec{x}, t)}{\partial t} \quad \vec{A} \mapsto \vec{A} - \vec{\nabla} f(\vec{x}, t) \quad \psi(\vec{x}, t) \mapsto e^{-i\frac{q}{\hbar} f(\vec{x}, t)} \psi(\vec{x}, t) \quad (4.39)$$

est manifeste. (Le vérifier!).

Inversement, la forme du couplage dans (4.37) est dictée par l'invariance (4.39). On parle de *couplage minimal* de la fonction d'onde ψ au champ électromagnétique. Noter encore que si plusieurs particules chargées sont présentes, chacune décrite par une fonction d'onde ψ_a , la transformation (4.39) s'applique encore avec une seule fonction $f(\vec{x}, t)$, la transformation de chaque fonction d'onde ψ_a étant dictée uniquement par sa charge q_a : on dit que le couplage minimal est "universel".

En fait l'invariance par (4.39) est intimement liée à la charge électrique et à sa conservation. Dans le cas particulier où f est indépendante de $\vec{x}$ et de t , (et donc V et $\vec{A}$ invariants), l'invariance de la physique par le groupe $U(1)$ agissant sur la fonction d'onde selon $\psi_a \mapsto e^{-iq_a f/\hbar} \psi_a$ est reliée (par le théorème de Noether) à la conservation de la charge totale $Q = \sum_a q_a$.

Ces considérations sur l'invariance de jauge, sur le couplage minimal etc, ont été généralisées par les physiciens C.N. Yang et R. Mills (1954) à des groupes plus généraux que le groupe $U(1)$ et non abéliens (voir Exercice B.2 ci-dessous). Cette généralisation est extraordinairement profonde et fructueuse, puisqu'elle est à la base de toutes les théories modernes de physique des particules. Le "modèle standard" qui décrit trois des quatre interactions fondamentales de la nature (forte, faible et électromagnétique) repose sur ce principe, étendant ainsi la théorie de Maxwell aux interactions nucléaires. La gravitation, qui est la quatrième interaction fondamentale, est elle-même fondée sur des principes d'invariance dans un esprit très proche de l'invariance de jauge (invariance de la Relativité Générale par changements de coordonnées).

4.4 Exercices en théorie classique des champs

A. Théorie classique des champs, Lagrangien, théorème de Noether, etc

On suppose que l'action d'une théorie classique des champs s'écrit

$$S = \int dt \int d^3x \mathcal{L}(\phi(\vec{x}, t), \partial_\mu \phi(\vec{x}, t); \vec{x}, t)$$

c'est-à-dire comme l'intégrale d'un *lagrangien* ne dépendant que du champ ϕ et de ses dérivées premières. Dans la suite, $x = \{x^0, x^1, x^2, x^3\}$ désigne les coordonnées de temps et d'espace $(ct, \vec{x})$, même si la théorie considérée n'est pas invariante relativiste, et on note $\partial_\mu = \frac{\partial}{\partial x^\mu}$. Noter qu'à ce stade, on autorise une dépendance explicite de $\mathcal{L}$ en x , par exemple à travers un champ extérieur dépendant de x . On suppose que le champ ϕ et ses dérivées s'annulent assez vite à l'infini d'espace et de temps pour assurer l'intégrabilité de $\mathcal{L}$.

a. Montrer que le principe de stationnarité de l'action S conduit aux équations d'Euler-Lagrange

$$\frac{\partial \mathcal{L}}{\partial \phi(x)} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi(x))} \right) = 0. \quad (\text{E} - \text{L})$$

b. On suppose maintenant que le lagrangien $\mathcal{L}$ est invariant par l'action d'un groupe de Lie G sur le champ ϕ , et que la transformation infinitésimale du champ s'écrit $\delta \phi(x) = \alpha_s T_s(\phi(x))$. Montrer en utilisant les équations d'E-L que la variation de $\mathcal{L}$ s'écrit

$$\delta \mathcal{L} = \alpha_s \partial_\mu \left(\frac{\partial \mathcal{L}(\phi, \partial \phi(x))}{\partial \partial_\mu \phi} \right) T_s(\phi(x))$$

et que son annulation pour tout α_s implique que le *courant de Noether*

$$j_s^\mu(x) = \left(\frac{\partial \mathcal{L}(\phi, \partial \phi)}{\partial \partial_\mu \phi(x)} \right) T_s(\phi(x))$$

a une divergence nulle, $\partial_\mu j_s^\mu = 0$.

Noter qu'il y a autant de tels courants de Noether qu'il y a de paramètres α_s indépendants dans le groupe, soit $\dim G$.

c. Montrer que cette condition implique la conservation de la charge associée à j^μ : $Q(t) = \int d^3x j^0(\vec{x}, t)$, à savoir $\frac{d}{dt}Q(t) = 0$. Par abus de langage, on dit que le courant j^μ est “conservé”.

d. On applique cela au cas du lagrangien d’un champ scalaire complexe, $\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi - \frac{\lambda}{2}(\phi^* \phi)^2$. Écrire les équations d’E-L pour ce lagrangien. Montrer que $\mathcal{L}$ est invariant par la transformation $\phi(x) \mapsto e^{-iq\alpha} \phi(x)$, (α indépendant de x !), dont on écrira la version infinitésimale. En déduire l’expression du courant de Noether. Vérifier qu’il est bien de divergence nulle en utilisant les équations d’E-L. Quelle est la charge conservée correspondante? Quel est le groupe? Un tel champ complexe est associé à une particule chargée, électriquement par exemple, et la charge de Noether conservée est alors la charge électrique.

B. Invariance de jauge

B.1. Dans le lagrangien précédent du champ complexe (ϕ, ϕ^*) , on effectue le remplacement $\partial_\mu \mapsto D_\mu := \partial_\mu + iqA_\mu(x)$, où A_μ est un nouveau champ. Montrer que $\mathcal{L}$ est maintenant invariant par les transformations simultanées de ϕ et de A_μ dépendant d’une fonction $\alpha(x)$

$$\phi(\vec{x}, t) \mapsto e^{-iq\alpha(\vec{x}, t)} \phi(\vec{x}, t) \quad A_\mu(\vec{x}, t) \mapsto A_\mu(\vec{x}, t) + \partial_\mu \alpha(\vec{x}, t) . \quad (4.40)$$

Le champ $A_\mu(x)$ s’interprète comme le champ (potentiel quadri-vecteur) électromagnétique. On ajoute à l’action un terme qui décrit sa dynamique sous la forme $\Delta\mathcal{L} = -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)^2$, lui aussi invariant par (4.40). C’est la manifestation de l’invariance de jauge de l’électrodynamique dans ce contexte lagrangien. En outre le lagrangien peut impliquer plusieurs champs chargés (c’est-à-dire complexes) $\phi_1, \phi_2 \dots$ de charges $q_1, q_2, \dots$, tous couplés au champ A_μ par leur *dérivée covariante* D_μ : on parle de “couplage minimal” de ces champs au champ A_μ . Ce couplage minimal est “universel” en ce sens qu’il ne dépend que de la charge q du champ ϕ . Le principe d’invariance de jauge a donc dicté la forme de l’interaction matière-champ électromagnétique.

B.2 Lagrangien de Yang-Mills.

Selon l’idée brillante de C.N. Yang et R. Mills (1954), on peut généraliser les considérations précédentes d’invariance de jauge, de couplage minimal etc à d’autres groupes de symétrie que les transformations du groupe $U(1)$ de l’électrodynamique.

On considère un champ $\vec{A}_\mu$ qui est à la fois un vecteur à trois composantes (dans un espace abstrait $\mathcal{E}$ de dimension 3 que nous ne précisons pas ici) et un *quadrivecteur* de Lorentz qui porte donc un indice μ prenant 4 valeurs, comme le champ électromagnétique $A_\mu = (V, A_i)$, cf § 4.3. On construit alors le lagrangien

$$\mathcal{L} = -\frac{1}{4} \vec{F}_{\mu\nu} \vec{F}^{\mu\nu} \quad \text{avec} \quad \vec{F}_{\mu\nu} = (\partial_\mu \vec{A}_\nu - \partial_\nu \vec{A}_\mu - \vec{A}_\mu \wedge \vec{A}_\nu)$$

– Montrer que ce lagrangien est invariant par la transformation infinitésimale

$$\delta \vec{A}_\mu(\vec{x}, t) = \partial_\mu \vec{\alpha}(\vec{x}, t) + \vec{A}_\mu(\vec{x}, t) \wedge \vec{\alpha}(\vec{x}, t)$$

qui généralise le cas “abélien” de l’électrodynamique.

– Montrer que pour α indépendant de x , la transformation précédente est une rotation infinitésimale des vecteurs $\vec{A}$ et $\vec{F}$ dans l’espace abstrait $\mathcal{E}$.

Ce champ de Yang-Mills peut alors être couplé à des champs de matière, “chargés” comme l’est le champ ϕ du A. ci-dessus, c’est-à-dire se transformant d’une façon définie sous l’action de ces rotations. Cette idée, étendue à d’autres groupes, est à la base de la construction du modèle standard de la physique des particules, cf chap. 6.

C. ★ Conservation du tenseur énergie-impulsion. Transformations conformes

On démontre que dans une théorie lagrangienne (classique) des champs, les invariances par changement de coordonnées –translations, rotations, dilatations, . . .– sont associées à la conservation de courants construits à partir du *tenseur énergie-impulsion* $\Theta_{\mu\nu}$. Plus précisément, nous admettrons que si le lagrangien n’a pas de dépendance explicite en x (pas de champ extérieur), la théorie est invariante par translation, il existe un tenseur $\Theta_{\mu\nu}$ “conservé” : $\partial^\mu \Theta_{\mu\nu} = 0$, comme conséquence des équations d’E-L, et l’invariance par un changement de coordonnées infinitésimal arbitraire $x^\mu \mapsto x'^\mu = x^\mu + \delta\alpha^\mu(x)$ se traduit par $\delta S = \int d^d x (\partial_\mu \alpha_\nu(x)) \Theta^{\mu\nu}(x)$.

a) Montrer que l’invariance par rotation (de l’espace-temps) découle de la symétrie du tenseur $\Theta_{\mu\nu} = \Theta_{\nu\mu}$.

b) Montrer que l’invariance par dilatation découle de $\Theta^\mu_\mu = 0$ (trace nulle).

c) Dans la suite on se place dans une théorie euclidienne (et non pas dotée de la métrique minkovskienne). On considère alors la transformation géométrique suivante, composée d’une “inversion” par rapport à l’origine (attention, inversion en un sens différent de celui de parité utilisé ailleurs dans le cours !) $x^\mu \mapsto \frac{x^\mu}{x^2}$, d’une translation par a^μ et d’une nouvelle inversion. Montrer que la forme infinitésimale d’une telle transformation est $\delta\alpha^\mu(x) := \delta x^\mu = x^2 \delta a^\mu - 2(\delta a \cdot x)x^\mu$.

d) Montrer alors que si Θ satisfait les propriétés précédentes de symétrie et de trace nulle, l’action est automatiquement invariante par les transformations infinitésimales $\delta\alpha^\mu$.

e) Ces dernières transformations, ainsi que les rotations et dilatations, sont des *transformations conformes*, qui conservent les angles. On démontre qu’en dimension d supérieure ou égale à 3, toutes les transformations conformes sont des composées des transformations précédentes, translations, rotations, dilatation et les “transformations conformes spéciales” de la question c). Calculer la dimension de ce groupe conforme en dimension d .

On vient donc de démontrer le résultat suivant : dans une théorie (lagrangienne) des champs classique, l’invariance par translation, rotation et dilatation implique nécessairement celle par toutes les transformations conformes. C’est en fait ce qui se passe dans une théorie des champs décrivant un point critique, cf chap. 3. Le résultat s’étend à $d = 2$ où il est infiniment plus contraignant, puisque les transformations conformes infinitésimales forment alors une algèbre de Lie de dimension infinie, pourquoi ?

Chapitre 5

Représentations des groupes

Nous avons vu au chapitre 2 l'action du groupe $O(3)$ sur les vecteurs de l'espace à trois dimensions $\mathbb{R}^3$ et sur les tenseurs de cet espace. Le concept de représentation va généraliser cette situation à l'action d'un groupe G quelconque dans un espace vectoriel quelconque. Dans ce cours on se restreindra à des espaces de dimension finie n , qu'on peut donc toujours voir comme $\mathbb{R}^n$ ou $\mathbb{C}^n$ ¹.

Ce chapitre va présenter quelques éléments de base de la théorie mathématique des représentations des groupes. Pour simplifier la discussion et les notations, l'analyse du paragraphe 5.2 se concentrera sur les groupes finis, mais les résultats obtenus s'étendent aux groupes de Lie "compacts" tels $U(1)$ et $SO(3)$ d'usage très courant pour le physicien.

5.1 Définition et propriétés générales des représentations

5.1.1 Définitions de base

Rappelons que l'ensemble des applications linéaires inversibles d'un espace de dimension n dans lui-même forme un groupe, appelé groupe linéaire et noté $GL(n)$, cf § 4.1.5. Sous forme matricielle, c'est le groupe des matrices $n \times n$ inversibles. On écrit à l'occasion $GL(n, \mathbb{R})$ ou $GL(n, \mathbb{C})$ pour préciser si on travaille sur les nombres réels ou les complexes.

On dit qu'un groupe G admet une représentation (réelle, resp. complexe) de dimension n si à tout élément g de G on peut associer un élément $D(g)$ de $GL(n)$ de telle façon qu'aux opérations de groupe dans G correspondent les opérations de groupe dans $GL(n)$ ². Autrement dit

$$\begin{aligned} \forall g \in G \quad & g \mapsto D(g) \in GL(n) \\ \forall g, g' \in G \quad & D(g.g') = D(g).D(g') \end{aligned} \tag{5.1}$$

1. Noter cependant que la physique peut nécessiter la considération d'espaces de dimension infinie, comme les espaces de fonctions de carré intégrable $L^2(\mathbb{R}^d)$ rencontrés comme espaces de Hilbert des états en mécanique quantique.

2. En langage mathématique, on dit qu'il y a un *homomorphisme* de G dans $GL(n)$

d'où découle immédiatement, en prenant $g' = e$ puis $g' = g^{-1}$, que

$$\begin{aligned} D(e) &= I \\ \forall g \in G \quad D(g^{-1}) &= (D(g))^{-1} \end{aligned}$$

où e est l'élément neutre dans G et I désigne l'opérateur identité dans $\text{GL}(n)$. La représentation qui à tout $g \in G$ associe le nombre 1 (considéré comme $\in \text{GL}(1, \mathbb{R})$) est appelée *triviale* ou *représentation identité*; elle est de dimension 1.

Dans l'espace $\mathbb{R}^n$ (ou $\mathbb{C}^n$), on peut choisir une base e_i , $i = 1, \dots, n$, et associer à tout $g \in G$ la matrice représentative de $D(g)$:

$$D(g)e_j = e_i \mathcal{D}_{ij}(g) \quad (5.2)$$

Comme (presque) partout dans la suite de ces notes, on a adopté ici la “convention de sommation sur les indices répétés” : la sommation sur $1 \leq i \leq n$ est implicite dans (5.2). La disposition des indices (i : indice de ligne, j indice de colonne) est dictée par la loi (5.1). En effet, on a bien

$$\begin{aligned} D(g.g')e_k &= e_i \mathcal{D}_{ik}(g.g') \\ &= D(g) (D(g')e_k) = D(g)e_j \mathcal{D}_{jk}(g') \\ &= e_i \mathcal{D}_{ij}(g) \mathcal{D}_{jk}(g') \\ \text{donc} \quad \mathcal{D}_{ik}(g.g') &= \mathcal{D}_{ij}(g) \mathcal{D}_{jk}(g') . \end{aligned}$$

Comme on sait bien, si on change de base $e'_i = e_j V_{ji}$ avec une matrice inversible, donc $e_i = \sum_j e'_j V_{ji}^{-1}$, la matrice du même opérateur D dans la nouvelle base est $\mathcal{D}' = V^{-1} \mathcal{D} V$, puisque $D e'_i = D e_{i'} V_{i'i} = e_{j'} \mathcal{D}_{j'i'} V_{i'i} = e'_j V_{jj'}^{-1} \mathcal{D}_{j'i'} V_{i'i} = e'_j (V^{-1} \mathcal{D} V)_{ji}$.

Exemples : Le groupe $\text{SO}(2)$ des rotations dans le plan admet une représentation de dimension deux, avec des matrices

$$\mathcal{D}(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (5.3)$$

qui décrivent les rotations d'angle θ autour de l'origine. Vérifier que la condition (5.1) est bien satisfaite, grâce à des identités trigonométriques simples.

On verra au chapitre 7 que les représentations (“irréductibles” et “inéquivalentes”, voir ci-dessous) du groupe $\text{SO}(3)$ et du groupe $\text{SU}(2)$ qui lui est apparenté sont caractérisées par le *spin* j entier ou demi-entier...

5.1.2 Représentations équivalentes. Caractères

Soient D et D' deux représentations de dimension n d'un groupe G . Ces représentations sont dites *équivalentes* s'il existe un opérateur linéaire V inversible

$$\forall g \in G \quad D'(g) = V^{-1} D(g) V \quad (5.4)$$

On voit que les représentations D et D' sont nécessairement de même dimension : les matrices représentatives de D et D' sont des matrices carrées, reliées par une transformation d'équivalence et peuvent être considérées comme différant par un changement de base. Il n'y a donc pas lieu de distinguer fondamentalement deux représentations équivalentes et on ne s'intéressera dans la suite qu'à des représentations *inéquivalentes*.

On appelle *caractère* d'une représentation de dimension finie la trace de l'opérateur $D(g)$:

$$\chi(g) = \text{tr } D(g) . \quad (5.5)$$

C'est une fonction de G dans $\mathbb{R}$ ou $\mathbb{C}$. Le caractère est indépendant du choix de base dans E et deux représentations équivalentes ont le même caractère puisque $\text{tr } D'(g) = \text{tr } V^{-1} D(g) V = \text{tr } D(g)$. Le caractère prend aussi la même valeur pour les différents éléments d'une même *classe de conjugaison*, c'est-à-dire tous les éléments de la forme $g' = hgh^{-1}$ pour un g donné et h quelconque dans G . En effet

$$\chi(hgh^{-1}) = \text{tr } D(hgh^{-1}) = \text{tr } (D(h)D(g)D(h)^{-1}) = \text{tr } D(g) = \chi(g)$$

où on a utilisé la propriété (5.1). On dit que le caractère est une *fonction de classe*.

Noter que tout élément du groupe appartient à une classe et une seule : on dit que les classes forment une *partition* du groupe. Dans un groupe fini, le nombre de classes est nécessairement fini, on les notera dans la suite C_i .

Exemples :

- dans un groupe abélien, les classes sont constituées d'un seul élément, pourquoi ?
- dans le groupe S_n des permutations de n objets, on démontre que les classes rassemblent toutes les permutations ayant la même décomposition en cycles ; ainsi le groupe S_3 des permutations de 3 objets a trois classes, celle de l'identité, notée $[1^3]$, la classe $[1\ 2]$ des 3 transpositions (un cycle de longueur 2, un de longueur 1) et la classe $[3]$ des 2 permutations cycliques (un cycle de longueur 3) ;
- dans le groupe de rotation du cube (cf chap. 1), qui a 24 éléments, il existe 5 classes distinctes constituées des 8 rotations d'angle $\pm 2\pi/3$ autour des 4 diagonales, des 6 rotations de π autour des 6 axes passant par les milieux des côtés, des 6 rotations de $\pm\pi/2$ et des 3 de π autour des 3 axes passant par les centres des faces, et de l'identité (voir chap. 1, § 1.3.3 et TD 5) ;
- la classe d'une rotation R de $SO(3)$ rassemble toutes les rotations d'axe $\vec{u}$ quelconque mais de même angle de rotation θ que R , comme on le verra au chapitre 7 et au TD5. Dans la représentation de dimension 3 (celle sur les vecteurs étudiée au Chapitre 2), le caractère de la

représentation, c'est-à-dire la trace de la matrice $R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$, est

$$\chi(R(\theta)) = 1 + 2 \cos \theta , \quad (5.6)$$

qui ne dépend effectivement que de θ et pas de l'axe de rotation $\vec{u}$. Dans le même ordre d'idées, la rotation-réflexion $S(\phi)$ introduite au chap. 1, composition d'une rotation $R(\phi)$ par une réflexion-miroir dans un plan orthogonal à l'axe Δ de rotation, a pour matrice $S(\phi) =$

$\begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & -1 \end{pmatrix}$ dans un repère où Δ est le 3ème axe de coordonnées. Elle a donc pour caractère

$$\chi(S(\phi)) = -1 + 2 \cos \phi. \quad (5.7)$$

Deux cas particuliers sont celui d'une pure réflexion σ , qui correspond à $\phi = 0$, donc $\chi(\sigma) = 1$, et celui de l'inversion I , $\phi = \pi$, donc $\chi(I) = -3$.

Comme on l'a déjà mentionné, les représentations "irréductibles" de $SO(3)$ sont caractérisées par le spin j . On se rappelle des cours de Mécanique Quantique (et on reverra au chap. 7) que la rotation d'angle θ autour de l'axe z est *représentée* par la matrice $R_z(\theta) = \exp -i\theta J_z = \text{diag}(e^{-i\theta m})_{-j \leq m \leq j}$, dans la base où $J_z = \text{diag}(j, j-1, \dots, -j)$. On calcule alors le caractère de la représentation de spin j

$$\chi_j(\theta) = \sum_{m=-j}^j e^{-im\theta} = \frac{\sin(j + \frac{1}{2})\theta}{\sin \frac{1}{2}\theta}. \quad (5.8)$$

On retrouve bien sûr (5.6) pour $j = 1$.

Nous ferons un usage fréquent des expressions de ces caractères.

On notera encore que le caractère, évalué pour l'élément identité du groupe, fournit la dimension de la représentation ($= n$ dans nos notations)

$$\chi(e) = \dim D = n. \quad (5.9)$$

5.1.3 Représentations réductibles et irréductibles

Considérons une certaine représentation D d'un groupe G . Supposons qu'on a trouvé une base dans laquelle son expression matricielle $\mathcal{D}(g)$ se décompose en matrices-blocs de taille $n_1 \times n_1$, $n_1 \times n_2$, $n_2 \times n_1$ et $n_2 \times n_2$, ($n_1 + n_2 = n$, dimension de la représentation), selon

$$\forall g \in G \quad \mathcal{D}(g) = \begin{pmatrix} \mathcal{D}_1(g) & \mathcal{D}'(g) \\ 0 & \mathcal{D}_2(g) \end{pmatrix}. \quad (5.10)$$

Cela signifie que les vecteurs n'ayant que les n_1 premières composantes non nulles se transforment "entre eux", sans se mélanger avec ceux n'ayant que les n_2 dernières composantes non nulles : en effet si x_1 est un vecteur colonne de dimension n_1

$$\mathcal{D}(g) \begin{pmatrix} x_1 \\ 0 \end{pmatrix} = \begin{pmatrix} \mathcal{D}_1(g)x_1 \\ 0 \end{pmatrix}.$$

Autrement dit l'espace E_1 engendré par les n_1 premiers vecteurs de base est *invariant* (sous l'action des opérateurs $D(g)$). Noter qu'en général pour un vecteur colonne x_2 n'ayant que les

dernières n_2 composantes non nulles, $\mathcal{D}(g) \begin{pmatrix} 0 \\ x_2 \end{pmatrix} = \begin{pmatrix} \mathcal{D}'(g)x_2 \\ \mathcal{D}_2(g)x_2 \end{pmatrix}$, qui mélange les deux types de composantes. Autrement dit le sous-espace supplémentaire E_2 de E_1 n'est en général pas invariant.

Inversement si E_1 est un sous-espace invariant “non-trivial”³ de la représentation D , on peut trouver une base telle que (5.10) y soit vraie. On dit qu'une représentation ayant cette propriété est *réductible*. Si en outre $\mathcal{D}'(g) = 0$ pour tout g , le sous-espace supplémentaire E_2 est lui aussi invariant, on dit que la représentation est *complètement réductible* et qu'elle est la “somme directe” des deux représentations D_1 et D_2 . Enfin si aucun sous-espace invariant n'existe (et donc qu'il est impossible de trouver une base où (5.10) tient), on dit que la représentation est *irréductible*.

On va voir que dans les cas les plus intéressants en pratique, toute représentation réductible est complètement réductible et peut donc se décomposer en somme directe de représentations irréductibles. Il est donc naturel, et profitable comme on verra, de se restreindre alors à l'étude des représentations irréductibles.

Il faut encore souligner l'importance du corps de base dans la discussion de l'irréductibilité. C'est ainsi que la représentation (5.3) qui est irréductible sur un espace vectoriel sur $\mathbb{R}$ ne l'est pas sur $\mathbb{C}$: au prix d'un changement de base on peut la récrire comme $\begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix}$. Exercice : écrire le changement de base (complexe !) qui fait passer de la forme (5.3) à cette forme diagonale.

Exemple : tenseurs de rang 2 de $\text{SO}(3)$

Au chapitre 2, § 1.2, on a étudié les tenseurs de rang 2 de l'espace $\mathbb{R}^3$, décrits par leurs $3^2 = 9$ composantes X^{ij} , $i, j = 1, 2, 3$, et se transformant selon $X^{i_1 i_2} \mapsto X'^{i_1 i_2} = O_{i_1 j_1} O_{i_2 j_2} X^{j_1 j_2}$, (somme sur les indices répétés!), ce qu'on peut noter de façon abrégée $X' = (O \otimes O)X$. Vérifions tout d'abord que ces tenseurs forment une représentation de $\text{SO}(3)$ au sens précédent. Si deux rotations de matrices O_1 et O_2 sont appliquées successivement, le tenseur X se transforme selon $X \mapsto X' = O_1 \otimes O_1 X \mapsto O_2 \otimes O_2 X' = (O_2 \otimes O_2)(O_1 \otimes O_1)X = (O_2 \cdot O_1 \otimes O_2 \cdot O_1)X$ qui est bien la propriété (5.1) de composition d'une représentation. Exercice : vérifier cette assertion en détaillant les composantes.

On remarque que l'action de la représentation est très simple sur les trois objets suivants

- (i) le tenseur identité I de composantes δ_{ij} est invariant : $\delta_{i_1 i_2} \mapsto O_{i_1 j_1} O_{i_2 j_2} \delta_{j_1 j_2} = \delta_{i_1 i_2}$ par l'orthogonalité des matrices O : c'est la représentation identité, de dimension 1 ;
- (ii) un tenseur antisymétrique $X_{i_1 i_2} = -X_{i_2 i_1}$ se transforme en un tenseur X' antisymétrique (le vérifier) ;
- (iii) de même un tenseur symétrique $X_{i_1 i_2} = X_{i_2 i_1}$ se transforme en un tenseur X' symétrique ;
- (iv) compte tenu du point (i), un tenseur symétrique de trace nulle $X_{i_1 i_2} = X_{i_2 i_1}$, $\text{tr } X = \sum X_{ii} = 0$ se transforme en un tenseur X' symétrique de trace nulle.

3. c'est-à-dire différent de 0 et de E tout entier

Autrement dit, les tenseurs symétriques de trace nulle, les tenseurs antisymétriques et les tenseurs proportionnels au tenseur identité se transforment indépendamment, et forment donc chacun une représentation du groupe $SO(3)$ (dont on peut démontrer qu'elle est irréductible).

Écrivons maintenant l'identité suivante, vraie pour tout tenseur de rang 2

$$X = \frac{1}{2}(X + X^T) + \frac{1}{2}(X - X^T) = \frac{1}{3}(\text{tr } X)I + \frac{1}{2}\left(X + X^T - \frac{2}{3}(\text{tr } X)I\right) + \frac{1}{2}(X - X^T)$$

où $X^T = \{X_{ji}\}$ est le tenseur transposé. On a décomposé le tenseur X en la somme d'un tenseur multiple de l'identité, d'un tenseur (de rang 2) symétrique de trace nulle, et d'un tenseur (de rang 2) antisymétrique. Selon les observations précédentes, chaque terme se transforme indépendamment des autres. La représentation portée par les tenseurs de rang 2 est donc complètement réductible.

Autre façon de dire cela : le produit tensoriel $\vec{x} \otimes \vec{y}$ de deux vecteurs $\vec{x}$ et $\vec{y}$ se décompose en trois termes se transformant de façon indépendante sous l'action des rotations de $SO(3)$, le produit scalaire invariant $\vec{x} \cdot \vec{y}$, le produit vectoriel $\vec{x} \wedge \vec{y}$, et le produit tensoriel symétrique de trace nulle, $\vec{x} \otimes \vec{y} + \vec{y} \otimes \vec{x} - \frac{2}{3}\vec{x} \cdot \vec{y} I$.

Les lecteurs les plus perspicaces auront reconnu dans ce qui précède un avatar d'un calcul fait en mécanique quantique : la "composition" de deux spins 1 (produit tensoriel de deux vecteurs à 3 composantes) peut se décomposer en somme d'un spin 0 (invariant), d'un spin 1 (objet vectoriel "dual" du tenseur de rang 2 antisymétrique) et d'un spin 2, le tenseur symétrique de rang de trace nulle, à 5 composantes. La somme des dimensions est bien $3 \times 3 = 1 + 3 + 5$.

5.1.4 Représentations unitaires

On considère un espace E de dimension n ($E = \mathbb{R}^n$ ou $\mathbb{C}^n$), dont les vecteurs seront écrits avec les notations de la physique quantique $|x\rangle$. Cet espace est supposé muni d'un produit scalaire $\langle x|y\rangle$ "défini positif"⁴. Une représentation d'un groupe G dans l'espace E est dite *unitaire* si pour tout $g \in G$, l'opérateur $D(g)$ est unitaire, c'est-à-dire si pour tout g l'adjoint $D^\dagger(g)$ est l'inverse de $D(g)$. La condition s'écrit donc $D(g).D^\dagger(g) = I$, ou encore $D^\dagger(g).D(g) = I$. Rappelons que l'opérateur $D(g)$ préserve alors le produit scalaire des vecteurs. En effet pour tout élément $g \in G$ et pour toute paire de vecteurs $|x\rangle, |y\rangle$ de E , et en notant $|D(g)x\rangle = D(g)|x\rangle, |D(g)y\rangle = D(g)|y\rangle$ leurs transformés par l'action de $D(g)$, on a

$$\langle D(g)x|D(g)y\rangle = \langle x|D(g)^\dagger D(g)|y\rangle = \langle x|y\rangle \quad (5.11)$$

$$\text{puisque } D(g)^\dagger D(g) = I. \quad (5.12)$$

En utilisant les propriétés de définition d'une représentation on a aussi

$$D(g^{-1}) = D^{-1}(g) = D^\dagger(g). \quad (5.13)$$

Comme on le verra au chapitre suivant, ces représentations unitaires sont très utiles au physicien, en particulier dans le cadre de la mécanique quantique.

4. petit abus de langage : on veut dire que la forme quadratique associée $\langle x|x\rangle$ est définie positive : $\langle x|x\rangle \geq 0$ et $\langle x|x\rangle = 0$ ssi $x = 0$.

Bornons-nous pour le moment à des considérations mathématiques. On a les deux propriétés importantes suivantes :

(i) *Toute représentation d'un groupe fini sur un espace doté d'un produit scalaire défini positif est "unitarisable", c'est-à-dire équivalente à une représentation unitaire.*

La preuve en est assez simple. Considérons une représentation D d'un groupe fini G et formons

$$Q = \sum_{g' \in G} D^\dagger(g') D(g') \quad (5.14)$$

qui satisfait

$$D^\dagger(g) Q D(g) = \sum_{g' \in G} D^\dagger(g' \cdot g) D(g' \cdot g) = \sum_{g' \cdot g \in G} D^\dagger(g' \cdot g) D(g' \cdot g) = \sum_{g'' \in G} D^\dagger(g'') D(g'') = Q \quad (5.15)$$

où on a remplacé $\sum_{g'}$ par $\sum_{g' \cdot g}$ puisque dans l'un et l'autre cas, la somme court sur tous les éléments du groupe ("lemme de réarrangement"). L'opérateur Q est auto-adjoint $Q = Q^\dagger$, il est défini positif, ce qui signifie que $\langle x | Q | x \rangle = \sum_g \| D(g) | x \rangle \|^2 > 0$ pour $| x \rangle \neq 0$. On peut alors écrire Q sous la forme

$$Q = V^\dagger V \quad (5.16)$$

avec V inversible. Par exemple, on effectue la diagonalisation de l'opérateur auto-adjoint Q par une matrice unitaire, $Q = U \Lambda^2 U^\dagger$, avec Λ diagonale réelle, ce qui permet d'en extraire la "racine carrée" $V = U \Lambda U^\dagger$. L'opérateur V permet alors de définir une représentation D' équivalente à D et unitaire :

$$D'(g) = V D(g) V^{-1} \quad (5.17)$$

$$\begin{aligned} D'^\dagger(g) D'(g) &= V^{\dagger-1} D^\dagger(g) V^\dagger V D(g) V^{-1} \\ &= V^{\dagger-1} D^\dagger(g) Q D(g) V^{-1} = V^{\dagger-1} Q V^{-1} = I . \end{aligned} \quad (5.18)$$

ce qui établit la propriété.

On démontre que cette propriété s'étend aux groupes continus "compacts" (comme les groupes orthogonaux $O(n)$ et $SO(n)$ ou unitaires $U(n)$ et $SU(n)$, au contraire du groupe des translations $\mathbb{R}^d$).

(ii) *Toute représentation unitaire est soit irréductible, soit complètement réductible.*

Autrement dit, la situation mentionnée au § 5.1.3 d'une représentation réductible sans être complètement réductible, ne peut se produire.

En effet soit E_1 un sous-espace invariant, le sous-espace E_2 orthogonal à E_1 est lui-même invariant puisque pour tout $g \in G$, $x \in E_1$ et $y \in E_2$ on a

$$\langle x | D(g) y \rangle = \langle D(g^{-1}) x | y \rangle = 0 \quad (5.19)$$

ce qui prouve que $| D(g) y \rangle \in E_2$.

Comme corollaire des deux propriétés précédentes, toute représentation réductible d'un groupe fini ou d'un groupe compact est (équivalente à) une représentation unitaire et complètement réductible. Il suffit donc pour nous de construire et de classifier les représentations unitaires irréductibles.

5.1.5 Lemme de Schur

Si D et D' sont deux représentations irréductibles d'un groupe G dans l'espace $\mathbb{C}^n$, on a l'important

Lemme de Schur. *S'il existe un opérateur V tel que $\forall g \in G, VD(g) = D'(g)V$, alors ou bien $V = 0$, ou bien V est inversible et les représentations D et D' sont équivalentes.*

Nous admettrons ce lemme, dont nous ferons un grand usage dans la suite.

Sa preuve consiste à étudier le *noyau* de V , c'est-à-dire le sous-espace des x tels que $Vx = 0$, et l'*image* de V , sous-espace des $y = Vx$, et de montrer que ce sont des sous-espaces invariants de D et de D' respectivement. L'hypothèse d'irréductibilité conduit alors au résultat.

Ce lemme a deux conséquences importantes

Corollaire 1. *Si D est une représentation irréductible d'un groupe G dans l'espace $\mathbb{C}^n$, tout opérateur V commutant avec tous les représentants du groupe, $\forall g \in G, VD(g) = D(g)V$, est un multiple de l'identité, $V = \lambda I$.*

En effet, sur $\mathbb{C}$, V a au moins une valeur propre λ (qui est non nulle puisque V est inversible par le lemme de Schur). L'opérateur $V - \lambda I$ commute aussi avec D , mais il a une valeur propre nulle, il n'est donc pas inversible donc est nul.

Ce corollaire va nous être extrêmement utile en mécanique quantique, comme on verra.

On en a en fait déjà rencontré une application dans l'étude du groupe des rotations $SO(3)$ dans le cours de Mécanique Quantique. On a montré là que si J_x, J_y, J_z sont les générateurs infinitésimaux (non commutants) du groupe, $\vec{J}^2 = J_x^2 + J_y^2 + J_z^2$ commute, lui, avec tous les générateurs, donc aussi avec les matrices de rotations finies. Le corollaire nous dit que dans toute représentation irréductible du groupe $SO(3)$, $\vec{J}^2$ est un multiple de l'identité, autrement dit que toute représentation irréductible de $SO(3)$ peut être repérée par le nombre j tel que $\vec{J}^2 = j(j+1)I$, cf le Chap. 7 où on montrera que ce j doit être entier.

Corollaire 2. *Une représentation irréductible complexe d'un groupe abélien est nécessairement de dimension 1.*

En effet, soit $g' \in G$, $D(g')$ commute avec tous les $D(g)$. Donc (corollaire 1) $D(g') = \lambda(g')I$. La représentation se décompose en $\dim D$ copies de la représentation de dimension 1 : $g \mapsto \lambda(g)$, et l'irréductibilité impose que $\dim D = 1$.

Insistons sur l'importance de considérer des représentations complexes dans ces deux corollaires. C'est le caractère "algébriquement clos" de $\mathbb{C}$, c'est-à-dire la propriété de toute équation algébrique d'y avoir au moins une racine, par opposition à $\mathbb{R}$, qui est déterminant. La représentation sur $\mathbb{R}^2$ du groupe $SO(2)$ par les matrices $\mathcal{D}(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ vient fournir des contre-exemples aux deux propositions précédentes : toute matrice $\mathcal{D}(\alpha)$ commute avec $\mathcal{D}(\theta)$ mais n'est pas de la forme λI avec λ réel (si $\alpha \neq 0, \pi$) et la représentation est irréductible, quoique de dimension deux.

5.2 Représentations des groupes finis

Dans cette section, nous allons nous intéresser aux représentations des groupes finis sur le corps des complexes $\mathbb{C}$. La plupart des résultats qu'on va obtenir sont basés sur le fait qu'on peut effectuer la sommation sur les éléments du groupe. Ces résultats pourront se généraliser par la suite à des groupes infinis, pourvu qu'on puisse y donner un sens à cette sommation.

5.2.1 Orthogonalité et complétude des représentations

Un petit rappel sera peut-être utile pour la suite : considérons un ensemble de m vecteurs $(X^{(1)}, \dots, X^{(m)})$ de $\mathbb{C}^n$. Dans une base orthonormée, leurs composantes sont $X_i^{(a)}$, $i = 1, \dots, n$, $a = 1, \dots, m$. Supposons qu'ils satisfont les deux familles de relations

$$\sum_{i=1}^n X_i^{(a)} X_i^{(a')*} = \delta_{aa'} \quad (5.20)$$

$$\sum_{a=1}^m X_i^{(a)} X_{i'}^{(a)*} = \delta_{ii'}. \quad (5.21)$$

La première relation exprime que les m vecteurs sont orthogonaux, donc indépendants. Dans un espace de dimension n , cela n'est possible que si $m \leq n$. La deuxième relation exprime une propriété de complétude : les $X^{(a)}$ forment un système de générateurs. En effet, tout vecteur Y de l'espace, de composantes Y_i peut s'écrire $Y_i = \sum_{i'=1}^n Y_{i'} \delta_{ii'} = \sum_{i'} Y_{i'} \sum_{a=1}^m X_i^{(a)} X_{i'}^{(a)*} = \sum_{a=1}^m X_i^{(a)} (X^{(a)}, Y)$ et donc $Y = \sum_{a=1}^m X^{(a)} (X^{(a)}, Y)$, Y se décompose bien sur les X , cqfd. Cela n'est possible que si $m \geq n$, et en définitive, les deux relations (5.20, 5.21) qui s'interprètent donc comme des propriétés d'orthogonalité et de complétude des X ne sont possibles que si $n = m$, et les $X^{(a)}$ forment une *base* de $\mathbb{C}^n$.

Soit G un groupe fini d'ordre (= nombre d'éléments) $|G|$, désignons ses représentations irréductibles inéquivalentes par un indice supérieur : $D^{(\rho)}$, et leur dimension par n_ρ ; leurs matrices $\mathcal{D}_{\alpha\beta}^{(\rho)}$ peuvent être supposées unitaires d'après le résultat du paragraphe 5.1.4. On va aussi s'intéresser aux caractères $\chi^{(\rho)} = \text{tr } D^{(\rho)}$. Ces caractères sont des fonctions de classe (cf. § 5.1.2) ; notons C_i les classes, $|C_i|$ leur nombre d'éléments et $\chi_i^{(\rho)} = \chi^{(\rho)}(g)|_{g \in C_i}$ la valeur que prend le caractère de la représentation ρ dans la classe C_i . On a bien sûr $\sum_i |C_i| = |G|$ puisque les classes forment une partition du groupe G .

On démontre alors le

Théorème 1. : Les matrices $\mathcal{D}_{\alpha\beta}^{(\rho)}$ satisfont les propriétés d'orthogonalité et de complétude suivantes

$$\frac{1}{|G|} \sum_g \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\alpha'\beta'}^{(\rho')*}(g) = \frac{1}{n_\rho} \delta_{\rho\rho'} \delta_{\alpha\alpha'} \delta_{\beta\beta'} \quad (5.22)$$

$$\sum_{\rho, \alpha, \beta} \frac{n_\rho}{|G|} \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\alpha\beta}^{(\rho)*}(g') = \delta_{g, g'}. \quad (5.23)$$

et les caractères satisfont les relations d'orthogonalité et de complétude

$$\frac{1}{|G|} \sum_g \chi^{(\rho)}(g) \chi^{(\rho')*}(g) = \delta_{\rho\rho'} \quad \text{ou encore} \quad \frac{1}{|G|} \sum_i |C_i| \chi_i^{(\rho)} \chi_i^{(\rho')*} = \delta_{\rho\rho'} \quad (5.24)$$

$$\frac{|C_i|}{|G|} \sum_\rho \chi_i^{(\rho)} \chi_j^{(\rho)*} = \delta_{ij}. \quad (5.25)$$

La propriété d'orthogonalité (5.22) découle assez simplement du lemme de Schur. Nous l'admettrons (voir Exercice 1). Celle des caractères (5.24) s'obtient alors en prenant la trace $\sum_{\alpha=\beta, \alpha'=\beta'}$ de (5.22). Les formules de complétude (5.23, 5.25) sont plus délicates à obtenir.

Exercice : vérifier que le membre de gauche de (5.23) peut s'écrire aussi $\sum_{\rho} \frac{n_{\rho}}{|G|} \chi^{(\rho)}(g.g'^{-1})$.

La formule de complétude (5.23) est importante car elle nous apprend que toute fonction $f(g)$ d'un élément du groupe peut s'exprimer comme une somme (finie) de contributions des différentes représentations irréductibles

$$f(g) = \sum_{g'} f(g') \delta_{g,g'} = \sum_{\rho, \alpha, \beta} \frac{n_{\rho}}{|G|} \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \sum_{g'} \mathcal{D}_{\alpha\beta}^{(\rho)*}(g') f(g') = \sum_{\rho, \alpha, \beta} \frac{n_{\rho}}{|G|} \mathcal{D}_{\alpha\beta}^{(\rho)}(g) f_{\alpha\beta}^{(\rho)}, \quad (5.26)$$

où $f_{\alpha\beta}^{(\rho)} = \sum_{g'} \mathcal{D}_{\alpha\beta}^{(\rho)*}(g') f(g')$. De même, la formule de complétude (5.25) nous apprend que toute fonction de classe $F(C_i)$ (cf. § 5.1.2) peut se développer comme une somme sur les caractères

$$F(C_i) = \sum_j F(C_j) \delta_{ij} = \frac{|C_i|}{|G|} \sum_{\rho} \chi_i^{(\rho)} \sum_j \chi_j^{(\rho)*} F(C_j) = \frac{|C_i|}{|G|} \sum_{\rho} \chi_i^{(\rho)} F^{(\rho)}. \quad (5.27)$$

Exercice. Montrer que si deux éléments g et g' de G ont même caractère $\chi^{(\rho)}$ pour tout ρ , alors ces deux éléments sont conjugués $g \sim g'$.

Remarques et conséquences

(i) Le cas le plus simple où la propriété d'orthogonalité est déjà familière est celui du groupe cyclique $\mathbb{Z}_p$, un groupe fini abélien. Dans ce cas, les représentations irréductibles (complexes) sont de dimension 1 (corollaire 2 du lemme de Schur), elles sont indexées par un entier $l = 0, 1, \dots$ et s'écrivent (en notation multiplicative où le groupe $\mathbb{Z}_p$ est l'ensemble des racines p -ièmes de l'unité $z_k = \exp 2\pi i \frac{k}{p}$)

$$D^{(l)} : \quad z_k \mapsto z_k^l \quad l = 0, 1, \dots, p-1. \quad (5.28)$$

(Que $z \mapsto z^l$ soit une représentation est clair : $(zz')^l = z^l z'^l$, que cela épuise toutes les représentations du groupe ne l'est pas mais va découler du théorème 2 plus bas.) La propriété d'orthogonalité s'écrit alors

$$\frac{1}{p} \sum_k z_k^l z_k^{l'*} = \frac{1}{p} \sum_{k=0}^{p-1} e^{i \frac{2\pi}{p} k(l-l')} = \delta_{ll'}, \quad (5.29)$$

un résultat classique (somme géométrique). Les classes se réduisent à un seul élément, et la propriété de complétude s'écrit donc

$$\frac{1}{p} \sum_{l=0}^{p-1} z_k^l z_{k'}^{l*} = \frac{1}{p} \sum_{l=0}^{p-1} e^{i \frac{2\pi}{p} l(k-k')} = \delta_{kk'}, \quad (5.30)$$

c'est-à-dire une identité de même forme que la précédente, mais exprimant ici une propriété différente.

(ii) La propriété (5.22) est importante parce qu'elle suffit souvent à construire ou à compléter une table des caractères des représentations irréductibles d'un groupe fini donné. Considérons par exemple le cas du groupe $\mathbb{Z}_2$, constitué des éléments 1 et -1 . On connaît la représentation

triviale qui associe à tout élément la valeur 1. Une deuxième représentation, elle aussi nécessairement de dimension 1 par le lemme de Schur, doit associer la valeur $\mathcal{D}(1) = \chi(1) = \dim D = 1$ à l'élément identité et une valeur x à -1 . Par (5.22) x est fixé sans ambiguïté à -1 . Comme il y a au plus deux représentations irréductibles non équivalentes (comme on va le démontrer dans un instant), on a la Table

$\downarrow$ repr. \ éléments $\rightarrow$	1	-1
id	1	1
ϵ	1	-1

Cette discussion peut être répétée dans des cas moins triviaux (cf ci-dessous et TD).

(iii) Au vu de (5.24) et (5.25), on peut appliquer l'argument du début de ce paragraphe aux $\left(\frac{|C_i|}{|G|}\right)^{\frac{1}{2}} \chi_i^{(\rho)}$. Donc

Corollaire 1. *Le nombre m de représentations irréductibles inéquivalentes (le nombre de valeurs de ρ) est égal au nombre de classes (le nombre de valeurs de i).*

(iv) On peut appliquer aussi ce même argument à l'ensemble des éléments de matrice $\mathcal{D}_{\alpha\beta}^{(\rho)}(g) \sqrt{\frac{n_\rho}{|G|}}$ quand g "parcourt" le groupe G . Ils peuvent être considérés comme les composantes (indexées par ρ et $\alpha, \beta = 1, \dots, n_\rho$) de $|G|$ vecteurs. Les relations (5.22) et (5.23) expriment la complétude et l'orthogonalité de ces vecteurs, on a donc $|G| = \sum_\rho \sum_{\alpha, \beta=1}^{n_\rho} 1 = \sum_\rho n_\rho^2$ d'où le

Théorème 2 : *Les dimensions n_ρ satisfont l'égalité*

$$\sum_\rho n_\rho^2 = |G|. \quad (5.31)$$

Exemples : 1. Dans un groupe abélien, le nombre de représentations irréductibles sur $\mathbb{C}$ est égal à l'ordre. Cela est une conséquence triviale du fait que ces représentations sont de dimension 1 (Corollaire 2 du lemme de Schur). Ainsi dans le groupe $\mathbb{Z}_p$ étudié plus haut, le nombre de représentations irréductibles est bien p .

2. Considérons à nouveau le groupe (de rotations) du cube. On a vu plus haut que ses 24 éléments se répartissent en 5 classes. Par le Corollaire 1, il a donc 5 représentations irréductibles distinctes. Par le Théorème 2, leurs dimensions doivent satisfaire $\sum n_\rho^2 = 24$ ce qui se trouve être très contraignant et n'admet qu'une seule solution : les n_ρ prennent les valeurs 1,1,2,3,3.

5.2.2 Conséquences

(i) Toute représentation D étant complètement réductible, on peut la décomposer en représentations irréductibles selon

$$D = \oplus_\rho m_\rho D^{(\rho)} \quad (5.32)$$

ce qui signifie qu'il existe une base où on peut écrire

$$\mathcal{D}(g) = \begin{pmatrix} \mathcal{D}^{(\rho_1)} & 0 & 0 \\ 0 & \mathcal{D}^{(\rho_2)} & 0 \\ 0 & 0 & \ddots \end{pmatrix}$$

avec peut-être une multiplicité (entière) m_ρ d'apparition de la représentation ρ , et (en prenant la trace) son caractère s'écrit

$$\chi = \sum_{\rho} m_{\rho} \chi^{(\rho)} . \quad (5.33)$$

Grâce aux formules d'orthogonalité des caractères, les multiplicités peuvent se calculer par la formule

$$m_{\rho} = \frac{1}{|G|} \sum_i |C_i| \chi_i \chi_i^{(\rho)*} = \frac{1}{|G|} \sum_g \chi(g) \chi^{(\rho)*}(g) . \quad (5.34)$$

(ii) Utilisons la relation (5.33) pour calculer la “norme carrée” d'un caractère quelconque, définie par

$$\|\chi\|^2 = \frac{1}{|G|} \sum_{g \in G} |\chi(g)|^2 = \frac{1}{|G|} \sum_i |C_i| \chi_i \chi_i^* \stackrel{(5.24)}{=} \sum_{\rho} m_{\rho}^2 . \quad (5.35)$$

C'est donc un entier supérieur ou égal à un, l'égalité à 1 étant satisfaite si et seulement si la représentation considérée est irréductible. Cela fournit donc un critère simple pour décider si une représentation dont on connaît les caractères est ou non irréductible. Voir comme illustration l'exercice 2 sur les groupes diédraux.

(iii) Dans tous ces calculs, il est extrêmement utile de disposer de la table des valeurs que prennent les différents caractères irréductibles pour les différentes classes. Dans la construction de ces *tables de caractères*, les relations (5.24-5.25) sont très importantes. A titre illustratif, étudions le cas du groupe $\mathcal{S}_3$. On sait qu'il a trois classes (cf. § 5.1.2), correspondant aux produits de cycles $[3]$, $[1\ 2]$ et $[1^3]$. Il a donc trois représentations irréductibles inéquivalentes, dont nous connaissons déjà deux, de dimension 1, qui existent dans tous les $\mathcal{S}_n$: la représentation identité d'une part, et la représentation ϵ qui associe à toute permutation sa *signature*⁵. Il doit donc exister une troisième représentation, de dimension 2 en vertu de (5.31). Son caractère prend donc la valeur 2 pour la classe $[1^3]$ (classe de l'identité). Les relations (5.24-5.25) permettent de déterminer sans difficulté les éléments manquants de la troisième ligne de la Table

$\downarrow$ Repr. $\rho \setminus$ Classes $C_i \rightarrow$	$[1^3]$	$[1\ 2]$	$[3]$
identité = $\{3\}$	1	1	1
$\epsilon = \{1^3\}$	1	-1	1
$\{2, 1\}$	2	0	-1
$ C_i $	1	3	2

Les notations $\{3\}$ etc se réfèrent à la théorie générale des représentations des groupes symétriques (“tableaux d'Young”); elles ne seront pas expliquées dans ce cours. Noter qu'on lit dans la première colonne de cette table la dimension n_ρ de la représentation (égale à la valeur

5. Rappelons qu'à toute permutation $\sigma \in \mathcal{S}_n$ on associe le signe $\epsilon_\sigma = (-1)^\#$ où $\#$ désigne un nombre de transpositions qui font passer de la permutation de départ à σ . Ce signe ϵ_σ , appelé *signature* de la permutation σ , est ce qui apparaît dans la formule du déterminant d'une matrice A de taille $n \times n$: $\det A = \sum_{\sigma \in \mathcal{S}_n} \epsilon_\sigma \prod_i A_{i\sigma(i)}$. On démontre que $\epsilon_\sigma \epsilon_\tau = \epsilon_{\sigma\tau}$ ce qui prouve bien que la signature est une représentation de dimension 1 du groupe $\mathcal{S}_n$.

du caractère pour l'identité), tandis que la dernière ligne donne le nombre d'éléments $|C_i|$ de chaque classe.

On verra en exercice et en TD d'autres exemples de construction de tables de caractères de groupes finis.

5.2.3 Un groupe continu : $U(1)$

Comme on l'a mentionné à plusieurs reprises, les considérations précédentes sur les groupes finis s'étendent à une large classe de groupes continus. Comme prototype de groupe continu particulièrement simple, considérons le groupe $U(1)$, groupe multiplicatif des nombres complexes de module 1

$$\begin{aligned} U(1) &= \{z, |z| = 1\} \\ &= \{e^{i\alpha}, (2\ell - 1)\pi < \alpha \leq (2\ell + 1)\pi\} \end{aligned} \quad (5.36)$$

où le choix de la détermination de l'angle α modulo 2π (le choix de l'entier ℓ) ne doit pas importer pour la paramétrisation du cercle. C'est un groupe abélien et ses représentations irréductibles sont donc de dimension 1. Elles sont faciles à trouver (exercice 5)

$$D^{(k)}(e^{i\alpha}) = \chi^{(k)}(e^{i\alpha}) = e^{ik\alpha} \quad (5.37)$$

et on requiert que $k \in \mathbb{Z}$ pour assurer que la représentation est univaluée quand α change de détermination $\alpha \rightarrow \alpha + 2\pi n$. Plus généralement, les fonctions sur le groupe (ici les fonctions de classe) sont les fonctions définies (univaluées) sur le cercle, c'est-à-dire les fonctions périodiques de α . Sous des hypothèses adéquates de régularité, elles admettent un développement convergent en série de Fourier

$$f(z = e^{i\alpha}) = \sum_{n \in \mathbb{Z}} f_n e^{in\alpha} = \sum_n f_n z^n \quad (5.38)$$

qui n'est autre que le développement de f sur les caractères (5.37).

On notera que le groupe $U(1)$ se distingue du groupe abélien ("non compact") $\mathbb{R}$ qui décrit l'addition des phases α si α est autorisé à varier sur tout $\mathbb{R}$. La représentation de l'élément $\alpha \in \mathbb{R}$ y est toujours donnée par l'exponentielle $e^{ix\alpha}$, mais x est maintenant arbitraire (réel si la représentation est unitaire).

Les caractères (5.37) du groupe $U(1)$ satisfont des relations d'orthogonalité et de complétude parallèles à celles démontrées plus haut pour les groupes finis. Ce qui remplace la sommation finie $\frac{1}{|G|} \sum_{g \in G}$ sur les éléments du groupe est l'intégrale sur le cercle

$$\int_{-\pi}^{\pi} \frac{d\alpha}{2\pi} .$$

Les caractères satisfont alors

$$\begin{aligned} \int_{-\pi}^{\pi} \frac{d\alpha}{2\pi} \chi^{(k)}(e^{i\alpha}) \chi^{(k')*}(e^{i\alpha}) &= \int_{-\pi}^{\pi} \frac{d\alpha}{2\pi} e^{i(k-k')\alpha} = \delta_{kk'} \\ \sum_{k \in \mathbb{Z}} \chi^{(k)}(e^{i\alpha}) \chi^{(k)*}(e^{i\beta}) &= \sum_{k \in \mathbb{Z}} e^{i(\alpha-\beta)k} = 2\pi \delta_P(\alpha - \beta) . \end{aligned} \quad (5.39)$$

La deuxième relation est une identité familière dans la transformation de Fourier. La “fonction” δ_P (en fait une distribution, appelée parfois “peigne de Dirac” ou fonction III – cyrillique, prononcer cha) qui apparaît ici est 2π -périodique et identifie α et β modulo 2π

$$\delta_P(\alpha - \beta) = \sum_{n \in \mathbb{Z}} \delta(\alpha - \beta + 2\pi n) . \quad (5.40)$$

Le groupe $\text{SO}(2)$ des rotations dans le plan est isomorphe au groupe $\text{U}(1)$. Noter que si on s'intéresse à des représentations irréductibles réelles, leur dimension n'est plus égale à 1 (sauf pour la représentation identité!) mais à 2

$$\begin{aligned} \mathcal{D}^{(k)}(\alpha) &= \begin{pmatrix} \cos k\alpha & -\sin k\alpha \\ \sin k\alpha & \cos k\alpha \end{pmatrix} & k \in \mathbb{N}^* \\ \chi^{(k)}(\alpha) &= 2 \cos k\alpha \end{aligned} \quad (5.41)$$

Noter aussi que ce groupe $\text{SO}(2)$ est isomorphe au groupe des rotations autour d'un axe de $\mathbb{R}^3$, appelé C_∞ en cristallographie, cf chap. 1 et 2. C'est le groupe de symétrie de rotation d'une molécule diatomique de type A-B.

Ce qu'on vient de dire pour le groupe $\text{U}(1)$ peut s'étendre à tout groupe compact (cf. chap. 4, § 4.1.5) : on peut y définir une mesure d'intégration (mesure de Haar) et donner un sens à des formules d'orthogonalité comme (5.22) où une intégrale remplace la somme sur les éléments du groupe, comme on vient de le faire pour le groupe $\text{U}(1)$. On verra au chapitre 7 (§ 7.2 et 7.3) comment cela se réalise dans les groupes $\text{SO}(3)$ et $\text{SU}(2)$.

5.3 Produits directs de groupes ou de représentations ; décomposition de Clebsch-Gordan

L'expression “produit direct” s'applique à deux situations bien distinctes à ne pas confondre, celle du produit $G = G_1 \times G_2$ de deux groupes, et celle du produit (direct ou tensoriel) de deux espaces de représentations d'un même groupe G . Ces deux situations sont rencontrées en physique.

5.3.1 Produits directs de groupes et leurs représentations

On a rencontré au paragraphe 1.3.3 la notion de groupe produit de deux groupes : on a vu que le groupe complet O_h (rotations et réflexions) du cube est le produit (direct) du groupe O des rotations par le groupe $\mathbb{Z}_2$ engendré par l'inversion σ par rapport au centre du cube. Plus généralement, on dit que le groupe G est le produit direct de deux groupes G_1 et G_2 , et on écrit $G = G_1 \times G_2$ si on peut écrire tout élément g de G sous la forme $g = (g_1, g_2)$, avec $g_1 \in G_1$, $g_2 \in G_2$, et si la loi de composition de G s'exprime par $g.g' = (g_1, g_2).(g'_1, g'_2) = (g_1.g'_1, g_2.g'_2)$. En particulier $g = (g_1, e_2).(e_1, g_2)$ mais aussi $g = (e_1, g_2).(g_1, e_2)$.

Si G_1 et G_2 sont d'ordre fini, il en est de même de G et on a $|G| = |G_1||G_2|$. Montrons que les représentations irréductibles de G s'obtiennent simplement à partir de celles de G_1 et

G_2 . Soient D_1, D_2 deux représentations irréductibles de G_1 et G_2 , respectivement, n_1, n_2 leurs dimensions. On construit alors la représentation *produit tensoriel* de D_1 et D_2

$$D = D_1 \otimes D_2 \quad D(g=(g_1, g_2)) = D_1(g_1) \otimes D_2(g_2) . \quad (5.42)$$

De façon plus explicite, si des bases $(e_1)_i, i = 1, \dots, n_1$ et $(e_2)_j, j = 1, \dots, n_2$ ont été choisies dans les espaces des représentations D_1 et D_2 , les éléments de matrice de D sont

$$\mathcal{D}_{ij;i'j'}(g) = (\mathcal{D}_1)_{ii'}(g_1)(\mathcal{D}_2)_{jj'}(g_2) .$$

Sous cette forme, la vérification que la condition (5.1) est bien satisfaite par D est élémentaire. On démontre en utilisant le lemme de Schur que la représentation $D = D_1 \otimes D_2$ est irréductible si les représentations D_1 et D_2 le sont. Le caractère de la représentation D est simplement le produit des caractères

$$\chi(g) = \text{tr } D(g) = \chi_1(g_1)\chi_2(g_2) .$$

Montrons maintenant que l'on obtient bien ainsi toutes les représentations irréductibles de G . Supposons qu'on ait la liste des représentations irréductibles $D_1^{(\rho)}$ de G_1 et celle $D_2^{(\sigma)}$ de G_2 . Il suffit alors d'utiliser la relation (5.31) pour se convaincre que l'ensemble des représentations produits tensoriels $D_1^{(\rho)} \otimes D_2^{(\sigma)}$, de dimensions $n_{\rho\sigma} = n_\rho^{(1)} n_\sigma^{(2)}$, fournit bien toutes les représentations irréductibles de $G = G_1 \times G_2$. En effet

$$\sum_{\rho} (n_\rho^{(1)})^2 = |G_1|, \quad \sum_{\sigma} (n_\sigma^{(2)})^2 = |G_2| \implies \sum_{\rho, \sigma} (n_{\rho\sigma})^2 = |G_1||G_2| = |G| , \quad (5.43)$$

ce qui montre qu'aucune autre représentation irréductible inéquivalente n'est possible pour G .

Revenant à l'exemple de départ $O_h = O \times \mathbb{Z}_2$, nous voyons qu'il suffit d'avoir construit la table des caractères de O pour obtenir celle de O_h . Chaque représentation irréductible D_ρ de O donne lieu à deux représentations irréductibles de O_h que nous pouvons indexer par $(\rho, +)$ ou $(\rho, -)$. Chaque classe $\mathcal{C}_i$ de O donne lieu à deux classes de O_h , notées $+\mathcal{C}_i$ et $-\mathcal{C}_i$, et les caractères de O_h se lisent

$$\chi_{(\rho, \pm)}(+\mathcal{C}_i) = \chi_\rho(\mathcal{C}_i) \quad \chi_{(\rho, \pm)}(-\mathcal{C}_i) = \pm \chi_\rho(\mathcal{C}_i) . \quad (5.44)$$

Application : voir la discussion des molécules diatomiques au chapitre suivant.

5.3.2 Produit tensoriel de représentations

On a examiné au § 5.1.3 comment les tenseurs de rang 2 de $\text{SO}(3)$ se décomposent en représentations irréductibles. C'est aussi la situation qu'on rencontre en Mécanique Quantique, quand on connaît la transformation des composantes d'un système et qu'on étudie comment le système composé se transforme (système de deux particules de spin j_1 et j_2 par exemple).

En général soient E_1 et E_2 deux espaces vectoriels portant des représentations D_1 et D_2 d'un groupe G . L'espace produit tensoriel $E = E_1 \otimes E_2$ est l'espace engendré par les combinaisons linéaires de "produits" (tensoriels) d'un élément de E_1 et d'un élément de E_2 : $z = \sum_a x^{(a)} \otimes y^{(a)}$.

L'espace E porte lui même une représentation, notée $D = D_1 \otimes D_2$, *produit tensoriel* (ou *produit direct*) des représentations D_1 et D_2 . Sur l'élément z ci-dessus

$$D(g)z = \sum_a D_1(g)x^{(a)} \otimes D_2(g)y^{(a)} . \quad (5.45)$$

Sous forme matricielle, on écrira $\mathcal{D}_{ij;i'j'}(g) = \mathcal{D}_{1ii'}(g)\mathcal{D}_{2jj'}(g)$ avec les indices i, i' se référant à l'espace E_1 et j, j' à E_2 . On vérifie immédiatement que le caractère de la représentation D est le produit des caractères χ_1 et χ_2 de D_1 et D_2

$$\chi(g) = \chi_1(g)\chi_2(g) . \quad (5.46)$$

En particulier en évaluant cette relation pour $g = e$, on a pour des représentations de dimension finie

$$\dim D = \dim(E_1 \otimes E_2) = \dim E_1 \cdot \dim E_2 = \dim D_1 \cdot \dim D_2 \quad (5.47)$$

comme il est bien connu pour un produit tensoriel.

Décomposition de Clebsch-Gordan

La représentation produit direct de deux représentations irréductibles D et D' n'est en général pas irréductible. Si elle est complètement réductible (comme c'est le cas pour les représentations unitaires qui vont nous intéresser au premier chef), on effectue la *décomposition de Clebsch-Gordan* en représentations irréductibles

$$D \otimes D' = \oplus_j D_j \quad (5.48)$$

où au second membre on somme sur un nombre fini de représentations. On peut préférer à (5.48) une autre écriture qui indique lesquelles des représentations inéquivalentes $D^{(\rho)}$ apparaissent, et avec quelle multiplicité

$$D \otimes D' = \oplus_\rho m_\rho D^{(\rho)} . \quad (5.49)$$

Les entiers $m_\rho = \frac{1}{|G|} \sum_g \chi_D \chi_{D'} \chi^{(\rho)*}$ (cf (5.34)) sont non négatifs. Les équations (5.48) et (5.49) impliquent des règles simples sur les caractères et les dimensions

$$\chi_D \cdot \chi_{D'} = \sum_j \chi_j = \sum_\rho m_\rho \chi^{(\rho)} \quad (5.50)$$

$$\dim D \cdot \dim D' = \sum_j \dim D_j = \sum_\rho m_\rho \dim D^{(\rho)} . \quad (5.51)$$

Exemple : le produit tensoriel de deux copies de l'espace euclidien de dimension 3 ne forme pas une représentation irréductible du groupe des rotations. C'est l'exemple considéré plus haut au § 5.1.3.

5.4 Décomposition des représentations d'un groupe sur celles d'un sous-groupe

Une situation rencontrée fréquemment en géométrie ou en physique est celle où un groupe de symétrie G est réduit à (on dit aussi “brisé en”) un de ses sous-groupes H . C'est par exemple le cas où dans un problème initial invariant par rotation, on introduit un champ électrique vertical à symétrie cylindrique qui brise le groupe $G = O(3)$ en son sous-groupe $H = C_{\infty v}$. C'est le cas aussi de la symétrie de rotation d'un atome réduite à un sous-groupe fini si l'atome est inséré dans un cristal, voir la discussion du “champ cristallin” à la fin du chap. 7. Si D est une représentation de G , c'en est aussi une de H , puisque la propriété (5.1) s'applique à tous $g, g' \in H \subset G$. Mais si D est une représentation irréductible de G , rien n'assure qu'elle l'est encore pour H . Se pose donc le problème de calculer la décomposition de D en représentations irréductibles de H . Là encore, les techniques de caractères sont utiles. Si on dispose des tables de caractères de G et de H , la simple application des formules (5.33, 5.34) fournit la multiplicité de la représentation irréductible de H indexée par ρ dans la représentation de caractère χ de G .

Exemple : pour discuter comment un niveau de moment orbital l d'un atome se scinde dans un champ cristallin de groupe G , il suffit de disposer de la table de caractères de G , du caractère (5.8) de la représentation de spin $j = l$ et de le projeter sur les caractères irréductibles de G , voir exercice 6.

5.5 Applications physiques

5.5.1 Modes de vibration des molécules

Nous revenons au problème présenté au chapitre 2, § 2.5, qui est de déterminer les $N = 3n - 6$ ou $3n - 5$ modes normaux (ou propres) de vibration des noyaux d'une molécule, cf (2.18)

$$H = \frac{1}{2} \sum_{1 \leq i \leq N} (\dot{Q}_i^2 + \omega_i^2 Q_i^2). \quad (2.18)$$

Si la molécule possède un groupe de symétrie G (un des groupes ponctuels étudiés au chapitre 1), ce groupe agit linéairement sur les coordonnées q_i des noyaux. Il agit donc aussi linéairement sur les coordonnées Q_i qui sont fonctions linéaires des q_j , autrement dit, les modes normaux Q_i forment une représentation D (de dimension N) du groupe G . Cette représentation est en général réductible. Inversement des modes normaux se transformant selon une représentation irréductible correspondent à une fréquence propre donnée ω . La dimension de la représentation irréductible $D^{(\rho)}$ fournit donc la multiplicité de la fréquence ω_ρ . En général, et sauf “dégénérescence accidentelle”, deux représentations irréductibles distinctes correspondent à des fréquences propres différentes⁶. Si une représentation irréductible $D^{(\rho)}$ apparaît avec une multiplicité m_ρ , on aura m_ρ fréquences $\omega_{\rho,i}$ distinctes, ($i = 1, \dots, m_\rho$), chacune de multiplicité $\dim D^{(\rho)}$. La théorie des groupes va donc nous permettre d'accéder assez facilement à la multiplicité des fréquences propres de vibration des molécules, sans avoir à déterminer effectivement ce que sont les modes normaux.

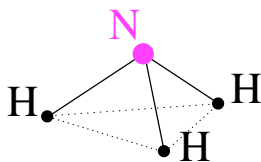
6. Exception : les représentations complexes conjuguées. Si une représentation irréductible complexe apparaît dans un problème physique, par essence réel, la représentation complexe conjuguée apparaît aussi avec la même multiplicité et la même fréquence propre.

Il nous faut donc décomposer la représentation totale D portée par les N modes de la molécule en représentations irréductibles de G . Pour cela on va calculer le caractère de D pour les différents éléments du groupe G et le projeter selon les formules (5.33, 5.34). Les éléments du groupe G sont soit des rotations R d'angle θ , soit des réflexions-miroirs σ , soit des “rotations-réflexions” S d'angle ϕ (cf Chap 1), soit l'inversion I par rapport à un centre de symétrie éventuel de la molécule, soit l'identité e . Comme le caractère de chacune de ces transformations est indépendant de la base où on l'évalue, on peut revenir à son action sur les noyaux de la molécule, en prenant bien soin d'éliminer à nouveau les contributions des modes zéros (cf Chap 2). Seuls les noyaux invariants par la transformation contribuent aux éléments *diagonaux* de l'opérateur D , donc à son caractère. En se rappelant les expressions (5.6, 5.7), on trouve ([9], § 100)

$$\begin{aligned}
 \chi(R(\theta)) &= (N_C - 2)(1 + 2 \cos \theta) && \text{s'il y a } N_C \text{ noyaux sur l'axe de rotation} \\
 \chi(S(\phi)) &= N_S(-1 + 2 \cos \phi) && \text{s'il y a } N_S \text{ noyaux sur l'axe de rotation-réflexion} \\
 \chi(\sigma) &= N_\sigma && \text{s'il y a } N_\sigma \text{ noyaux dans le plan de la réflexion-miroir } \sigma \\
 \chi(I) &= -3N_I && \text{si } N_I (=0 \text{ ou } 1) \text{ noyau est invariant par l'inversion } I \\
 \chi(e) &= N = 3n - 6 && \text{(pour une molécule non linéaire)}
 \end{aligned}
 \tag{5.52}$$

où le -2 dans $N_C - 2$ prend en compte les mouvements de translation et de rotation globales à retrancher comme modes zéros.

Exemple : molécule NH_3 . Le groupe de symétrie est le groupe C_{3v} , d'ordre 6, fait de 3 classes : l'identité e , C_3 : 2 rotations d'angle $\pm 2\pi/3$ et σ_v : 3 réflexions dans les plans bissecteurs des liaisons H-H ; la table de caractères de C_{3v} est la suivante (les notations pour les représentations irréductibles sont traditionnelles...)



↓ Repr. \ Classes →	e	C_3	σ_v
A_1	1	1	1
A_2	1	1	-1
E	2	-1	0
$ C_i $	1	2	3

(Noter que cette table reproduit, aux notations près, celle donnée plus haut pour le groupe $\mathcal{S}_3$ de permutations de trois objets. S'agit-il d'une coïncidence?). On calcule alors selon les formules précédentes $\chi(e) = N = 6$, $\chi(C_3) = 0$ puisque $\theta = 2\pi/3$ et $\chi(\sigma_v) = N_\sigma = 2$, d'où en utilisant (5.34) les multiplicités $m_{A_1} = \frac{1}{6}(6+0+3 \times 2) = 2$, $m_{A_2} = 0$ et $m_E = 2$. On a donc deux fréquences simples (de multiplicité 1) a priori distinctes correspondant au mode normal A_1 ; ces modes normaux attachés à la représentation identité conservent la symétrie de la molécule : ils décrivent d'une part les vibrations radiales de même amplitude des trois H, de l'autre la vibration selon l'axe vertical du N. On a par ailleurs deux autres fréquences correspondant aux modes normaux de type E , chacun venant avec une multiplicité deux.

Répéter cette discussion de NH_3 en ne considérant que les seules rotations (groupe $\mathbb{Z}_3 \cong C_3$). Comment le résultat se compare-t-il à celui qu'on vient d'obtenir ?

Cette analyse peut être répétée pour de nombreuses molécules dotées de symétries variées, voir les références [9], [14], [5].

Exercice : Discuter le cas des molécules linéaires, pour lesquelles selon Landau et Lifschitz “il n'est pas besoin de mettre en œuvre la théorie des groupes”. Montrer que pour une molécule linéaire de n atomes, il y a $n - 1$ fréquences propres *a priori* de multiplicité 1 et $(n - 2)$ autres, de multiplicité 2, [9] page 440. Exemple, acétylène C_2H_2 .

La théorie des groupes permet donc de déterminer assez facilement les multiplicités des modes propres de vibration. Une discussion plus poussée permet aussi de trouver les formes des coordonnées normales en termes des coordonnées internes, par projection sur les représentations irréductibles, voir [5] p 105.

Une autre discussion, de nature quantique et combinatoire, permet aussi de calculer les multiplicités des niveaux d'énergie de vibration, [9] p § 101. Si la fréquence propre ω_ρ vient avec la multiplicité f_ρ ($= \dim D^{(\rho)}$ ou $= 2 \dim D^{(\rho)}$ si $D^{(\rho)}$ est complexe), et si ces f_ρ modes sont dans des états de vibration $v_{\rho,i} \in \mathbb{N}$, $i = 1, \dots, f_\rho$, et $v_\rho = \sum_i v_{\rho,i}$, l'énergie totale de vibration est alors

$$E_{vib} = \hbar \sum_{\rho,i} \omega_\rho (v_{\rho,i} + \frac{1}{2}) = \hbar \sum_{\rho,i} \omega_\rho (v_\rho + \frac{f_\rho}{2}). \quad (5.53)$$

Selon un argument combinatoire classique, il y a $\frac{(v_\rho + f_\rho - 1)!}{v_\rho! (f_\rho - 1)!}$ façons d'obtenir le niveau $v_\rho + \frac{f_\rho}{2}$ avec les f_ρ oscillateurs de fréquence ω_ρ , et la multiplicité du niveau E_{vib} est donc

$$\prod_{\rho} \frac{(v_\rho + f_\rho - 1)!}{v_\rho! (f_\rho - 1)!},$$

le produit courant sur toutes les fréquences propres distinctes.

5.5.2 Vibration des cristaux

Des techniques similaires s'appliquent à l'étude des modes normaux de vibration dans les cristaux, ou phonons. Une complication non triviale par rapport au cas des molécules vient du fait que le groupe d'invariance du cristal, ou groupe d'espace, est infini. Sa réduction à un “groupe quotient” fini, appartenant à la liste des 32 groupes ponctuels mentionnés au chapitre 1, est une des techniques courantes. Nous renvoyons à la littérature, en particulier à [5], chap VI, § V, pour une introduction, et à D. L. Rousseau, R. P. Bauman, S. P. S. Porto, *Normal mode determination in crystals*, *Journal of Raman Spectroscopy*, Volume 10, (1981), 253-290, pour une discussion approfondie.

5.6 Exercices

1. ★ Démonstration des formules d'orthogonalité des $\mathcal{D}$.

Avec les notations du § 5.2.1, construire la matrice

$$V = \frac{1}{|G|} \sum_{g'} \mathcal{D}^{(\rho)}(g') M \mathcal{D}^{(\rho')\dagger}(g') \quad (5.54)$$

où M une matrice quelconque de dimension $n_\rho \times n_{\rho'}$. Montrer qu'elle satisfait

$$V \mathcal{D}^{(\rho')}(g) = \mathcal{D}^{(\rho)}(g) V \quad (5.55)$$

Lui appliquer alors le lemme de Schur dans les deux cas $\rho \neq \rho'$ et $\rho = \rho'$ et identifier la contribution de $M_{\beta\beta'}$ dans l'élément de matrice $V_{\alpha\alpha'}$ pour obtenir (5.22).

2. Groupes diédraux

On étudie les groupes diédraux D_p , d'ordre $2p$, définis au chapitre 1.

a) Groupe D_2 . Montrer que D_2 est un groupe abélien. Combien de représentations irréductibles (a priori complexes) a-t-il ? Les construire comme application de la propriété (5.35).

b) Groupes diédraux D_p . On considère les matrices

$$A = \begin{pmatrix} \epsilon_j & 0 \\ 0 & \epsilon_j^* \end{pmatrix} \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

où $\epsilon_j = \exp 2i\pi \frac{j}{p}$, et on représente les rotations autour de l'axe d'ordre p par $I, A, \dots, A^{p-1}$, et ces rotations composées avec la rotation autour de l'axe orthogonal par $B, BA, \dots, BA^{p-1}$. En calculant la norme du caractère de cette représentation, montrer qu'elle est irréductible pour $1 \leq j \leq p-1$, et $j \neq p/2$ si p est pair. Que peut-on dire de celle pour $j = p/2$ (p pair) ? Montrer qu'en se restreignant à $1 \leq j < p/2$, on a des représentations irréductibles inéquivalentes. En déduire que D_p admet pour p impair $(p-1)/2$ représentations irréductibles de dimension 2 et deux de dimension 1, et pour p pair $p/2 - 1$ représentations irréductibles de dimension 2 et quatre de dimension 1.

3. Montrer que les relations d'orthogonalité de la section 5.2.1 impliquent les formules suivantes pour un groupe fini :

$$\frac{1}{|G|} \sum_{g \in G} \chi^{(\rho)}(g \cdot g_1 \cdot g^{-1} \cdot g_2) = \frac{1}{n_\rho} \chi^{(\rho)}(g_1) \chi^{(\rho)}(g_2) ,$$

et

$$\frac{1}{|G|} \sum_{g \in G} \chi^{(\rho)}(g \cdot g_1) \chi^{(\sigma)}(g^{-1} \cdot g_2) = \frac{\delta_{\rho, \sigma}}{n_\rho} \chi^{(\rho)}(g_1 \cdot g_2) .$$

4. On considère le groupe constitué des éléments $\pm e, \pm i, \pm j$ et $\pm k$ satisfaisant les relations $i^2 = j^2 = k^2 = -e, i \cdot j = k$, etc par permutation cyclique. C'est le *groupe des quaternions* Q . Montrer qu'on peut trouver une représentation de dimension 2 de Q en termes des matrices de Pauli $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ et $\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. En déduire que Q a cinq représentations irréductibles et cinq classes. Que sont ces classes ? Quelles sont les valeurs possibles des caractères pour les classes de i, j ou k ? Dresser la table de caractères de Q .

On considère ensuite le groupe T engendré par les matrices 2×2 : $u = i\sigma_2$ et $v = \sigma_1$. Quel est son ordre, que sont ses classes ? Montrer que T et Q ne sont pas isomorphes. Dresser la table de caractères de T . Qu'en conclut-on ?

5. Représentations du groupe $U(1)$.

On cherche toutes les fonctions f dérivables satisfaisant

$$f(x)f(y) = f(x+y) . \quad (5.56)$$

- Montrer que f satisfait $f'(x) = kf(x)$ où k est une constante arbitraire et en déduire que $f(x) = \exp kx$.
- Appliquer ce résultat à la détermination des représentations complexes irréductibles du groupe $U(1)$.

6. Décomposition d'une représentation de $SO(3)$ en représentations irréductibles de son sous-groupe O .

On étudie la décomposition de la représentation de dimension 5 de $SO(3)$ en représentations irréductibles du groupe du cube O . On rappelle que $\chi_{j=2}(\theta) = \sin(5\theta/2)/\sin \theta/2$.

- a) Calculer la valeur que prend $\chi_{j=2}$ dans les différentes classes du groupe O .
- b) En utilisant la table des caractères de O construite en TD, déterminer combien de représentations irréductibles de O apparaissent dans cette décomposition.
- c) Déterminer quelles sont ces représentations.

Chapitre 6

Symétries en Mécanique Quantique

6.1 Théorème de Wigner

6.1.1 Théorème de Wigner.

Dans un système quantique, les états (purs) sont représentés par les vecteurs $|\psi\rangle$ normalisés d'un espace de Hilbert $\mathcal{H}$, ou plus précisément par ces vecteurs à une phase près, $|\psi\rangle \sim |\psi\rangle e^{i\alpha}$, ce qu'on appelle les *rayons*. Les “observables” A sont des opérateurs auto-adjoints sur $\mathcal{H}$, $A = A^\dagger$. Les quantités observables sont les $|\langle \phi|A|\psi\rangle|$. Supposons qu'il existe une transformation inversible (une “bijection”) g du système (états *et* observables) qui laissent inchangées ces quantités, c'est-à-dire

$$|\phi\rangle \rightarrow |^g\phi\rangle, |\psi\rangle \rightarrow |^g\psi\rangle, \quad A \rightarrow {}^gA \quad \text{telle que} \quad |\langle \phi|A|\psi\rangle| = |\langle {}^g\phi|{}^gA|^g\psi\rangle|.$$

On démontre alors le théorème suivant

Théorème de Wigner. *Si une bijection entre les rayons et les opérateurs auto-adjoints d'un espace de Hilbert $\mathcal{H}$ préserve les modules des “valeurs moyennes”*

$$|\langle \phi|A|\psi\rangle| = |\langle {}^g\phi|{}^gA|^g\psi\rangle|, \quad (6.1)$$

alors cette bijection est réalisée par un opérateur $U(g)$, linéaire ou antilinéaire, unitaire sur $\mathcal{H}$, et unique à une phase près, c'est-à-dire

$$|^g\psi\rangle = U(g)|\phi\rangle, \quad {}^gA = U(g)AU^\dagger(g); \quad U(g)U^\dagger(g) = U(g)^\dagger U(g) = I. \quad (6.2)$$

Rappelons d'abord ce qu'on entend par opérateur antilinéaire. Un tel opérateur satisfait

$$U(\lambda|\phi\rangle + \mu|\psi\rangle) = \lambda^*U|\phi\rangle + \mu^*U|\psi\rangle \quad (6.3)$$

et son adjoint est défini par

$$\langle \phi|U^\dagger|\psi\rangle = \langle U\phi|\psi\rangle^* = \langle \psi|U\phi\rangle, \quad (6.4)$$

de façon à être compatible avec la linéarité :

$$\langle \lambda \phi | U^\dagger | \psi \rangle = \lambda^* \langle \phi | U^\dagger | \psi \rangle . \quad (6.5)$$

S'il est en outre unitaire, on a

$$\langle \psi | \phi \rangle^* = \langle \phi | \psi \rangle = \langle \phi | U^\dagger U | \psi \rangle = \langle U \phi | U \psi \rangle^* , \quad (6.6)$$

donc $\langle U \phi | U \psi \rangle = \langle \psi | \phi \rangle$.

Nous admettrons ce théorème de Wigner, dont la preuve est un peu laborieuse : voir par exemple [12] tome 2, p. 540, ou [16].

Une fois déterminée la transformation des états par l'opérateur $U(g)$, on connaît celle des observables ${}^g A = U(g) A U^\dagger(g)$ de façon à avoir

$$\begin{aligned} \langle {}^g \phi | {}^g A | {}^g \psi \rangle &= \langle U \phi | U A U^\dagger | U \psi \rangle \\ &= \langle \phi | U^\dagger U A U^\dagger U | \psi \rangle^\# \\ &= \langle \phi | A | \psi \rangle^\# \end{aligned} \quad (6.7)$$

avec $\# =$ rien ou $*$ selon que U est linéaire ou antilinéaire.

Le cas antilinéaire n'a pas qu'un intérêt académique. Il se rencontre dans l'étude du renversement du sens du temps, voir plus bas, § 6.2.3.

Les transformations d'un système quantique, c'est-à-dire les bijections du théorème de Wigner, forment un groupe G : si g_1 et g_2 sont de telles bijections, leur composition $g_1 g_2$ en est une aussi, ainsi que g_1^{-1} etc. Cette transformation $g_1 g_2$ est réalisée sur les états soit par $U(g_1)U(g_2)$ soit par $U(g_1 g_2)$ et le théorème de Wigner nous dit que ces deux expressions doivent coïncider à une phase près. Les opérateurs $U(g)$ qu'on va supposer linéaires jusqu'à nouvel ordre satisfont donc

$$U(g_1)U(g_2) = e^{i\zeta(g_1, g_2)} U(g_1 g_2) \quad (6.8)$$

et forment ce qu'on appelle une *représentation à une phase près* de G .

La mécanique quantique nous force donc à élargir un peu le cadre de la théorie des représentations. Mais on montre par exemple que les représentations à une phase près du groupe des rotations $SO(3)$ se ramènent à des représentations à un signe près : on a d'une part les vraies représentations, avec des représentations irréductibles décrites par un spin j entier ; et de l'autre les représentations irréductibles à un signe près de spin demi-entier. Nous y reviendrons au chapitre suivant.

6.1.2 Invariance d'un système quantique

Supposons maintenant que sous l'action d'un certain groupe de transformations G , le système est invariant, en ce sens que sa dynamique, contrôlée par son hamiltonien H , est inchangée. Supposons aussi pour simplifier que H est indépendant du temps. On va écrire

$$H = U(g) H U^\dagger(g) \quad (6.9)$$

ou encore

$$[H, U(g)] = 0 . \quad (6.10)$$

On définit donc une *invariance* (ou *symétrie*) d'un système quantique sous l'action d'un groupe G comme l'existence d'une représentation projective unitaire (linéaire ou antilinéaire) de ce groupe dans l'espace des états, qui commute avec l'hamiltonien.

On sait aussi que l'hamiltonien H est le générateur infinitésimal de l'évolution temporelle. Si à l'instant $t = 0$ le système est dans un état (pur) décrit par le vecteur $|\psi_0\rangle$, il est décrit à tout instant t par $|\psi(t)\rangle$ solution de $i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle$ avec la condition initiale $|\psi(0)\rangle = |\psi_0\rangle$, donc, avec l'hypothèse faite que H ne dépend pas du temps, $|\psi(t)\rangle = \exp(-iHt/\hbar) |\psi_0\rangle$. L'évolution temporelle des états est donc fournie par l'opérateur unitaire

$$\mathcal{U}(t) = \exp \frac{-i}{\hbar} Ht , \quad (6.11)$$

et l'hypothèse d'invariance (6.10) peut se redire sous la forme

$$[\mathcal{U}(t), U(g)] = 0 . \quad (6.12)$$

L'évolution temporelle commute avec l'action du groupe d'invariance.

- Ce formalisme est très général et couvre des situations très variées : les transformations laissant le système invariant peuvent être de natures très différentes, transformations géométriques –translations, rotations, réflexions, inversion, dilatation, etc–, permutations de composantes du système, “symétries internes” ou de jauge, ...

- L'hypothèse d'invariance implique l'existence de lois de conservation. En effet toute observable $\mathcal{F}$ fonction des $U(g)$ commute avec H , donc est une quantité conservée

$$i\hbar \frac{d\mathcal{F}(U(g))}{dt} = [\mathcal{F}(U(g)), H] = 0 \quad (6.13)$$

et chacune de ses valeurs propres est un “bon nombre quantique”, c'est-à-dire conservé : si le système appartient à un sous-espace propre de $\mathcal{F}$ au temps $t = 0$, il y demeure lors de son évolution dans le temps. Si G est un groupe de Lie, pour g une transformation infinitésimale et si T désignent les générateurs infinitésimaux dans la représentation considérée,

$$U(g) = I + i \delta\alpha^j T_j$$

(où on choisit les T hermitiens pour avoir U unitaire), les T_j sont des observables commutant avec H , donc des quantités conservées.

On peut montrer que ces opérateurs T_i qui réalisent dans la théorie quantique les opérations infinitésimales du groupe G forment d'un point de vue mathématique une “représentation de l'algèbre de Lie $\mathfrak{g}$ ”, c'est-à-dire qu'ils satisfont les relations de commutation

$$[T_i, T_j] = i C_{ij}^k T_k \quad (6.14)$$

de cette algèbre de Lie (avec un “ i ” parce qu'on a fait un choix d'opérateurs hermitiens). Le nombre maximal de ces opérateurs qu'on peut diagonaliser simultanément, donc des quantités conservées qu'on pourra fixer, dépend de la structure de $\mathfrak{g}$ et de ces relations de commutation.

Exemples.

- Groupe des translations d’espace ou de temps $\longrightarrow \vec{P}$, H impulsion et énergie sont conservés et simultanément observables ;
- groupe des rotations $\text{SO}(3) \longrightarrow \vec{J}$ moment cinétique conservé, mais $[J_i, J_j] = i\epsilon_{ijk}J_k$ et $[\vec{J}^2, J_i] = 0$, une seule composante J_i observable et conservée, en plus de $\vec{J}^2$ lui aussi conservé ; donc deux “bons nombres quantiques”, j et m ;
- sous-groupe des rotations $\text{SO}(2)$ autour de l’axe $Oz \longrightarrow$ composante J_z de $\vec{J}$ conservée.

Un autre exemple sur lequel nous revenons plus bas concerne l’inversion d’espace (la parité). Considérons l’opération P d’inversion d’espace, $\vec{x} \mapsto -\vec{x}$. Dans l’espace de Hilbert d’une théorie quantique, elle est représentée par un certain opérateur unitaire $U(P)$. Comme $U(P)^2 = I$, les valeurs propres de $U(P)$ sont ± 1 , et ses vecteurs propres, qui engendrent des représentations de dimension 1, sont indexés par cette valeur propre nommée *parité*. Si l’hamiltonien est invariant par P , c’est-à-dire commute avec $U(P)$, on peut diagonaliser simultanément H et $U(P)$, c’est-à-dire trouver des états stationnaires (états propres de H) de parité donnée. Les états propres d’énergie ont donc une parité fixée. Autre manifestation de cette commutation, si le système a été préparé à l’instant $t = 0$ dans un état de parité donnée, l’évolution temporelle le fera rester dans cette parité. C’est en ce sens que la parité est une quantité conservée, ou “un bon nombre quantique”. Par abus de langage, on donne souvent le nom de parité à la transformation P .

• Mais l’hypothèse d’invariance (6.10) faite plus haute a une autre conséquence, d’application fréquente et importante. L’espace de Hilbert des états $\mathcal{H}$ qui “porte une représentation” du groupe G peut être décomposé en représentations irréductibles, $\mathcal{H} = \oplus E^{(\rho)}$ et $U(g) = \oplus D^{(\rho)}$.

- Évaluons la relation de commutation $[U(g), H] = 0$ entre deux représentations irréductibles. Le lemme de Schur nous apprend alors que H n’a d’éléments de matrice non nuls qu’entre des représentations irréductibles équivalentes.
- Si toutes les représentations irréductibles inéquivalentes $D^{(\rho)}$ sont de multiplicité 1, dans chaque espace $E^{(\rho)}$, l’Hamiltonien est multiple de l’identité en vertu du Corollaire 1 du lemme de Schur, cf § 5.1.5. On a donc dans ce cas une information complète sur la nature du spectre : les espaces propres sont les $E^{(\rho)}$ et les multiplicités des valeurs propres de H sont égales à $\dim E^{(\rho)}$ ¹.
- Si certaines représentations $D^{(\rho)}$ apparaissent avec une multiplicité m_ρ supérieure à 1, le lemme de Schur n’interdit pas que H ait des éléments de matrice entre les espaces $E^{(\rho,i)}$, $i = 1, \dots, m_\rho$; il faut encore diagonaliser H dans la somme de ces espaces $\oplus_{i=1}^{m_\rho} E^{(\rho,i)}$, ce qui est tout de même plus simple que le problème de diagonalisation dans l’espace $\mathcal{H}$ de départ. On verra plus bas que le théorème de Wigner-Eckart permet de réduire encore la complexité de ce dernier calcul. La théorie des groupes va donc nous simplifier considérablement la tâche...

Exemple ([14], p 37). Considérons le problème d’un électron placé dans le champ de trois protons fixés aux trois sommets d’un triangle équilatéral. Le problème admet la symétrie du groupe (de rotations et réflexions) du triangle : il s’agit d’un groupe d’ordre 6, qui a trois

1. Il peut arriver que la multiplicité d’une valeur propre de H soit plus élevée, soit à cause de l’existence d’un groupe de symétrie plus grand que G , soit parce que certaines représentations viennent en paires complexes conjuguées, soit pour une autre raison et on parle alors de “dégénérescence accidentelle”, cf la note 6 du chap. 5.

représentations inéquivalentes de dimensions 1, 1 et 2 (on a bien $1^2 + 1^2 + 2^2 = 6$), cf discussion au § 5.5.1. Sur la seule base de la théorie des groupes, nous apprenons immédiatement et sans aucun calcul que le spectre d'énergie de l'électron est constitué de niveaux de multiplicité 1 ou 2.

- Levées de dégénérescence. Règles de sélection.

Considérons un système invariant par l'action d'un groupe G . Ses niveaux d'énergie ont été classés selon les représentations irréductibles de G . Supposons maintenant qu'une perturbation H' soit ajoutée à l'hamiltonien, qui a pour effet de briser le groupe G en un sous-groupe G' . Quelles levées de dégénérescence, quelles transitions entre états sont rendues possibles par l'introduction de H' ?

Si le groupe d'invariance est réduit de G à G' , les niveaux d'énergie se classent maintenant en représentations irréductibles de G' , et la première question revient à : comment une représentation irréductible de G se scinde-t-elle en représentations de G' ? Pour cette question aussi, les techniques utilisant les caractères que nous avons introduites au chapitre précédent (§ 5.4) s'avèrent utiles.

Un exemple d'intérêt physique est celui où G est le groupe $O(3)$ et G' le groupe ponctuel d'un cristal et où on étudie la levée de dégénérescence (partielle) des niveaux d'énergie d'un atome placé dans un cristal. On appelle *champ cristallin* cet effet. Voir le problème en fin de chapitre 7, § 7.6.2. Ou encore, le groupe de symétrie initial $O(3)$ peut être brisé par l'introduction d'un champ électrique $\vec{E}$ à symétrie cylindrique, qui va laisser le sous-groupe d'invariance $C_{\infty v}$, cf § 2.3.3.

6.2 Symétries discrètes en physique quantique : parité, renversement du temps, ...

6.2.1 Transformation d'un état quantique sous l'effet de la parité

Intéressons-nous d'abord à la transformation d'un état quantique sous l'effet de l'inversion/parité, $P : \vec{x} \mapsto -\vec{x}$. Selon le théorème de Wigner, cette transformation est réalisée par un opérateur unitaire $U(P)$. Comme la parité est "involutive" (c'est-à-dire de carré égal à la transformation identité), le groupe qu'elle engendre est le groupe $\mathbb{Z}_2$ (cf chap. 1, §1.2). Ses représentations irréductibles sont de dimension 1 par le lemme de Schur. On s'attend à ce que dans une représentation irréductible $U(P) = \eta_P$, avec $\eta_P = 1$ ou $= -1$ ².

Pour un état à une particule $|\phi\rangle$ de parité donnée, on écrit

$$|\phi'\rangle = U(P)|\phi\rangle = \eta_P|\phi\rangle \quad (6.15)$$

et sa fonction d'onde

$$\phi(\vec{x}) = \langle \vec{x} | \phi \rangle = \langle \vec{x} | U^\dagger(P) U(P) | \phi \rangle = \langle -\vec{x} | \phi' \rangle = \phi'(-\vec{x}) = \eta_P \phi(-\vec{x}) , \quad (6.16)$$

2. La possibilité d'avoir des représentations à une phase près élargit les possibilités à des représentations dans lesquelles $U(P) = e^{i\alpha}$, mais nous ne considérerons pas cette possibilité ici.

est une fonction paire ou impaire, selon le signe de la *parité intrinsèque* η_P de la particule considérée. Cette parité intrinsèque a un certain arbitraire et peut parfois être redéfinie, voir des exemples ci-dessous.

L'opérateur $U(P)$ ne commute pas avec l'opérateur impulsion $\vec{p}$ (mais anticommute!) : on ne peut trouver de vecteur propre commun à $U(P)$ et à $\vec{p}$. Agissant sur un état d'impulsion $\vec{p}$ et de spin s_z fixés, $U(P)$ renverse le sens de $\vec{p}$ mais n'affecte pas s_z (rappelons que classiquement, le moment cinétique est un *pseudo*-vecteur : il est invariant par inversion/parité). On écrit donc pour la transformation d'un état à une particule d'impulsion $\vec{p}$ et de spin s_z donnés

$$U(P)|\vec{p}, s_z\rangle = \eta_P |-\vec{p}, s_z\rangle. \quad (6.17)$$

Molécule diatomique

D'autres symétries $\mathbb{Z}_2$ que la parité peuvent être à l'œuvre : considérons ainsi les symétries d'une molécule diatomique. Soit Oz l'axe de symétrie de la molécule. Le groupe d'invariance est $C_{\infty v}$, cf. Chap. 1 et 2. L'invariance par rotation autour de Oz implique que la composante $\Lambda = |L_z|$ du moment angulaire est conservée et on note Σ, Π, Δ etc les états correspondant à $\Lambda = 0, 1, 2, \dots$. En raison de l'invariance de la molécule par réflexion dans tout plan passant par Δ , on peut distinguer les états pairs ou impairs dans cette réflexion : cela est noté par un indice supérieur $\pm$; enfin si les deux atomes de la molécule sont identiques, le groupe est $D_{\infty h}$ et on peut distinguer les états pairs (g) ou impairs (u) dans l'inversion des coordonnées électroniques. Un état sera finalement noté $^{2S+1}\Pi_u^+$, ou Σ_g^+ etc. Empiriquement le fondamental de la molécule a la symétrie maximale, c'est un état $^1\Sigma^+$ ou $^1\Sigma_g^+$. Voir [9, 14].

Règles de sélection relatives à la parité

La conservation de la parité a d'importantes conséquences par exemple en physique atomique, sur les transitions possibles ou interdites entre niveaux sous l'effet d'un rayonnement électromagnétique, en physique nucléaire ou en physique des particules sur les désintégrations possibles, etc. Illustrons cela par quelques exemples en anticipant un peu sur la discussion du chapitre suivant sur le rôle des rotations en physique quantique.

Règles de sélection en physique atomique. Transitions dipolaires.

Le moment dipolaire électrique $\vec{d}$ d'une distribution de charges est défini classiquement par $\vec{d} = \sum_i q_i \vec{r}_i$. C'est un vecteur, donc impair par parité. En mécanique quantique, ce moment dipolaire est promu au statut d'opérateur et il garde son caractère vectoriel, $U(P)\vec{d}U^\dagger(P) = -\vec{d}$, avec les règles de sélection qui en découlent. Ainsi, si on évalue sa valeur moyenne entre deux états de parité fixée, $\langle \psi | \vec{d} | \phi \rangle$ est non nulle seulement si les parités des deux états sont opposées. En effet on écrit

$$\langle \psi | \vec{d} | \phi \rangle = \langle \psi U^\dagger(P) | (U(P)\vec{d}U^\dagger(P)) | U(P)\phi \rangle = -\eta_\psi \eta_\phi \langle \psi | \vec{d} | \phi \rangle \quad (6.18)$$

qui est donc nul si les parités η_ψ et η_ϕ des deux états sont identiques. Cela est important quand on étudie l'effet du couplage d'un système de charges à un champ électrique : ce couplage est représenté par un terme $\vec{d} \cdot \vec{E}$ dans l'hamiltonien. Au premier ordre de la théorie des perturbations, l'amplitude de transition entre deux états est donnée par $\langle \psi | \vec{d} | \phi \rangle$ et s'annule donc (à cet ordre) pour des états de même parité.

Quand le moment dipolaire s'annule, en raison des symétries de la distribution de charges, ou a une contribution nulle, comme on vient de le voir, on est amené à considérer les *moments multipolaires* d'ordre plus élevé, moment quadrupolaire, etc. Nous y reviendrons au chapitre 7, § 7.5.1.

Effets Stark et Zeeman

Effet de la symétrie miroir ou de l'inversion d'espace sur un champ électrique, un champ magnétique, cf chap. 2. Levée de dégénérescence des niveaux. Voir le cours de mécanique quantique.

Parité des noyaux

La parité d'un état de moment orbital l est $(-1)^l$ comme on sait (et comme on verra au chapitre suivant). Ainsi le deutéron (noyau de deutérium) qui est essentiellement dans l'état 3S_1 (c'est-à-dire $l = 0$, spin 1)³ a une parité $+1$. En général on peut assigner un spin et une parité donnés à tout noyau. Pour déterminer la parité, on fait appel soit à l'expérience soit à un modèle théorique comme le *modèle en couches* qui attache un moment angulaire l_i bien défini à chaque nucléon. La parité est alors $(-1)^{\sum_i l_i}$ (par convention, la parité intrinsèque d'un nucléon, proton ou neutron, est $+1$). Voir une illustration plus bas avec le noyau Li^7 .

Noyaux déformés ellipsoïdaux

Les noyaux peuvent avoir une forme sphérique, auquel cas le spectre est invariant par rotation, les niveaux d'énergie sont indépendants du moment cinétique et on n'observe pas de "spectre de rotation". C'est le cas des noyaux "légers" (jusqu'à $A \approx 150$), mais aussi des noyaux "magiques" possédant des nombres de protons ou de neutrons égaux à 2, 8, 20, 28, 50, 82, et 126, comme le prédit le modèle en couches. Un noyau comme le plomb ${}^{208}\text{Pb}$ qui a $Z = 82$ (et donc $N = 126$ neutrons) est dit "doublement magique", ce qui explique sa stabilité et sa sphéricité en dépit de sa masse élevée. Les noyaux lourds (actinides...) sont en général déformés, la situation la plus simple étant celle d'un ellipsoïde de révolution, soit allongé en forme de cigare, soit aplati en galette. C'est ce qui se produit pour les noyaux "pair-pair" comme ${}^{238}\text{U}$, qui a $Z = 92$. Dans ce cas, on conserve une invariance de rotation autour de l'axe de révolution, et on s'attend dans une image classique à observer un spectre de rotation avec des énergies de la forme $\bar{L}^2/2I$ où I est le moment d'inertie autour d'un axe orthogonal à l'axe de symétrie. Quantiquement cela devient $\hbar^2 l(l+1)/2I$ et l'état de moment angulaire l est dégéré $2l+1$ fois. La forme ellipsoïdale supposée est compatible avec une symétrie par inversion, donc avec la conservation de la parité $(-1)^l$. On s'attendrait donc à observer au dessus de l'état fondamental 0^+ une suite d'états $1^-, 2^+, 3^-, \dots$. On observe bien expérimentalement des énergies de transition $E_{l+2} - E_l \propto 2l+3$ compatibles avec ce modèle et avec la conservation de la parité, mais pour les seules valeurs de l pair. On montre en effet que seuls subsistent les états à l pair, donc la suite d'états pairs $0^+, 2^+, 4^+, \dots$. Voir [15] chap. VII.

Réactions $\text{Li}^7 + p$.

Dans la réaction $\text{Li}^7 + p \rightarrow \text{Be}^8 + \gamma$, on observe un pic dans la section efficace quand le proton a une énergie de 0,44 MeV : on parle de *résonance* et cela correspond à la formation d'un état excité Be^{8*} qui se désintègre ensuite en $\text{Be}^8 + \gamma$. Cette résonance ne se manifeste pas dans la réaction $\text{Li}^7 + p \rightarrow \alpha + \alpha$. En effet la parité de l'état final est $+1$, puisqu'elle s'identifie à l'échange des deux particules α , qui sont des bosons. Mais dans l'état initial, le noyau Li^7 avec 4 nucléons sur la couche $l = 0$ et 3 sur la couche $l = 1$ a une parité $(-1)^{\sum_i l_i} = -1$. La réaction de production du Be^{8*} a lieu à basse énergie, dans l'onde s, donc la parité de l'état Be^{8*} est -1 . Cela interdit à la réaction $\text{Li}^7 + p \rightarrow \text{Be}^{8*} \rightarrow \alpha + \alpha$ de se produire.

Mésons pseudoscalaires.

Le pion π est un "méson pseudoscalaire" venant en trois états de charge possibles π^+, π^0, π^- ; méson signifie particule de masse intermédiaire (ici $M_{\pi}c^2 \approx 140 \text{ MeV}$)⁴ par opposition aux "baryons", particules lourdes comme le proton ou le neutron ($Mc^2 \approx 1 \text{ GeV}$); "pseudoscalaire" signifie qu'il a un spin nul, mais une parité intrinsèque $\eta_{\pi} = -1$. Pour se convaincre de ce dernier fait, on étudie la réaction suivante, voir TD 7

$$\pi^- d \rightarrow n n .$$

Dans les tables du "Particle Data Group" http://pdg.lbl.gov/2013/listings/contents_listings.html qui collectent toutes les propriétés des particules connues, chaque particule sujette à l'interaction forte vient avec (entre autres) l'indication J^P qui donne son spin et sa parité.

Bien d'autres exemples pourraient être donnés du rôle de la parité en physique.

3. La notation classique en spectroscopie est $2S+1 L_J$.

4. On rappelle que l'électron-volt, eV, est l'unité d'énergie égale à $1,6 \cdot 10^{-19}$ Joule, $1 \text{ MeV} = 10^6 \text{ eV}$, $1 \text{ GeV} = 10^9 \text{ eV}$, 1 TeV (tera-eV) = 10^{12} eV , l'énergie du LHC au CERN est de 7 TeV par faisceau...

6.2.2 Brisure de la symétrie de parité en physique des particules

Dans les années 50, des expériences déconcertantes avec d'autres mésons pseudoscalaires, appelés mésons K ($M_K \approx 495 \text{ MeV}/c^2$), ont conduit T.D. Lee et C.N. Yang à remettre en cause le principe de conservation de la parité. Si la parité est bien conservée par les interactions nucléaires fortes, responsables de la cohésion des noyaux et des interactions entre "hadrons"⁵, elle ne l'est pas dans les interactions faibles comme dans la désintégration β

$$n \rightarrow p + e^- + \nu.$$

L'expérience paradoxale de 1956 portait sur deux types de désintégrations des mésons K, soit en 2 pions, soit en 3 pions. Compte tenu de la parité -1 du pion, voir ci-dessus, on semblait obtenir alors deux valeurs opposées pour la parité du K!

Lee et Yang ont alors fait l'hypothèse que l'interaction faible n'est pas invariante par l'inversion d'espace, et la parité n'est donc pas un bon nombre quantique dans ces interactions. Cette hypothèse considérée comme révolutionnaire à l'époque (la Nature distingue la gauche de la droite!) a été rapidement vérifiée par l'expérience historique de Mme Wu (1957) sur la désintégration β d'un noyau de Co

$$^{60}\text{Co} \rightarrow ^{60}\text{Ni} e^- \bar{\nu}.$$

En utilisant des noyaux de Co polarisés (par un champ magnétique intense, puis maintenus à très basse température pour éviter la dépolarisation par agitation thermique) et en observant l'asymétrie "avant-arrière" de l'électron émis par rapport au plan orthogonal à la polarisation, on vérifie l'hypothèse de Lee et Yang (voir TD 7).

Dans la théorie moderne qui unifie interactions faibles et électromagnétiques, dite théorie de Glashow–Salam–Weinberg, une violation de la parité dans les interactions faibles en induit une dans les phénomènes électromagnétiques, et donc dans certaines expériences de physique atomique. Cet effet très ténu et prédit au milieu des années 70 a été vérifié par des expériences très délicates (précision de l'ordre de 10^{-13} !!) impliquant des transitions entre raies du Césium [M.-A. Bouchiat et C. Bouchiat, au LKB de l'ENS : pour une revue du sujet par ces auteurs, voir Rep. Prog. Phys 60 (1997) 1351.]

6.2.3 Renversement du temps

En mécanique classique, l'opération de renversement du temps est simple à comprendre : on imagine qu'un film est tourné d'un processus physique ou naturel. Le processus obtenu par renversement du temps est ce qu'on observe en visualisant le film à l'envers. Pour des processus élémentaires, diffusion de particules, réflexion de lumière, phénomènes électro-magnétiques simples, etc, le processus obtenu par renversement du temps est aussi physiquement vraisemblable et possible que le processus initial. On dira que les lois microscopiques de la mécanique classique, de l'optique, de l'électromagnétisme, etc sont invariantes par renversement du temps. Pour des processus plus complexes, chocs destructifs, croissance de la matière vivante, et autres phénomènes dits irréversibles en thermodynamique par exemple, les choses sont différentes, et

5. Par définition, les hadrons sont les particules sensibles aux interactions fortes. Les baryons (proton et neutron) et leurs résonances, les mésons sont des hadrons. Les électrons, neutrinos et autres "leptons" ne le sont pas.

l'on entre dans les questions difficiles de la création d'irréversibilité à partir de lois microscopiques réversibles, “flèche du temps”, “démon de Maxwell” etc. Nous ne nous étendrons pas davantage sur ces questions.

La question qui nous occupe ici est la formulation du renversement du temps et son invariance dans le contexte de la mécanique quantique. Comme on l'a annoncé plus haut, ce problème offre la situation nouvelle où la transformation T est réalisée par un opérateur *antilinéaire* plutôt que linéaire.

En effet cette opération T laisse l'opérateur position $\vec{x}$ inchangé, mais change le signe des vitesses, donc de l'impulsion $\vec{p}$

$$\begin{aligned}\vec{x}' &= U(T)\vec{x}U^\dagger(T) = \vec{x} \\ \vec{p}' &= U(T)\vec{p}U^\dagger(T) = -\vec{p}.\end{aligned}\tag{6.19}$$

Les relations de commutation canoniques $[x_j, p_k] = i\hbar\delta_{jk}$ ne sont compatibles avec cette transformation que si $U(T)$ est antilinéaire

$$\begin{aligned}[x'_j, p'_k] &= [x_j, -p_k] = -[x_j, p_k] = -i\hbar\delta_{jk} \\ &= U(T)[x_j, p_k]U^\dagger(T) = U(T)i\hbar\delta_{jk}U^\dagger(T).\end{aligned}\tag{6.20}$$

Sur les conséquences de l'invariance par renversement du temps, telle la dégénérescence de Kramers voir [12], t 2, p576.

Depuis une quarantaine d'années, on dispose d'une indication indirecte que l'invariance par renversement du temps est brisée au niveau microscopique.

6.2.4 Brisure de la symétrie de renversement du temps en physique des particules

Dans le cadre d'une théorie qui unifie relativité restreinte et mécanique quantique, la *théorie quantique des champs*, on démontre un théorème dit CPT, selon lequel la composition de trois opérations discrètes, la parité P , le renversement du temps T et la *conjugaison de charge* C , opération qui consiste à changer toute particule en son anti-particule, est une symétrie fondamentale des lois de la Nature. (Comme pour tout “théorème” en physique –contrairement aux mathématiques–, la validité de ce théorème repose sur des hypothèses qui pourraient être remises en question un jour...) Or en 1964, l'étude du système des deux particules $K^0, \bar{K}^0$ conjuguées l'une de l'autre, a révélé une brisure faible mais non nulle de l'invariance par CP , le produit de deux de ces trois opérations. Si le théorème CPT est bien vrai, le système $K^0, \bar{K}^0$ fournit donc une preuve indirecte de la violation de l'invariance par renversement du temps.

Le système $K^0, \bar{K}^0$: quelques indications.

Considérons l'opération C de “conjugaison de charge”, qui associe à toute particule son antiparticule (de charge opposée, si cette charge est non nulle). Cette opération C est représentée par un opérateur $U(C)$ qui est involutif, $U(C)^2 = I$, donc de valeurs propres ± 1 . En fait, $U(C)$ pas plus que $U(P)$ n'est une symétrie des interactions faibles. Par exemple, les neutrinos, leptons de masse ≈ 0 et de charge électrique nulle, sensibles

aux seules interactions faibles, sont observés dans un état d’“hélicité gauche”⁶, tandis que leurs antiparticules, antineutrinos, possèdent l’hélicité droite. L’existence de ces ν et $\bar{\nu}$ est donc une manifestation de la brisure de P et de C , mais est compatible avec une symétrie par le produit PC . On suppose donc dans un premier temps que le produit $U(CP) := U(C)U(P)$ est une symétrie de toutes les interactions. On peut alors construire des états propres, pairs ou impairs, de $U(CP)$. Ainsi pour le système des deux mésons K neutres, $K^0, \bar{K}^0$, de parité -1 , ces états propres sont

$$K_1^0 = \frac{1}{\sqrt{2}}(K^0 - \bar{K}^0) \quad K_2^0 = \frac{1}{\sqrt{2}}(K^0 + \bar{K}^0) \quad (6.21)$$

avec η_{CP} égal à $+1$ et -1 , respectivement. Le mode de désintégration le plus fréquent de ces kaons neutres est en deux pions. Mais la désintégration en deux pions n’est possible que pour K_1^0 , l’état à $CP = +1$. En effet l’état final $\pi^0\pi^0$ ou celui $\pi^+\pi^-$ à $l = 0$ (par conservation du moment cinétique, les pions et kaons étant de spin nul) et parité $+1$, il est pair par $U(C)$, donc $CP = 1$. Sous l’hypothèse d’invariance par CP , l’état K_2^0 ne peut se désintégrer en deux pions, mais seulement en 3 pions, ou $\pi^+e^-\nu_e$, etc. Or dans une expérience de 1964, J. W. Cronin, V. L. Fitch, (tous deux prix Nobel 1980), J. H. Christenson et R. Turlay ont observé des désintégrations $K_2^0 \rightarrow 2\pi$. La conclusion est que CP n’est pas une symétrie, ni donc non plus T .

Depuis l’expérience sur les $K^0, \bar{K}^0$, d’autres manifestations de la brisure (ou “violation”) de CP en physique des particules ont été découvertes. L’origine de la brisure de CP demeure un sujet mal compris. Ses implications cosmologiques –pourquoi l’Univers semble-t-il fait de matière plutôt que d’antimatière?– sont aussi l’objet d’intenses spéculations. . .

6.3 Symétries de translation, théorème de Bloch

Revenons au cas d’un électron dans le potentiel périodique créé par les ions d’un cristal, cf chapitre 2, § 4. L’hamiltonien est supposé invariant par les translations $\vec{t}$ du réseau $\mathcal{R}$, représentées par des opérateurs unitaires $U(\vec{t})$,

$$U(\vec{t})\psi(x) = \psi(\vec{x} + \vec{t}) . \quad (6.22)$$

Il s’agit d’un groupe abélien, ses représentations irréductibles complexes sont de dimension 1 et de forme exponentielle (Corollaire 2 du lemme de Schur et § 5.2.3). Notons $\psi_{\vec{k}}(\vec{x})$ une fonction se transformant selon une telle représentation irréductible indexée par $\vec{k}$,

$$\forall \vec{t} \in \mathcal{R} \quad U(\vec{t})\psi_{\vec{k}}(\vec{x}) = e^{i\vec{k} \cdot \vec{t}} \psi_{\vec{k}}(\vec{x}) . \quad (6.23)$$

Pour une telle fonction d’onde $\psi_{\vec{k}}$, dite *fonction d’onde de Bloch*, on a donc

$$\psi_{\vec{k}}(\vec{x} + \vec{t}) = e^{i\vec{k} \cdot \vec{t}} \psi_{\vec{k}}(\vec{x}) . \quad (6.24)$$

De façon équivalente, (6.24) implique que le rapport $u_{\vec{k}}(\vec{x}) := \psi_{\vec{k}}(\vec{x})/e^{i\vec{k} \cdot \vec{x}}$ est périodique, $u_{\vec{k}}(\vec{x} + \vec{t}) = u_{\vec{k}}(\vec{x})$, et le théorème de Bloch affirme donc que l’hamiltonien admet un système complet de fonctions propres satisfaisant

$$\psi_{\vec{k}}(\vec{x}) = e^{i\vec{k} \cdot \vec{x}} u_{\vec{k}}(\vec{x}) \quad (6.25)$$

6. ce qui signifie que leur spin $\frac{1}{2}$ est colinéaire et opposé à leur impulsion

avec $u_{\vec{k}}(\vec{x})$ une fonction arbitraire périodique sur le réseau $\mathcal{R}$. Selon l'observation faite au chapitre 2, les vecteurs $\vec{k}$ qui indexent les représentations irréductibles du groupe du réseau et donc les fonctions de Bloch, peuvent être choisis dans la 1ère zone de Brillouin. Nous admettrons que toute fonction d'onde peut se développer sur ces fonctions d'onde de Bloch.

On se reportera pour une discussion des applications de cet important théorème de Bloch au cours de Physique de la matière condensée ou à [1].

6.4 Application à la physique moléculaire

Au chapitre précédent nous avons examiné quelles conséquences l'existence d'une symétrie de la molécule a sur les niveaux d'énergie de vibration des noyaux. Examinons maintenant ce qu'il en est sur les niveaux d'énergie électroniques. Dans une molécule polyatomique, les niveaux d'énergie électroniques sont donnés par la solution de l'équation de Schrödinger des électrons dans le potentiel créé par les noyaux (et les autres électrons). Nous supposons à nouveau que la configuration d'équilibre de ces noyaux a une certaine symétrie.

Le traitement du problème utilise plusieurs approximations de nature différente. On néglige le mouvement des noyaux et on ne considère que celui des N électrons, et on néglige les interactions mutuelles et les spins de ces derniers. On suppose encore que seuls les électrons des couches périphériques participent à la liaison moléculaire, et pour simplifier l'écriture, on se bornera au cas d'un seul électron par noyau. Dans la molécule, chacun de ces électrons est supposé décrit par une fonction d'onde indépendante du temps, et on fait l'hypothèse qu'une bonne approximation de cette fonction d'onde est obtenue par combinaison linéaire des fonctions d'onde supposées connues de l'électron au voisinage du i -ème noyau. Autrement dit, soit $v(\vec{r} - \vec{r}_i)$ le potentiel créé par le i -ème noyau placé en $\vec{r}_i$ sur un électron en $\vec{r}$. (Pour simplifier les notations, on suppose les N noyaux identiques). Soit $f_i(\vec{r}) = \psi(\vec{r} - \vec{r}_i)$ une fonction d'onde normalisée d'un électron au voisinage du i -ème noyau, $i = 1, \dots, N$, solution de

$$\left(-\hbar^2 \frac{\vec{\nabla}^2}{2m} + v(\vec{r} - \vec{r}_i) - e_0 \right) \psi(\vec{r} - \vec{r}_i) = 0 . \quad (6.26)$$

On considère alors l'espace E des combinaisons linéaires des f_i :

$$f(\vec{r}) = \sum_{i=1}^N c_i f_i(\vec{r}) \quad (6.27)$$

et on cherche une bonne approximation de la fonction d'onde dans cet espace (c'est la méthode LCAO, *linear combination of atomic orbitals*). L'hamiltonien auquel est soumis chaque électron s'écrit

$$H = \left(-\hbar^2 \frac{\vec{\nabla}^2}{2m} + \sum_i v(\vec{r} - \vec{r}_i) \right) . \quad (6.28)$$

Les coefficients c_i sont alors déterminés en cherchant les minima de $\frac{\langle f|H|f \rangle}{\langle f|f \rangle}$. En négligeant les recouvrements des fonctions d'onde f_i et f_j si $i \neq j$, on a $\langle f|f \rangle \approx \sum_i |c_i|^2$ et $\langle f|H|f \rangle = \sum_{i,j} c_i^* c_j h_{ij}$, où h est la matrice $N \times N$: $h_{ij} = \langle f_i|H|f_j \rangle$. Le problème de minimisation revient à résoudre $\sum_j h_{ij} c_j = \mathcal{E} c_i$, c'est-à-dire à diagonaliser la matrice h . C'est ce problème qui peut être simplifié par la théorie des groupes si la configuration $\{\vec{r}_i\}$ des noyaux admet des propriétés de symétrie.

Soit g un élément d'un groupe de symétrie G fini laissant invariante (à une permutation près) la configuration des noyaux $\vec{r}_i$. Le groupe G est donc un groupe de symétrie de l'hamiltonien. On peut faire agir G sur les fonctions d'onde de l'espace E selon

$$D(g)f(\vec{r}) = f(g^{-1}\vec{r}) . \quad (6.29)$$

L'espace E porte donc une représentation D de G ,

$$D(g)|f_i\rangle = \sum_j |f_j\rangle \mathcal{D}_{ji}(g) \quad (6.30)$$

qu'on connaît explicitement une fois connues les fonctions f_i . L'invariance de l'hamiltonien H sous l'action de G se traduit dans l'espace E par la commutation $[D, h] = 0$. On décompose alors E en sous-espaces invariants se transformant par des représentations irréductibles, $D = \oplus m_\rho D^{(\rho)}$. Le lemme de Schur nous dit que h n'a pas d'élément de matrice entre des états relatifs à des représentations irréductibles non équivalentes et que dans chaque représentation irréductible $D^{(\rho)}$ de multiplicité $m_\rho = 1$ et de dimension n_ρ , h est un multiple de l'identité : sa valeur propre, c'est-à-dire le niveau d'énergie dans notre approximation LCAO, est connue sans qu'il soit besoin de procéder à une diagonalisation, c'est simplement $\langle f_\rho | H | f_\rho \rangle / \langle f_\rho | f_\rho \rangle$, et elle vient avec une multiplicité (une "dégénérescence") égale à n_ρ . En revanche, si une représentation irréductible apparaît avec une multiplicité supérieure à 1, h peut avoir des éléments de matrice entre des composantes relatives aux différentes copies de cette représentation. Dans ce cas, il reste à diagonaliser h dans ces sous-espaces invariants. On peut utiliser pour ce faire le théorème dit de Wigner-Eckart (cf chap. suivant) ; si l'indice i désigne les différentes copies de la représentation ρ , $i = 1, \dots, m_\rho$

$$\langle \rho \alpha i | H | \rho \alpha' i' \rangle = \delta_{\alpha \alpha'} \langle \rho i | H | \rho i' \rangle \quad (6.31)$$

et il suffit de diagonaliser la matrice réduite de taille $m_\rho \times m_\rho$. Pour résumer, si on a su construire les combinaisons de f_i se transformant par les représentations irréductibles du groupe G , on a ramené la détermination des niveaux à la diagonalisation de matrices de dimension $m_\rho \times m_\rho$.

Illustrons cette méthode dans le cas simple de la molécule de benzène C_6H_6 dans laquelle les atomes de carbone forment un hexagone régulier. Le groupe de symétrie est donc le groupe noté traditionnellement D_{6h} d'ordre 24, produit direct du groupe diédral D_6 par le groupe engendré par la réflexion par rapport au plan de la molécule. Il va suffire en fait de discuter les conséquences de l'invariance par D_6 . On ne s'intéresse qu'aux 6 électrons mobiles dits π ; leurs fonctions d'onde atomiques ψ sont telles que par action du groupe

$$f_i(\vec{r}) = \psi(\vec{r} - \vec{r}_i) \rightarrow D(g)f_i(\vec{r}) = f_i(g^{-1}\vec{r}) = \sum_j f_j(\vec{r}) D_{ji}(g) = f_{i'}(\vec{r}) \quad (6.32)$$

si $g\vec{r}_i = \vec{r}_{i'}$. (En général, il pourrait apparaître une phase entre Df_i et $f_{i'}$, ou s'il y a plusieurs électrons sur la même orbitale autour du i -ème noyau, une combinaison linéaire de leurs fonctions d'onde). Désignons de 1 à 6 les sommets de l'hexagone ; le groupe D_6 est engendré par la rotation R d'angle $\pi/3$ qui permute $1 \rightarrow 2, \dots, 6 \rightarrow 1$ et la rotation de π autour de l'axe 1-4 : $1 \rightarrow 1, 2 \leftrightarrow 6, 3 \leftrightarrow 5, 4 \rightarrow 4$. L'équation (6.32) nous apprend comment ces deux transformations sont représentées sur les fonctions d'onde $f_i, i = 1, \dots, 6$:

$$\mathcal{R} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad \mathcal{S} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (6.33)$$

Le groupe D_6 admet deux représentations de dimension 2 et quatre de dimensions 1 (cf exercice 2 du chap. 5). La représentation portée par l'espace E des combinaisons linéaires des f_i se décompose en deux représentations de dimension 1 et deux de dimension 2 : on pourrait utiliser les méthodes générales de calcul de ces décompositions, une fois dressée la table des caractères de D_6 ; en fait il est simple de vérifier que les combinaisons suivantes se transforment selon de telles représentations

$$\begin{aligned} D^{(1)} \quad F_1 &= (f_1 + f_2 + f_3 + f_4 + f_5 + f_6) \\ D^{(2)} \quad \begin{cases} F_2 &= (2f_1 + f_2 - f_3 - 2f_4 - f_5 + f_6) \\ F_3 &= (f_1 + 2f_2 + f_3 - f_4 - 2f_5 - f_6) \end{cases} \\ D^{(3)} \quad \begin{cases} F_4 &= (2f_1 - f_2 - f_3 + 2f_4 - f_5 - f_6) \\ F_5 &= (f_1 - 2f_2 + f_3 + f_4 - 2f_5 + f_6) \end{cases} \\ D^{(4)} \quad F_6 &= (f_1 - f_2 + f_3 - f_4 + f_5 - f_6) \end{aligned} \quad (6.34)$$

Ces représentations étant inéquivalentes fournissent des états propres d'énergie. Pour les calculer supposons comme précédemment que les recouvrements des fonctions f_i et f_j , $i \neq j$, sont négligeables, et que les seuls éléments de matrice de H sont entre fonctions f_i sur le même atome ou sur des atomes voisins : $\langle f_i | H | f_i \rangle = \mathcal{E}_0$ et $\langle f_i | H | f_j \rangle = \delta$ si i et j sont voisins. On trouve alors

$$\mathcal{E}^{(1)} : \mathcal{E}^{(2)} : \mathcal{E}^{(3)} : \mathcal{E}^{(4)} = (\mathcal{E}_0 + 2\delta) : (\mathcal{E}_0 + \delta) : (\mathcal{E}_0 - \delta) : (\mathcal{E}_0 - 2\delta) .$$

Comme $\delta < 0$, cela montre que l'énergie croît selon

$$\mathcal{E}^{(1)} < \mathcal{E}^{(2)} < \mathcal{E}^{(3)} < \mathcal{E}^{(4)} .$$

En utilisant le fait que chaque niveau peut accueillir deux électrons de spins opposés, l'état fondamental de la molécule est donc obtenu en plaçant les 6 électrons π sur les niveaux (1) et (2).

6.5 Symétries de la physique des particules

Le “modèle standard” (MS) de la physique des particules contemporaine a été élaboré sur une période de 25 ans environ, suite à une série de découvertes expérimentales et de percées théoriques. Il est maintenant très largement confirmé par l'expérience, et la toute récente découverte au LHC du *boson de Higgs* vient parachever l'édifice. Ce qui suit est un bref survol de quelques concepts-clés, du point de vue des symétries. Il ne constitue ni une revue historique ni un exposé systématique du “MS”.

Des groupes de symétrie interviennent à trois niveaux et de trois façons différentes dans le MS :

1. les interactions fortes sont sujettes à une symétrie exacte, dite symétrie de couleur, et la théorie qui les décrit, la *chromodynamique quantique*, (QCD), est une théorie de jauge de groupe $SU(3)$, le groupe des matrices unitaires 3×3 . Les particules sensibles aux interactions fortes, les hadrons, sont des états liés des constituants fondamentaux, les quarks et les gluons. Ces derniers sont associés à des représentations du groupe de jauge $SU(3)$, respectivement de dimension 3 pour les quarks et 8 pour les gluons. Un fait expérimental mal compris encore par les théoriciens est que les seuls hadrons observés portent la représentation identité de $SU(3)$. C'est ce qu'on appelle le *confinement des quarks et des gluons* (ou de la couleur). Pour expliquer ce phénomène, il est fondamental que le groupe de jauge $SU(3)$ soit une symétrie absolue de la théorie ;
2. les interactions faibles et électromagnétiques sont unifiées au sein d'un “secteur” du MS. Là encore une théorie de jauge est à l'œuvre, basée cette fois sur le groupe $SU(2) \times U(1)$. Les particules qui sont soumises aux interactions électro-faibles (quarks, leptons, bosons de jauge mais pas les gluons de la QCD), sont classées dans des représentations du groupe $SU(2) \times U(1)$. Au contraire de la QCD, la symétrie de jauge $SU(2) \times U(1)$ est brisée spontanément selon un mécanisme qui généralise ce qu'on a vu au chap. 3. Cela est fondamental pour assurer que toutes les particules observées (sauf le photon) ont une masse. La manière la plus simple de déclencher le mécanisme de brisure spontanée (en termes simples, de créer dans la théorie un potentiel en chapeau mexicain comme au chap. 3) est d'introduire un multiplet de champs de bosons scalaires (mécanisme Brout–Englert–Higgs ou BEH, voir Chap. 3), dont seul subsiste finalement une particule de spin

	mass →	charge →	spin →
QUARKS			
u (up)	$\approx 2.3 \text{ MeV}/c^2$	$2/3$	$1/2$
c (charm)	$\approx 1.275 \text{ GeV}/c^2$	$2/3$	$1/2$
t (top)	$\approx 173.07 \text{ GeV}/c^2$	$2/3$	$1/2$
d (down)	$\approx 4.8 \text{ MeV}/c^2$	$-1/3$	$1/2$
s (strange)	$\approx 95 \text{ MeV}/c^2$	$-1/3$	$1/2$
b (bottom)	$\approx 4.18 \text{ GeV}/c^2$	$-1/3$	$1/2$
LEPTONS			
e (electron)	$0.511 \text{ MeV}/c^2$	-1	$1/2$
μ (muon)	$105.7 \text{ MeV}/c^2$	-1	$1/2$
τ (tau)	$1.777 \text{ GeV}/c^2$	-1	$1/2$
ν_e (electron neutrino)	$< 2.2 \text{ eV}/c^2$	0	$1/2$
ν_μ (muon neutrino)	$< 0.17 \text{ MeV}/c^2$	0	$1/2$
ν_τ (tau neutrino)	$< 15.5 \text{ MeV}/c^2$	0	$1/2$
GAUGE BOSONS			
g (gluon)	0	0	1
γ (photon)	0	0	1
Z boson	$91.2 \text{ GeV}/c^2$	0	1
W boson	$80.4 \text{ GeV}/c^2$	± 1	1
H (Higgs boson)	$\approx 126 \text{ GeV}/c^2$	0	0

FIGURE 6.1 – Petit tableau des particules fondamentales du MS Source Wikipedia

0, dite boson de Higgs. Contrairement à ce qui a été dit au chap. 3, la brisure spontanée d’une symétrie continue *dans une théorie de jauge* n’engendre pas de particules de masse nulle mais au contraire est la cause de la génération de masses. Finalement de la symétrie $SU(2) \times U(1)$ initiale ne subsiste qu’une symétrie de jauge exacte par un groupe $U(1)$ (un mélange du facteur $U(1)$ de $SU(2) \times U(1)$ et d’un sous-groupe $U(1)$ de $SU(2)$), qui est le groupe de jauge familier de l’électrodynamique ;

- les paires de quarks et celles de leptons introduits dans cette construction s’organisent en plusieurs familles ou “générations”. On connaît actuellement trois familles et on estime que l’histoire s’arrête là. Les particules de ces familles manifestent des symétries **approchées**. C’est ainsi que si on ne considère que les quarks u and d de la première famille et les hadrons qui en sont des états liés (proton, neutron, pions), une symétrie $SU(2)$ dite d’*isospin* (voir fin du chap. 7) apparaît. Cette symétrie est approximative, elle est brisée par les interactions électromagnétiques. Si on incorpore aussi un ou deux des quarks de la génération suivante (s et c), on trouve des traces d’une symétrie plus sévèrement brisée, respectivement $SU(3)$ ou $SU(4)$. Ces “symétries de saveur” ne doivent pas être confondues avec les précédentes : les groupes $SU(2)$, $SU(3)$ etc de saveur ne sont **pas** les groupes de jauge rencontrés précédemment.

Au total on voit que les trois manifestations possibles de symétries apparaissent dans le MS : symétrie de jauge exacte (symétries $SU(3)$ de couleur et $U(1)$ de l’électrodynamique) ; symétrie $SU(2) \times U(1)$ brisée spontanément ; symétries de saveur brisées “explicitement” par des termes d’interaction dans le hamiltonien de la théorie. Pour résoudre quelques problèmes résiduels du MS, certains ont proposé des extensions *supersymétriques* du modèle standard. La supersymétrie, une nouvelle forme de symétrie, a la particularité d’introduire des degrés de liberté fermioniques dans le groupe de symétrie, avec des générateurs infinitésimaux satisfaisant maintenant des relations d’*anticommutation*, etc. Mais cela est une autre histoire...

Chapitre 7

Rotations à trois dimensions. Le groupe $SO(3)$

7.1 Rotations de $\mathbb{R}^3$, le groupe $SO(3)$, ses générateurs infinitésimaux

7.1.1 Le groupe $SO(3)$, groupe à trois paramètres

Dans l'espace euclidien à trois dimensions, toute rotation est définie par l'angle ψ de rotation autour d'un axe de vecteur directeur unitaire $\hat{n}$, et les rotations associées à $(\hat{n}, \psi)$ et à $(-\hat{n}, -\psi)$ sont identiques. On notera $R_{\hat{n}}(\psi)$ cette rotation. De façon très explicite, on écrit $\vec{x} = \vec{x}_{\parallel} + \vec{x}_{\perp} = (\vec{x} \cdot \vec{n})\vec{n} + (\vec{x} - (\vec{x} \cdot \vec{n})\vec{n})$ et $\vec{x}' = \vec{x}_{\parallel} + \cos \psi \vec{x}_{\perp} + \sin \psi \vec{n} \wedge \vec{x}_{\perp}$, d'où la formule d'O. Rodrigues

$$\vec{x}' = R_{\hat{n}}(\psi)\vec{x} = \cos \psi \vec{x} + (1 - \cos \psi)(\vec{x} \cdot \vec{n})\vec{n} + \sin \psi (\vec{n} \wedge \vec{x}) . \quad (7.1)$$

Le fait qu'une rotation de $\mathbb{R}^n$ laisse un axe invariant est vrai en toute dimension n *impair*. Soit R une matrice de $SO(n)$, $R.R^T = I$, $\det R = 1$. Alors on peut écrire $\det(R - I) = \det R(I - R^T) = (1)^n \det R \det(R - I) = -\det(R - I)$ si n est impair, donc $\det(R - I) = 0$, R a un vecteur propre de valeur propre 1, cqfd.

Comme un vecteur unitaire $\hat{n}$ dans $\mathbb{R}^3$ dépend de deux paramètres, par exemple l'angle θ qu'il fait avec l'axe Oz et l'angle ϕ que fait sa projection dans le plan Ox, Oy avec l'axe Ox (voir figure 1) un élément de $SO(3)$ est paramétrisé par 3 variables continues. On prendra ainsi

$$0 \leq \theta \leq \pi, \quad 0 \leq \phi < 2\pi, \quad 0 \leq \psi \leq \pi . \quad (7.2)$$

$SO(3)$ est donc un groupe de dimension 3. Pour la rotation d'axe $\hat{n}$ colinéaire à l'axe Oz , on a la matrice

$$\mathcal{R}_z(\psi) = \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (7.3)$$

tandis qu'autour des axes Ox et Oy

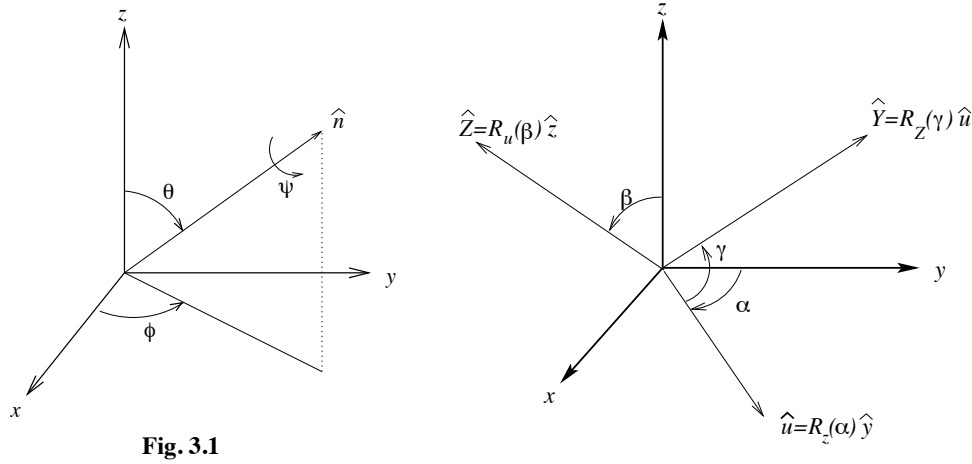


Fig. 3.1

FIGURE 7.1 – Deux paramétrisations des rotations : par le vecteur $\hat{n}$ et l'angle ψ ; par les angles d'Euler

$$\mathcal{R}_x(\psi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & -\sin \psi \\ 0 & \sin \psi & \cos \psi \end{pmatrix} \quad \mathcal{R}_y(\psi) = \begin{pmatrix} \cos \psi & 0 & \sin \psi \\ 0 & 1 & 0 \\ -\sin \psi & 0 & \cos \psi \end{pmatrix}. \quad (7.4)$$

Une relation que nous allons abondamment utiliser est que

$$RR_{\hat{n}}(\psi)R^{-1} = R_{\hat{n}'}(\psi) \quad (7.5)$$

où $\hat{n}'$ est le transformé de $\hat{n}$ par la rotation R , $\hat{n}' = R\hat{n}$. Exercice : la vérifier, soit en examinant l'action des deux membres sur le vecteur $\hat{n}'$, cf TD 5, soit en utilisant (7.1). Inversement toute rotation d'angle ψ autour d'un vecteur $\hat{n}'$ peut se mettre sous la forme (7.5) : dans le langage introduit au chap. 5 (§ 5.1.2), on dira que les “classes de conjugaison” du groupe $SO(3)$ sont caractérisées par l'angle ψ .

Une autre description fait appel aux *angles d'Euler* : étant donné un repère orthonormé (Ox, Oy, Oz) , toute rotation peut être considérée comme résultant de la composition d'une rotation d'angle α autour de Oz qui amène Oy en Ou , suivie d'une rotation d'angle β autour de Ou amenant Oz en OZ , et enfin d'une rotation d'angle γ autour de OZ (voir figure 2). On prend donc $0 \leq \alpha \leq 2\pi$, $0 \leq \beta \leq \pi$, $0 \leq \gamma \leq 2\pi$ et on écrit

$$R(\alpha, \beta, \gamma) = R_Z(\gamma)R_u(\beta)R_z(\alpha) \quad (7.6)$$

mais selon (7.5)

$$R_Z(\gamma) = R_u(\beta)R_z(\gamma)R_u^{-1}(\beta) \quad \text{et} \quad R_u(\beta) = R_z(\alpha)R_y(\beta)R_z^{-1}(\alpha)$$

d'où en reportant dans (7.6)

$$R(\alpha, \beta, \gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma). \quad (7.7)$$

où on a utilisé le fait que $R_z(\alpha)R_z(\gamma)R_z^{-1}(\alpha) = R_z(\gamma)$ car les rotations autour d'un même axe commutent (elles forment un sous-groupe abélien, isomorphe à $\text{SO}(2)$). Cette méthode permet de comparer les deux paramétrisations précédentes et d'en tirer les relations entre (θ, ϕ, ψ) et (α, β, γ) .

Exercice : En utilisant (7.5), écrire la matrice R qui amène le vecteur unitaire $\hat{z}$ porté par Oz sur le vecteur unitaire $\hat{n}$, puis l'expression de $R_{\hat{n}}(\psi)$ en termes de R_y et R_z , et en déduire les relations entre θ, ϕ, ψ et les angles d'Euler.

7.1.2 Générateurs infinitésimaux de $\text{SO}(3)$

Comme on l'a vu au chapitre 4, dans un groupe de Lie comme $\text{SO}(3)$, une étude des éléments au voisinage de l'identité, ou des générateurs infinitésimaux, suffit à reconstruire tout le groupe. Étudions donc ce que sont les générateurs infinitésimaux de $\text{SO}(3)$.

Pour un angle $d\psi$ très petit (infinitésimal), $R_{\hat{n}}(d\psi)$ diffère de l'opérateur identité par un terme d'ordre $d\psi$, coefficient d'un opérateur appelé *générateur infinitésimal*. Dans ce chapitre, nous suivons l'usage en écrivant les générateurs infinitésimaux des rotations comme des opérateurs hermitiens $J = J^\dagger$. Ainsi on écrit au premier ordre en $d\psi$

$$R_{\hat{n}}(d\psi) = (I - id\psi J_{\hat{n}}) \quad (7.8)$$

où $J_{\hat{n}}$ est le générateur de ces rotations¹, une matrice hermitique 3×3 , de façon à garantir l'unitarité (ou simplement comme ici, l'orthogonalité) de $R_{\hat{n}} : R_{\hat{n}} \cdot R_{\hat{n}}^\dagger = I$.

Pour trouver l'expression explicite de ces générateurs infinitésimaux, il est utile de les considérer soit comme des matrices 3×3 agissant sur les composantes des vecteurs $\vec{x}$ de $\mathbb{R}^3$, soit comme des opérateurs différentiels agissant sur les fonctions de $\vec{x}$. Pour une rotation infinitésimale $R_{\hat{n}}(d\psi)$ agissant sur $\vec{x}$, le vecteur transformé $\vec{x}'$ est tel que $\vec{x}' - \vec{x}$ est orthogonal à $\vec{x}$ et à $\hat{n}$, donc s'écrit

$$\vec{x}' = \vec{x} + d\psi \hat{n} \wedge \vec{x}, \quad (7.9)$$

(une version infinitésimale de (7.1)). Une fonction scalaire de $\vec{x}$ se transforme, cf chap. 2, selon $f'(\vec{x}') = f(\vec{x})$ soit pour une rotation infinitésimale

$$f'(\vec{x}) = f(R^{-1}\vec{x}) = f(\vec{x} - d\psi \hat{n} \wedge \vec{x}) = (1 - d\psi \hat{n} \cdot \vec{x} \wedge \vec{\nabla})f(\vec{x}) = (1 - id\psi \hat{n} \cdot \vec{J})f(\vec{x}) \quad (7.10)$$

avec $J_{\hat{n}} = -i\hat{n} \cdot \vec{x} \wedge \frac{\partial}{\partial \vec{x}}$ soit encore

$$\boxed{J_{\hat{n}} = \hat{n} \cdot \vec{J} \quad \text{avec} \quad \vec{J} = -i\vec{x} \wedge \frac{\partial}{\partial \vec{x}}} \quad (7.11)$$

Remarque. Dans le cours de mécanique quantique, on s'est intéressé à l'opérateur de moment cinétique orbital $\vec{L}$ qui diffère du présent générateur $\vec{J}$ par un facteur $\hbar$: $\vec{L} = \hbar\vec{J}$. Attention ! $\vec{J}$, opérateur de moment angulaire total du cours de mécanique quantique, ne doit pas être confondu avec le présent $\vec{J}$.

1. Ne pas confondre $J_{\hat{n}}$ indexé par le vecteur $\hat{n}$, avec J_k , k ième composante de $\vec{J}$. La relation entre les deux va être donnée ci-dessous.

Par ailleurs, introduisons les trois matrices de base J_1 , J_2 et J_3 qui décrivent les rotations infinitésimales autour des axes correspondants. De la version infinitésimale de (7.3) on tire

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad J_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \quad J_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (7.12)$$

ce qu'on peut exprimer par une formule unique

$$(J_k)_{ij} = -i\epsilon_{ijk} \quad (7.13)$$

à l'aide du tenseur complètement antisymétrique ϵ_{ijk} . Les trois matrices J_i , $i = 1, 2, 3$ satisfont les très importantes relations de commutation suivantes

$$[J_i, J_j] = i\epsilon_{ijk}J_k \quad (7.14)$$

qu'on vérifie bien entendu aussi sur la forme (7.11). Par exemple (7.11) nous donne $J_1 = J_x = -i(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y})$, $J_2 = J_y = -i(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z})$ d'où on tire $[J_1, J_2] = iJ_3$.

Les générateurs infinitésimaux étant déterminés, comment reconstruit-on les rotations finies? On va adapter au groupe SO(3) ce qu'on a vu au § 4.1.3. Les rotations $R_{\hat{n}}(\psi)$ autour d'un axe $\hat{n}$ donné forment un sous-groupe (on parle de *sous-groupe à un paramètre*); ces matrices commutent, le sous-groupe est commutatif (ou *abélien*). Par la propriété de groupe,

$$R_{\hat{n}}(\psi + d\psi) = R_{\hat{n}}(d\psi)R_{\hat{n}}(\psi) = (I - id\psi J_{\hat{n}})R_{\hat{n}}(\psi) , \quad (7.15)$$

ou encore

$$\frac{\partial R_{\hat{n}}(\psi)}{\partial \psi} = -iJ_{\hat{n}}R_{\hat{n}}(\psi) \quad (7.16)$$

équation différentielle, qui, compte tenu de $R(0) = I$, s'intègre en

$$R_{\hat{n}}(\psi) = e^{-i\psi J_{\hat{n}}} . \quad (7.17)$$

Comme l'a montré (7.11), on peut toujours écrire

$$J_{\hat{n}} = \sum_k n_k J_k \quad \text{donc} \quad R_{\hat{n}}(\psi) = e^{-i\psi \sum_k n_k J_k} . \quad (7.18)$$

Attention ! En raison de la non-commutation des générateurs infinitésimaux J_1 , J_2 et J_3 , cf (7.14), la propriété "bien connue" de la fonction exponentielle $e^a e^b = e^{a+b}$ ne s'applique pas : il ne faut pas écrire $e^{-i\psi \sum_k n_k J_k} \stackrel{?}{=} \prod_k e^{-i\psi n_k J_k}$.

Remarque. Noter que l'équation (7.5) implique que

$$R e^{-i\psi J_{\hat{n}}} R^{-1} = e^{-i\psi R J_{\hat{n}} R^{-1}} = e^{-i\psi J_{\hat{n}'}} \quad (7.19)$$

avec $\hat{n}' = R\hat{n}$, donc

$$R J_{\hat{n}} R^{-1} = J_{\hat{n}'} . \quad (7.20)$$

7.1.3 Algèbre de Lie $\mathfrak{so}(3)$

Récapitulons : nous venons d'introduire l'algèbre de commutation des générateurs infinitésimaux (ou algèbre de Lie) du groupe $\mathrm{SO}(3)$, notée $\mathfrak{so}(3)$. Elle est définie par les relations (7.14). On utilise aussi beaucoup les trois combinaisons

$$J_z \equiv J_3, \quad J_+ = J_1 + iJ_2, \quad J_- = J_1 - iJ_2. \quad (7.21)$$

Il est immédiat de calculer

$$[J_3, J_+] = J_+ \quad (7.22a)$$

$$[J_3, J_-] = -J_- \quad (7.22b)$$

$$[J_+, J_-] = 2J_3. \quad (7.22c)$$

On vérifie aussi que l'opérateur

$$\vec{J}^2 = J_1^2 + J_2^2 + J_3^2 = J_3^2 + J_3 + J_- J_+ = J_3^2 - J_3 + J_+ J_- \quad (7.23)$$

commute avec tous les J

$$[\vec{J}^2, J_i] = 0. \quad (7.24)$$

$\vec{J}^2$ est l'*opérateur de Casimir*. Comme on l'a vu, nous sommes intéressés aux “représentations unitaires”, où les générateurs J_i , $i = 1, 2, 3$ sont hermitiques, donc

$$J_i^\dagger = J_i, \quad i = 1, 2, 3 \quad J_\pm^\dagger = J_\mp. \quad (7.25)$$

L'importance de l'opérateur de Casimir $\vec{J}^2$ est liée au lemme de Schur : $\vec{J}^2$ commutant avec tous les générateurs infinitésimaux, il commute avec leur exponentielle, donc avec les opérateurs de rotation, et ce, dans toute représentation. Dans une représentation irréductible, $\vec{J}^2$ est donc un multiple de l'identité.

7.2 Représentations de $\mathfrak{so}(3)$ et de $\mathrm{SO}(3)$

On a vu dans les chapitres précédents que l'action du groupe $\mathrm{SO}(3)$ sur les vecteurs et les tenseurs de l'espace $\mathbb{R}^3$ fournissait un exemple de représentation. Nous allons maintenant nous intéresser à la construction des représentations générales de $\mathrm{SO}(3)$. Pour les besoins de la physique, en particulier de la mécanique quantique, on a surtout besoin de représentations unitaires, dans lesquelles les matrices de représentation sont unitaires. En fait, comme on le verra, il suffit d'étudier la façon dont sont représentés les éléments du groupe au voisinage de l'identité, c'est-à-dire étudier les représentations des générateurs infinitésimaux de $\mathrm{SO}(3)$.

Il suffit donc pour trouver les représentations unitaires du groupe $\mathrm{SO}(3)$ de trouver les représentations par des matrices hermitiques de son algèbre de Lie $\mathfrak{so}(3)$.

7.2.1 Représentations de l'algèbre $\mathfrak{so}(3)$

Définition

Il convient d'abord de définir ce qu'on entend par *représentation de l'algèbre de Lie $\mathfrak{so}(3)$* . Par définition, une représentation de l'algèbre de Lie $\mathfrak{so}(3)$ est un ensemble de matrices J_i que nous prendrons hermitiennes, de taille $N \times N$ a priori arbitraire, satisfaisant les relations de commutation (7.14) ou de façon équivalente, (7.22).

On peut démontrer et nous admettrons qu'il y a une relation (presque) biunivoque entre les représentations du groupe de Lie $\text{SO}(3)$ et celles de son algèbre de Lie $\mathfrak{so}(3)$: les générateurs infinitésimaux dans une représentation du groupe fournissent une représentation de $\mathfrak{so}(3)$, et inversement par exponentiation des générateurs infinitésimaux dans une représentation de $\mathfrak{so}(3)$, on reconstruit la représentation de $\text{SO}(3)$. Ce dernier point réserve toutefois une surprise...

Algèbres $\mathfrak{so}(3)$ et $\mathfrak{su}(2)$

Comme on sait bien, les trois matrices de Pauli

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (7.26)$$

satisfont la relation $[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k$, autrement dit les matrices $\frac{1}{2}\sigma_i$ satisfont (7.14) et forment donc une représentation de dimension 2 de l'algèbre $\mathfrak{so}(3)$. Mais ces matrices forment aussi une base des générateurs infinitésimaux (hermitiens et de trace nulle) du groupe $\text{SU}(2)$ des matrices unitaires de déterminant 1. On en conclut que les groupes $\text{SU}(2)$ et $\text{SO}(3)$ ont la même algèbre de Lie (les mathématiciens disent que les algèbres de Lie $\mathfrak{su}(2)$ et $\mathfrak{so}(3)$ sont isomorphes). Et l'on voit que les deux groupes vont être liés de façon très étroite : quand on exponentie une représentation de $\mathfrak{so}(3) \simeq \mathfrak{su}(2)$, construit-on une représentation du groupe $\text{SO}(3)$ ou de $\text{SU}(2)$? Nous reviendrons plus bas sur cette question.

Représentations de $\mathfrak{so}(3)$ de spin j

Procédons à la construction des représentations de l'algèbre $\mathfrak{so}(3) \simeq \mathfrak{su}(2)$. Comme précédemment, $J_\pm$ et J_z désignent les représentants des générateurs infinitésimaux dans une certaine représentation. Ils satisfont aux relations de commutation (7.22) et d'hermiticité (7.25). La commutation des opérateurs J_z et $\vec{J}^2$ garantit qu'on peut en chercher des vecteurs propres communs. Comme on l'a indiqué à la fin du § 7.1.3, $\vec{J}^2$ est un multiple de l'opérateur identité, c'est-à-dire est diagonal avec une seule valeur propre dans toute représentation irréductible. Les valeurs propres des opérateurs hermitiques $\vec{J}^2$ et J_z étant réelles et $\vec{J}^2$ étant semi-défini positif, on peut toujours écrire leurs valeurs propres respectives sous la forme $j(j+1)$ et m , avec j réel positif ou nul, et on considère donc un vecteur propre commun $|j\ m\rangle$

$$\begin{aligned} \vec{J}^2 |j\ m\rangle &= j(j+1) |j\ m\rangle \\ J_z |j\ m\rangle &= m |j\ m\rangle. \end{aligned} \quad (7.27)$$

avec m un réel a priori arbitraire. Par abus de langage, on dira que $|jm\rangle$ est un “vecteur propre de valeurs propres (j, m) ”.

(i) Agissons avec J_+ et $J_- = J_+^\dagger$ sur $|jm\rangle$. Utilisant la relation $J_\pm J_\mp = \vec{J}^2 - J_z^2 \pm J_z$ (conséquence de (7.22)), on calcule la norme carrée de $J_\pm |jm\rangle$:

$$\begin{aligned}\langle jm|J_-J_+|jm\rangle &= (j(j+1) - m(m+1)) \langle jm|jm\rangle \\ &= (j-m)(j+m+1) \langle jm|jm\rangle \\ \langle jm|J_+J_-|jm\rangle &= (j(j+1) - m(m-1)) \langle jm|jm\rangle \\ &= (j+m)(j-m+1) \langle jm|jm\rangle.\end{aligned}\tag{7.28}$$

Ces normes carrées ne peuvent être négatives donc

$$\begin{aligned}(j-m)(j+m+1) &\geq 0 & : & -j-1 \leq m \leq j \\ (j+m)(j-m+1) &\geq 0 & : & -j \leq m \leq j+1\end{aligned}\tag{7.29}$$

qui impliquent

$$-j \leq m \leq j.\tag{7.30}$$

En outre $J_+|jm\rangle = 0$ si et seulement si $m = j$ et $J_-|jm\rangle = 0$ si et seulement si $m = -j$

$$J_+|jj\rangle = 0 \quad J_-|j-j\rangle = 0.\tag{7.31}$$

(ii) Si $m \neq j$, $J_+|jm\rangle$ est un vecteur non nul, vecteur propre de valeurs propres $(j, m+1)$. En effet

$$\begin{aligned}\vec{J}^2 J_+|jm\rangle &= J_+ \vec{J}^2 |jm\rangle = j(j+1) J_+|jm\rangle \\ J_z J_+|jm\rangle &= J_+(J_z+1)|jm\rangle = (m+1) J_+|jm\rangle.\end{aligned}\tag{7.32}$$

De même si $m \neq -j$, $J_-|jm\rangle$ est un vecteur propre (non nul) de valeurs propres $(j, m-1)$.

(iii) Considérons la suite des vecteurs

$$|jm\rangle, J_-|jm\rangle, J_-^2|jm\rangle, \dots, J_-^p|jm\rangle \dots$$

S'ils sont non nuls ils constituent des vecteurs propres de J_z de valeurs propres $m, m-1, m-2, \dots, m-p \dots$. Les valeurs propres autorisées de J_z étant bornées par (7.30), cette suite doit s'arrêter au bout d'un nombre fini d'étapes. Soit p l'entier tel que $J_-^p|jm\rangle \neq 0$, $J_-^{p+1}|jm\rangle = 0$. En vertu de (7.31), $J_-^p|jm\rangle$ est un vecteur propre de valeurs propres $(j, -j)$ donc $m-p = -j$ c'est-à-dire

$$(j+m) \text{ est entier.}\tag{7.33}$$

Opérant de même avec $J_+, J_+^2, \dots$ sur $|jm\rangle$, on est mené à la conclusion que

$$(j-m) \text{ est entier}\tag{7.34}$$

et par conséquent j et m sont simultanément entiers ou demi-entiers. Pour chaque valeur de j

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$$

m peut prendre les $2j + 1$ valeurs²

$$m = -j, -j + 1, \dots, j - 1, j. \quad (7.35)$$

Partant du vecteur $|j\ m = j\rangle$, (“vecteur de plus haut poids”), choisi de norme 1, on construit la base orthonormée $|j\ m\rangle$ par application répétée de J_- et on a

$$J_+|j\ m\rangle = \sqrt{j(j+1) - m(m+1)}|j\ m+1\rangle \quad (7.36a)$$

$$J_-|j\ m\rangle = \sqrt{j(j+1) - m(m-1)}|j\ m-1\rangle \quad (7.36b)$$

$$J_z|j\ m\rangle = m|j\ m\rangle. \quad (7.36c)$$

Ces $2j + 1$ vecteurs forment la base de la “représentation de spin j ” de l’algèbre. Comme on l’a mentionné plus haut, cette représentation de l’algèbre $\text{so}(3)$ s’étend en une représentation du groupe $\text{SO}(3)$... à une petite subtilité près dans le cas de spin demi-entier, comme nous allons voir.

7.2.2 Représentation de spin j du groupe $\text{SO}(3)$

Les matrices de la représentation de spin j sont telles que sous l’action de la rotation $R \in \text{SO}(3)$

$$|j\ m\rangle \mapsto D^j(R)|j\ m\rangle = |j\ m'\rangle \mathcal{D}_{m'm}^j(R). \quad (7.37)$$

Selon la paramétrisation $(\hat{n}, \psi, \text{angles d'Euler}, \dots)$, on écrira aussi $\mathcal{D}_{m'm}^j(\hat{n}, \psi)$, $\mathcal{D}_{m'm}^j(\alpha, \beta, \gamma)$, etc. Par (7.7), on a donc

$$\begin{aligned} \mathcal{D}_{m'm}^j(\alpha, \beta, \gamma) &= \langle j\ m' | D(\alpha, \beta, \gamma) | j\ m \rangle \\ &= \langle j\ m' | e^{-i\alpha J_z} e^{-i\beta J_y} e^{-i\gamma J_z} | j\ m \rangle \\ &= e^{-i\alpha m'} d_{m'm}^j(\beta) e^{-i\gamma m} \end{aligned} \quad (7.38)$$

où la matrice d^j est définie par

$$d_{m'm}^j(\beta) = \langle j\ m' | e^{-i\beta J_y} | j\ m \rangle. \quad (7.39)$$

Une formule explicite pour d^j sera donnée au paragraphe suivant. On a encore

$$\begin{aligned} \mathcal{D}_{m'm}^j(\hat{z}, \psi) &= e^{-i\psi m} \delta_{mm'} \\ \mathcal{D}_{m'm}^j(\hat{y}, \psi) &= d_{m'm}^j(\psi). \end{aligned} \quad (7.40)$$

Exercice : Calculer $\mathcal{D}^j(\hat{x}, \psi)$. (On pourra utiliser (7.5).)

On note que $\mathcal{D}^j(\hat{z}, 2\pi) = (-1)^{2j} I$, puisque $(-1)^{2m} = (-1)^{2j}$, et la propriété est vraie pour tout axe $\hat{n}$ par conjugaison (7.5)

$$\mathcal{D}^j(\hat{n}, 2\pi) = (-1)^{2j} I. \quad (7.41)$$

2. En fait, on vient de trouver une condition nécessaire sur les j, m . Le fait que tous ces j donnent effectivement des représentations va être vérifié au paragraphe suivant.

Cela peut paraître très surprenant, la rotation de 2π , donc une opération identité, étant représentée par la matrice $-I$ dans une représentation de spin demi-entier. L'interprétation est que les représentations de spin demi-entier sont des représentations "à une phase près", en fait ici à un signe près, de $\mathrm{SO}(3)$. La rotation identité (ou de 2π) est représentée soit par I soit par $-I$. Comme nous l'avons discuté au § 6.1.1, ces représentations à une phase près sont admises par la physique quantique.

On vérifie aussi que les matrices $\mathcal{D}^j$ sont de déterminant 1, (ou de façon équivalente, que les représentants des générateurs infinitésimaux sont de trace nulle). En effet si $\hat{n} = R\hat{z}$, $\mathcal{D}(\hat{n}, \psi) = \mathcal{D}(R)\mathcal{D}(\hat{z}, \psi)\mathcal{D}^{-1}(R)$, donc

$$\det \mathcal{D}(\hat{n}, \psi) = \det \mathcal{D}(\hat{z}, \psi) = \det e^{-i\psi J_z} = \prod_{m=-j}^j e^{-im\psi} = 1. \quad (7.42)$$

Enfin, puisque le caractère peut être évalué sur le représentant $\mathcal{D}(\hat{z}, \psi)$ de la classe de conjugaison

$$\chi_j(\psi) = \mathrm{tr} \mathcal{D}(\hat{z}, \psi) = \sum_{m=-j}^j e^{-im\psi} = \frac{\sin(j + \frac{1}{2})\psi}{\sin \frac{1}{2}\psi}, \quad (7.43)$$

comme annoncé plus haut en (5.8).

Il peut être utile d'écrire explicitement ces matrices dans les cas $j = \frac{1}{2}$ et $j = 1$. Le cas de $j = \frac{1}{2}$ est très simple, puisque

$$\begin{aligned} \mathcal{D}^{\frac{1}{2}}(R) &= U = e^{-i\frac{1}{2}\psi \hat{n} \cdot \vec{\sigma}} = \begin{pmatrix} \cos \frac{\psi}{2} - i \cos \theta \sin \frac{\psi}{2} & -i \sin \frac{\psi}{2} \sin \theta e^{-i\phi} \\ -i \sin \frac{\psi}{2} \sin \theta e^{i\phi} & \cos \frac{\psi}{2} + i \cos \theta \sin \frac{\psi}{2} \end{pmatrix} \\ &= e^{-i\frac{\alpha}{2}\sigma_3} e^{-i\frac{\beta}{2}\sigma_2} e^{-i\frac{\gamma}{2}\sigma_3} = \begin{pmatrix} \cos \frac{\beta}{2} e^{-\frac{i}{2}(\alpha+\gamma)} & -\sin \frac{\beta}{2} e^{-\frac{i}{2}(\alpha-\gamma)} \\ \sin \frac{\beta}{2} e^{\frac{i}{2}(\alpha-\gamma)} & \cos \frac{\beta}{2} e^{\frac{i}{2}(\alpha+\gamma)} \end{pmatrix} \end{aligned}$$

Pour $j = 1$, dans la base $|1, 1\rangle$, $|1, 0\rangle$ et $|1, -1\rangle$ où J_z est diagonale

$$J_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad J_+ = \sqrt{2} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad J_- = \sqrt{2} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad (7.44)$$

d'où

$$d^1(\beta) = e^{-i\beta J_y} = \begin{pmatrix} \frac{1+\cos \beta}{2} & -\frac{\sin \beta}{\sqrt{2}} & \frac{1-\cos \beta}{2} \\ \frac{\sin \beta}{\sqrt{2}} & \cos \beta & -\frac{\sin \beta}{\sqrt{2}} \\ \frac{1-\cos \beta}{2} & \frac{\sin \beta}{\sqrt{2}} & \frac{1+\cos \beta}{2} \end{pmatrix}. \quad (7.45)$$

Exercice : vérifier ces expressions en calculant J_y^2 , en notant que $J_y^3 = J_y$ et en construisant l'exponentielle $e^{-i\beta J_y}$ par son développement en série.

De $\mathrm{SO}(3)$ à $\mathrm{SU}(2)$

On voit aussi que dans cette construction des représentations de $\mathrm{SO}(3)$ par exponentiation des

représentations de son algèbre de Lie $so(3)$, on a été naturellement mené à l'étude du groupe $SU(2)$, puisque la représentation de spin $\frac{1}{2}$ de l'algèbre $so(3)$ fournit une matrice de $SU(2)$, unitaire et de déterminant 1 comme on vient de le voir. Comme on l'a vu, le groupe $SU(2)$ a une algèbre de Lie isomorphe à celle de $SO(3)$. Les représentations des deux algèbres sont donc les mêmes, et les représentations de spin j entier ou demi-entier sont des représentations (au sens ordinaire) du groupe $SU(2)$. La relation entre les groupes $SO(3)$ et $SU(2)$, ayant même structure locale (même algèbre de Lie) mais une structure globale (topologie) différente, est un sujet fort intéressant mais que nous ne pouvons poursuivre ici.

Irréductibilité

On démontre que la représentation de spin j est irréductible : elle n'admet pas de sous-espace invariant.

Sinon elle serait complètement réductible et il existerait nécessairement des opérateurs diagonaux par blocs, différents de l'identité et commutant avec les matrices de la représentation, en particulier avec les générateurs J_i . Or dans la base (7.36) toute matrice M commutant avec J_z est diagonale, $M_{mm'} = \mu_m \delta_{mm'}$, (le vérifier!), et la commutation avec J_+ force tous les μ_m à être égaux : la matrice M est multiple de l'identité et la représentation est bien irréductible.

On peut aussi se demander pourquoi l'étude des représentations de dimension finie que vous venons de construire suffit aux besoins du physicien, par exemple en mécanique quantique, où la scène se passe en général dans un espace de Hilbert de dimension infinie. On peut en fait appliquer les résultats du § 5.1.4 et affirmer que

Toute représentation de $SO(3)$ dans un espace de Hilbert est équivalente à une représentation unitaire, et donc est complètement réductible en une somme (finie ou infinie) de représentations irréductibles de dimension finie.

Orthogonalité et complétude des représentations irréductibles de $SO(3)$ et $SU(2)$

Comme on l'a déjà mentionné, les relations d'orthogonalité et de complétude du type (5.22, 5.23, 5.24) des éléments de matrice et des caractères des représentations irréductibles sont encore satisfaites dans un groupe de Lie compact comme $SO(3)$ ou $SU(2)$. Le seul point qui mérite d'être explicité est la mesure d'intégration "de Haar" sur le groupe, ou la forme de la distribution delta de Dirac correspondante. L'une et l'autre dépendent de la paramétrisation choisie pour le groupe. Dans les paramètres $\hat{n} = (\theta, \phi)$, ψ , on montre que, pour $U \in SU(2)$

$$d\mu(U) = \frac{1}{2} \sin^2 \frac{\psi}{2} \sin \theta \, d\psi \, d\theta \, d\phi \quad (7.46)$$

avec un volume

$$\text{vol}(SU(2)) = \int_{SU(2)} d\mu(U) = \frac{1}{2} \int_0^{2\pi} d\psi \sin^2 \frac{\psi}{2} \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi = 2\pi^2 \quad (7.47)$$

tandis que pour $SO(3)$, on a encore $d\mu(R) = \frac{1}{2} \sin^2 \frac{\psi}{2} \sin \theta \, d\psi \, d\theta \, d\phi$ mais comme ψ est restreint à $(0, \pi)$,

$$\text{vol}(SO(3)) = \pi^2. \quad (7.48)$$

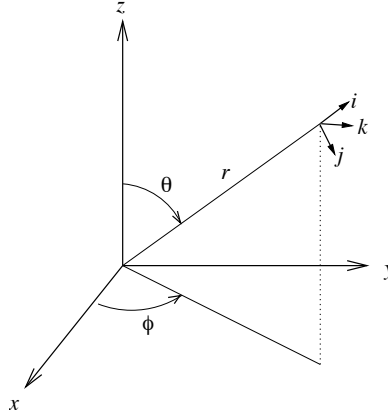


FIGURE 7.2 – Coordonnées sphériques : le vecteur unitaire $\vec{i}$ est radial, $\vec{j}$ est dans le plan vertical d'angle azimutal ϕ .

Dans ces conditions on a pour $U \in \text{SU}(2)$

$$(2j+1) \int \frac{d\mu(U)}{2\pi^2} \mathcal{D}_{mn}^j(U) \mathcal{D}_{m'n'}^{j'*}(U) = \delta_{jj'} \delta_{mm'} \delta_{nn'} \quad (7.49)$$

$$\sum_{j=0, \frac{1}{2}, \dots} \sum_{m, n=-j, \dots, j} (2j+1) \mathcal{D}_{mn}^j(U) \mathcal{D}_{mn}^{j*}(U') = 2\pi^2 \delta(U, U') \quad (7.50)$$

et pour les caractères $\chi_j(\psi) = \frac{\sin(j+\frac{1}{2})\psi}{\sin \frac{1}{2}\psi}$

$$\int_0^{2\pi} d\psi \sin^2 \frac{\psi}{2} \chi_j(\psi) \chi_{j'}(\psi) = \pi \delta_{jj'} \quad (7.51)$$

$$\sum_{j=0, \frac{1}{2}, \dots} \chi_j(\psi) \chi_j(\psi') = \frac{\pi}{\sin^2 \frac{\psi}{2}} \delta(\psi, \psi'), \quad (7.52)$$

avec $\delta(\psi, \psi') = \sum_{k \in \mathbb{Z}} (\delta(\psi - \psi' + k4\pi) + \delta(\psi + \psi' + k4\pi))$. Exercice : vérifier ces relations d'orthogonalité et de complétude, la deuxième en utilisant (5.40).

7.3 Coordonnées sphériques. Harmoniques sphériques.

Dans les coordonnées sphériques $\vec{r} = (r, \phi, \theta)$ (cf Fig. 7.2), on écrit le gradient

$$\vec{\nabla} = \vec{i} \frac{\partial}{\partial r} + \vec{j} \frac{1}{r} \frac{\partial}{\partial \theta} + \vec{k} \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi} \quad (7.53)$$

donc $\vec{r} \wedge \vec{\nabla} = \vec{j} \left(-\frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \right) + \vec{k} \frac{\partial}{\partial \theta}$ et en projetant sur les axes Ox, Oy, Oz , on trouve que les générateurs J_i s'écrivent

$$J_3 = -i \frac{\partial}{\partial \phi} \quad (7.54)$$

$$J_1 = -i \left[-\cos \phi \cotg \theta \frac{\partial}{\partial \phi} - \sin \phi \frac{\partial}{\partial \theta} \right] \quad (7.55)$$

$$J_2 = -i \left[-\sin \phi \cotg \theta \frac{\partial}{\partial \phi} + \cos \phi \frac{\partial}{\partial \theta} \right] \quad (7.56)$$

$$(7.57)$$

On vérifie que le laplacien s'écrit

$$\Delta = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} - \frac{1}{r^2} \vec{J}^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} - \frac{1}{r^2} \vec{J}^2. \quad (7.58)$$

L'opérateur de Casimir de SO(3), $\vec{J}^2 = J_1^2 + J_2^2 + J_3^2$, est au facteur $1/r^2$ près, la partie angulaire du laplacien en coordonnées sphériques. On note $\Delta_{S^2} = -\vec{J}^2$ ce "laplacien sur la sphère unité"

$$\Delta_{S^2} = -\vec{J}^2 = \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) \quad (7.59)$$

Les *harmoniques sphériques* $Y_l^m(\theta, \phi)$ sont des fonctions propres de $\vec{J}^2$ et J_z dans ces coordonnées, auxquelles on impose d'être monovaluées, ce qui n'a lieu que si l et m sont entiers. Dans le langage de la mécanique quantique, nous dirons que les états propres $|lm\rangle$ de $\vec{J}^2$ et J_z sont représentés par les fonctions d'onde

$$Y_l^m(\theta, \phi) = \langle \theta, \phi | l, m \rangle. \quad (7.60)$$

Les expressions des premiers Y_l^m peuvent être utiles

$$Y_0^0 = \frac{1}{\sqrt{4\pi}} \quad (7.61)$$

$$Y_1^0 = \sqrt{\frac{3}{4\pi}} \cos \theta \quad Y_1^{\pm 1} = \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi} \quad (7.62)$$

$$Y_2^0 = \sqrt{\frac{5}{16\pi}} (3 \cos^3 \theta - 1) \quad Y_2^{\pm 1} = \mp \sqrt{\frac{15}{8\pi}} \cos \theta \sin \theta e^{\pm i\phi} \quad Y_2^{\pm 2} = \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi}. \quad (7.63)$$

Ces fonctions $Y_l^m(\theta, \phi)$ sont dotées de nombreuses propriétés remarquables :

(i) Elles satisfont les équations différentielles

$$(\Delta_{S^2} + l(l+1)) Y_l^m = 0 \quad (7.64)$$

$$J_z Y_l^m = -i \frac{\partial}{\partial \phi} Y_l^m = m Y_l^m \quad (7.65)$$

et peuvent s'écrire

$$Y_l^m(\theta, \phi) = \frac{(-1)^l}{2^l l!} \sqrt{\frac{(2l+1)(l+m)!}{4\pi(l-m)!}} e^{im\phi} \sin^{-m} \theta \left(\frac{d}{d \cos \theta} \right)^{l-m} \sin^{2l} \theta . \quad (7.66)$$

(ii) Elles sont normalisées à 1 sur la sphère unité et plus généralement y satisfont des propriétés d'orthogonalité et de complétude

$$\int d\Omega Y_l^{m*} Y_{l'}^{m'} = \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta Y_l^{m*} Y_{l'}^{m'} = \delta_{ll'} \delta_{mm'} \quad (7.67a)$$

$$\begin{aligned} \sum_{l=0}^{\infty} \sum_{m=-l}^l Y_l^{m*}(\theta, \phi) Y_l^m(\theta', \phi') &= \delta(\Omega, \Omega') := \frac{\delta(\theta - \theta') \delta(\phi - \phi')}{\sin \theta} \\ &= \delta(\cos \theta - \cos \theta') \delta(\phi - \phi') \end{aligned} \quad (7.67b)$$

où la “fonction delta” de Dirac est adaptée à la mesure d'intégration $d\Omega = \sin \theta d\theta d\phi$. Eq. (7.67b) veut dire que toute fonction de (θ, ϕ) peut se développer sur les Y_l^m selon

$$f(\theta, \phi) = \sum_{l,m} f_{lm} Y_l^m(\theta, \phi) \quad \text{avec} \quad f_{lm} = \int d\Omega' Y_l^{m*}(\theta', \phi') f(\theta', \phi')$$

(iii) On peut considérer $Y_l^m(\theta, \phi)$ comme fonction du vecteur unitaire $\hat{n}$ d'angles directeurs θ, ϕ . Si le vecteur $\hat{n}$ est transformé en $\hat{n}'$ par la rotation R , on a

$$Y_l^m(\hat{n}') = Y_l^{m'}(\hat{n}) \mathcal{D}^l(R)_{m'm} \quad (7.68)$$

ce qui exprime que les Y_l^m se transforment comme des vecteurs de la représentation de spin l .

(iv) On vérifie sur l'expression ci-dessus les relations de symétrie en m

$$Y_l^{m*}(\theta, \phi) = (-1)^m Y_l^{-m}(\theta, \phi) \quad (7.69)$$

et de parité

$$Y_l^m(\pi - \theta, \phi + \pi) = (-1)^l Y_l^m(\theta, \phi) . \quad (7.70)$$

Noter qu'à $\theta = 0$, $Y_l^m(0, \phi)$ s'annule sauf pour $m = 0$, cf. (7.74, 7.76).

(v) Les harmoniques sphériques satisfont aussi des relations de récurrence de deux types : celles issues de l'action de $J_{\pm}$, opérateurs différentiels qui agissent selon (7.36)

$$e^{\pm i\phi} \left[\pm \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right] Y_l^m = \sqrt{l(l+1) - m(m \pm 1)} Y_l^{m \pm 1} \quad (7.71)$$

et celles provenant de la multiplication des représentations, (cf. plus bas, § 7.4)

$$\sqrt{2l+1} \cos \theta Y_l^m = \left(\frac{(l+m)(l-m)}{2l-1} \right)^{\frac{1}{2}} Y_{l-1}^m + \left(\frac{(l+m+1)(l-m+1)}{2l+3} \right)^{\frac{1}{2}} Y_{l+1}^m . \quad (7.72)$$

(vi) Les harmoniques sphériques donnent des éléments de matrice particuliers des matrices de rotations $\mathcal{D}$

$$\left[\frac{2l+1}{4\pi} \right]^{\frac{1}{2}} \mathcal{D}_{m0}^l(\phi, \theta, 0) = Y_l^{m*}(\theta, \phi) . \quad (7.73)$$

(vii) Finalement on introduit les *polynômes et fonctions de Legendre* $P_l(u)$ et $P_l^m(u)$ définis pour l entier et $u \in [-1, 1]$ par

$$P_l(u) = \frac{1}{2^l l!} \frac{d^l}{du^l} (u^2 - 1)^l \quad (7.74)$$

$$P_l^m(u) = (1 - u^2)^{\frac{1}{2}m} \frac{d^m}{du^m} P_l(u) \quad \text{pour } 0 \leq m \leq l . \quad (7.75)$$

Les harmoniques sphériques sont reliées aux fonctions de Legendre $P_l^m(\cos \theta)$ (pour $m \geq 0$) par

$$Y_l^m(\theta, \phi) = (-1)^m \left[\frac{(2l+1)(l-m)}{4\pi(l+m)} \right]^{\frac{1}{2}} P_l^m(\cos \theta) e^{im\phi} \quad (7.76)$$

donc

$$\mathcal{D}_{m0}^l(0, \theta, 0) = d_{m0}^l(\theta) = (-1)^m \left[\frac{(l-m)}{(l+m)} \right]^{\frac{1}{2}} P_l^m(\cos \theta) = \left(\frac{4\pi}{2l+1} \right)^{\frac{1}{2}} Y_l^{m*}(\theta, 0) . \quad (7.77)$$

En particulier, $d_{00}^l(\theta) = P_l(\cos \theta)$. En général, $d_{m'm}^l(\theta)$ est relié au polynôme de Jacobi

$$P_l^{(\alpha, \beta)}(u) = \frac{(-1)^l}{2^l l!} (1-u)^{-\alpha} (1+u)^{-\beta} \frac{d^l}{du^l} [(1-u)^{\alpha+l} (1+u)^{\beta+l}] \quad (7.78)$$

par

$$d_{m'm}^j(\theta) = \left[\frac{(j+m')!(j-m')!}{(j+m)!(j-m)!} \right]^{\frac{1}{2}} \left(\cos \frac{\theta}{2} \right)^{m+m'} \left(\sin \frac{\theta}{2} \right)^{m-m'} P_{j-m'}^{(m'-m, m'+m)}(\cos \theta) . \quad (7.79)$$

Les polynômes de Legendre $P_l(u)$ sont des polynômes orthogonaux sur l'intervalle $[-1, 1]$: $\int_{-1}^1 du P_l(u) P_{l'}(u) = \frac{2}{2l+1} \delta_{ll'}$. Ils satisfont les relations de récurrence $(2l+1)uP_l(u) = (l+1)P_{l+1} + lP_{l-1}$, ainsi que l'équation différentielle $(\Delta_{S^2} + l(l+1)) P_l(\hat{n} \cdot \hat{n}') = 0$, le laplacien agissant sur le vecteur unitaire $\vec{n}$. Les premiers P_l sont

$$P_0 = 1 \quad P_1 = u \quad P_2 = \frac{1}{2}(3u^2 - 1) \quad P_3 = \frac{1}{2}(5u^3 - 3u) , \dots \quad (7.80)$$

Vérifier que (7.74) implique $P_l(1) = 1$, $P_l(-x) = (-1)^l P_l(x)$. Une “fonction génératrice” des P_l est donnée par

$$\frac{1}{\sqrt{1-2ut+t^2}} = \sum_{l=0}^{\infty} t^l P_l(u) \quad (7.81)$$

avec comme conséquence que si $r' < r$, et θ l'angle entre $\vec{r}$ et $\vec{r}'$

$$\frac{1}{|\vec{r} - \vec{r}'|} = \sum_{l=0}^{\infty} \frac{r'^l}{r^{l+1}} P_l(\cos \theta) \quad (7.82)$$

formule importante pour le développement d'un potentiel coulombien.

Ces polynômes nous seront utiles par exemple pour les “développements en ondes partielles”, cf plus bas (7.110).

On a alors la “formule d'addition”

$$\frac{2l+1}{4\pi} P_l(\cos \theta) = \sum_{m=-l}^l Y_l^m(\hat{n}) Y_l^{m*}(\hat{n}') \quad (7.83)$$

où θ désigne l'angle entre les directions $\hat{n}$ et $\hat{n}'$. Cette formule peut se vérifier en démontrant que le membre de droite satisfait bien les équations différentielles satisfaites par P_l .

7.4 Produit tensoriel de représentations

7.4.1 Produit direct de représentations et l’“addition de moments angulaires”

Intéressons nous au produit de deux représentations de spin j_1 et j_2 et à leur décomposition sur des états de spin total donné (“décomposition en représentations irréductibles”). On part donc

de la représentation produit engendrée par les vecteurs

$$|j_1 m_1\rangle \otimes |j_2 m_2\rangle \equiv |j_1 m_1; j_2 m_2\rangle \quad \text{abrégé en} \quad |m_1 m_2\rangle \quad (7.84)$$

sur lesquels agissent les générateurs infinitésimaux sous la forme

$$\vec{J} = \vec{J}^{(1)} \otimes I^{(2)} + I^{(1)} \otimes \vec{J}^{(2)} . \quad (7.85)$$

L'indice supérieur indique sur quel espace agissent les opérateurs. Par abus de notation, on écrit souvent au lieu de (7.85)

$$\vec{J} = \vec{J}^{(1)} + \vec{J}^{(2)} \quad (7.85')$$

et (en Mécanique Quantique), on parle de l'“addition des moments angulaires” $J^{(1)}$ et $J^{(2)}$. Il s'agit donc de décomposer les vecteurs (7.84) sur une base de vecteurs propres de $\vec{J}$ et J_z . Comme $\vec{J}^{(1)2}$ et $\vec{J}^{(2)2}$ commutent entre eux et avec $\vec{J}^2$ et J_z , on peut chercher des vecteurs propres communs qu'on notera

$$|(j_1 j_2) J M\rangle \quad \text{ou plus simplement} \quad |J M\rangle \quad (7.86)$$

étant entendu qu'on s'est fixé la valeur de j_1 et j_2 . La question est donc double : quelles valeurs J et M peuvent-ils prendre et quelle est la matrice du changement de base $|m_1 m_2\rangle \rightarrow |J M\rangle$? En d'autres termes quelle est la décomposition (de Clebsch-Gordan) et quels sont les coefficients de Clebsch-Gordan ?

Les valeurs possibles de M , valeur propre de $J_z = J_z^{(1)} + J_z^{(2)}$ sont aisées à trouver

$$\begin{aligned} \langle m_1 m_2 | J_z | J M \rangle &= (m_1 + m_2) \langle m_1 m_2 | J M \rangle \\ &= M \langle m_1 m_2 | J M \rangle \end{aligned} \quad (7.87)$$

et la seule valeur de M telle que $\langle m_1 m_2 | J M \rangle \neq 0$ est donc

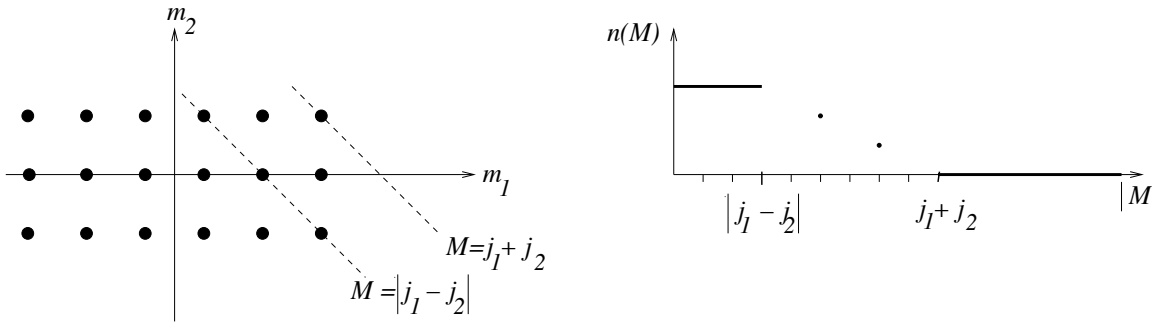
$$M = m_1 + m_2 . \quad (7.88)$$

A j_1, j_2 et M fixés, il y a autant de vecteurs indépendants ayant cette valeur de M qu'il y a de couples (m_1, m_2) satisfaisant (7.88), soit

$$n(M) = \begin{cases} 0 & \text{si } |M| > j_1 + j_2 \\ j_1 + j_2 + 1 - |M| & \text{si } |j_1 - j_2| \leq |M| \leq j_1 + j_2 \\ 2 \inf(j_1, j_2) + 1 & \text{si } 0 \leq |M| \leq |j_1 - j_2| \end{cases} \quad (7.89)$$

(voir Fig. 7.3 pour laquelle $j_1 = 5/2$ and $j_2 = 1$). Soit N_J le nombre de fois où la représentation de spin J apparaît dans la décomposition du produit des représentations de spin j_1 et j_2 . Les $n(M)$ vecteurs de valeur propre M pour J_z peuvent aussi s'interpréter comme provenant des N_J vecteurs $|J M\rangle$ pour les différentes valeurs de J compatibles avec cette valeur de M

$$n(M) = \sum_{J \geq |M|} N_J \quad (7.90)$$

FIGURE 7.3 – Les valeurs possibles de M et leur multiplicité

soit en retranchant membre à membre deux telles relations

$$\begin{aligned}
 N_J &= n(J) - n(J+1) \\
 &= 1 \quad \text{si et seulement si} \quad |j_1 - j_2| \leq J \leq j_1 + j_2 \\
 &= 0 \quad \text{sinon.}
 \end{aligned} \tag{7.91}$$

En conclusion, nous venons de démontrer que les $(2j_1 + 1)(2j_2 + 1)$ vecteurs (7.84) (à j_1 et j_2 fixés) peuvent se réexprimer en fonction des vecteurs $|J M\rangle$ où

$$\begin{aligned}
 J &= |j_1 - j_2|, |j_1 - j_2| + 1, \dots, j_1 + j_2 \\
 M &= -J, -J + 1, \dots, J.
 \end{aligned} \tag{7.92}$$

Noter qu'en définitive les multiplicités N_J valent 0 ou 1 ; c'est une particularité de $SU(2)$ que des multiplicités supérieures à 1 n'apparaissent pas dans la décomposition du produit de deux représentations irréductibles, c'est-à-dire ici de spin fixé.

Exercices : 1. Vérifier que la multiplicité totale est $\sum_{J=|j_1-j_2|}^{j_1+j_2} (2j_1 + 1)(2j_2 + 1)$ comme il se doit.

2. Vérifier que ces règles d'“addition des moments angulaires” sont bien ce qu'on obtient par la méthode générale du chapitre 5, en utilisant les caractères de $SU(2)$ et leur orthogonalité pour la mesure (7.51), autrement dit calculer $\int_0^{2\pi} \frac{d\psi}{\pi} \sin^2 \frac{\psi}{2} \chi_{j_1}(\psi) \chi_{j_2}(\psi) \chi_J(\psi)$.

7.4.2 Coefficients de Clebsch-Gordan, symboles 3-j

Le changement de base $|j_1 m_1; j_2 m_2\rangle \rightarrow |(j_1 j_2) J M\rangle$ s'effectue à l'aide des coefficients de Clebsch-Gordan (C.G.) $\langle (j_1 j_2); J M | j_1 m_1; j_2 m_2 \rangle$

$$|j_1 m_1; j_2 m_2\rangle = \sum_{J=|j_1-j_2|}^{j_1+j_2} \sum_{M=-J}^J \langle (j_1 j_2) J M | j_1 m_1; j_2 m_2 \rangle |(j_1 j_2) J M\rangle \tag{7.93}$$

$$|j_1 j_2; J M\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} \langle (j_1 j_2) J M | j_1 m_1; j_2 m_2 \rangle^* |j_1 m_1; j_2 m_2\rangle. \tag{7.94}$$

Leur valeur dépend en fait d'un choix de phase relative entre les vecteurs (7.84-7.86) ; la convention habituelle est que pour chaque valeur de J , on choisit

$$\langle j_1 m_1 = j_1; j_2 m_2 = J - j_1 | J M = J \rangle \quad \text{réel.} \quad (7.95)$$

Les autres vecteurs sont alors définis sans ambiguïté par (7.36) et on va montrer que tous les C.G. sont réels. Les C.G. satisfont des relations de récurrence conséquences de (7.36). Appliquant en effet $J_{\pm}$ aux deux membres de (7.93a), on obtient

$$\begin{aligned} \sqrt{J(J+1) - M(M \pm 1)} \langle (j_1 j_2) J M | j_1 m_1; j_2 m_2 \rangle \\ = \sqrt{j_1(j_1+1) - m_1(m_1 \pm 1)} \langle (j_1 j_2) J M \pm 1 | j_1 m_1 \pm 1; j_2 m_2 \rangle \\ + \sqrt{j_2(j_2+1) - m_2(m_2 \pm 1)} \langle (j_1 j_2) J M \pm 1 | j_1 m_1; j_2 m_2 \pm 1 \rangle \end{aligned} \quad (7.96)$$

qui permet à l'aide de la normalisation $\sum_{m_1, m_2} |\langle j_1 m_1; j_2 m_2 | (j_1 j_2) J M \rangle|^2 = 1$ et de la convention (7.95) de déterminer tous les C.G. Comme annoncé, ils sont clairement tous réels.

Les C.G. du groupe SU(2), qui décrivent un changement de base orthogonale, satisfont des propriétés d'orthogonalité et de complétude

$$\begin{aligned} \sum_{m_1=-j_1}^{j_1} \langle j_1 m_1; j_2 m_2 | (j_1 j_2) J M \rangle \langle j_1 m_1; j_2 m_2 | (j_1 j_2) J' M' \rangle &= \delta_{JJ'} \delta_{MM'} \\ \text{si } |j_1 - j_2| \leq J \leq j_1 + j_2 \\ \sum_{J=|j_1-j_2|}^{j_1+j_2} \langle j_1 m_1; j_2 m_2 | (j_1 j_2) J M \rangle \langle j_1 m'_1; j_2 m'_2 | (j_1 j_2) J M \rangle &= \delta_{m_1 m'_1} \delta_{m_2 m'_2} \\ \text{si } |m_1| \leq j_1, |m_2| \leq j_2. \end{aligned} \quad (7.97)$$

Noter que dans la première ligne, m_2 est fixé par la donnée de m_1 , à M donné ; et que dans la deuxième, M est fixé en termes de m_1 et de m_2 . Chaque relation n'implique donc qu'une seule somme.

Exercice. Montrer en appliquant une opération du groupe SU(2) aux deux membres de (7.93) que l'on peut décomposer le produit des matrices $\mathcal{D}^{j_1}$ et $\mathcal{D}^{j_2}$ sous la forme³

$$\mathcal{D}_{m_1 m'_1}^{j_1} \mathcal{D}_{m_2 m'_2}^{j_2} = \sum_{J, M, M'} \langle J M | j_1 m_1; j_2 m_2 \rangle^* \langle J M' | j_1 m'_1; j_2 m'_2 \rangle \mathcal{D}_{M M'}^J. \quad (7.98)$$

Plutôt que les coefficients de Clebsch-Gordan, on peut considérer un ensemble de coefficients équivalents, dits *symboles 3-j*. Ils sont définis par

$$\begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & -M \end{pmatrix} = \frac{(-1)^{j_1-j_2+M}}{\sqrt{2J+1}} \langle j_1 m_1; j_2 m_2 | (j_1 j_2) J M \rangle \quad (7.99)$$

3. Dans le cas présent de SO(3), où tous les coefficients de C-G sont réels, la conjugaison complexe ne s'impose pas, mais nous la laissons pour bien mettre en évidence la structure du développement.

et ont l'intérêt de jouir de propriétés de symétrie simples :

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$$

est invariant par permutation circulaire des trois colonnes et change par le signe $(-1)^{j_1+j_2+j_3}$ quand deux colonnes sont permutées ou quand on change les signes de m_1 , m_2 et m_3 . Le lecteur trouvera dans la littérature de nombreuses tables et formules explicites.

Contentons-nous de donner les valeurs pour les spins les plus bas

$$\frac{1}{2} \otimes \frac{1}{2} : \quad \begin{cases} |(\frac{1}{2}, \frac{1}{2})1, 1\rangle &= | \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2} \rangle \\ |(\frac{1}{2}, \frac{1}{2})1, 0\rangle &= \frac{1}{\sqrt{2}} (| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \rangle + | \frac{1}{2}, -\frac{1}{2}; \frac{1}{2}, \frac{1}{2} \rangle) \\ |(\frac{1}{2}, \frac{1}{2})0, 0\rangle &= \frac{1}{\sqrt{2}} (| \frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \rangle - | \frac{1}{2}, -\frac{1}{2}; \frac{1}{2}, \frac{1}{2} \rangle) \\ |(\frac{1}{2}, \frac{1}{2})1, -1\rangle &= | \frac{1}{2}, -\frac{1}{2}; \frac{1}{2}, -\frac{1}{2} \rangle \end{cases} \quad (7.100)$$

et

$$\frac{1}{2} \otimes 1 : \quad \begin{cases} |(\frac{1}{2}, 1)\frac{3}{2}, \frac{3}{2}\rangle &= | \frac{1}{2}, \frac{1}{2}; 1, 1 \rangle \\ |(\frac{1}{2}, 1)\frac{3}{2}, \frac{1}{2}\rangle &= \frac{1}{\sqrt{3}} (\sqrt{2}| \frac{1}{2}, \frac{1}{2}; 1, 0 \rangle + | \frac{1}{2}, -\frac{1}{2}; 1, 1 \rangle) \\ |(\frac{1}{2}, 1)\frac{3}{2}, -\frac{1}{2}\rangle &= \frac{1}{\sqrt{3}} (| \frac{1}{2}, \frac{1}{2}; 1, -1 \rangle + \sqrt{2}| \frac{1}{2}, -\frac{1}{2}; 1, 0 \rangle) \\ |(\frac{1}{2}, 1)\frac{3}{2}, -\frac{3}{2}\rangle &= | \frac{1}{2}, -\frac{1}{2}; 1, -1 \rangle \\ |(\frac{1}{2}, 1)\frac{1}{2}, \frac{1}{2}\rangle &= \frac{1}{\sqrt{3}} (-| \frac{1}{2}, \frac{1}{2}; 1, 0 \rangle + \sqrt{2}| \frac{1}{2}, -\frac{1}{2}; 1, 1 \rangle) \\ |(\frac{1}{2}, 1)\frac{1}{2}, -\frac{1}{2}\rangle &= \frac{1}{\sqrt{3}} (-\sqrt{2}| \frac{1}{2}, \frac{1}{2}; 1, -1 \rangle + | \frac{1}{2}, -\frac{1}{2}; 1, 0 \rangle) \end{cases} \quad (7.101)$$

On note sur le cas $\frac{1}{2} \otimes \frac{1}{2}$ la propriété que les vecteurs de spin total $j = 1$ sont symétriques dans l'échange des deux spins, celui de spin 0 antisymétrique. La propriété est générale : dans la composition de deux représentations de spin $j_1 = j_2$, les vecteurs résultants de spin $j = 2j_1, 2j_1 - 2, \dots$ sont symétriques, ceux de spin $2j_1 - 1, 2j_1 - 3, \dots$ sont antisymétriques. Cela est apparent sur l'expression (7.99) ci-dessus, compte tenu des propriétés annoncées des symboles 3- j .

Dans le même ordre d'idées, soit le produit complètement antisymétrique de $2j + 1$ copies d'une représentation de spin j . On peut montrer que cette représentation est de spin 0 (exercice suivant). (Cela a une conséquence en physique atomique, dans le remplissage des couches électroniques : une couche complète a un moment orbital total et un spin total nuls donc aussi un moment angulaire total nul.)

Exercice. On considère le produit complètement antisymétrique de $N = 2j + 1$ représentations de spin j . Montrer que cette représentation est engendrée par le vecteur $\epsilon_{m_1 m_2 \dots m_N} |j m_1, j m_2, \dots, j m_N\rangle$, qu'il est invariant par l'action de SU(2) et donc que la représentation construite est celle de spin $J = 0$.

7.4.3 Théorème de Wigner–Eckart

Selon le théorème de Wigner (chap. 6, § 1), la transformation d'un opérateur sur $\mathcal{H}$ sous l'action d'une rotation R de SO(3) obéit à : $A \rightarrow U(R)AU(R)^\dagger$. Supposons qu'on a un ensemble de tels

opérateurs, A_α , $\alpha = 1, 2, \dots$, qui se transforment linéairement les uns dans les autres dans ces transformations, c'est-à-dire qui forment une représentation :

$$A_\alpha \rightarrow U(R)A_\alpha U(R)^\dagger = \sum_{\alpha'} A_{\alpha'} \mathcal{D}_{\alpha'\alpha}(R) . \quad (7.102)$$

Si la représentation D est irréductible, les opérateurs A_α forment ce qu'on appelle un opérateur (ou “tenseur”) irréductible. Dans le cas qui nous occupe ici du groupe $SO(3)$ ou $SU(2)$, un tel opérateur irréductible a un spin j fixé et $2j+1$ composantes indexées par $m = -j, -j+1, \dots, j$. Par exemple, en physique atomique, l'opérateur moment cinétique $\vec{J}$ et l'opérateur moment dipolaire électrique $\sum_i q_i \vec{r}_i$ se transforment comme des vecteurs sous l'effet des rotations, donc $j = 1$, $m = -1, 0, 1$. Supposons donc que les A_m se transforment par la représentation irréductible D^j et appliquons les sur des états $|j'm'\rangle$ se transformant selon la représentation irréductible $D^{j'}$. L'état résultant se transforme selon

$$U(R)A_m|j'm'\rangle = U(R)A_m U(R)^\dagger U(R)|j'm'\rangle = \mathcal{D}_{m_1 m}^j(R) \mathcal{D}_{m'_1 m'}^{j'}(R) A_{m_1} |j'm'_1\rangle \quad (7.103)$$

c'est-à-dire selon le produit des représentations D^j et $D^{j'}$. On peut développer ce produit sur des représentations irréductibles comme on l'a vu plus haut en (7.98) :

$$\mathcal{D}_{m_1 m}^j(R) \mathcal{D}_{m'_1 m'}^{j'}(R) = \sum_{j'', m'', m''_1} \langle j'' m'' | j, m; j', m' \rangle \langle j'' m''_1 | j, m_1; j', m'_1 \rangle^* \mathcal{D}_{m'' m''_1}^{j''}(R) . \quad (7.104)$$

Les matrices des représentations satisfont les propriétés d'orthogonalité (7.49) et de décomposition (7.104). On peut donc écrire

$$\begin{aligned} \langle j'' m'' | A_m | j' m' \rangle &= \langle j'' m'' | U(R)^\dagger U(R) A_m | j' m' \rangle \\ &= \int_{SO(3)} \frac{d\mu(R)}{\text{vol}(SO(3))} \langle j'' m'' | U(R)^\dagger U(R) A_m | j' m' \rangle \\ &= \int_{SO(3)} \frac{d\mu(R)}{\text{vol}(SO(3))} \sum_{m_1, m'_1, m''_1} \mathcal{D}_{m'' m''_1}^{j''*}(R) \langle j'' m''_1 | A_{m_1} | j' m'_1 \rangle \mathcal{D}_{m_1 m}^j(R) \mathcal{D}_{m'_1 m'}^{j'}(R) \\ &= \frac{1}{2j''+1} \sum_{m_1, m'_1, m''_1} \langle j'' m'' | j, m; j', m' \rangle \langle j'' m''_1 | j, m_1; j', m'_1 \rangle^* \langle j'' m''_1 | A_{m_1} | j' m'_1 \rangle , \end{aligned} \quad (7.105)$$

avec la mesure $d\mu(R)$ et le volume $\text{vol}(SO(3)) = \pi^2$ discutés plus haut en (7.48). Notons

$$\langle j'' || A || j' \rangle = \frac{1}{2j''+1} \sum_{m_1, m'_1, m''_1} \langle j'' m''_1 | j, m_1; j', m'_1 \rangle^* \langle j'' m''_1 | A_{m_1} | j' m'_1 \rangle . \quad (7.106)$$

Il en découle que (théorème de Wigner–Eckart) :

$$\langle j'' m'' | A_m | j' m' \rangle = \langle j'' || A || j' \rangle \langle j'' m'' | j, m; j', m' \rangle \quad (7.107)$$

dans laquelle les éléments de matrice “réduits” $\langle . || A || . \rangle$ sont indépendants de m, m', m'' . L'élément de matrice du membre de gauche dans (7.107) s'annule si le coefficient de Clebsch–Gordan est nul (en particulier si la représentation j'' n'apparaît pas dans le produit de j et j'). Ce

théorème a de nombreuses conséquences en physique atomique et nucléaire, où il occasionne des “règles de sélection”. Voir par exemple le cas des opérateurs moments multipolaires électriques au paragraphe suivant.

Exercices.

1. Toujours pour le groupe $SO(3)$, soit A_m les composantes d'un opérateur vectoriel irréductible (par exemple, l'opérateur moment dipolaire du paragraphe suivant). Montrer en utilisant le théorème de Wigner–Eckart que

$$\langle j, m_2 | A_m | j, m_1 \rangle = \langle j, m_2 | J_m | j, m_1 \rangle \frac{\langle \vec{J} \cdot \vec{A} \rangle}{j(j+1)}$$

où $\langle \vec{J} \cdot \vec{A} \rangle$ désigne la valeur moyenne de $\vec{J} \cdot \vec{A}$ dans l'état j . En d'autres termes, on peut remplacer $\vec{A}$ par sa projection $\vec{J} \frac{\langle \vec{J} \cdot \vec{A} \rangle}{j(j+1)}$.

2. Reprendre la discussion du théorème de Wigner–Eckart en supposant que les représentations j' et j'' apparaissent avec des multiplicités $m_{j'}$ et $m_{j''}$ et montrer que

$$\langle j'', m''; i'' | A_m | j', m'; i' \rangle = \langle j'' m'' | j, m; j', m' \rangle \langle j''; i'' || A || j'; i' \rangle \quad i' = 1, \dots, m_{j'}, \quad i'' = 1, \dots, m_{j''}$$

En particulier pour un opérateur A invariant ($j = 0$), (comme un Hamiltonien dans le cas d'une symétrie),

$$\langle j'', m''; i'' | A | j', m'; i' \rangle = \delta_{j'j''} \delta_{m'm''} \langle j'; i'' || A || j'; i' \rangle \quad i', i'' = 1, \dots, m_{j'}$$

et le problème de diagonalisation est ramené à celui d'une matrice $m_{j'} \times m_{j'}$.

7.5 Quelques applications physiques

7.5.1 Moments multipolaires

On considère un potentiel créé par une distribution de charge statique $\rho(\vec{r})$

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{d^3r' \rho(\vec{r}')}{|\vec{r} - \vec{r}'|}$$

et on le développe sur les harmoniques sphériques en utilisant les identités (7.82) et (7.83). Il vient

$$\phi(\vec{r}) = \frac{1}{\epsilon_0} \sum_{l,m} \frac{1}{2l+1} \frac{Y_l^{m*}(\hat{n})}{r^{l+1}} Q_{lm} \quad (7.108)$$

où les Q_{lm} , définis par

$$Q_{lm} = \int d^3r' \rho(\vec{r}') r'^l Y_l^m(\hat{n}') \quad (7.109)$$

sont les *moments multipolaires* de la distribution de charge ρ . Par exemple, si $\rho(\vec{r}) = \rho(r)$ est invariant par rotation, seul Q_{00} est non nul, égal à la charge totale (à $1/\sqrt{4\pi}$ près)

$$Q_{00} = \frac{Q}{\sqrt{4\pi}} = \sqrt{4\pi} \int r^2 dr \rho(r) \quad \phi(r) = \frac{Q}{4\pi\epsilon_0 r}.$$

Les trois composantes de Q_{1m} reconstruisent le moment dipolaire $\int d^3r' \rho(\vec{r}') \vec{r}'$. Plus généralement, sous l'effet des rotations, les Q_{lm} forment les composantes d'un opérateur tensoriel se transformant selon la représentation de spin l (et cf. (7.70), de parité $(-1)^l$). En Mécanique Quantique, les Q_{lm} deviennent des opérateurs.

Exercice. En appliquant le théorème de Wigner–Eckart, montrer que $\langle jm_1 | Q_l^m | jm_2 \rangle$ n'est non nul que si $l \leq 2j$.

7.5.2 Etats propres de moment angulaire en Mécanique Quantique

Les harmoniques sphériques peuvent s'interpréter comme les fonctions d'onde dans les coordonnées θ, ϕ des états propres du moment angulaire $\vec{L} = \hbar \vec{J} = \hbar \vec{r} \wedge \vec{\nabla}$

$$Y_l^m(\theta, \phi) = \langle \theta, \phi | l, m \rangle$$

en analogie avec

$$e^{-i\vec{x} \cdot \vec{p}} = \langle \vec{x} | \vec{p} \rangle .$$

En particulier, dans un processus de collision décrit par un Hamiltonien invariant par rotation, un état d'impulsion initiale $\vec{p}_i$ selon l'axe des z , (c'est-à-dire $\theta = \phi = 0$), interagit avec un certain centre diffuseur et ressort dans un état d'impulsion $\vec{p}_f$, avec $|p_i| = |p_f| = p$, selon la direction $\hat{n} = (\theta, \phi)$. On écrit l'amplitude

$$\begin{aligned} \langle p, \theta, \phi | \mathcal{T} | p, 0, 0 \rangle &= \sum_{l'm'm'} Y_l^m(\theta, \phi) \langle p, l, m | \mathcal{T} | p, l', m' \rangle Y_{l'}^{m'*}(0, 0) \\ &= \sum_{l'm} Y_l^m(\theta, \phi) \langle p, l, m | \mathcal{T} | p, l', 0 \rangle Y_{l'}^{0*}(0, 0) \\ &= \sum_l \frac{2l+1}{4\pi} \mathcal{T}_l(p) P_l(\cos \theta) \end{aligned} \quad (7.110)$$

selon à nouveau la formule d'addition et $\langle plm | \mathcal{T} | pl'm' \rangle = \delta_{ll'} \delta_{mm'} \mathcal{T}_l(p)$ exprimant l'invariance par rotation. C'est le très utile développement en ondes partielles de l'amplitude de diffusion.

Se reporter aux ouvrages de Mécanique Quantique pour de nombreuses applications détaillées de ces principes.

7.5.3 L'isospin

Le groupe SU(2) n'intervient pas en Physique qu'en tant que (relié au) groupe de rotation de l'espace euclidien. Illustrons une autre de ses apparitions par la symétrie d'isospin. Il existe dans la nature un certain nombre de particules élémentaires présentant des propriétés voisines, mais différant par leur charge électrique. C'est le cas du proton et du neutron, de masses 938,28 MeV/c² et 939,57 MeV/c² respectivement, mais aussi du triplet de mésons pi, π^0 (masse 134,96 MeV/c²) et $\pi^\pm$ (139,57 MeV/c²), des mésons K etc. Il a été proposé que cela est la manifestation d'une symétrie brisée par les effets électromagnétiques. En l'absence d'interactions électromagnétiques, le proton et le neutron d'une part, les trois mésons π de l'autre seraient des particules de même nature, de même masse, différant seulement par un nombre quantique "interne", à la façon de deux électrons dotés de spins différents. En fait le groupe régissant cette symétrie est aussi SU(2), mais un SU(2) agissant dans un espace abstrait autre que l'espace usuel. On a donné le nom d'*isospin* ou *spin isotopique* au nombre quantique correspondant. Pour résumer, la proposition est donc qu'il existe un groupe SU(2) de symétrie de l'Hamiltonien des interactions fortes, et que les différentes particules sujettes à ces interactions forment des représentations de SU(2) : représentation d'isospin $I = \frac{1}{2}$ pour le nucléon (proton $I_z = +\frac{1}{2}$,

neutron $I_z = -\frac{1}{2}$), isospin $I = 1$ pour les pions ($\pi^\pm : I_z = \pm 1$, $\pi^0 : I_z = 0$) etc. L'isospin est donc un "bon nombre quantique", conservé dans ces interactions. Ainsi le processus (virtuel) $N \rightarrow N + \pi$, (N pour nucléon) important en physique nucléaire, est compatible avec les règles d'addition des isospins ($\frac{1}{2} \otimes 1$ "contient" $\frac{1}{2}$). Les différentes réactions $N + \pi \rightarrow N + \pi$ autorisées par la conservation de la charge électrique

$$\begin{array}{ll} p + \pi^+ \rightarrow p + \pi^+ & I_z = \frac{3}{2} \\ p + \pi^0 \rightarrow p + \pi^0 & I_z = \frac{1}{2} \\ \quad \rightarrow n + \pi^+ & " \\ p + \pi^- \rightarrow p + \pi^- & I_z = -\frac{1}{2} \\ \quad \rightarrow n + \pi^0 & " \\ n + \pi^- \rightarrow n + \pi^- & I_z = -\frac{3}{2} \end{array}$$

conservent aussi l'isospin total I et sa composante I_z mais l'hypothèse d'invariance par SU(2) d'isospin nous apprend d'avantage. Les éléments de matrice de transition des deux réactions dans le canal $I_z = -\frac{1}{2}$, par exemple, doivent être reliés par les règles d'addition de l'isospin. En inversant les relations (7.101), on obtient

$$\begin{aligned} |p, \pi^- \rangle &= \sqrt{\frac{1}{3}} |I = \frac{3}{2}, I_z = -\frac{1}{2} \rangle - \sqrt{\frac{2}{3}} |I = \frac{1}{2}, I_z = -\frac{1}{2} \rangle \\ |n, \pi^0 \rangle &= \sqrt{\frac{2}{3}} |I = \frac{3}{2}, I_z = -\frac{1}{2} \rangle + \sqrt{\frac{1}{3}} |I = \frac{1}{2}, I_z = -\frac{1}{2} \rangle \end{aligned} \quad (7.111)$$

tandis que pour $I_z = 3/2$

$$|p, \pi^+ \rangle = |I = \frac{3}{2}, I_z = \frac{3}{2} \rangle .$$

L'invariance d'isospin implique que $\langle I I_z | \mathcal{T} | I' I'_z \rangle = \mathcal{T}_I \delta_{II'} \delta_{I_z I'_z}$. (Exercice : le justifier). En calculant alors les éléments de matrice de l'opérateur de transition $\mathcal{T}$ entre ces différents états,

$$\begin{aligned} \langle p\pi^+ | \mathcal{T} | p\pi^+ \rangle &= \mathcal{T}_{3/2} \\ \langle p\pi^- | \mathcal{T} | p\pi^- \rangle &= \frac{1}{3} (\mathcal{T}_{3/2} + 2\mathcal{T}_{1/2}) \\ \langle n\pi^0 | \mathcal{T} | p\pi^- \rangle &= \frac{\sqrt{2}}{3} (\mathcal{T}_{3/2} - \mathcal{T}_{1/2}) \end{aligned} \quad (7.112)$$

on trouve que les amplitudes satisfont une relation

$$\sqrt{2} \langle n, \pi^0 | \mathcal{T} | p, \pi^- \rangle + \langle p, \pi^- | \mathcal{T} | p, \pi^- \rangle = \langle p, \pi^+ | \mathcal{T} | p, \pi^+ \rangle = \mathcal{T}_{3/2}$$

conséquence non triviale de l'invariance d'isospin, qui implique des inégalités triangulaires entre les modules carrés de ces amplitudes donc entre les sections efficaces de ces réactions

$$\begin{aligned} [\sqrt{\sigma(\pi^- p \rightarrow \pi^- p)} - \sqrt{2\sigma(\pi^- p \rightarrow \pi^0 n)}]^2 &\leq \sigma(\pi^+ p \rightarrow \pi^+ p) \leq \\ &\leq [\sqrt{\sigma(\pi^- p \rightarrow \pi^- p)} + \sqrt{2\sigma(\pi^- p \rightarrow \pi^0 n)}]^2 \end{aligned} \quad (7.113)$$

qui sont bien vérifiées expérimentalement.

Mieux, on constate qu'à une énergie d'environ 180 MeV, les sections efficaces (proportionnelles aux carrés des amplitudes) sont dans les rapports

$$\sigma(\pi^+p \rightarrow \pi^+p) : \sigma(\pi^-p \rightarrow \pi^0n) : \sigma(\pi^-p \rightarrow \pi^-p) = 9 : 2 : 1$$

ce qui indique qu'à cette énergie, la diffusion dans le canal d'isospin 3/2 est prédominante et signale en fait l'existence d'un état intermédiaire, particule très instable ou "résonance", notée Δ , d'isospin 3/2 donc avec quatre états de charge

$$\Delta^{++}, \Delta^+, \Delta^0, \Delta^-.$$

Cette particule a un spin 3/2 et une masse $M(\Delta) \approx 1230 \text{ MeV}/c^2$.

7.6 2 Problèmes

7.6.1 Molécule de fullerène

On se propose d'étudier la molécule C_{60} , récemment découverte, ses propriétés de symétrie et ses états électroniques.

A.

La structure de la molécule de C_{60} peut être obtenue en partant d'un icosaèdre régulier (cf. Fig. 1.6) et en tronquant chacune des pyramides à base pentagonale par un plan perpendiculaire à son axe qui coupe les arêtes au tiers de la distance au sommet. La figure résultante est constituée de 20 hexagones réguliers et de 12 pentagones réguliers, et de 60 sommets qui portent les atomes de carbone.

1. On s'intéresse tout d'abord au groupe I de symétrie de rotation de l'icosaèdre. Montrer qu'il est constitué de cinq classes de conjugaison qu'on caractérisera. Combien de représentations irréductibles possède-t-il ?

Montrer que par restriction à I , les représentations de spin entier de $SO(3)$ fournissent des représentations de I . Pour dresser la table de caractères, on considère les représentations de spin 0, 1 et 2 de $SO(3)$. Calculer leurs caractères pour les classes de I . On rappelle que $-2\cos\frac{\pi}{5}$ et $2\cos\frac{2\pi}{5}$ sont les racines de l'équation $z^2 + z - 1 = 0$. Montrer que ces représentations de I sont irréductibles ; on les notera **1**, **3** et **5**. La même méthode appliquée à la représentation de spin 3 de $SO(3)$ donne une représentation réductible qui se scinde en deux représentations de dimensions 3 et 4, qu'on notera $\tilde{\mathbf{3}}$ et **4** et dont on trouvera les caractères dans la table ci-après. Compléter cette table.

2. Le groupe d'invariance de la molécule C_{60} est le groupe noté I_h engendré par les rotations de I et la symétrie τ par rapport au centre de l'icosaèdre. Montrer que ce groupe a la structure d'un produit direct $I \otimes \mathbf{Z}_2$ et que chaque représentation D_ρ de I donne lieu à deux représentations $D_{\rho\pm}$ de I_h . Que sont leurs caractères $\chi_{\rho\pm}$?

$\downarrow \text{rep.} \backslash \text{classes} \rightarrow$	1	C_2	C_3	$C_5^{(1)}$	$C_5^{(2)}$
1					
3					
$\tilde{3}$	3	-1	0	$-2 \cos \frac{2\pi}{5}$	$2 \cos \frac{\pi}{5}$
4	4	0	1	-1	-1
5					

B. 1. Les six électrons de chaque carbone se répartissent en deux électrons sur une couche $1s$ interne, trois électrons de valence qui sont responsables des liaisons entre atomes voisins et un dernier qui va nous intéresser. C'est donc pour 60 électrons qu'il s'agit de trouver des niveaux d'énergie, au prix de certaines approximations qui vont en particulier nous faire négliger les interactions entre électrons.

Soit $v(\vec{r} - \vec{r}_i)$ le potentiel créé par l'ion C^+ placé en $\vec{r}_i$ sur un de ces électrons en $\vec{r}$. Soit $f_i(\vec{r}) = \psi(\vec{r} - \vec{r}_i)$ une fonction d'onde normalisée de cet électron au voisinage du i -ème atome ($i = 1, \dots, 60$), solution de

$$\left(-\hbar^2 \frac{\vec{\nabla}^2}{2m} + v(\vec{r} - \vec{r}_i) - e_0 \right) \psi(\vec{r} - \vec{r}_i) = 0 .$$

Dans toute la suite du problème, la fonction ψ sera supposée donnée et pour simplifier, on la supposera à symétrie sphérique autour de $\vec{r}_i$:

$$f_i(\vec{r}) = \psi(|\vec{r} - \vec{r}_i|) .$$

On considère l'espace E des fonctions d'onde à un électron obtenues par combinaisons linéaires des f_i ,

$$f(\vec{r}) = \sum_{i=1}^{60} c_i f_i(\vec{r}) .$$

L'hamiltonien auquel est soumis chaque électron s'écrit

$$H = \left(-\hbar^2 \frac{\vec{\nabla}^2}{2m} + \sum_i v(\vec{r} - \vec{r}_i) \right) .$$

On va chercher une approximation de ses états propres dans l'espace E en minimisant par rapport aux coefficients c_i le rapport $\frac{\langle f|H|f \rangle}{\langle f|f \rangle}$; on admettra que cela revient à diagonaliser la matrice 60×60

$$h_{ij} = \langle f_i | H | f_j \rangle \quad (7.114)$$

c'est-à-dire à résoudre $\sum_j h_{ij} c_j = \mathcal{E} c_i$. C'est ce dernier problème qui doit être simplifié par la théorie des groupes.

Montrer que E constitue l'espace d'une représentation D de I . Écrire la forme de la matrice de la représentation et évaluer son caractère pour les différentes classes. Montrer qu'elle est isomorphe à la représentation régulière. Quelle est sa décomposition en représentations irréductibles

$$D = \oplus n_\rho D_\rho ?$$

2. On examine maintenant la symétrie σ par rapport à un plan passant par les pôles N et S de l'icosaèdre et par deux sommets diamétralement opposés (voir Fig 1.6) : en termes de quels éléments de I et de $\mathbf{Z}_2$ s'exprime-t-elle ? Combien d'atomes de la molécule laisse-t-elle invariants ?

3. Montrer que l'espace E défini en 1) est aussi l'espace d'une représentation $\hat{D}$ de I_h dont on donnera la valeur du caractère $\hat{\chi}(g)$ en fonction du nombre d'éléments laissés invariants par la transformation g . Quelle est par exemple la valeur de ce caractère pour l'élément σ ?

On se propose de déterminer la décomposition de cette représentation de I_h en représentations irréductibles. Montrer que ce qui précède permet d'écrire

$$\hat{\chi}(g) = \sum_{\rho} n_{\rho+} \chi_{\rho+}(g) + n_{\rho-} \chi_{\rho-}(g) \quad (7.115)$$

pour tout élément g de I_h . Que vaut $n_{\rho+} + n_{\rho-}$? Montrer qu'on peut aisément trouver une combinaison de f_i invariante par I_h . Qu'en déduit-on sur les valeurs de $n_{1\pm}$? En évaluant la somme (7.115) pour g égal successivement à σ (cf. B 2), à τ (cf. A 2) et au produit de τ par une rotation de $2\pi/5$ autour d'un axe joignant une paire de sommets opposés de l'icosaèdre, montrer qu'on peut trouver un nombre suffisant de contraintes pour déterminer complètement les $n_{\rho\pm}$. Vérifier que la solution est

$$n_{3-} = n_{\bar{3}-} = n_{4-} = n_{5-} = 2 .$$

4. Montrer que l'hamiltonien H est invariant sous l'action de I_h et que cela implique la commutation de h (équation (7.114)) avec la matrice de la représentation $\hat{D}$. On suppose qu'on a su calculer les éléments de matrice de h dans une base de E correspondant à la décomposition de $\hat{D}$ en représentations irréductibles. Que peut-on dire de ces éléments de matrice de h ?

Montrer que la théorie des groupes réduit alors la détermination des niveaux de h à un problème de diagonalisation de petites matrices dont on précisera la taille. Le résultat du calcul est indiqué sur la Fig. 7.4, où les niveaux sont désignés par leur représentation $\rho\pm$, les différentes représentations de I étant notées a, b, c, d, e. Noter que deux niveaux désignés par c+ et e+ sont dégénérés. Au vu des résultats précédents, établir la correspondance entre ces a, b, ... e et les représentations étudiées plus haut et en déduire la multiplicité des niveaux.

C.

1. On place maintenant les 60 électrons qui nous occupent sur les niveaux d'énergie les plus bas du spectre obtenu au B 4), chaque niveau ayant deux états de spin. Écrire la forme générale de la fonction d'onde de ces électrons. Sous quelle représentation de I_h se transforme-t-elle ?

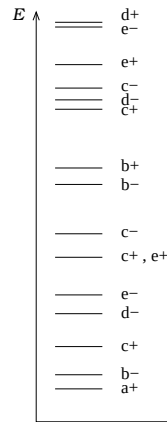


FIGURE 7.4 – Spectre des niveaux de la molécule de fullerène.

2. Le moment électrique dipolaire de la molécule se transforme comme l'opérateur position $\vec{r}$. En utilisant le théorème de Wigner-Eckart, montrer qu'il s'annule dans l'état fondamental étudié en D 1).

3. On considère maintenant l'ion C_{60}^- obtenu en ajoutant un électron sur le premier ou sur le deuxième niveau électronique non occupé du spectre. Que peut-on dire du moment dipolaire ?

7.6.2 Champ cristallin

Ce problème est consacré au déplacement et à la levée partielle de la dégénérescence des niveaux d'un atome (ou d'un ion) placé dans un cristal et soumis aux anisotropies du champ électrique créé par les ions de ce cristal.

1) Soit G un groupe, H un de ses sous-groupes. Montrer que toute représentation D de G fournit une représentation de H . Si D est irréductible comme représentation de G , peut-on affirmer qu'elle l'est aussi comme représentation de H ? Dans la situation qui nous occupe, G désigne le groupe de symétrie de l'Hamiltonien de l'atome isolé, que vous préciserez, H celui de l'atome dans le cristal, qu'on va étudier dans la suite. Montrer que les niveaux de l'atome considéré se scindent selon les représentations irréductibles de H .

2) On suppose six charges identiques q situées aux points de coordonnées $(x = \pm a, y = z = 0)$, $(x = 0, y = \pm a, z = 0)$ et $(x = y = 0, z = \pm a)$ et on considère le potentiel qu'elles créent au point $\vec{r} = (x, y, z)$. Montrer, sans calcul mais par de simples considérations de symétrie, que pour $r = |\vec{r}| \ll a$,

$$V(x, y, z) \approx \frac{q}{4\pi\epsilon_0 a} \left(6 + \alpha \frac{r^2}{a^2} + \frac{\beta}{a^4} (x^4 + y^4 + z^4 - \gamma r^4) + \dots \right)$$

où α , β et γ sont des constantes numériques qu'on ne cherchera pas à déterminer. (En fait, l'équation de Laplace satisfaite par V fixe $\alpha = 0$ et $\gamma = \frac{3}{5}$.)

Pouvez-vous caractériser géométriquement le groupe de symétrie de ce potentiel ?

Plus généralement, on considère le potentiel créé par des charges identiques situées aux sommets d'un réseau cubique régulier. (On ne se souciera pas des questions de convergence dans la définition de ce potentiel.) Discuter sans calcul le groupe de symétrie H de ce potentiel.

3) On considère d'abord les états à un électron de l'atome isolé. On néglige le spin des électrons. Montrer que ces états sont classifiés par leur moment orbital ℓ . Quelle est la dégénérescence des niveaux d'énergie ? Dans la suite du problème on imagine que l'atome est placé sur un nœud du réseau considéré au 2), et soumis au potentiel créé par les autres charges. Montrer que ses niveaux sont maintenant classifiés par des représentations du groupe H . Pour notre propos, on peut oublier l'inversion d'espace (parité) et ne considérer que le sous-groupe de rotations H' de H .

4) Montrer que H' est d'ordre 24. Il a 5 classes qu'on note E , C_2 , C_3 , C_4 et C'_2 , d'ordres respectifs 1, 6, 8, 6 et 3 ; dans cette notation traditionnelle en cristallographie, l'indice k de la classe C_k indique qu'il s'agit de rotations de $\pm \frac{2\pi}{k}$. Pouvez-vous préciser les axes des rotations de ces différentes classes ? (ce dernier point n'interviendra pas dans la suite).

Combien de représentations irréductibles inéquivalentes ce groupe possède-t-il ? On fournit la table de caractères incomplète suivante, dont vous remplirez les cases vides : A_1, A_2, E, T_1 et T_2 sont des notations conventionnelles en physique moléculaire pour ces représentations ; en particulier, A_1 désigne la représentation identité.

$\downarrow$ rep. \ classes $\rightarrow$	E	C_2	C_3	C_4	C'_2
A_1					
A_2					
E	2	0	-1	0	
T_1	3	-1	0	1	
T_2	3	1	0	-1	

5) Calculer la valeur des caractères des représentations de $SO(3)$ de moment orbital ℓ pour les différentes classes de H' . Montrer que la décomposition des représentations de $SO(3)$ selon les représentations irréductibles de H' peut être déterminée grâce à ces caractères. Calculer effectivement comment se scindent les représentations de moment $0 \leq \ell \leq 3$, c'est-à-dire comment les niveaux de départ se scindent en sous-niveaux. En admettant que pour chaque ℓ , les énergies des sous-niveaux s'ordonnent, le cas échéant, selon $\mathcal{E}_{A_2} < \mathcal{E}_{T_1} < \mathcal{E}_{T_2} < \mathcal{E}_E$, dessiner schématiquement la façon dont les niveaux $0 \leq \ell \leq 3$ se scindent en sous-niveaux, en indiquant à chaque fois la dégénérescence.

Chapitre 8

Invariance relativiste. Les groupes de la Relativité

8.1 Principes de la Relativité restreinte. Conséquences physiques

Au début du XX^{ième} siècle, diverses expériences paraissant en contradiction avec les principes de base de la physique conduisent à une refonte de ces principes. Ce sont en particulier les expériences négatives essayant de mettre en évidence la vitesse de la Terre par rapport à l’“éther” (Michelson-Morley). Cette réflexion théorique, après les travaux de Lorentz et Poincaré, culmine dans l’œuvre d’Einstein, dont nous rappelons les grandes lignes.

8.1.1 Référentiels, synchronisation des horloges et simultanéité.

Notions de référentiel : système de coordonnées spatiales et temporelle. Les coordonnées spatiales d’un événement sont mesurées dans un repère (disons orthogonal) par des coordonnées x, y, z (ou x^1, x^2, x^3) à l’aide de “règles” ; sa coordonnée temporelle t est mesurée par une “horloge” au repos dans le référentiel. Règles et horloge sont bien entendu idéalisées, supposées parfaites, ...

Notion de référentiel inertiel, dans lequel une particule matérielle libre (c’est-à-dire soumise à aucune force) garde une vitesse constante (a un mouvement uniforme). On va voir que dès qu’on connaît un tel référentiel¹, tout autre s’en déduisant par une translation uniforme, à vitesse constante, l’est aussi. Les lois de transformation des coordonnées d’un référentiel à l’autre vont devoir être précisées.

Synchronisation des horloges, simultanéité de deux événements. On admet qu’on peut attacher des horloges à chaque point d’un référentiel et les *synchroniser*. Selon Einstein, deux horloges en A et B sont synchronisées si pour un rayon lumineux émis de A vers B , et

1. comme le fameux système de coordonnées “lié à des étoiles fixes”

immédiatement réfléchi vers A , les temps t_A , t_B et t'_A où le rayon atteint A ou B satisfont

$$t_B - t_A = t'_A - t_B .$$

Noter le rôle opérationnel que la lumière joue à ce point dans la définition de la théorie. Dans la conception de la physique classique, la notion de synchronisation sous-entend l'existence de signaux de vitesse de propagation infinie : le temps t est le même en tout point du référentiel.

Cette définition de la synchronisation permet alors de parler d'événements simultanés **dans un référentiel donné** : deux tels événements ont même valeur de t .

8.1.2 Postulats de la Relativité

1. Principe de relativité. Les lois de la physique sont les mêmes dans tout référentiel inertiel. Il y a déjà un principe de relativité en physique classique. Les lois de la *mécanique* y sont les mêmes dans deux référentiels inertiels. c'est-à-dire obtenus l'un par rapport à l'autre par une transformation du groupe de Galilée $\vec{x} \rightarrow \vec{x}' = \mathcal{R}\vec{x} + \mathbf{a} - \vec{v}t$, $t' = t + \Delta t$. Par exemple la loi fondamentale de la dynamique, $\mathbf{F} = m\vec{\ddot{x}}$ est la même dans tout référentiel inertiel, à condition que m soit constante et que $\mathbf{F}$ se transforme comme un vecteur. Ce n'est pas le cas des lois de l'électrodynamique ... Le postulat d'Einstein étend donc ce principe de relativité à toutes les lois de la Physique.

Ce principe de relativité va être réexprimé par la suite comme une invariance des théories physiques sous l'action de transformations de coordonnées. On parle alors d'invariance relativiste ou d'invariance de Lorentz.

2. Vitesse de la lumière. Tout rayon lumineux se propage dans tout référentiel inertiel R avec la même vitesse c , qu'il ait été émis par une source au repos ou en mouvement dans R .

Cette constance de la vitesse de la lumière est une idée neuve d'Einstein, en rupture avec ses contemporains qui s'ingéniaient (en vain) à mettre en évidence l'influence d'un mouvement d'une source par rapport à l'éther sur la vitesse de propagation de la lumière. La valeur de cette constante est

$$c = 299\,792\,458 \text{ ms}^{-1} .$$

Cette valeur est *exacte*. En effet depuis le début des années 60, le mètre est *défini* comme étant la longueur du trajet parcouru dans le vide par la lumière de certains lasers pendant la fraction $1/c$ de seconde. (Actuellement "la seconde est la durée de 9 192 631 770 périodes de la radiation correspondant à la transition entre les niveaux hyperfins $F = 3$ et $F = 4$ de l'état fondamental $6 \text{ S } \frac{1}{2}$ de l'atome de césium 133".) Le choix optimal des radiations à utiliser continue de faire l'objet de travaux en métrologie.

Un ordre de grandeur à retenir : 1 année lumière (1 ly) $\approx 10^{16}$ m (plus précisément $0.9461 \cdot 10^{16}$ m).

8.1.3 Transformations de Lorentz.

Soient deux référentiels R et R' , le second étant doté d'une vitesse $\vec{v}$ par rapport au premier. Quelle transformation supposée *linéaire* des coordonnées $(x, y, z, t) \rightarrow (x', y', z', t')$ est compatible avec les deux principes précédents ?

On suppose qu'au temps $t = t' = 0$, l'origine $x' = y' = z' = 0$ du référentiel R' passe par celle $x = y = z = 0$ de R avec la vitesse $\vec{v}$. Supposons qu'une onde lumineuse sphérique est émise à l'origine au temps $t = t' = 0$. Selon le principe de relativité, les deux observateurs

doivent observer le même phénomène : aux instants ultérieurs, cette onde est localisée sur la sphère d'équation

$$x^2 + y^2 + z^2 = c^2 t^2 \quad (8.1)$$

dans R , et sur $x'^2 + y'^2 + z'^2 = c^2 t'^2$ dans R' , avec la même vitesse c (deuxième postulat). La transformation cherchée doit donc préserver le cône futur $\mathcal{C}_+$ d'équation

$$C(x, y, z, t) := x^2 + y^2 + z^2 - c^2 t^2 = 0, \quad t \geq 0$$

Proposition Les transformations linéaires cherchées préservent non seulement le cône futur $\mathcal{C}_+$ mais aussi la forme quadratique $x^2 + y^2 + z^2 - c^2 t^2$.

Preuve². Soit $Q(x, y, z, t)$ la forme quadratique obtenue en substituant x', y', z', t' fonctions linéaires de x, y, z, t dans la forme $x'^2 + y'^2 + z'^2 - c^2 t'^2$. Quelles sont les formes quadratiques Q telles que $Q = 0$ contienne le cône $x^2 + y^2 + z^2 - c^2 t^2 = 0$?

Lemme Une forme quadratique $Q(x, y, z, t)$ telle que $Q = 0$ contient le cône futur $x^2 + y^2 + z^2 - c^2 t^2 = 0, t \geq 0$ est proportionnelle à la forme $C(x, y, z, t)$

$$Q(x, y, z, t) = K(x^2 + y^2 + z^2 - c^2 t^2),$$

avec K une constante.

Preuve du lemme. Donnons en la preuve en dimension d'espace d : $\vec{x} \in \mathbf{R}^d$. On écrit $Q(\vec{x}, t) = A_{ij}x^i x^j + 2tB_i x^i + Ct^2$ en explicitant les différents termes³. La forme $Q(\vec{x}, t)$ s'annule si $(\vec{x}, t) \in \mathcal{C}_+$, c'est-à-dire $t = \frac{1}{c}\sqrt{\vec{x}^2}$,

$$\forall \vec{x} \in \mathbf{R}^d \quad Q(\vec{x}, \frac{1}{c}\sqrt{\vec{x}^2}) = 0$$

et on trouve donc $A_{ij}x^i x^j + \frac{2}{c}B_i x^i \sqrt{\vec{x}^2} + (\frac{1}{c})^2 C|\vec{x}|^2 = 0$ pour tout $\vec{x}$, ce qui n'est possible que si le terme 'croisé' est nul, $B_i = 0$, tandis que $A_{ij}x^i x^j = -(\frac{1}{c})^2 C|\vec{x}|^2$. La forme Q s'écrit donc

$$Q(\vec{x}, t) = -\frac{C}{c^2}(\vec{x}^2 - c^2 t^2)$$

ce qui établit le lemme.

Revenant à la proposition, cela implique que les transformations linéaires cherchées préservent la forme quadratique $x^2 + y^2 + z^2 - c^2 t^2$, à un facteur près indépendant des coordonnées, ne pouvant donc dépendre que de la vitesse $\vec{v}$

$$x'^2 + y'^2 + z'^2 - c^2 t'^2 = \varphi(\vec{v})(x^2 + y^2 + z^2 - c^2 t^2).$$

Finalement, le facteur $\varphi(\vec{v})$ qui satisfait $\varphi(\vec{v})\varphi(-\vec{v}) = 1$ (par composition de la transformation avec son inverse), $\varphi(\vec{v}) = \varphi(-\vec{v})$ (par symétrie de retournement des axes $\vec{x} \rightarrow -\vec{x}$), $\varphi(\mathbf{0}) = 1$ et est une fonction continue de $\vec{v}$, vaut 1.

2. Je dois cette preuve à Michel Bauer.

3. Ici et dans toute la suite, convention de sommation sur les indices répétés.

Les transformations des coordonnées cherchées sont donc les transformations linéaires laissant invariantes la forme quadratique $x^2 + y^2 + z^2 - c^2 t^2$. Elles forment un groupe, le groupe de Lorentz. Nous avons donc le

Théorème *Les transformations linéaires de coordonnées entre deux référentiels inertiels sont les transformations de Lorentz, qui préservent la forme quadratique $x^2 + y^2 + z^2 - c^2 t^2$.*

Remarques et exercices.

- 1. Plutôt que l’argument géométrique précédent, on peut préférer l’argument physique d’Einstein (voir aussi Panofsky et Phillips) qui *construisent* les transformations de Lorentz par une suite d’expériences de pensée.
- 2. Peut-on relâcher l’hypothèse de linéarité? Autrement dit, y-a-t-il des transformations non linéaires $x \rightarrow x'$ (non singulières, c’est-à-dire inversibles en tout point), qui préservent le cône? y-en-a-t-il qui préservent la forme quadratique? (éléments de réponse dans [Wg, p27].)

Donnons l’expression explicite de ces transformations⁴. Elles comprennent bien sûr les rotations d’espace usuelles : $\vec{x} \rightarrow \mathcal{R}_{\vec{n}}(\theta)\vec{x}$, $t \rightarrow t$ est la rotation d’angle θ autour d’un axe $\vec{n}$. Ainsi la rotation d’axe $\hat{z}$

$$\mathcal{R}_{\hat{z}}(\theta) \left\{ \begin{array}{l} x' = x \cos \theta + y \sin \theta \\ y' = -x \sin \theta + y \cos \theta \\ z' = z \quad t' = t \end{array} \right. . \quad (8.2)$$

Les “transformations de Lorentz spéciales” (ou *boosts*), qui décrivent un changement de coordonnées entre deux référentiels animés d’une vitesse relative $\vec{v} = v\vec{n}$, disons de R' par rapport à R , en sont l’analogie en géométrie hyperbolique. Ainsi pour une vitesse le long de l’axe des x

$$L_{\hat{x}}(\beta) \left\{ \begin{array}{l} x' = x \cosh \beta - ct \sinh \beta \\ ct' = -x \sinh \beta + ct \cosh \beta \\ y' = y, \quad z' = z \end{array} \right. \quad (8.3)$$

où l’angle hyperbolique β , appelé aussi *rapidité*, est défini par

$$\cosh \beta = \gamma := \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \tanh \beta = \frac{v}{c} . \quad (8.4)$$

Ces relations prennent donc aussi la forme

$$\begin{aligned} x' &= \gamma(x - vt) & y' &= y & z' &= z & t' &= \gamma\left(t - \frac{v}{c^2}x\right) \\ \text{ou inversement} & & & & & & & \\ x &= \gamma(x' + vt') & y &= y' & z &= z' & t &= \gamma\left(t' + \frac{v}{c^2}x'\right) . \end{aligned} \quad (8.5)$$

En général pour une vitesse $\vec{v}$ de R' par rapport à R quelconque

$$\vec{x}'_{\parallel} = \gamma(\vec{x}_{\parallel} - \vec{v}t) \quad \vec{x}'_{\perp} = \vec{x}_{\perp} \quad t' = \gamma\left(t - \frac{1}{c^2}\vec{v} \cdot \vec{x}\right) \quad (8.6)$$

$$\text{avec} \quad \vec{x}_{\parallel} := \vec{v} \frac{\vec{v} \cdot \vec{x}}{v^2} \quad \vec{x}_{\perp} := \vec{x} - \vec{x}_{\parallel} , \quad (8.7)$$

4. ou tout au moins de celles qui sont “connectées à l’identité”, voir la discussion plus bas.

où par définition $\vec{x}_{\parallel}$ est colinéaire à $\vec{v}$ tandis que $\vec{x}_{\perp}$ lui est orthogonal, et donc satisfait $\vec{x}_{\parallel} \cdot \vec{x}_{\perp} = 0$.)

Nous reviendrons plus bas sur ces transformations de Lorentz (§ 2.1).

8.1.4 Conséquences des transformations de Lorentz. Cinématique relativiste

Première conséquence des transformations de Lorentz : *le temps perd son caractère absolu*. En particulier, la simultanéité de deux événements s est une notion qui est liée à un référentiel donné : pour deux événements s , on peut avoir $\Delta t = 0$ dans R mais $\Delta t' \neq 0$ dans R' .

- *Dilatation des temps*. Une horloge placée (disons à l'origine) dans un référentiel inertiel R' en mouvement paraît retarder par un facteur γ quand elle est observée dans R : selon les formules (8.5), les battements successifs en $t' = 0, 1, 2, \dots$ secondes paraissent dans R avoir lieu en $t = 0, \gamma, 2\gamma, \dots$. Bien sûr, il y a symétrie et une horloge au repos dans R paraît retarder vue de R' .

Cette dilatation du temps s'observe couramment en physique des particules. Les muons $\mu^{\pm}$, (des leptons chargés, c'est-à-dire des analogues environ 207 fois plus massifs de l'électron et du positron $e^{\pm}$), sont produits dans la haute atmosphère par les rayons cosmiques (dont les protons créent par collisions avec les noyaux de l'air des mésons π , qui se désintègrent très rapidement en $\mu^{\pm}$). Au repos, ces $\mu^{\pm}$ se désintègrent rapidement, avec une demi-vie de $2.19 \cdot 10^{-6}$ s. Pendant un temps aussi court, une particule ultrarelativiste, de vitesse $\approx c \approx 3 \cdot 10^8$ m parcourt environ 600 m. Or ces muons constituent une bonne part du rayonnement cosmique reçu sur terre, au niveau de la mer. Cela est rendu possible par la dilatation de leur temps de vie par un facteur γ . Pour $v = (1 - \epsilon)c$, $\epsilon \ll 1$, $\gamma \approx 1/\sqrt{2\epsilon}$; exemple $\epsilon = 5 \cdot 10^{-3}$, $\gamma = 10$.

D'autres vérifications expérimentales très précises sont obtenues au laboratoire avec des mesures d'effet Doppler sur des ions ... (cf revue de C.M. Will citée dans les références).

- *Notion de temps propre*. La notion de dilatation du temps peut aussi s'appliquer à un mouvement non uniforme. Soit un point M animé d'un mouvement donné dans un référentiel R "au repos" ; sa trajectoire est donnée par $\vec{x}(t)$ dans les coordonnées de R . On suppose M doté d'une horloge qui lui est liée. Pendant un temps dt infinitésimal, la position de M dans R change de $d\vec{x}$, et on peut assimiler le mouvement de M à une translation de vitesse $\vec{v}(t) = d\vec{x}(t)/dt$. On peut donc attacher de façon instantanée un référentiel inertiel $R'(t)$ à M , doté par rapport à R de la vitesse $\vec{v}(t)$: on pourrait parler du référentiel inertiel tangent à la trajectoire de M . Pendant ce temps infinitésimal, l'horloge liée à M avance de dt' (et sa coordonnée ne change pas, $d\vec{x}' = 0$). Calculant l'élément de longueur invariante dans R et dans R' , on trouve

$$ds^2 = c^2 dt'^2 = c^2 dt^2 - d\vec{x}^2 = (c^2 - \vec{v}^2) dt^2$$

donc

$$d\tau := \frac{1}{c} ds = dt' = dt \sqrt{1 - \frac{\vec{v}^2}{c^2}} < dt . \quad (8.8)$$

C'est une nouvelle manifestation du ralentissement de l'horloge en mouvement par rapport à une horloge au repos. Le temps t' de l'horloge liée au mouvement, qu'on notera désormais τ ,

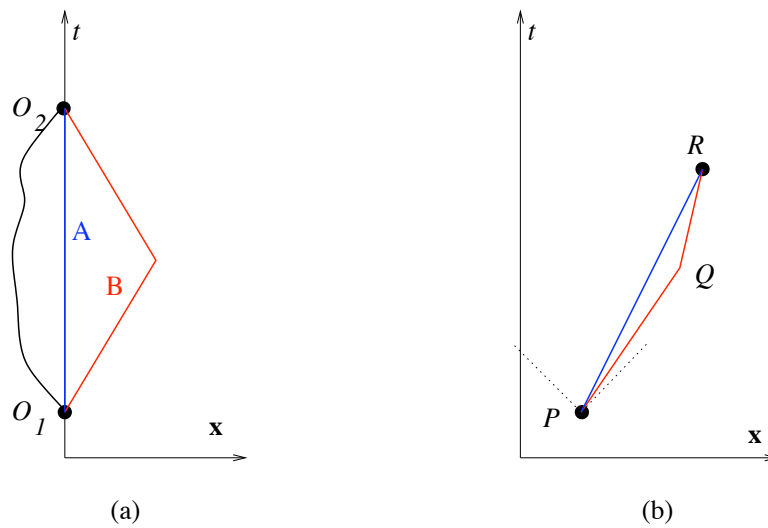


FIGURE 8.1 – (a) : trois trajectoires de genre temps entre O_1 et O_2 : la ligne droite donne le chemin le plus long en temps! (b) : l’inégalité triangulaire inversée pour des points P, Q, R de séparations mutuelles de genre temps.

est appelé le *temps propre* du mobile M . En intégrant le long de la trajectoire parcourue entre les temps (de R) t_1 et t_2 , on voit que

$$\Delta\tau = \tau_2 - \tau_1 = \int_{t_1}^{t_2} \sqrt{1 - \frac{v^2}{c^2}} dt \leq (t_2 - t_1) .$$

On a donc obtenu le résultat que l’intervalle de temps propre d’un mobile est toujours inférieur à l’intervalle de temps mesuré par un observateur au repos.

Si le mobile parcourt une trajectoire et recroise celle de l’observateur au repos, on peut comparer les laps de temps écoulés entre les deux croisements, et on trouve $\Delta\tau < t_2 - t_1$. Par conséquent, dans la géométrie hyperbolique de l’espace-temps, la trajectoire rectiligne de l’observateur au repos fournit le temps le plus long entre deux “points” (événements), supérieur à celui de toute trajectoire curviligne, voir figure~1. Autrement dit, parmi toutes les trajectoires entre deux points O_1 et O_2 , le mouvement uniforme donne celle qui maximise $\int_1^2 d\tau$ (ou encore $\int_1^2 ds$) (cf ci-après le principe variationnel d’action).

• *Le “paradoxe” des jumeaux de Langevin.* Ce fameux paradoxe se rattache au même ordre d’idées. Le jumeau A est au repos, à l’origine du référentiel R . À l’instant $t = 0$, son jumeau B le quitte à bord d’un engin spatial se déplaçant à une vitesse v d’ordre c . Au bout d’un temps t' (mesuré dans son référentiel), B fait “demi-tour” et repart immédiatement à la même vitesse vers la Terre, qu’il atteint au bout du temps $2t'$ (temps de R'). Quel est le temps de l’aller-retour du point de vue de A ? Pour A, l’horloge de B retarde d’un facteur γ et le temps écoulé mesuré dans R est donc $\Delta t = 2t'\gamma$. A a vieilli plus vite que son jumeau astronaute ! Pour $v = 0.6 c$, donc $\gamma = 1.25$, et $t' = 2$ ans, on trouve $\Delta t' = 4$ ans, $\Delta t = 2\gamma t' = 5$ ans.

Jusque là, tout est en accord avec la discussion précédente de dilatation du temps. Le

paradoxe (apparent !) consiste à inverser les rôles de A et B. Dans le référentiel de B, A s'éloigne à la vitesse $-\vec{v}$, etc. Pourquoi devrait-il y avoir une dissymétrie entre A et B ?

Il n'y a pas symétrie parce que A est un observateur inertiel tandis que B ne l'est pas. Comme on l'a discuté plus haut, la ligne brisée parcourue par B entre ses deux rencontres avec A fournit un intervalle de temps inférieur à celui de la ligne droite suivie par A (voir figure 1(a)).

• *Contraction des distances.* On parle aussi de contraction des règles, ou de contraction de Lorentz. C'est la contrepartie spatiale du phénomène de dilatation du temps. Une "règle" de longueur L au repos dans le référentiel R' a pour coordonnées de ses extrémités x'_A et $x'_B = x'_A + L$, mesurées au même temps $t'_A = t'_B$. Dans le référentiel R , on mesure en un même temps $t_A = t_B$ des coordonnées x_A, x_B , et selon la première équation (8.5), $x' = \gamma(x - vt)$, donc

$$L = |x'_A - x'_B| = \gamma|x_A - x_B| \quad \text{donc} \quad |x_A - x_B| = L/\gamma < L. \quad (8.9)$$

Il y a donc *contraction* de la règle, quand elle est observée dans le référentiel R .

En physique des particules, le phénomène se manifeste dans les expériences de très haute énergie. Des particules "ultrarelativistes" (c'est-à-dire accélérées à des vitesses proches de c), qu'on décrirait au repos comme de petites sphères, apparaissent plutôt comme des galettes aplaties dans leurs collisions. . .

La description de certains phénomènes impliquant des hadrons (protons, neutrons, . . .) de très haute énergie fait appel à la notion de *partons*, sous-constituants de ces hadrons, dotés de mouvements internes à ce hadron. Dans l'image où le hadron est aplati par l'effet précédent, seules comptent les impulsions *transverses* des partons par rapport à celle du hadron.

• *Composition des vitesses.* En différentiant (8.6) on trouve immédiatement la loi de transformation des vitesses $\vec{V} = \frac{d\vec{x}}{dt} \mapsto \vec{V}' = \frac{d\vec{x}'}{dt'}$

$$\vec{V}'_{\parallel} = \frac{\vec{V}_{\parallel} - \vec{v}}{1 - \frac{\vec{v} \cdot \vec{V}}{c^2}}, \quad \vec{V}'_{\perp} = \sqrt{1 - \frac{v^2}{c^2}} \frac{\vec{V}_{\perp}}{1 - \frac{\vec{v} \cdot \vec{V}}{c^2}}. \quad (8.10)$$

On en déduit que $\vec{V}^2 = \vec{V}_{\perp}^2 + \vec{V}_{\parallel}^2$ est tel que

$$\left(1 - \frac{\vec{V}^2}{c^2}\right) \left(1 - \frac{\vec{v} \cdot \vec{V}}{c^2}\right)^2 = \left(1 - \frac{v^2}{c^2}\right) \left(1 - \frac{\vec{V}^2}{c^2}\right),$$

une relation compatible avec le fait que c est la limite supérieure des vitesses observables, et que la vitesse de la lumière est c quel que soit le référentiel d'observation (Postulat 2).

Application. Considérons l'expérience de Fizeau où la lumière se propage dans un milieu d'indice n , au repos dans un référentiel R' . La vitesse de la lumière observée dans R' est donc $V' = c/n$. Si R' se déplace avec une vitesse $\vec{v}$ (disons le long des x) par rapport à R , la vitesse observée dans R est (composante x)

$$V = \frac{(V' + v)}{(1 + \frac{v}{cn})} \approx \frac{c}{n} + v(1 - \frac{1}{n^2}) + O(v^2).$$

C'est là une formule trouvée (par un raisonnement tout différent !) par Fresnel (1818) et vérifiée expérimentalement par Fizeau (1851).

8.2 Espace de Minkowski. Groupe de Lorentz $O(3,1)$.

8.2.1 Métrique de l'espace de Minkowski. Groupes de Lorentz et de Poincaré

On peut résumer ce qui précède en disant que l'espace à trois dimensions de la physique classique (augmenté d'un temps universel) a été remplacé par un *espace-temps*, espace réel de dimension 4, doté d'une métrique indéfinie de signature $(+, -, -, -)$, l'espace de Minkowski $\mathcal{M} = \mathbf{R}^4$. Les vecteurs V de cet espace sont appelés *quadrivecteurs*. Dans un repère orthonormé, le produit scalaire de deux quadrivecteurs V et W de coordonnées $V = (V^0, \vec{V})$ et $W = (W^0, \vec{W})$ est donc ⁵

$$(V, W) = V.W = V^0 W^0 - \vec{V}.\vec{W} = g_{\mu\nu} V^\mu W^\nu \quad (8.11)$$

avec le tenseur métrique ⁶

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \quad (8.12)$$

Un cas particulièrement important de quadrivecteur est le vecteur position x avec des coordonnées $x^0 = ct$ et $\vec{x}$. En prenant la norme carrée de sa différentielle dx , on obtient le carré de l'élément de longueur

$$ds^2 = (dx, dx) = c^2 dt^2 - d\vec{x}^2 = g_{\mu\nu} dx^\mu dx^\nu . \quad (8.13)$$

La signature indéfinie de cette métrique a des conséquences géométriques inhabituelles, comme l'inégalité triangulaire inversée déjà mentionnée plus haut (§ 1.4). Anticipant un peu sur une terminologie précisée au § 2.3, si P , Q et R sont trois points d'espace-temps séparés par des intervalles de genre temps, avec Q dans le "cône futur" de P , et R dans ceux de P et Q , voir figure 1(b), alors leurs distances minkowskienne satisfont $d(P, R) \geq d(P, Q) + d(Q, R)$.

Le groupe d'invariance du produit scalaire (8.11) est le groupe de Lorentz $\mathcal{L} = O(1, 3)$, c'est-à-dire le sous-groupe du groupe linéaire, $V \mapsto \Lambda V$, constitué des matrices réelles $\Lambda = (\Lambda^\mu_\rho)$ satisfaisant

$$\Lambda^T g \Lambda = g , \quad (8.14)$$

ou, en composantes, $V^\mu \mapsto \Lambda^\mu_\rho V^\rho$,

$$\Lambda^\mu_\rho \Lambda^\nu_\sigma g_{\mu\nu} = g_{\rho\sigma} . \quad (8.15)$$

5. On utilisera la notation abrégée $V.W$ toutes les fois qu'elle est non ambiguë, sinon (V, W) . La composante en caractère gras $\vec{V}$ ou surmontée d'une flèche, $\vec{V}$, désigne la partie spatiale du vecteur, de coordonnées cartésiennes V^i , $i = 1, 2, 3$. Par convention, les indices grecs se réfèrent aux coordonnées minkowskienne, $\mu, \nu, \dots = 0, 1, 2, 3$, les indices latins aux coordonnées spatiales $i, j, \dots = 1, 2, 3$.

6. La notation $g_{\mu\nu}$ pour ce tenseur est remplacée par $\eta_{\mu\nu}$ dès qu'on aborde la relativité générale : g désigne alors la métrique non triviale couplée à la matière, tandis que η désigne la métrique minkowskienne de référence.

Il contient comme sous-groupe le groupe de rotations $O(3)$ de l'espace $\mathbb{R}^3$. Il contient aussi toute transformation de Lorentz spéciale de rapidité β le long d'une direction $\vec{n}$ ($\vec{n}$ un vecteur unitaire) (on rappelle que $\vec{x}_{\parallel} := \vec{x} \cdot \vec{n} \vec{n}$, $\vec{x}_{\perp} = \vec{x} - \vec{x}_{\parallel}$)

$$L_{\vec{n}}(\beta) \begin{cases} \vec{x}'_{\parallel} &= \vec{x}_{\parallel} \cosh \beta - x_0 \vec{n} \sinh \beta \\ x'_0 &= -\vec{x}_{\parallel} \cdot \vec{n} \sinh \beta + x_0 \cosh \beta \\ \vec{x}'_{\perp} &= \vec{x}_{\perp} , \end{cases} \quad (8.16)$$

une formule qui généralise (8.3).

Les transformations de Lorentz spéciales forment-elles un groupe? Montrer que les transformations de Lorentz spéciales de vecteur directeur $\vec{n}$ fixé forment un groupe dont les rapidités se composent de façon additive : $L_{\vec{n}}(\beta)L_{\vec{n}}(\beta') = L_{\vec{n}}(\beta')L_{\vec{n}}(\beta) = L_{\vec{n}}(\beta + \beta')$.

Le groupe de Lorentz a quatre composantes connexes (on dit aussi “nappes”). En effet, en prenant le déterminant de (8.14), on apprend que $\det \Lambda = \pm 1$ et en prenant l'élément de matrice 00 de (8.15), on lit $1 = g_{\mu\nu} \Lambda^\mu_0 \Lambda^\nu_0 = (\Lambda^0_0)^2 - \sum_i (\Lambda^i_0)^2$, donc $|\Lambda^0_0| > 1$, Λ^0_0 positif ou négatif. Les quatre nappes correspondent aux quatre choix de signes de $(\det \Lambda, \Lambda^0_0)$, la nappe $(+, +)$ constituant le sous-groupe $\mathcal{L}^{\uparrow}_+$ des transformations de Lorentz “propres, orthochrones”.

La parité d'espace $\Lambda_P^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$, le renversement du temps $\Lambda_T = -\Lambda_P$ et leur

produit $\Lambda_P \Lambda_T = -I$ fournissent des représentants des trois autres nappes; en les composant avec un élément quelconque de $\mathcal{L}^{\uparrow}_+$, on obtient l'élément générique de chacune de ces nappes. Toute transformation de $\mathcal{L}^{\uparrow}_+$ s'écrit comme produit d'une rotation “ordinaire” de $SO(3)$ par une transformation de Lorentz spéciale.

Le groupe de Poincaré $\mathcal{P}$, appelé aussi groupe de Lorentz inhomogène, est le groupe engendré par $\mathcal{L}$ et les translations d'espace-temps. Plus précisément, c'est le “produit semi-direct” $\mathbf{R}^4 \rtimes \mathcal{L}$ du groupe $\mathbf{R}^4$ des translations par $\mathcal{L}$, puisque si on note (a, Λ) la transformation $x' = \Lambda x + a$, la loi de composition est (le vérifier)

$$(a', \Lambda') \cdot (a, \Lambda) = (a' + \Lambda' a, \Lambda' \Lambda) ,$$

avec une action de $\mathcal{L}$ dans le groupe des translations $\mathbf{R}^4$. Quel est l'inverse de (a, Λ) ?

Nous reviendrons par la suite sur ces groupes, leurs propriétés et leurs “représentations”.

8.2.2 Vecteurs et tenseurs contravariants et covariants.

La métrique indéfinie impose un peu de soin dans la manipulation des quadrivecteurs et de leurs composantes. Les objets se transformant selon

$$V^\mu \mapsto \Lambda^\mu{}_\nu V^\nu \quad (8.17)$$

sont appelés vecteurs *contravariants*, tandis que

$$U_\mu \mapsto (\Lambda^{-1})^\nu{}_\mu U_\nu = ((\Lambda^T)^{-1})^\nu{}_\mu U_\nu \quad (8.18)$$

définit un vecteur *covariant*. Un objet tensoriel peut être doté d'indices des deux types, covariant et contravariant. (Exemples : le tenseur de Kronecker δ_μ^ν , égal à 1 si $\mu = \nu$, à 0 sinon, est un tenseur invariant (pourquoi?). Le tenseur métrique $g_{\mu\nu}$ est un tenseur deux fois covariant, en fait invariant par définition de $O(1, 3)$ (cf (8.15)).

Si V est un vecteur contravariant de composantes V^μ , $V_\mu := g_{\mu\nu}V^\nu$ est sa version covariante, et on peut écrire le produit scalaire $V.W = V_\mu W^\mu = V_0W^0 + V_1W^1 + V_2W^2 + V_3W^3$. Attention au changement de signe des composantes spatiales $V_i = -V^i$, $i = 1, 2, 3$, tandis que $V_0 = V^0$. Inversement on passe des composantes covariantes aux contravariantes par multiplication par l'inverse de g , noté $g^{\mu\nu}$:

$$g_{\mu\nu}g^{\nu\lambda} = \delta_\mu^\lambda, \quad (8.19)$$

$$V^\mu = g^{\mu\nu}V_\nu. \quad (8.20)$$

On vérifie immédiatement que $g^{\mu\nu}$ est un tenseur deux fois contravariant, égal dans une base orthonormée à $g_{\mu\nu}$.

Exemples de vecteur covariant : le gradient

$$\partial = (\partial_\mu \equiv \frac{\partial}{\partial x^\mu}) = (\frac{\partial}{\partial x^0}, \vec{\nabla}) \quad (8.21)$$

ou, (cf infra) le potentiel électromagnétique A .

La notation avec indices covariants et contravariants permet d'identifier aisément les invariants (sous l'action des transformations (8.17) et (8.18)). Ainsi si $A(x)$ se transforme comme un quadrivecteur, $\partial A \equiv \partial_\mu A^\mu = \frac{1}{c} \frac{\partial A^0}{\partial t} + \vec{\nabla} \cdot \vec{A}$ est invariant, de même que $A_\mu(x)dx^\mu$, et donc aussi $\int_{s_1}^{s_2} A_\mu(x(s)) \frac{dx^\mu}{ds} ds$.

• *Tenseur complètement antisymétrique $\epsilon^{\mu\nu\rho\sigma}$ de Levi-Civita.* Le tenseur complètement antisymétrique à trois indices ϵ^{ijk} , invariant par l'action du groupe des rotations $SO(3)$, a un analogue à 4 dimensions, noté $\epsilon^{\mu\nu\rho\sigma}$. Il est défini par sa propriété de complète antisymétrie (donc toute composante avec deux indices identiques s'annule) et par la donnée de $\epsilon^{0123} = 1$. Noter que sa version covariante est telle que $\epsilon_{0123} = -1$, et que $\epsilon^{0ijk} = \epsilon^{ijk} = \epsilon_{ijk}$.

Exercices 1. Montrer que le tenseur $\epsilon^{\mu\nu\rho\sigma}$ est en fait *invariant* par $\mathcal{L}_+^\uparrow$, c'est-à-dire que $\Lambda_\mu^\mu \Lambda_\nu^\nu \Lambda_\rho^\rho \Lambda_\sigma^\sigma \epsilon^{\mu'\nu'\rho'\sigma'} = \epsilon^{\mu\nu\rho\sigma}$ si $\Lambda \in \mathcal{L}_+^\uparrow$. (Pour cela, se rappeler la définition du déterminant). 2. Soit q un quadrivecteur. Montrer que le tenseur

$$q_\mu \epsilon_{\nu\rho\sigma\tau} + q_\nu \epsilon_{\rho\sigma\tau\mu} + q_\rho \epsilon_{\sigma\tau\mu\nu} + q_\sigma \epsilon_{\tau\mu\nu\rho} + q_\tau \epsilon_{\mu\nu\rho\sigma}$$

est complètement antisymétrique en $(\mu, \nu, \rho, \sigma, \tau)$. Que peut-on dire d'un tel tenseur dans un espace à quatre dimensions? Les identités de ce type, parfois appelées "circulantes", sont utiles quand on cherche à construire une base d'invariants. 3. De même qu'on a les identités $\epsilon^{ijk}\epsilon_{i'j'k} = \delta_i^{i'}\delta_j^{j'} - \delta_i^{j'}\delta_j^{i'}$, $\epsilon^{ijk}\epsilon_{i'jk} = 2\delta_i^{i'}$ et $\epsilon^{ijk}\epsilon_{ijk} = 6$, on a $\epsilon^{\mu\nu\rho\sigma}\epsilon_{\mu\nu\rho\sigma} = -24$, $\epsilon^{\mu\nu\rho\sigma}\epsilon_{\mu'\nu\rho\sigma} = -6\delta_\mu^{\mu'}$. Que valent $\epsilon^{\mu\nu\rho\sigma}\epsilon_{\mu'\nu'\rho\sigma}$? $\epsilon^{\mu\nu\rho\sigma}\epsilon_{\mu'\nu'\rho'\sigma}$? Trouver une expression déterminantale pour $\epsilon^{\mu\nu\rho\sigma}\epsilon_{\mu'\nu'\rho'\sigma'}$.

8.2.3 Géométrie de l'espace de Minkowski, temps propre, lignes d'univers.

Dans la suite "vecteur" signifie quadrivecteur.

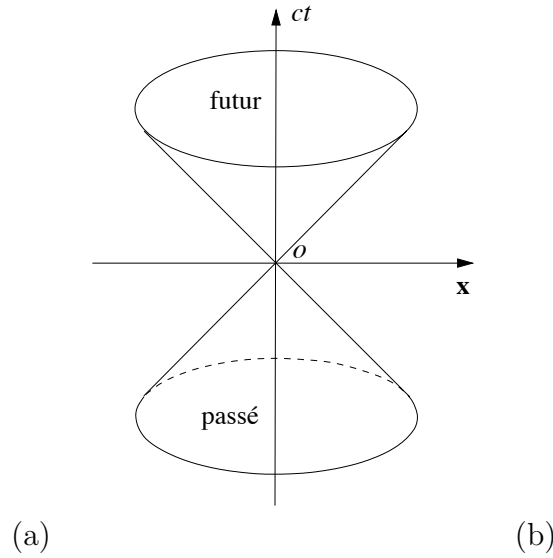


FIGURE 8.2 – Cône de lumière.

Un vecteur de norme carrée > 0 , $= 0$, < 0 est dit respectivement de genre temps, lumière et espace.

Tout vecteur non nul W orthogonal à un vecteur V de genre temps est de genre espace. Si $V \cdot W = 0$ et $V^2 > 0$, on peut en effet trouver un repère où $V = (V^0, \mathbf{0})$, et alors nécessairement $W^0 = 0$, donc $W^2 = -\vec{W}^2 < 0$.

La distance minkowskienne entre deux points A et B de coordonnées x et y est définie par $d^2(A, B) = (x - y)^2 = (x^0 - y^0)^2 - (\vec{x} - \vec{y})^2$. Plaçons un de ces points à l'origine. Le *cône de lumière* (light cone) sépare l'ensemble des points séparés de l'origine par un intervalle de genre espace de ceux séparés par un intervalle de genre temps. Ces notions sont invariantes par transformation de Lorentz. On distingue le cône futur du cône passé.

Une notion importante est celle de *causalité*. Selon le principe que la vitesse de la lumière est la plus grande qu'on puisse communiquer à un corps matériel ou à un signal, un événement ayant lieu par exemple à l'origine O de l'espace-temps ne peut influencer causalement que les points de l'espace-temps situés dans le cône futur de O , et ne peut avoir reçu d'influence causale que des points situés dans son cône passé (voir figure 2). Autrement dit, il est certain que deux événements x et y séparés par un intervalle de genre espace ne peuvent pas s'être influencés causalement dans le passé, ni s'influencer dans le futur, puisque cela nécessiterait un signal se propageant à une vitesse $|\vec{x} - \vec{y}|/(x_0 - y_0) > c$.

Attention, si le cône est globalement invariant sous l'effet des transformations de Lorentz, seules les transformations telles que $\Lambda_0^0 > 0$ préservant le sens du temps conservent le cône futur...

Dans tout espace (pseudo-)riemannien, donc doté d'une métrique $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ dans des coordonnées x^μ , avec $g := \det g_{\mu\nu}$, et $g^{\lambda\mu}$ le tenseur inverse de $g_{\mu\nu}$, on peut construire un

laplacien et un élément de volume invariants par changement de coordonnées

$$\Delta = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^\mu} g^{\mu\nu} \sqrt{|g|} \frac{\partial}{\partial x^\nu} \quad (8.22)$$

$$dV = \sqrt{|g|} \prod_{\mu} dx^\mu . \quad (8.23)$$

C'est le cas de l'espace de Minkowski, pour lequel le laplacien, aussi appelé le d'Alembertien et traditionnellement noté $\square$, est simplement

$$\square = g^{\mu\nu} \partial_\mu \partial_\nu = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \quad (8.24)$$

avec Δ le laplacien usuel à 3 dimensions. C'est bien sûr l'opérateur différentiel familier dans l'équation des ondes. Le d'Alembertien et l'élément de volume $dV = c dt d^3\vec{x}$ sont invariants par les transformations de Lorentz.

8.3 Formalisme lagrangien et relativité. Cinématique relativiste

On a rappelé plus haut au Chapitre 4 ce que sont le principe d'action et les équations d'Euler-Lagrange. Examinons maintenant comment ce principe peut être formulé en accord avec l'invariance relativiste.

8.3.1 Particule libre relativiste

Utilisant des coordonnées d'espace-temps $x = (x^\mu)$, cherchons quelle action ne dépendant que de x et de sa dérivée première et invariante relativiste décrit une particule libre. Il est clair que le choix de lagrangien $L = \text{const.} \frac{d\tau}{dt}$ où la constante est fixée par la limite non-relativiste, remplit les deux conditions précédentes :

$$L = -mc^2 \frac{d\tau}{dt} = -mc^2 \sqrt{1 - \frac{\vec{v}^2}{c^2}} \quad (8.25)$$

dont le développement à petit v donne bien

$$L \approx -mc^2 + \frac{1}{2} m \vec{v}^2 + \dots$$

c'est-à-dire le lagrangien classique non-relativiste, à une constante additive sans importance près. L'action entre les temps t_1 et t_2 est donc $I = \int_{t_1}^{t_2} dt L = -mc \int_{s_1}^{s_2} ds$. Cette action est invariante de Lorentz, ds l'étant. Le signe $-$ garantit que cette action est *minimale* pour un mouvement uniforme entre deux points (cf discussion plus haut au sujet du temps propre). Sa variation conduit à l'équation du mouvement

$$\frac{d^2 x^\mu}{ds^2} = 0 , \quad (8.26)$$

qui est bien satisfaite par un mouvement uniforme.

Remarques

- - Montrer que cette action peut se mettre sous la forme $\int mc ds = \int d^4x \frac{ds}{dt} \mu(\vec{x})$ à condition d'écrire $m = \int d^3\vec{x} \mu(x)$ en termes d'une densité de masse. En discuter l'invariance de Lorentz.
- - Montrer que l'action $S = -\frac{1}{2}mc \int_1^2 ds(dx/ds)^2$, où s est un paramètre le long de la trajectoire, conduit à la bonne équation du mouvement, et que s s'identifie alors à la longueur invariante le long de la trajectoire.

À partir de (8.25), l'impulsion relativiste et l'énergie se calculent comme en (4.19-4.21) :

$$\vec{p} = \frac{\partial L}{\partial \vec{v}} = \frac{m\vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \quad E = \vec{p} \cdot \vec{v} - L = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} . \quad (8.27)$$

Le développement de E dans la limite non relativiste est important

$$E = mc^2 + \frac{1}{2}m\vec{v}^2 + \dots \quad (8.28)$$

puisqu'il présente le terme d'énergie de repos plus le terme classique d'énergie cinétique.

Les expressions (8.27) coïncident avec celles qu'on obtient à partir du *quadrivecteur vitesse* u^μ . Ce dernier est défini par dérivation de x^μ par rapport au temps propre

$$u^\mu = \frac{dx^\mu}{d\tau} = \left(c \frac{dt}{d\tau}, \frac{d\vec{x}}{d\tau} \right) = \frac{dt}{d\tau} (c, \vec{v}) . \quad (8.29)$$

Calculons sa norme carrée : $u^2 = c^2$ puisque $(d\tau/dt)^2 = 1 - v^2/c^2$. Le quadrivecteur vitesse u est de genre temps et de longueur constante. Le quadrivecteur accélération défini par $du^\mu/d\tau$ lui est donc orthogonal, et par conséquent de genre espace (cf début du § 2.3). On définit finalement la quadri-impulsion

$$p^\mu = mu^\mu = \left(\frac{1}{c} E, \vec{p} \right) \quad (8.30)$$

en accord avec les expressions ci-dessus de $\vec{p}$ et E . Cette quantité est de grande importance physique, puisque l'énergie et l'impulsion totales d'un système isolé sont conservées (dans un référentiel donné).

Récapitulons : pour une particule libre relativiste

$$p^\mu = mu^\mu = \left(\frac{1}{c} E, \vec{p} \right) = \frac{(mc, m\vec{v})}{\sqrt{1 - \vec{v}^2/c^2}} \quad (8.31)$$

$$\vec{p} = \frac{E\vec{v}}{c^2} \quad (8.32)$$

$$p^2 = p_\mu p^\mu = m^2 c^2 = \frac{E^2}{c^2} - \vec{p}^2 . \quad (8.33)$$

La dernière relation exprime que la quadri-impulsion est sur la "couche de masse" (*mass shell*), c'est-à-dire sur l'hyperboloïde $p^2 = m^2 c^2$. On peut encore relire cette dernière relation comme $E^2 = (mc^2)^2 + (c\vec{p})^2$, une autre manifestation du bilan d'énergie entre énergie de masse et énergie cinétique.

En physique des particules, les énergies sont couramment mesurées en unités d'électrons-volts (rappelons qu'1 eV = $1.6 \cdot 10^{-19}$ Joule) et dans ses multiples keV, MeV, GeV, TeV, ... en attendant le PeV (Peta-eV = 10^{15} eV). ... Selon les expressions ci-dessus, une impulsion peut alors se mesurer en XeV/c, une masse en XeV/c². Par abus de langage, on oublie souvent le facteur c et on parle par exemple de masse de ≈ 91 GeV pour le boson intermédiaire Z ...

• *Élément de volume sur la couche de masse.* La couche de masse $p^2 = m^2 c^2$ est elle-même une variété à 3 dimensions dans l'espace des impulsions. Géométriquement, c'est un hyperboloïde à deux nappes, $p^0 = \pm \sqrt{\vec{p}^2 + m^2 c^2}$. Une particule physique est dotée d'une énergie positive, donc on se restreint en général à la nappe $p^0 > 0$. Il est important pour la suite de connaître l'élément de volume naturel (c'est-à-dire invariant de Lorentz) sur cette nappe. On peut soit réduire la forme $d^4 p \delta_+(p^2 - m^2 c^2)$, où δ_+ ne garde que la racine $p^0 > 0$, soit procéder comme plus haut, en calculant l'élément de longueur $(dp)^2$ sur la couche de masse et l'élément de volume associé. Montrer que l'une et l'autre méthode conduisent à

$$d\tilde{p} := \frac{mc}{2p^0} d^3 \vec{p} = \frac{mc}{2\sqrt{\vec{p}^2 + m^2 c^2}} d^3 \vec{p}, \quad (8.34)$$

à une constante multiplicative près, choisie ici pour respecter la dimension $[p]^3$. On peut ensuite aussi écrire $d^3 \vec{p} = d^2 \Omega_p |\vec{p}|^2 d|\vec{p}|$, réexprimer $d|\vec{p}|$ en termes de dE , etc.

• *Cas des particules de masse nulle.* Si $m \rightarrow 0$, les situations physiquement intéressantes correspondent à $v \rightarrow c$, $\vec{p}$ et E fixés. Ce qui remplace (8.31) est donc

$$m = 0 \quad p^\mu = \left(\frac{1}{c} E, \vec{p} \right) \quad (8.35)$$

$$\vec{p} = \frac{E \vec{v}}{c^2}, \quad |\vec{v}| = c \quad (8.36)$$

$$p^2 = p_\mu p^\mu = 0 = \frac{E^2}{c^2} - \vec{p}^2. \quad (8.37)$$

8.3.2 Cinématique relativiste : collisions et désintégration de particules

Centre de masses

Considérons un ensemble de points matériels ("particules") de masses m_i , $i = 1, \dots, N$, dotés d'énergies et d'impulsions $(E_i, \vec{p}_i)$ dans un certain référentiel, dit repère du laboratoire (*laboratory frame*). Notons $P^\mu = \sum_i p^\mu$ la (quadri-)impulsion totale. C'est un vecteur de genre temps futur si tous les p_i le sont (pourquoi?). On appelle repère du centre de masses un référentiel inertiel dans lequel les nouvelles énergies et impulsions $(E'_i, \vec{p}'_i)$ sont telles que $\sum_i \vec{p}'_i = 0$. L'énergie $E' := \sum_i E'_i$ est l'"énergie totale dans le centre de masses". Pour la calculer, faisons appel au fait que par changement de référentiel, $P^\mu \rightarrow P'^\mu$ de même norme carrée.

$$\frac{1}{c^2} E'^2 = P'^2 = P^2 = \frac{1}{c^2} \left(\sum_i E_i \right)^2 - \left(\sum_i \vec{p}_i \right)^2 =: s \quad (8.38)$$

(la notation $s = (\sum_i p_i)^2$ est traditionnelle). L'énergie totale dans le centre de masses est donc $c\sqrt{s}$.

Désintégration de particules

Considérons le cas d'une particule de masse M se désintégrant en deux particules de masses m_1 et m_2 . Si P, p_1, p_2 désignent les quadri-impulsions des trois particules, on a $P = p_1 + p_2$.

On se place d'abord dans le repère où la particule initiale est au repos, qui est aussi le repère du centre de masse du système final. **Attention**, pour simplifier les expressions, on va faire $c = 1$ dans le reste de ce paragraphe.

État initial $P^\mu = (M, \mathbf{0})$, état final $p_1^\mu = (E_1, \vec{p})$, $p_2^\mu = (E_2, -\vec{p})$. Conservation de l'énergie-impulsion $E_1 + E_2 = M$, avec $\vec{p}^2 = E_1^2 - m_1^2 = E_2^2 - m_2^2$. Éliminant E_2 entre ces deux équations conduit à

$$E_1 = \frac{M^2 + m_1^2 - m_2^2}{2M} \quad E_2 = \frac{M^2 + m_2^2 - m_1^2}{2M} \quad (8.39)$$

$$\vec{p}^2 = \frac{M^4 + m_1^4 + m_2^4 - 2m_1^2 m_2^2 - 2M^2 m_1^2 - 2M^2 m_2^2}{4M^2} \quad (8.40)$$

et donc tous les éléments cinématiques en dehors de la direction de $\vec{p}$ sont déterminés.

Si maintenant la particule se désintègre "en vol", avec une certaine vitesse $\vec{v}$ et une énergie E dans le "repère du laboratoire", on a $|\vec{v}| = \sqrt{1 - M^2/E^2}$ selon les expressions (8.31). Choissant l'axe des x le long de $\vec{v}$, on applique les formules (8.5) avec $\gamma = E/M$ pour déterminer l'expression de E_i^{lab} et $\vec{p}_i^{\text{lab}}$ dans le laboratoire à partir de celles de E et $\vec{p}$ dans le centre de masses

$$E_i^{\text{lab}} = \frac{E}{M} E_i + p_{i,x} \sqrt{\frac{E^2}{M^2} - 1} \quad (8.41)$$

$$p_{1,x}^{\text{lab}} = \frac{E}{M} p_x + E_1 \sqrt{\frac{E^2}{M^2} - 1} \quad p_{1,z}^{\text{lab}} = p_z \quad (8.42)$$

$$p_{2,x}^{\text{lab}} = -\frac{E}{M} p_x + E_2 \sqrt{\frac{E^2}{M^2} - 1} \quad p_{2,z}^{\text{lab}} = -p_z \quad (8.43)$$

$$(8.44)$$

Supposons que la particule se désintègre dans le plan x, y selon la direction θ , donc $p_z = 0$, et donc aussi $p_{i,z}^{\text{lab}} = 0$, $i = 1, 2$, et $p_x = p \cos \theta$, $p_y = p \sin \theta$; soit θ^{lab} l'angle d'émission de la particule 1 par rapport à la particule incidente, donc $p_{1,x}^{\text{lab}} = p^{\text{lab}} \cos \theta^{\text{lab}}$, $p_{1,y}^{\text{lab}} = p^{\text{lab}} \sin \theta^{\text{lab}}$. Les formules précédentes fournissent θ^{lab} en termes de θ et de E

$$\text{tg} \theta^{\text{lab}} = \frac{p \sin \theta}{\frac{E}{M} p \cos \theta + E_1 \sqrt{\frac{E^2}{M^2} - 1}}.$$

Si la distribution angulaire est uniforme en θ (dans le centre de masses), $dn/d\cos \theta = \frac{1}{2}N$, N le nombre total de désintégrations observées, la distribution angulaire en θ^{lab} se ramène à un "simple" jacobien

$$\frac{dn}{d\cos \theta^{\text{lab}}} = \frac{d\cos \theta}{d\cos \theta^{\text{lab}}} \frac{dn}{d\cos \theta} = \frac{1}{2}N \frac{d\cos \theta}{d\cos \theta^{\text{lab}}}.$$

Collisions à deux corps

On considère un processus de diffusion

$$1 + 2 \rightarrow 3 + 4$$

avec quatre particules de masses m_i et de quadri-impulsions p_i , $i = 1, \dots, 4$. Soit $P = p_1 + p_2 = p_3 + p_4$ l'impulsion totale conservée. Si on se place dans le référentiel du centre de masses, où $P = (E = \sqrt{s}, \mathbf{0})$, on est ramené à la discussion de la désintégration.

Combien d'invariants relativistes indépendants peut-on construire avec les quatre quadri-impulsions p_i^μ , compte tenu des conditions $p_i^2 = m_i^2$ ("la particule i est sur sa couche de masse") ? On a déjà vu $s = (p_1 + p_2)^2 = (p_3 + p_4)^2$. On définit les deux autres "variables de Mandelstam"

$$t = (p_1 - p_3)^2 = (p_2 - p_4)^2 \quad (8.45)$$

$$u = (p_1 - p_4)^2 = (p_2 - p_3)^2. \quad (8.46)$$

Montrer que $s + t + u = \sum_i m_i^2$ et que l'on ne peut construire d'autre invariant indépendant. Relier t à l'angle de diffusion (entre les particules 1 et 3) dans le repère du centre de masses.

Bibliographie sommaire

Les deux articles originaux d'A. Einstein, *Annalen der Physik*, vol XVII (1905), pages 891-921, et vol XVIII (1905) p. 639-641 se lisent très bien ! Ils sont traduits et annotés dans les *Œuvres choisies d'Albert Einstein* éditées par F. Balibar, O. Darrigol et B. Jech, Seuil/CNRS 1994. Ils sont également reproduits dans le fascicule *Einstein 1905-2005*, Séminaire Poincaré 9/04/2005, qui contient aussi des articles variés sur l'histoire (O. Darrigol) et les tests de la relativité (C.M. Will, gr-qc/0504085), la géométrie de l'espace de Minkovski, (J. Bros et U. Moschella), etc.

Pour une discussion détaillée des propriétés du groupe de Lorentz, voir le livre récent d'Éricourgoulhon, *Relativité restreinte. Des particules à l'astrophysique*, (EDP Sciences / CNRS Éditions).

Index

- abélien, 11, 84
- algèbre de Lie, 65, 118
- algèbre de Lie $\mathfrak{so}(3)$, 117, 118
- Bloch, fonction d'onde de, 108
- Bloch, théorème de, 46, 108
- Bravais, 16
- Brillouin, zone de, 46
- brisure spontanée de symétrie, 49, 112
- Brout–Englert–Higgs, mécanisme de, 55
- caractère, 79, 85
- caractères de $\mathrm{SO}(3)$, 80, 121
- caractères, orthogonalité des, 85
- Casimir, opérateur de, 117
- champ cristallin, 138
- champ moyen, 57
- chiralité, 9
- chromodynamique quantique, 111
- classe de conjugaison, 79
- Clebsch–Gordan, coefficients de, 128
- Clebsch–Gordan, décomposition de, 92
- coefficients $3j$, 129
- compacité d'un empilement, 28
- compact, groupe, 67, 83
- confinement, 111
- conjugaison de charge, 107
- CP, violation de, 107
- cristaux liquides, 29, 56
- cube, groupe du, 13, 139
- cubiques, réseaux, 24
- Curie, principe de, 43
- Curie, température de, 50
- déplacement, 7
- dextrogyre, 9
- diédral, groupe, 11, 14, 96
- diffraction cristalline, 44
- dilatation, invariance par, 53
- dimension critique supérieure, 59
- dimension d'un groupe de Lie, 66
- dimension d'une algèbre de Lie, 66
- Euler, angles d', 114
- Euler–Lagrange, équations d', 68
- exposant critique, 53
- ferroélectrique, 42
- ferroïque, 42
- ferromagnétisme, 50
- fonction de classe, 79
- frises, 21
- fullerène, 4, 15, 135
- générateur infinitésimal, 62, 67
- générateur infinitésimal des rotations, 115
- Ginzburg–Landau, approximation de, 56
- glissement, 21
- Goldstone, bosons de, 54
- graphène, 4, 25
- graphite, 25
- groupe, 1
- groupe compact, 67, 83
- groupe d'espace, 16, 17
- groupe de Lie, 65
- groupe linéaire, 65
- groupe orthogonal $\mathrm{O}(n)$, 65
- groupe ponctuel, 16, 17
- groupe spécial orthogonal $\mathrm{SO}(n)$, 65
- groupe spécial unitaire $\mathrm{SU}(n)$, 65
- groupe unitaire $\mathrm{U}(n)$, 65
- groupes d'espace à 2d, 19
- groupes d'espace à 3d, 25
- groupes ponctuels à 2d, 18

- groupes ponctuels à 3d, 13, 24
- Haar, mesure de, pour $SO(3)$ et $SU(2)$, 122
- Hamilton, équations d', 69
- hamiltonien, 69
- hamiltonien, formalisme, 72
- harmoniques sphériques, 124
- Heisenberg, modèle de, 54
- Hermann–Mauguin, notation de, 21, 23
- hexagonal, réseau, 24
- Higgs, 55
- Higgs, boson de, 112
- icosaèdre, groupes d'invariance, 13, 135
- impropre, transformation, 8
- inversion, 8
- Ising, modèle d', 51
- isométrie, 7
- isospin, 112, 133
- jauge, invariance de, 73, 111
- Lagrange, fonction de, 68
- lagrangien, formalisme, 68
- LCAO, méthode, 109
- Legendre, fonctions et polynômes de, 125
- lévogyre, 9
- lois de conservation, 70, 101
- Lorentz, groupe de, 148
- Lorentz, transformations de, 142
- maille, 16
- Maxwell, équations de, 36
- Miller, indices de, 45
- Minkowski, espace de, 150
- modèle standard, 55, 111
- mode normal de vibration, 48, 93
- modes de vibration des molécules, 93
- Nambu–Goldstone, bosons de, 54
- nanotubes de C, 30
- Noether, théorème de, 74
- orthogonalité, relations d', 85
- paramagnétisme, 50
- parité, 8, 102, 103
- parité, brisure de la symétrie de, 106
- partition, 79
- pavages à 2d, 30
- pérovskite, 31
- piézoélectricité, 42
- Poincaré, groupe de, 148
- Poisson, crochet de, 69
- polarisation, 9
- polyèdres réguliers, 2, 11
- pouvoir rotatoire, 9
- produit direct de représentations, 91, 126
- produit direct de groupes, 90
- produit tensoriel de représentations, 91, 126
- pseudoscalaires, pseudovecteurs, 39
- pyroélectricité, 42
- quartz, 2, 9, 27
- quasi-cristaux, 29
- quaternions, 96
- racémique, 9
- réflexion-miroir, 8
- règles de sélection, 103
- renversement du temps, 106
- représentation, 77
- représentation à une phase près, 100
- représentation d'une algèbre de Lie, 118
- représentation de $SO(3)$, 117
- représentation de spin j , 120
- représentation irréductible, 80
- représentation réductible, 80
- représentation unitaire, 82
- représentations équivalentes, 78
- réseau réciproque, 45
- réseau de Bravais, 16
- réseau hexagonal, 24
- réseau primitif, 19, 22
- réseaux cubiques, 24
- réseaux de Bravais à 2d, 18, 23
- réticulaire, plan, 45
- Rodrigues, formule d'O., 113
- rotation-réflexion, 14, 79
- rotations, groupe des, 10, 113

saveurs, 112
Schönflies, notation de, 14
Schur, lemme de, 84
signature d'une permutation, 88
solides platoniciens, 2
spontanée, brisure, 49
symétrie interne, 55
système réticulaire, 16

tenseur, 36
tenseur énergie-impulsion, 76
tenseur piézoélectrique, 42
tenseur pour les rotations, 36
tétraèdre, groupe du, 12
théorème CPT, 107
théorème de Wigner–Eckart, 130
théorie quantique des champs, 107
transformations actives, passives, 33

von Laue, 44
Voronoi, cellule de, 46, 48

Wigner, théorème de, 99
Wigner–Eckart, théorème de, 130
Wigner–Seitz, cellule de, 46, 48

Yang et Mills, théorie de, 75