

SYMETRIES ET TRANSITIONS DE PHASE : LA THEORIE DE LANDAU

J. ZINN-JUSTIN*

*CEA/IRFU, Centre de Saclay,
F91191 Gif-sur-Yvette cédex, FRANCE*

RÉSUMÉ

Ces notes, basées sur un cours donné à l'Université de Cergy-Pontoise en 2000-2001 (Saclay preprint T01/019), tentent d'introduire la théorie de Landau des phénomènes critiques en insistant sur le rôle des symétries. Dans l'introduction le rôle des symétries est souligné et donc dans une première partie la structure mathématique associée de groupe de symétrie est décrite et les propriétés de quelques groupes utiles en physique seront présentées de façon élémentaire.

Le rôle des groupes de symétrie en physique est d'abord illustré par des exemples simples tirés de la mécanique analytique classique et de la mécanique quantique non-relativiste.

Ensuite une analyse de la marche au hasard sur réseaux est présentée du point de vue des symétries.

Enfin le sujet des transitions de phase est abordée. Une grande classe de transitions de phase est caractérisée par une brisure spontanée de symétrie, une notion qui est expliquée en détail. La théorie de Landau des phénomènes critiques est introduite. Il est montré sur des exemples simples comment la théorie de Landau combinée avec les propriétés de symétrie conduit à des prédictions de comportements universels des quantités thermodynamiques au voisinage d'une transition de phase continue.

*email : jean.zinn-justin@cea.fr

Table des matières

1 Introduction : Le rôle des symétries en physique	1
2 Symétries : la notion de groupe	3
2.1 Groupes de matrices. Algèbres. Algèbres de Lie	7
3 Groupes discrets. Groupe finis	11
3.1 Groupes de translations sur réseau	12
3.2 Groupe groupe symétrique ou des permutations	12
3.3 Transformations linéaires du réseau cubique général	14
Exercices	15
4 Groupes continus : groupes de Lie et isométries du plan	17
4.1 Rotations et réflexions du plan	17
4.2 Le groupe $U(1)$	20
4.3 Générateur. Exponentiation	22
4.4 Représentations	22
Exercices	28
5 Groupes de Lie : isométries en dimension trois et spin	29
5.1 Rotations et groupe orthogonal	29
5.2 Représentation adjointe	32
5.3 Les représentations matricielles de $SO(3)$	32
5.4 Groupe $SO(3)$ et matrices σ_i	34
5.5 Le groupe $SU(2)$	35
5.6 Représentations irréductibles de l'algèbre de Lie de $SU(2)$	40
6 Groupes et algèbres de Lie plus généraux : exemples orthogonaux et unitaires 44	44
6.1 Groupes et algèbres de Lie	44
6.2 Représentations, représentation adjointe	46
6.3 Groupes orthogonaux	48
6.4 $SO(N)$: Algèbre de Lie	49
6.5 Groupes unitaires	50
6.6 Algèbre de Lie de $SU(N)$	52
6.7 Représentations: représentation adjointe	54
6.8 Un exemple : le groupe $SU(3)$	56
6.9 Représentations de $SU(3)$	57
7 Produit tensoriel. Tenseurs.	60
7.1 Groupe symétrique et tenseurs : réduction des représentations	62
7.2 Groupe linéaire général et analyse tensorielle	64
8 Symétries en mécanique lagrangienne : théorème de Noether	67
8.1 Symétries. Lois de conservation	67
8.2 Mécanique hamiltonienne	70
8.3 Théorie classique des champs	74
8.4 Tenseurs : un exemple, le moment d'inertie	75
9 Symétries en mécanique quantique	77
9.1 Evolution dans le temps. Quantités conservées	77
9.2 Transformations et symétries	79
10 Mouvement brownien ou marche au hasard : Limite continue	84
11 Transitions de phase et brisure spontanée de symétrie	90

11.1	Brisure spontanée de symétrie	91
11.2	Potentiel thermodynamique. Transformation de Legendre	94
11.3	Symétries continues et modes de Goldstone	97
11.4	Modèle quasi-gaussien ou théorie de Landau : universalité	98
11.5	Fonction de corrélation spin–spin	105
	Appendices. Points divers	108
A1	Solide rigide libre	108
A2	Oscillateur harmonique quantique et algèbre de Lie	109

1 Introduction : Le rôle des symétries en physique

Les quelques remarques qui suivent ont pour but de motiver l'utilité de l'étude des symétries des systèmes physiques, et donc des groupe de symétries.

L'émergence de la notion de symétrie. Beaucoup d'objets naturels macroscopiques ont des propriétés de symétrie, en particulier les êtres vivants, mais aussi les cristaux, les astres proches. Ce sont des symétries discrètes, réflexion, rotations discrètes ou continues (cercle, sphère).

La symétrie comme principe utile ou esthétique. Citons l'invention de la roue. La symétrie aide à l'équilibre dans les constructions. Les symétries sont idéalisées en principe esthétique, par exemple dans l'architecture, les objets d'art.

Symétrie : Un principe de la nature? La nature doit-elle être esthétique? Une réponse positive et nécessairement subjective à cette question conduit à une recherche de symétries explicites ou cachées. Notons qu'une telle démarche est parfois couronnée de succès, mais peut aussi conduire à des théories fausses ou aberrantes (le mouvement des planètes à partir de cercles par Ptolémée, les planètes associées aux solides réguliers par Képler).

Très rapidement on prend conscience que les objets macroscopiques ne peuvent jamais être complètement symétriques au moins par des symétries spatiales à cause du nombre très grand de leurs constituants, et ceci conduit à rechercher la symétrie dans la nature d'une autre façon.

Mécanique rationnelle : La symétrie n'est plus nécessairement une propriété des objets naturels mais une propriété des lois de la nature. Ce sont les conditions aux limites qui peuvent briser les symétries. Par exemple dans la mécanique newtonienne les lois de la nature sont invariantes par translation d'espace et de temps, insensible à un déplacement rectiligne uniforme. Il n'y a pas de direction privilégiée dans l'espace, et l'Univers est isotrope à grande échelle (ceci reste en accord avec les observations récentes, par exemple le satellite WMAP a vérifié l'isotropie du rayonnement du fond du ciel à une précision de 10^{-5}). Ceci conduit à une notion d'invariance par rotation. L'ensemble de ces symétries a reçu le nom de groupe de Galilée de la mécanique newtonienne.

Symétrie et problèmes solubles. Une symétrie implique une réduction effective du nombre de degrés de liberté. Par exemple en gravitation newtonienne une sphère homogène par couches est équivalente à un point matériel. De façon générale en mécanique lagrangienne à toute symétrie continue est associée une constante du mouvement. Ainsi à la symétrie par translation du temps correspond la conservation de l'énergie, à la symétrie par translation d'espace correspond la conservation de l'impulsion ou quantité de mouvement, à l'invariance par rotation correspond la conservation du moment cinétique. De même en mécanique quantique à toute symétrie est associé un opérateur qui commute avec l'hamiltonien, les valeurs moyennes de cet opérateur étant alors indépendantes du temps.

Ces lois de conservations rendent certains problèmes solubles. Ces problèmes solubles peuvent être soit des approximations de la réalité soit des problèmes

modèles qui permettent d'étudier des effets physiques qualitatifs.

Symétrie et principe dynamique. Ce que nous ne décrivons pas ici c'est la symétrie conçue comme principe dynamique telle qu'elle apparaît dans les équations de Maxwell, dans la théorie des interactions fondamentales, et qui est à l'origine de la plupart des interactions. Elle prend dans ce cas le nom d'invariance de jauge (voir modèle Standard des Interactions Fondamentales, Relativité Générale).

Supersymétrie. Il existe aussi un type de symétrie généralisé que nous n'abordons pas, connu sous le nom de supersymétrie, d'utilité technique dans certains problèmes de mécanique statistique, et qui a été proposé de façon spéculative comme une symétrie (approchée) possible des interactions fondamentales à l'échelle microscopique où elle relie bosons et fermions.

Les symétries : un rôle croissant en physique. Le XIX^{ème} et XX^{ème} siècle ont vu la notion de symétrie jouer un rôle croissant en physique. On peut trouver à cela nombre de raisons. Evoquons en quelques-unes :

(i) Etudes des objets microscopiques, atomes, noyaux, quarks et leptons (fondamentaux?). On imagine qu'ils ont de moins en moins de structure interne, et ceci suggère qu'ils ont de plus en plus de symétrie (principe d'esthétique?). Par ailleurs la mécanique quantique, à cause du principe de superposition, amène à développer les états sur des bases d'entités simples.

(ii) La structure de l'espace-temps se révèle moins triviale que celle supposée dans la mécanique newtonienne : la théorie de la Relativité restreinte repose sur l'invariance par les groupes de Lorentz et de Poincaré.

(iii) Un grand nombre de systèmes nouveaux présentant des transitions de phase a été observé. Une compréhension plus détaillée de la notion de transition de phase a mis en évidence le rôle des symétries et a introduit le concept de brisure spontanée d'une symétrie.

Bien entendu la prise de conscience croissante de l'importance des symétries en physique a été facilitée par le développement mathématique parallèle de la théorie des groupes.

Conclusion. Ainsi aucune compréhension profonde de la physique contemporaine n'est possible sans une prise en compte du rôle des symétries. À la notion intuitive de symétrie est associée la notion mathématique de groupe, que nous allons rappeler. Les éléments du groupe correspondent aux opérations physiques qui permettent de vérifier l'existence de la symétrie, une translation, une rotation....

2 Symétries : la notion de groupe

L'exposé qui suit n'a pas comme objectif de faire de la *théorie des groupes*, un domaine mathématique par ailleurs extraordinairement riche, mais de rappeler quelques notions essentielles et d'examiner les propriétés simples de quelques groupes directement utiles. Il aura donc un contenu beaucoup plus descriptif que déductif.

Groupe : définition. On considère un ensemble G muni d'une loi de composition interne, appelée en général produit, telle qu'à tout couple ordonné $g_1, g_2 \in G$ on associe un élément de G noté alors multiplicativement $g_1 g_2$. Pour que le produit donne à G une structure de groupe, il faut qu'il satisfasse à quelques règles simples que nous associons intuitivement à la notion de symétrie :

(i) Le produit est associatif :

$$(g_1 g_2) g_3 = g_1 (g_2 g_3) \quad \forall g_1, g_2, g_3.$$

On peut donc effectuer les produits sans se soucier de l'ordre dans lequel on les effectue.

(ii) Il existe un élément *neutre* dans G , que nous noterons en général $\mathbf{1}$, c'est à dire un élément qui satisfait

$$\mathbf{1}g = g\mathbf{1} = g \quad \forall g,$$

et dont on vérifie qu'il est unique.

(iii) Tout élément $g \in G$ a un *inverse*, que nous noterons en général g^{-1} , tel que

$$gg^{-1} = g^{-1}g = \mathbf{1},$$

dont on vérifie aussi qu'il est unique.

Groupes commutatifs ou abéliens. En général le produit $g_1 g_2$ est différent du produit $g_2 g_1$. Les groupes spéciaux pour lesquels

$$g_1 g_2 = g_2 g_1 \quad \forall g_1, g_2,$$

sont appelés *commutatifs* ou *abéliens*. Des exemples particuliers sont fournis par le groupe des translations, ou le groupe des rotations du plan. Dans ce cas la loi de composition est parfois notée additivement, et l'élément neutre $\mathbf{0}$. Les autres groupes sont donc non commutatifs ou encore *non-abéliens*.

Sous-groupe. Un sous-groupe de G est un sous-ensemble H non vide de G tel que la loi de composition de G induite sur H lui donne une structure de groupe. Ceci est équivalent aux deux conditions :

$$g_1 \text{ et } g_2 \in H \Rightarrow g_1 g_2 \in H, \quad g \in H \Rightarrow g^{-1} \in H,$$

les deux conditions impliquant $\mathbf{1} \in H$.

Centre d'un groupe. Le centre C d'un groupe est l'ensemble des éléments du groupe G qui commutent avec tout G . Le centre a une structure de groupe, en effet

$$h_1, h_2 \in C \Rightarrow h_1 h_2 \in C, \quad h \in C \Leftrightarrow hG = Gh \Rightarrow Gh^{-1} = h^{-1}G.$$

Produit de groupes. Soient G et H deux groupes. Nous considérons l'ensemble produit formé par les couples (g, h) , $g \in G$, $h \in H$. Nous définissons le produit de deux éléments par

$$(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1 h_2).$$

Ce produit définit une loi de groupe et le groupe est noté $G \times H$. Dans le cas où $H \equiv G$ on notera le produit G^2 , ou plus généralement le produit de p groupes G^p .

Générateurs d'un groupe. Une notion qui nous sera très utile est la notion de générateurs d'un groupe.

(i) Un ensemble L de générateurs d'un groupe G est un sous-ensemble de G , tel que tout élément de G puisse être écrit comme un produit d'éléments de L .

(ii) L'ensemble est appelé minimal si aucun élément de L ne puisse être écrit comme un produit des autres éléments de L . Par défaut, quand nous parlerons d'ensemble de générateurs nous le supposons *minimal*.

Pour étudier des propriétés de symétrie il est souvent suffisant d'étudier l'action des générateurs du groupe.

Représentation ou homomorphisme de groupe. Soient G un groupe et H un ensemble muni d'une loi de composition interne. Une application $\mathcal{R}(G)$ de G dans H s'appelle une représentation si quels que soient g_1 et g_2 appartenant à G le produit $\mathcal{R}(g_1)\mathcal{R}(g_2)$ est défini, et

$$\mathcal{R}(g_1)\mathcal{R}(g_2) = \mathcal{R}(g_1 g_2).$$

Alors l'image $\mathcal{R}(G)$ est un groupe, d'élément neutre $\mathcal{R}(1)$ en effet

$$\mathcal{R}(g) = \mathcal{R}(g1) = \mathcal{R}(g)\mathcal{R}(1) = \mathcal{R}(1g) = \mathcal{R}(1)\mathcal{R}(g).$$

De même si

$$h = \mathcal{R}(g) \Rightarrow h^{-1} = \mathcal{R}(g^{-1}).$$

On utilisera par la suite aussi le terme représentation de G pour désigner le groupe des éléments appartenant à $\mathcal{R}(G)$.

Si l'application \mathcal{R} est bijective, c'est à dire que réciproquement tout élément de $\mathcal{R}(G)$ est associé à un élément unique de G , les groupes G et $\mathcal{R}(G)$ sont dits isomorphes. Nous noterons l'isomorphisme de deux groupes G et H :

$$G \sim H.$$

Un isomorphisme de G dans G est un automorphisme. Un exemple simple est le suivant : Soit x un élément fixe de G . La représentation \mathcal{R} qui à tout élément g associe

$$\mathcal{R}(g) = xgx^{-1}$$

est un automorphisme.

On rencontrera aussi des applications bijectives qui ne sont pas des isomorphismes parce qu'elles renversent l'ordre du produit,

$$\mathcal{R}(g_1g_2) = \mathcal{R}(g_2)\mathcal{R}(g_1).$$

Un exemple est l'application $g \mapsto g^{-1}$ et donc $(g_1g_2)^{-1} = g_2^{-1}g_1^{-1}$. Le groupe ainsi obtenu est appelé groupe opposé.

Dans le cas de groupes de matrices la transposition et la conjugaison hermitienne partagent cette propriété. Il est clair cependant qu'une telle réflexion n'engendre aucune structure de groupe nouvelle.

Ces définitions quelque peu abstraites vont être illustrées plus loin de la manière suivante :

Un groupe peut être entièrement décrit par ses éléments et leurs produits, comme nous le ferons de façon explicite pour le groupe symétrique ou des permutations. On parle alors parfois de groupe *abstrait*. Mais nous exhiberons toujours une ou plusieurs représentations du groupe dans l'ensemble des matrices à coefficients réels ou complexes. Une au moins de ces représentations sera un isomorphisme, ce qui permettra une définition alternative "concrète" du groupe comme un groupe de matrices.

Le fait qu'un même groupe puisse avoir plusieurs représentations en termes de matrices est un phénomène très général qui conduit par exemple à la théorie des représentations des groupes de Lie.

Sous-groupes distingués (ou invariants) et groupes quotients. Soit H un sous-groupe d'un groupe G . On définit dans G la relation d'équivalence suivante

$$x, y \in G : x \equiv y \Leftrightarrow x^{-1}y \in H.$$

On vérifie que c'est bien une relation d'équivalence, c'est à dire qu'elle est réflexive, symétrique et transitive :

$$\begin{aligned} x \equiv x &\Leftrightarrow x^{-1}x = \mathbf{1} \in H \\ x \equiv y &\Rightarrow y \equiv x \Leftrightarrow x^{-1}y \in H \Rightarrow y^{-1}x = (x^{-1}y)^{-1} \in H \\ x \equiv y \text{ et } y \equiv z &\Rightarrow x \equiv z \Leftrightarrow x^{-1}y \in H \text{ et } y^{-1}z \in H \\ &\Rightarrow x^{-1}z = x^{-1}yy^{-1}z = x^{-1}z \in H. \end{aligned}$$

Tous les éléments appartenant à une même classe sont obtenus à partir de l'un d'entre eux en le multipliant à droite par tous les éléments de H : $y = xh$.

Comme il est également possible de définir une autre relation d'équivalence par

$$x, y \in G : x \equiv y \Leftrightarrow yx^{-1} \in H,$$

on distinguera les deux cas, quand c'est nécessaire, en parlant de classes à gauche suivant H dans le premier cas et à droite dans le deuxième cas. Comme toutes les propriétés sont symétriques nous ne nous intéresserons ici qu'au premier cas.

L'ensemble des classes d'équivalence est appelé *ensemble quotient* et noté G/H .

Le sous-groupe H est dit *distingué* ou *invariant* si pour tout $g \in G$ et tout $h \in H$

$$ghg^{-1} \in H. \quad (2.1)$$

Cette relation est équivalente à

$$\forall g \in G \text{ et } h \in H \quad \exists h' \in H \text{ tel que } hg = gh', \quad (2.2)$$

c'est à dire que les classes à droite sont aussi classes à gauche.

Exemple.

- (i) Tout sous-groupe d'un groupe abélien est distingué.
- (ii) Le *centre* d'un groupe (l'ensemble des éléments qui commutent avec tout le groupe) est un sous-groupe distingué.
- (iii) Soit G un groupe et $\mathcal{R}(G)$ une représentation. L'ensemble H des éléments h de G tels que $\mathcal{R}(h) = \mathcal{R}(\mathbf{1})$ forme un sous-groupe distingué de G . En effet
 - (i) C'est un sous-groupe :

$$h_1, h_2 \in H \Rightarrow \mathcal{R}(h_1 h_2) = \mathcal{R}(h_1) \mathcal{R}(h_2) = \mathcal{R}(\mathbf{1}) \Rightarrow h_1 h_2 \in H,$$

$$h \in H \Rightarrow \mathcal{R}(h^{-1}) = (\mathcal{R}(h))^{-1} = \mathcal{R}(\mathbf{1}) \Rightarrow h^{-1} \in H.$$

Il satisfait la condition (2.1)

$$h \in H \Rightarrow \mathcal{R}(ghg^{-1}) = \mathcal{R}(g) \mathcal{R}(h) \mathcal{R}(g^{-1}) = \mathcal{R}(g) \mathcal{R}(g^{-1}) = \mathcal{R}(\mathbf{1}).$$

Par ailleurs la condition

$$\mathcal{R}(gh) = \mathcal{R}(h'g) = \mathcal{R}(g),$$

montre l'application $G/H \mapsto \mathcal{R}(G)$ est bijective et donne donc à l'ensemble quotient G/H une structure de groupe.

La condition (2.1) est donc une condition est nécessaire pour que l'ensemble quotient G/H ait une structure un groupe.

Condition suffisante. Montrons que la condition (2.1) est suffisante en l'utilisant sous la forme (2.2). On considère l'application $\mathcal{R}(G)$ du groupe G sur l'ensemble des classes G/H . Si H est un sous-groupe distingué l'application

$$\mathcal{R}(g_1) \mathcal{R}(g_2) = \mathcal{R}(g_1 g_2),$$

est une application interne dans G/H . En effet quel que soit le représentant d'après (2.2)

$$\mathcal{R}(g_1 h_1) \mathcal{R}(g_2 h_2) = \mathcal{R}(g_1 h_1 g_2 h_2) = \mathcal{R}(g_1 g_2 h'_1 h_2) = \mathcal{R}(g_1 g_2).$$

Elle définit donc une représentation de G .

La condition ‘ H est un sous-groupe distingué de G ’ est donc une condition nécessaire et suffisante pour que l’ensemble quotient G/H ait une structure de groupe.

Générateurs. Pour vérifier la propriété (2.1) il suffit de la vérifier pour un ensemble de générateur de G et H . En effet à $g \in G$ fixé

$$gh_1g^{-1} \in H \text{ et } gh_2g^{-1} \in H \Rightarrow gh_1g^{-1}gh_2g^{-1} = gh_1h_2g^{-1} \in H.$$

Donc il suffit de vérifier cette propriété pour les générateurs de H .

Par ailleurs si pour tout $h \in H$:

$$g_1hg_1^{-1} \in H \text{ et } g_2hg_2^{-1} \in H \Rightarrow g_2g_1hg_1^{-1}g_2^{-1} \in H.$$

2.1 Groupes de matrices. Algèbres. Algèbres de Lie

Groupe de matrices. Pour qu’un ensemble \mathcal{M} de matrices muni du produit ordinaire des matrices soit un groupe il faut :

(i) Que le produit soit une loi de composition interne à l’ensemble \mathcal{M} :

$$\forall \mathbf{m}_1, \mathbf{m}_2 \in \mathcal{M} \Rightarrow \mathbf{m}_1\mathbf{m}_2 \in \mathcal{M}.$$

Le produit de matrices est automatiquement associatif.

(ii) Toute matrice de \mathcal{M} doit avoir un inverse, ce qui est équivalent à la condition que son déterminant ne s’annule pas. De plus l’inverse doit appartenir à \mathcal{M} ,

$$\forall \mathbf{m} \in \mathcal{M} \quad \det \mathbf{m} \neq 0 \quad \text{et} \quad \mathbf{m}^{-1} \in \mathcal{M}.$$

L’inverse est alors automatiquement inverse à droite et à gauche.

Ces deux propriétés entraîne que la matrice unité appartienne à \mathcal{M} . C’est l’élément neutre.

Les groupes formés de toutes les matrices inversibles $N \times N$ réelles et complexes sont notés $GL(N, \mathbb{R})$ (pour général linéaire) et $GL(N, \mathbb{C})$, respectivement.

Quelques isomorphismes. Si \mathbf{s} est une matrice fixe inversible, alors la transformation qui à toute matrice \mathbf{g} du groupe G associe la matrice

$$\mathbf{g} \mapsto \mathbf{s}^{-1}\mathbf{g}\mathbf{s},$$

est un isomorphisme de groupe. En effet

$$\mathbf{s}^{-1}\mathbf{g}_1\mathbf{s}\mathbf{s}^{-1}\mathbf{g}_2\mathbf{s} = \mathbf{s}^{-1}\mathbf{g}_1\mathbf{g}_2\mathbf{s} \dots$$

On dit de deux représentations d’un groupe reliées par un tel isomorphisme qu’elles sont équivalentes.

Considérons l'application qui à toute matrice d'un groupe associe la matrice inverse transposée

$$\mathcal{R}(\mathbf{g}) = (\mathbf{g}^{-1})^T = (\mathbf{g}^T)^{-1}.$$

C'est un isomorphisme puisque l'application est biunivoque et respecte le produit :

$$\mathcal{R}(\mathbf{g}_1)\mathcal{R}(\mathbf{g}_2) = (\mathbf{g}_1^{-1})^T(\mathbf{g}_2^{-1})^T = (\mathbf{g}_2^{-1}\mathbf{g}_1^{-1})^T = ((\mathbf{g}_1\mathbf{g}_2)^{-1})^T = \mathcal{R}(\mathbf{g}_1\mathbf{g}_2).$$

Pour des matrices complexes la conjugaison complexe est un isomorphisme

$$(\mathbf{g}_1\mathbf{g}_2)^* = \mathbf{g}_1^*\mathbf{g}_2^*.$$

Enfin le produit de ces deux derniers isomorphismes donne l'isomorphisme $(\mathbf{g}^{-1})^\dagger$.

Déterminants. Si les matrices \mathbf{g}_1 et \mathbf{g}_2 sont deux éléments d'un groupe G , alors

$$\det(\mathbf{g}_1\mathbf{g}_2) = \det \mathbf{g}_1 \det \mathbf{g}_2, \quad \det \mathbf{1} = 1, \quad \det \mathbf{g}^{-1} \det \mathbf{g} = 1.$$

Donc les déterminants des matrices d'un groupe forment également un groupe, sous-groupe du groupe multiplicatif des réels $\mathbb{R}^* \equiv \{\mathbb{R}\} - 0$ ou des complexes $\mathbb{C}^* \equiv \{\mathbb{C}\} - 0$, représentation abélienne du groupe G .

Considérons le sous-ensemble H des matrices de déterminant 1. Cet ensemble n'est pas vide puisqu'il contient la matrice unité. Il forme un sous-groupe de G puisque

$$\det \mathbf{g}_1 = 1 \text{ et } \det \mathbf{g}_2 = 1 \Rightarrow \det(\mathbf{g}_1\mathbf{g}_2) = 1, \quad \det \mathbf{g} = 1 \Rightarrow \det \mathbf{g}^{-1} = 1.$$

C'est un sous-groupe distingué de G . En effet,

$$\forall \mathbf{g} \in G \text{ et } \mathbf{h} \in H : \det(\mathbf{g}^{-1}\mathbf{h}\mathbf{g}) = 1 \Rightarrow \mathbf{g}^{-1}\mathbf{h}\mathbf{g} \in H.$$

L'ensemble quotient G/H des classes forme un groupe isomorphe au groupe des déterminants.

Dans le cas de matrices quelconques réelles et complexes les groupes des matrices de déterminant 1 sont notés $SL(N, \mathbb{R})$ et $SL(N, \mathbb{C})$, sous-groupes de $GL(N, \mathbb{R})$ et $GL(N, \mathbb{C})$, respectivement. À titre anecdotique (pour nous ici) on note aussi $SL(N, \mathbb{Z})$ le sous-groupe des matrices à coefficients entiers de déterminant 1 (les matrices d'un groupe de matrices à coefficients entiers ont nécessairement déterminant ± 1).

Transformations linéaires. Nous rappelons qu'une transformation linéaire \mathcal{T} agissant sur un espace vectoriel \mathcal{V} réel ou complexe a la propriété suivante : Soient \mathbf{x} et \mathbf{y} deux éléments de \mathcal{V} . Pour toute combinaison linéaire $\lambda\mathbf{x} + \mu\mathbf{y}$ de \mathbf{x} et \mathbf{y} ($\lambda, \mu \in \mathbb{R}$ ou \mathbb{C}),

$$\mathcal{T}(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda\mathcal{T}(\mathbf{x}) + \mu\mathcal{T}(\mathbf{y}).$$

En mécanique quantique les espaces vectoriels sont complexes et l'on rencontre de plus des transformations antilinéaires qui impliquent une conjugaison complexe supplémentaire

$$\mathcal{T}(\lambda \mathbf{x} + \mu \mathbf{y}) = \lambda^* \mathcal{T}(\mathbf{x}) + \mu^* \mathcal{T}(\mathbf{y}).$$

Nous ne les considérerons que rarement ici.

Pour associer une matrice à une transformation linéaire, on introduit une base de l'espace vectoriel $\{e_i\}$, $i = 1, \dots, N$. Tout vecteur v peut alors s'écrire

$$v = \sum_{i=1}^N e_i x_i,$$

où les composantes x_i sont des nombres réels ou complexes.

Par ailleurs

$$\mathcal{T}(e_i) = \sum_{j=1}^N e_j T_{ji},$$

où T_{ij} sont les éléments d'une matrice $N \times N$ que nous notons \mathbf{T} . Dans ces conditions

$$\mathcal{T}\left(\sum_j e_j x_j\right) = \sum_{ij=1}^N e_i T_{ij} x_j.$$

Les composantes $\{x_i\}$ sont transformées par \mathcal{T} en $\mathbf{T}x$.

Si un ensemble de transformations linéaires forment un groupe G , l'application qui à tout élément \mathcal{T} de G associe la matrice \mathbf{T} correspondante est une représentation, comme on le vérifie en calculant le produit de deux transformations $\mathcal{T}_1 \mathcal{T}_2$.

Représentations irréductibles. Considérons un groupe de transformations linéaires agissant sur un espace vectoriel \mathcal{V} de dimension N . Ce groupe a une représentation $\mathcal{R}(G)$ qui est un groupe de matrices $N \times N$.

La notion de réduction d'une représentation généralise alors aux groupes de matrices la notion de diagonalisation d'une matrice. S'il existe un sous-espace \mathcal{V}' de \mathcal{V} ($\mathcal{V}' \neq \mathcal{V}$) invariant par toutes les transformations du groupe G :

$$\forall g \in G \quad \forall v \in \mathcal{V}' \quad : \quad gv \in \mathcal{V}',$$

la représentation $\mathcal{R}(G)$ est dite réductible. Dans ce cas en choisissant une base dans \mathcal{V}' , on obtient par la construction précédente une autre représentation $\mathcal{R}'(G)$ qui est un groupe de matrices $N' \times N'$, $N' < N$.

Dans le cas contraire, s'il n'existe aucun sous-espace invariant la représentation est dite irréductible.

Si la représentation initiale est réductible et s'il est possible d'écrire l'espace vectoriel \mathcal{V} comme une somme directe d'espaces vectoriels

$$\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2 \dots,$$

correspondant tous à des représentations irréductibles (une condition qui généralise la notion de matrice diagonalisable), on a dit qu'on a réduit la représentation initiale en une somme de représentations irréductibles.

Notons que de même qu'il existe des matrices non diagonalisables, une représentation peut admettre un sous-espace invariant sans être décomposable en représentations irréductibles.

Notons enfin que si une représentation est irréductible pour un sous-groupe, elle est irréductible pour le groupe tout entier. Réciproquement si une représentation est réductible pour un groupe, elle est réductible pour tout sous-groupe.

Algèbre. Un espace vectoriel \mathcal{V} sur \mathbb{R} (ou \mathbb{C}) muni d'une loi de composition interne, appelée produit, qui est distributive par rapport à la structure d'espace vectoriel, est appelée algèbre :

$$\begin{aligned}(\lambda \mathbf{v}_1 + \mu \mathbf{v}_2) \mathbf{w} &= \lambda \mathbf{v}_1 \mathbf{w} + \mu \mathbf{v}_2 \mathbf{w} \\ \mathbf{w}(\lambda \mathbf{v}_1 + \mu \mathbf{v}_2) &= \lambda \mathbf{w} \mathbf{v}_1 + \mu \mathbf{w} \mathbf{v}_2\end{aligned}$$

quels que soient $\mathbf{v}_1, \mathbf{v}_2, \mathbf{w} \in \mathcal{V}$ et $\lambda, \mu \in \mathbb{R}$ (ou \mathbb{C}).

Une base de l'espace vectoriel constitue un ensemble de générateurs de l'algèbre.

Le produit ordinaire des matrices induit une structure d'algèbre associative (dans le sens que le produit est associatif) sur les matrices. Les notions de représentations et de représentations irréductibles se généralisent aux groupes et algèbres de matrice.

Algèbre de Lie. Les algèbres de Lie jouent un rôle très important dans l'étude des groupes continus. Le produit de Lie de deux matrices \mathbf{A}, \mathbf{B} est défini comme le commutateur de ces matrices, noté ici \wedge pour le distinguer du produit ordinaire,

$$\mathbf{A} \wedge \mathbf{B} \equiv [\mathbf{A}, \mathbf{B}] \equiv \mathbf{AB} - \mathbf{BA}.$$

Ce produit induit une structure d'algèbre (non associative) appelée *algèbre de Lie*.

Le produit de Lie est antisymétrique :

$$\mathbf{A} \wedge \mathbf{B} = -\mathbf{B} \wedge \mathbf{A},$$

et satisfait une identité importante et facile à vérifier,

$$(\mathbf{A} \wedge \mathbf{B}) \wedge \mathbf{C} + (\mathbf{B} \wedge \mathbf{C}) \wedge \mathbf{A} + (\mathbf{C} \wedge \mathbf{A}) \wedge \mathbf{B} = 0, \quad (2.3)$$

appelée identité de Jacobi.

3 Groupes discrets. Groupe finis

Empiriquement les groupes que nous rencontrerons peuvent se regrouper en plusieurs classes que nous allons brièvement évoquer et illustrer par des exemples sur lesquels nous reviendrons. Les groupes discrets sont tels qu'on puisse en numérotter les éléments, c'est à dire établir une correspondance entre les éléments de G et les entiers naturels.

Les groupes finis sont les groupes discrets qui ont un nombre fini d'éléments. Le nombre d'éléments d'un groupe fini s'appelle *l'ordre du groupe*.

Un sous-groupe d'un groupe fini divise le groupe en classes. Chaque classe contient un nombre d'éléments égal à l'ordre du sous-groupe, et donc l'ordre du sous-groupe divise l'ordre du groupe.

Une conséquence simple est la suivante. Soit N l'ordre du groupe. Considérons un élément g de G et ses puissances successives $1, g, g^2, \dots, g^N$. Cet ensemble a $N + 1 > N$ éléments et donc deux éléments au moins sont égaux. Dans ces conditions il existe $n \leq N$ tel que

$$g^n = 1.$$

Les éléments $1, g, g^2, \dots, g^{n-1}$ forment un sous-groupe d'ordre n et donc n divise N l'ordre du groupe.

Groupes finis et groupe symétrique (ou des permutations). Soit un groupe d'ordre N . Numérotons les éléments du groupe de 1 à N . Considérons le produit $g_1 g_2$ de deux éléments de G . À g_1 fixé quand g_2 décrit tout le groupe $g_1 g_2$ décrit aussi tout le groupe mais dans un ordre différent. On obtient donc une permutation des entiers de $1, \dots, N$. On peut donc associer à tout élément g une permutation de façon biunivoque. Au produit de deux éléments est associé le produit des permutations correspondantes. Cette application établit donc un isomorphisme entre le groupe G et un sous-groupe du groupe des permutations.

Ceci démontre que tout groupe fini est isomorphe à un sous-groupe du groupe symétrique \mathfrak{S}_N ou groupe des permutations de N objets, que nous étudierons plus loin.

Représentation régulière. Pour un groupe fini G d'ordre N , on obtient une représentation en termes de matrices de la manière suivante : on considère une base de N vecteurs dans l'espace vectoriel \mathbb{R}^N . À chaque vecteur de base on associe un élément g du groupe et on note ce vecteur $|g\rangle$. On considère alors l'application linéaire $\mathcal{R}(g)$ de \mathbb{R}^N dans \mathbb{R}^N définie par son action sur les vecteurs de base :

$$\mathcal{R}(g) |g_i\rangle = |gg_i\rangle. \quad (3.1)$$

La représentation $G \mapsto \mathcal{R}(G)$, où $\mathcal{R}(G)$ est le groupe formé par ces applications linéaires, est un isomorphisme. De plus à tout élément du groupe est associée une permutation des vecteurs de base, et donc une matrice $N \times N$, ne comprenant qu'un seul 1 et des zéros dans chaque ligne et colonne.

Nous allons donner quelques exemples simples et utiles de groupes discrets.

3.1 Groupes de translations sur réseau

Le groupe additif des entiers est commutatif : l'élément neutre est 0 et l'inverse au sens des groupes l'opposé. C'est un groupe qui contient un nombre infini d'éléments, les éléments de \mathbb{Z} . C'est aussi le groupe des translations sur le réseau uni-dimensionnel des points de coordonnée entière. Ce groupe a deux générateurs les éléments $+1$ et -1 son opposé.

Ce groupe peut être généralisé à d dimensions. On obtient alors le groupe des translations du réseau des points de coordonnées entières à d dimensions. Si

$$\mathbf{x} \equiv \{x_1, x_2, \dots, x_d\}, \quad \mathbf{y} \equiv \{y_1, y_2, \dots, y_d\} \in \mathbb{Z}^d$$

le produit au sens de la loi de groupe est la somme des deux vecteurs $\mathbf{x} + \mathbf{y}$

$$\mathbf{x} + \mathbf{y} \equiv \{x_1 + y_1, x_2 + y_2, \dots, x_d + y_d\}.$$

Ce groupe qui se note \mathbb{Z}^d est évidemment abélien. Il a $2d$ générateurs.

Groupe de translations sur réseau fini périodique. Si nous nous restreignons à l'ensemble des entiers modulo N , que nous noterons $\mathbb{Z}_N \equiv \mathbb{Z}/N\mathbb{Z}$, l'addition engendre une structure de groupe commutatif fini à N éléments. Ce groupe a un seul générateur $+1$.

Le groupe des réflexions par rapport à une droite du plan n'a que deux éléments $\{R, \mathbf{1}\}$ et $R^2 = \mathbf{1}$: R est donc son propre inverse. Ce groupe est isomorphe à \mathbb{Z}_2 .

Le groupe des rotations dans le plan d'angle multiple entier de $2\pi/N$ est également un groupe abélien fini, isomorphe à \mathbb{Z}_N . Les éléments sont de la forme ω^n , $n = 1, \dots, N$, où ω est le générateur du groupe, la rotation d'angle $2\pi/N$.

Le groupe \mathbb{Z}_N est donc aussi le groupe des translations sur un réseau unidimensionnel fini avec conditions aux limites périodiques. De nouveau ces notions se généralisent en plusieurs dimensions. Le groupe de symétrie de translation du réseau carré de dimension $N_1 \times N_2$ avec conditions aux limites périodiques (qui a la topologie d'un tore) est $\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}$.

3.2 Groupe symétrique ou des permutations

Un exemple important de groupe fini non-abélien est fourni par le groupe symétrique ou groupe des (c'est à dire de toutes les) permutations de N objets, noté \mathfrak{S}_N , pour $N \geq 3$ (\mathfrak{S}_2 est isomorphe à \mathbb{Z}_2). Le groupe symétrique joue un rôle important

en mécanique quantique, puisqu'il permet de classer les fonctions d'onde à N particules par leurs propriétés de symétrie, et donc de construire les fonctions d'onde appropriées à des systèmes de bosons ou de fermions,

en mathématique puisque par exemple il contient tout groupe fini comme sous-groupe; puisqu'il intervient dans la décomposition des produits tensoriels de représentations en représentations irréductibles.

Le groupe \mathfrak{S}_3 . Le groupe \mathfrak{S}_3 est aussi le groupe de symétrie du triangle équilatéral. Il est engendré par deux éléments a, b , qui sont des transpositions au sens des permutations,

$$a [1 2 3] = [2 1 3] \quad b [1 2 3] = [1 3 2],$$

et du point de vue géométrique des réflexions par rapport à deux médiatrices du triangle. Ces éléments a, b satisfont donc

$$a^2 = 1, \quad b^2 = 1. \quad (3.2)$$

Leur produit ab qui est une permutation circulaire, et géométriquement une rotation d'angle $2\pi/3$ du plan, satisfait donc

$$(ab)^3 = 1. \quad (3.3)$$

Les relations (3.2,3.3) suffisent à caractériser le groupe \mathfrak{S}_3 . Ce groupe n'est pas abélien car

$$ab [1 2 3] = a [1 3 2] = [3 1 2], \quad ba [1 2 3] = b [2 1 3] = [2 3 1].$$

Les $3!$ différents éléments sont

$$1, \quad a, \quad b, \quad ab, \quad aba, \quad abab.$$

Ce groupe est isomorphe au groupe des matrices 2×2 engendré par

$$a = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad b = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}. \quad (3.4)$$

Il est un sous-groupe du groupe orthogonal à deux dimensions $O(2)$ que nous étudierons plus loin.

Groupe \mathfrak{S}_N : générateurs. On vérifie que les transpositions du groupe des permutations notées $t_i, i = 1, \dots, N-1$, où t_i permute l'élément i avec l'élément $i+1$, satisfont

$$t_i^2 = 1, \quad t_i t_j = t_j t_i \quad \text{pour} \quad |i - j| > 1 \quad \text{et} \quad t_i t_{i+1} t_i = t_{i+1} t_i t_{i+1}.$$

Réciproquement on vérifie que le groupe engendré par les éléments t_i qui satisfont ces relations et leurs produits a $N!$ éléments au plus. Donc les transpositions forment un ensemble de générateurs et ces relations caractérisent le groupe symétrique.

Le groupe \mathfrak{S}_4 est le groupe de symétrie du tétraèdre, sous-groupe fini du groupe orthogonal à trois dimensions $O(3)$.

Le groupe \mathfrak{S}_N est le groupe de symétrie du *simplexe*, un polytope régulier de l'espace à N dimensions qui généralise le triangle et le tétraèdre.

3.3 Transformations linéaires du réseau cubique général

Le réseau cubique général est le réseau des points de coordonnées entières dans l'espace de dimension quelconque N . Le groupe de symétrie du réseau cubique correspondant à des transformations linéaires (ce qui exclut les translations déjà mentionnées) est un groupe fini, que nous notons \mathfrak{C}_N . Il peut être défini de la manière suivante : agissant sur le point de coordonnées x_1, x_2, \dots, x_N il engendre les points ayant les coordonnées permutées et avec tous les changements de signe possibles :

$$(x_1, x_2, \dots, x_N) \xrightarrow{\text{par } \mathfrak{C}_N} (\pm x_{P_1}, \pm x_{P_2}, \dots, \pm x_{P_N}),$$

où $\{P_1, P_2, \dots, P_N\}$ est une permutation quelconque des entiers $\{1, 2, \dots, N\}$, et les signes sont tous indépendants. Il a donc $2^N N!$ éléments. C'est un sous-groupe de \mathfrak{S}_{2N} . Il est engendré par les $N-1$ générateurs de \mathfrak{S}_N et une réflexion supplémentaire, par exemple

$$(x_1, x_2, \dots, x_N) \mapsto (-x_1, x_2, \dots, x_N).$$

Ce groupe de symétrie est aussi le groupe de symétrie de la maille élémentaire du réseau, un carré ($N = 2$), un cube ($N = 3$) ou plus généralement un hypercube en dimension N .

Groupe du carré. Pour illustrer cette définition nous examinons d'abord l'exemple de la dimension deux et donc de \mathfrak{C}_2 . Agissant sur le point de coordonnées x_1, x_2 d'un point du réseau il le transforme en

$$(\pm x_1, \pm x_2) \quad \text{et} \quad (\pm x_2, \pm x_1).$$

Ce groupe a 8 éléments. C'est aussi un groupe de symétrie du réseau carré. Il peut alors être engendré par deux réflexions : une réflexion par rapport à une diagonale que nous notons a ,

$$a(x_1, x_2) = (x_2, x_1), \tag{3.5}$$

et une réflexion b par rapport à un axe de symétrie parallèle aux axes, par exemple

$$b(x_1, x_2) = (-x_1, x_2). \tag{3.6}$$

Alors

$$a^2 = 1, \quad b^2 = 1, \quad (ab)^4 = 1. \tag{3.7}$$

De façon générale tous les polygones réguliers sont associés à des groupes finis, sous-groupes du groupe $O(2)$. Ils peuvent être réalisés sous forme de matrices réelles 2×2 ou d'opérations sur les nombres complexes. Les groupes du triangle, du carré, de l'hexagone sont aussi les groupes de symétries des réseaux dont ils forment la maille élémentaire.

Groupe du cube. Le groupe \mathfrak{C}_3 est aussi un groupe de symétrie du réseau cubique. Il peut être défini comme le groupe qui change le point de coordonnées (x_1, x_2, x_3) en

$$(\pm x_1, \pm x_2, \pm x_3) \quad \text{ainsi que les trois permutations de } (x_1, x_2, x_3).$$

Il est engendré par les deux transpositions t_1, t_2 de \mathfrak{S}_3

$$t_1(x_1, x_2, x_3) = (x_2, x_1, x_3), \quad t_2(x_1, x_2, x_3) = (x_1, x_3, x_2), \quad (3.8)$$

et une réflexion r qui par exemple change le signe de la première coordonnée

$$r(x_1, x_2, x_3) = (-x_1, x_2, x_3). \quad (3.9)$$

Il compte 48 éléments.

Notons que le groupe de symétrie du cube est aussi celui de son dual l'octaèdre. Les autres solides (polyèdres) réguliers (dodécaèdre et son dual l'icosaèdre) ont comme groupe de symétrie un sous-groupe fini de $O(3)$ de 120 éléments.

Exercices

Exercice 3.1

L'ensemble des deux matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

forme-t-il un groupe?

Exercice 3.2

Le groupe \mathfrak{S}_3 .

a- Vérifier que deux réflexions du triangle équilatéral engendrent le groupe \mathfrak{S}_3 .

b- Vérifier les relations (3.2,3.3) et en déduire que a, b sont un ensemble de générateurs de \mathfrak{S}_3 .

c -Identifier les éléments correspondant à des permutations circulaires, et montrer qu'ils forment un sous-groupe. À quelles transformations géométriques du triangle correspondent-elles? Montrer que le sous-groupe est distingué.

d- Montrer que les deux matrices (3.4)

$$a = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad b = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix}$$

satisfont les relations (3.2,3.3), et sont donc les générateurs d'un groupe de matrices isomorphe à \mathfrak{S}_3 .

e- Calculer les déterminants des six matrices. Relier la valeur du déterminant au sous-groupe des permutations circulaires. Que peut-on dire des valeurs propres?

Exercice 3.3

Vérifier que les éléments a, b définis par (3.5) et (3.6) satisfont les relations (3.7). En déduire qu'ils forment un ensemble de générateurs du groupe \mathfrak{C}_2 du carré.

Trouver une représentation du groupe par un ensemble de matrices 2×2 . Calculer les déterminants.

Exercice 3.4

Vérifier que les éléments (3.8,3.9) sont des générateurs du groupe du cube \mathfrak{C}_3 . Trouver une représentation en termes de matrices 3×3 .

4 Groupes continus : groupes de Lie et isométries du plan

Les groupes de Lie constituent une famille excessivement riche. Elle contient les groupes dont les éléments peuvent être paramétrés en termes d'un nombre fini de paramètres réels. Par ailleurs ce sont des groupes topologiques, ce qui donne un sens à parler de deux éléments proches.

Plutôt que définir de façon plus précise ce qu'est un groupe de Lie, structure dont nous n'aurons pas besoin dans toute sa généralité, nous allons de nouveau illustrer cette notion par des exemples. Dans cette section nous allons d'abord donner des exemples de groupes commutatifs puis non commutatifs, correspondant aux isométries du plan.

Translations. Le groupe des translations de la droite réelle est un groupe continu et commutatif. Les éléments sont paramétrés par des réels, et le groupe est simplement le groupe additif des réels. De façon plus générale on définit les translations de l'espace euclidien \mathbb{R}^n par une collection de n réels. On dit que ces groupes sont des groupes de Lie abéliens non compacts au sens de la topologie. En physique les translations dans le temps jouent aussi un rôle distinct mais sont mathématiquement identiques.

Un exemple un peu plus intéressant par sa structure est fourni par le groupe des rotations–réflexions du plan. Combiné avec le groupe des translations (ce n'est pas un produit de groupes), il engendre le groupe des isométries du plan, c'est à dire des transformations du plan qui ne changent pas les distances.

4.1 Rotations et réflexions du plan

Les rotations du plan peuvent être paramétrées par un angle θ , c'est à dire un réel modulo 2π . Elles sont représentées par des matrices $\mathbf{g}(\theta)$ réelles 2×2

$$\mathbf{g}(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (4.1)$$

On vérifie

$$\mathbf{g}(\theta_1)\mathbf{g}(\theta_2) = \mathbf{g}(\theta_1 + \theta_2) = \mathbf{g}(\theta_2)\mathbf{g}(\theta_1).$$

Cette loi de multiplication combinée avec l'identité $\mathbf{g}(\theta) = \mathbf{g}(\theta + 2\pi)$ exprime la propriété que le groupe des rotations en dimension deux est isomorphe au groupe $\mathbb{R}/2\pi\mathbb{Z}$ des translations sur le cercle, et constitue la définition abstraite du groupe.

La relation d'orthogonalité

$$\mathbf{g}^T(\theta)\mathbf{g}(\theta) = \mathbf{1},$$

(\mathbf{g}^T est la matrice transposée) traduit la propriété que les rotations conservent la norme des vecteurs.

On vérifie aussi que $\det \mathbf{g}(\theta) = 1$ ce qui implique que la transformation conserve l'élément d'aire et appliquée à un repère cartésien conserve son orientation.

Sous cette forme le groupe est noté $SO(2)$, ce qui signifie le groupe des matrices orthogonales de déterminant unité.

Matrices de Pauli. Pour discuter les groupes des rotations à deux et trois dimensions, il est commode d'introduire les trois matrices de Pauli

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (4.2)$$

qui satisfont

$$\sigma_i = \sigma_i^\dagger, \quad \text{tr } \sigma_i = 0,$$

et dont les produits peuvent s'écrire

$$\begin{aligned} \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \mathbf{1}, \\ \sigma_1\sigma_2 = -\sigma_2\sigma_1 = i\sigma_3, \quad \sigma_2\sigma_3 = -\sigma_3\sigma_2 = i\sigma_1, \quad \sigma_3\sigma_1 = -\sigma_1\sigma_3 = i\sigma_2, \end{aligned}$$

ce qui se résume en

$$\sigma_i\sigma_j = \delta_{ij}\mathbf{1} + i \sum_k \epsilon_{ijk}\sigma_k, \quad (4.3)$$

où ϵ_{ijk} est le symbole complètement antisymétrique

$$\epsilon_{jik} = -\epsilon_{ijk}, \quad \epsilon_{ikj} = -\epsilon_{ijk}, \quad \epsilon_{123} = 1. \quad (4.4)$$

Une conséquence de l'identité (4.3) sont les relations de commutation

$$[\sigma_i, \sigma_j] = 2i \sum_k \epsilon_{ijk}\sigma_k, \quad (4.5)$$

qui impliquent que les matrices σ_i sont les générateurs d'une algèbre de Lie de dimension trois.

La matrice unité $\mathbf{1} \equiv \sigma_0$ et les trois matrices σ_i sont orthogonales au sens de la trace

$$\text{tr } \sigma_\mu \sigma_\nu^\dagger = \text{tr } \sigma_\mu \sigma_\nu = 2\delta_{\mu\nu}, \quad \mu, \nu = 0, 1, 2, 3.$$

Les matrices 2×2 complexes forment un espace vectoriel de dimension quatre. Les trois matrices de Pauli et la matrice unité étant linéairement indépendantes forment une base de cet espace. Toute matrice \mathbf{M} peut donc s'écrire

$$\mathbf{M} = z_0\mathbf{1} + \sum_{i=1}^3 z_i\sigma_i, \quad z_0, z_1, z_2, z_3 \in \mathbf{C}.$$

Dans ces conditions

$$\text{tr } \mathbf{M} = 2z_0.$$

On en déduit que, sur \mathbf{C} , les matrices de Pauli forment une base du sous-espace des matrices de trace nulle.

Si \mathbf{M} est hermitienne

$$\mathbf{M}^\dagger = z_0^* + \sum_{i=1}^3 z_i^* \sigma_i = \mathbf{M} = z_0 \mathbf{1} + \sum_{i=1}^3 z_i \sigma_i.$$

Donc les quatre coefficients z_i sont réels. Sur \mathbb{R} les quatre matrices $\mathbf{1}, \{\sigma_i\}$ forment une base de l'espace des matrices hermitiennes, et les trois matrices de Pauli forment une base pour les matrices hermitiennes de trace nulle.

Groupes $O(2)$ et $SO(2)$. Un élément $\mathbf{g}(\theta)$ de $SO(2)$ peut alors s'écrire

$$\mathbf{g}(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \mathbf{1} \cos \theta - i \sigma_2 \sin \theta. \quad (4.6)$$

Une rotation ne change pas la longueur d'un vecteur.

De façon générale les matrices R qui laissent la longueur de tout vecteur \mathbf{v} du plan invariant satisfont

$$|R\mathbf{v}| = |\mathbf{v}| \Rightarrow \mathbf{v}^2 = \mathbf{v} R^T R \mathbf{v} \quad \forall \mathbf{v},$$

où R^T est la transposée de la matrice R . Ceci entraîne

$$R^T R = \mathbf{1},$$

relation qui définit une matrice orthogonale. Les matrices orthogonales forment un groupe noté $O(2)$.

Par ailleurs, prenant le déterminant des deux membres, on obtient

$$\det R^T \det R = (\det R)^2 = 1 \Rightarrow \det R = \pm 1.$$

Comme nous l'avons montré de façon générale en section 2.1, les matrices de déterminant 1 forment un sous-groupe. De plus on note que si R_1 et R_2 sont deux éléments du groupe

$$\det R_1 = 1 \text{ et } \det R_2 = 1 \Rightarrow \det(R_1 R_2) = \det R_1 \det R_2 = 1.$$

Nous reconnaissons le groupe $SO(2)$ sous-groupe de $O(2)$ qui correspond aux transformations qui respectent l'orientation d'un repère cartésien.

Pour engendrer tout le groupe $O(2)$ il suffit d'introduire un élément de $O(2)$ supplémentaire, qui est une réflexion par rapport à une droite, par exemple l'élément

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \Rightarrow \sigma_3 \sigma_3^T = \sigma_3^2 = \mathbf{1}, \quad \det \sigma_3 = -1.$$

En effet un élément quelconque de $O(2)$ soit appartient à $SO(2)$, soit son produit par σ_3 a déterminant 1,

$$\det R = -1 \Rightarrow \det(R\sigma_3) = \det R \det \sigma_3 = 1,$$

et donc appartient à $SO(2)$. Tout élément \mathbf{r} de $O(2)$ qui n'appartient pas à $SO(2)$ peut donc s'écrire

$$\mathbf{g}(\theta)\sigma_3 = \sigma_3 \cos \theta + \sigma_1 \sin \theta = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}.$$

Notons par ailleurs que

$$\mathbf{r} = \sigma_3 \mathbf{g}(\theta) \sigma_3 = \mathbf{1} \cos \theta + i\sigma_2 \sin \theta = \mathbf{g}(-\theta), \quad (4.7)$$

et donc $O(2)$ qui est aussi appelé groupe des rotations–réflexions, à la différence de $SO(2)$, n'est pas abélien.

De plus toute matrice $\mathbf{r} = \mathbf{g}(\theta)\sigma_3$ satisfait

$$\mathbf{r}^2 = (\mathbf{g}(\theta)\sigma_3)^2 = \mathbf{1}, \quad \mathbf{r}\mathbf{g}(\varphi)\mathbf{r} = \mathbf{g}(-\varphi). \quad (4.8)$$

et donc c'est une réflexion, par rapport à la droite d'angle $\theta/2$. Ces deux propriétés définissent du point de vue du groupe abstrait une réflexion.

Remarques.

(i) Comme $\sigma_1 = \sigma_1^{-1}$, la relation (4.7) établit un isomorphisme assez évident entre le groupe des éléments $\mathbf{g}(\theta)$ et $\mathbf{g}(-\theta)$, composé des rotations d'angle opposé.

(ii) On vérifie que le centre de $O(2)$ se réduit aux matrices $\pm\mathbf{1}$ qui forment un sous-groupe \mathbb{Z}_2 . L'ensemble quotient $O(2)/\mathbb{Z}_2$ est donc un groupe, le groupe des réflexions et rotations d'angle 2θ .

4.2 Le groupe $U(1)$

Nous pouvons aussi représenter les points du plan par un nombre complexe $z = x_1 + ix_2$. Une rotation de centre l'origine est alors représentée par une multiplication par des nombres complexes de module 1,

$$z \mapsto z_\theta = e^{i\theta} z, \Rightarrow |z| = |z_\theta|.$$

Le produit de deux éléments du groupe est simplement

$$e^{i\theta_1} e^{i\theta_2} = e^{i(\theta_1+\theta_2)},$$

et sous cette forme on parle du groupe $U(1)$. Les groupes $SO(2)$ et $U(1)$ sont isomorphes.

Notons que les éléments $e^{-i\theta}$ forment une autre représentation, isomorphe puisque obtenue par conjugaison complexe, du groupe $U(1)$,

$$e^{-i\theta_1} e^{-i\theta_2} = e^{-i(\theta_1+\theta_2)}.$$

Dans la représentation du plan par des nombres complexes, la réflexion \mathbf{r} des relations (4.8) peut être représentée par une transformation antilinéaire, la

conjugaison complexe $z \mapsto z^*$. Elle laisse zz^* invariant. Son carré est l'identité, et par ailleurs

$$(e^{i\theta} z^*)^* = e^{-i\theta} z.$$

La conjugaison complexe satisfait donc la relation (4.8) et du point des matrices réelles correspond au choix $\mathbf{r} = \sigma_3$.

La conjugaison complexe transforme la représentation initiale en sa complexe conjuguée. Si l'on veut qu'elle soit une opération interne il faut doubler l'espace de représentation et agir sur des éléments de \mathbb{C}^2 , comme nous allons le voir.

Représentation complexe de $O(2)$. Une matrice orthogonale est une matrice unitaire réelle. Toute matrice \mathbf{U} telle que

$$\mathbf{U}^\dagger \mathbf{U} = \mathbf{1},$$

est appelée matrice unitaire. Elle conserve la longueur des vecteurs complexes :

$$\|\mathbf{v}\|^2 = \mathbf{v}^\dagger \cdot \mathbf{v} = |v_1|^2 + |v_2|^2, \quad \|\mathbf{U}\mathbf{v}\|^2 = \mathbf{v}^\dagger \mathbf{U}^\dagger \mathbf{U} \mathbf{v} = \|\mathbf{v}\|^2.$$

On appelle $U(2)$ le groupe des matrices unitaires 2×2 . Donc $O(2)$ est un sous-groupe de $U(2)$.

De même le groupe $SO(2)$ est un sous-groupe du sous-groupe $SU(2)$ de $U(2)$, correspondant aux matrices unitaires de déterminant 1.

Une matrice de $SO(2)$ peut être diagonalisée par une transformation unitaire. En fait comme $\mathbf{g}(\theta)$ s'exprime uniquement en fonction de σ_2 il suffit de diagonaliser σ_2 . Ceci peut être réalisé en utilisant la matrice unitaire

$$U = \frac{1}{\sqrt{2}} (\mathbf{1} + i\sigma_1).$$

En effet,

$$U (\cos \theta - i \sin \theta \sigma_2) U^\dagger = \cos \theta + i \sin \theta \sigma_3 = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}. \quad (4.9)$$

Nous avons pu diagonaliser toutes les matrices du groupe simultanément par la même transformation. Par conséquent la représentation complexe de $SO(2)$ est *réductible*. Cela signifie que l'espace des vecteurs complexes à deux dimensions se décompose en deux sous-espaces invariants par tous les éléments du groupe $SO(2)$: si on agit sur un vecteur $(z_1, 0)$ par tous les éléments du groupe on obtient toujours un vecteur dont la deuxième composante s'annule,

$$\mathbf{g}(\theta)(z_1, 0) = e^{i\theta}(z_1, 0),$$

et de même pour $(0, z_2)$

$$\mathbf{g}(\theta)(0, z_2) = e^{-i\theta}(0, z_2).$$

Nous avons obtenu deux copies isomorphes du groupe $U(1)$ (et donc $SO(2)$), correspondant à la représentation de définition et sa complexe conjuguée.

Par contre pour engendrer le groupe $O(2)$ il faut rajouter une réflexion, par exemple σ_3 dans la base initiale qui maintenant devient

$$U\sigma_3U^\dagger = \sigma_2.$$

Nous voyons que cette matrice n'est pas diagonale. L'espace \mathbb{C}^2 ne se décompose plus. On dit que la représentation complexe de $O(2)$ est *irréductible*.

4.3 Générateur. Exponentiation

On peut prendre comme distance entre deux éléments de $SO(2)$ ou $U(1)$ le minimum de $|\theta_1 - \theta_2|$. Avec cette distance un élément $e^{i\theta}$ de $U(1)$ tend vers l'identité quand $\theta \rightarrow 0$. Par ailleurs tout élément de $U(1)$ peut être obtenu par un produit d'éléments appartenant à un voisinage arbitrairement petit de l'identité puisque

$$e^{i\theta} = \left(e^{i\theta/n}\right)^n.$$

Toute la structure du groupe peut donc être caractérisée par le voisinage de l'identité

$$e^{i\theta} = 1 + i\theta + O(\theta^2),$$

et la propriété de groupe. Le coefficient du terme linéaire en θ , ici la multiplication par i , suffit à caractériser l'action du groupe.

La même propriété est nécessairement vraie pour le groupe $SO(2)$ qui lui est isomorphe. Pour $\theta \rightarrow 0$,

$$\mathbf{g}(\theta) = \mathbf{1} - i\theta\sigma_2 + O(\theta^2).$$

La propriété de groupe conduit alors à une propriété d'exponentiation. Un élément quelconque du groupe peut être écrit

$$\mathbf{g}(\theta) = e^{\theta \mathbf{T}} = e^{-i\sigma_2\theta} \equiv \sum_{k=0}^{\infty} \frac{\theta^k}{k!} (-i\sigma_2)^k, \quad (4.10)$$

où le développement en puissances de θ définit ce que nous appelons exponentielle d'une matrice (la somme étant convergente au sens de la norme usuelle des matrices). La matrice $\mathbf{T} = -i\sigma_2$, $\mathbf{T}^2 = -\mathbf{1}$, est le générateur d'une algèbre à laquelle appartiennent tous les éléments du groupe de Lie $SO(2)$. Notons que seul le groupe $SO(2)$ peut être obtenu par exponentiation. En effet par continuité tous les éléments au voisinage de l'identité ont pour déterminant l'unité, et donc aussi tous leurs produits. Pour engendrer le groupe $O(2)$ il faut ajouter un générateur qui est une réflexion.

Remarque. La propriété que tout le groupe est engendré par les éléments voisins de l'identité est très importante. En effet elle entraîne que l'invariance d'une quantité ou plus généralement ses propriétés de transformation peuvent être étudiées en faisant des transformations 'infinitésimales'.

4.4 Représentations

Dans la suite nous nous intéressons aux propriétés de transformation de polynômes dans les composantes de vecteurs sur lesquels agissent une représentation d'un groupe. Comme les transformations sont linéaires, les polynômes de degré différent ne se mélangent pas. On peut donc se limiter aux polynômes homogènes. Par ailleurs les polynômes homogènes de degré n peuvent être considérés comme des restrictions de formes multilinéaires de n vecteurs dont il suffit donc d'étudier les propriétés de transformation.

Transformation de formes bilinéaires. Nous illustrons ce sujet avec l'exemple de formes bilinéaires dans deux vecteurs \mathbf{x}, \mathbf{y} sur lesquels agissent les groupes $SO(2)$ et $O(2)$.

Soit \mathbf{M} une matrice réelle 2×2 de composantes M_{ij} et considérons dans un repère donné la forme bilinéaire réelle

$$\mathcal{M}(\mathbf{x}, \mathbf{y}) = \sum_{i,j} M_{ij} x_i y_j,$$

où x_i, y_i sont les composantes des deux vecteurs.

Nous faisons maintenant un changement de repère, correspondant à une rotation de matrice \mathbf{R} de composantes R_{ij} :

$$x_i = \sum_j R_{ij} x'_j, \quad y_i = \sum_j R_{ij} y'_j.$$

On cherche la transformation $\mathbf{M} \mapsto \mathbf{M}'$ telle que la quantité \mathcal{M} soit invariante :

$$\mathcal{M}(\mathbf{x}, \mathbf{y}) = \mathcal{M}'(\mathbf{x}', \mathbf{y}') = \sum_{i,j} M'_{ij} x'_i y'_j.$$

Substituant (et utilisant $\mathbf{R}^T \mathbf{R} = \mathbf{1}$) on trouve

$$M_{ij} = \sum_{k,l} R_{ik} R_{jl} M'_{kl}. \quad (4.11)$$

Cette relation définit une transformation linéaire sur l'espace vectoriel de dimension quatre des matrices réelles 2×2 . On peut introduire une notation symbolique pour écrire cette transformation :

$$\mathbf{M} = (\mathbf{R} \otimes \mathbf{R}) \mathbf{M}',$$

où chaque matrice \mathbf{R} agit sur un indice de \mathbf{M}' . On vérifie que pour deux rotations successives,

$$\mathbf{M} = (\mathbf{R}_1 \otimes \mathbf{R}_1) \mathbf{M}', \quad \mathbf{M}' = (\mathbf{R}_2 \otimes \mathbf{R}_2) \mathbf{M}'' ,$$

la résultat est

$$\mathbf{M} = (\mathbf{R}_1 \mathbf{R}_2 \otimes \mathbf{R}_1 \mathbf{R}_2) \mathbf{M}'' .$$

L'application $\mathbf{R} \mapsto \mathbf{R} \otimes \mathbf{R}$ est une représentation de dimension quatre du groupe $O(2)$. On appelle l'opération \otimes *produit tensoriel* et on dit que M_{ij} se transforme comme un *tenseur* de rang deux par le groupe $O(2)$.

Si nous rangeons les composantes M_{ij} dans un ordre canonique, par exemple $M_{11}, M_{12}, M_{21}, M_{22}$, la transformation correspondant à la rotation (4.1), $\mathbf{R} = \mathbf{g}(\theta)$, s'écrit comme une matrice 4×4 :

$$\begin{pmatrix} M_{11} \\ M_{12} \\ M_{21} \\ M_{22} \end{pmatrix} = \begin{pmatrix} \cos^2 \theta & -\cos \theta \sin \theta & -\cos \theta \sin \theta & \sin^2 \theta \\ \cos \theta \sin \theta & \cos^2 \theta & -\sin^2 \theta & -\cos \theta \sin \theta \\ \cos \theta \sin \theta & -\sin^2 \theta & \cos^2 \theta & -\cos \theta \sin \theta \\ \sin^2 \theta & \cos \theta \sin \theta & \cos \theta \sin \theta & \cos^2 \theta \end{pmatrix} \begin{pmatrix} M'_{11} \\ M'_{12} \\ M'_{21} \\ M'_{22} \end{pmatrix}. \quad (4.12)$$

Une question naturelle se pose cette représentation est-elle irréductible?

Comme le tenseur \mathbf{M} n'a que deux indices, on peut organiser les quatre composantes comme une matrice 2×2 (ce qui rend l'exercice facile mais dangereux car il peut conduire à confondre matrice et tenseur). Toute matrice 2×2 peut être écrite de façon unique comme la somme d'une matrice antisymétrique \mathbf{A} , une matrice symétrique \mathbf{S} de trace nulle, et un multiple de la matrice unité $a_0 \mathbf{1}$

$$\mathbf{M} = \mathbf{A} + \mathbf{S} + a_0 \mathbf{1} \Leftrightarrow \mathbf{A} = \frac{1}{2}(\mathbf{M} - \mathbf{M}^T), \quad \mathbf{S} = \frac{1}{2}(\mathbf{M} + \mathbf{M}^T) - \frac{1}{2} \mathbf{1} \operatorname{tr} \mathbf{M}.$$

Par ailleurs, en dimension deux,

$$a_0 = \frac{1}{2} \operatorname{tr} \mathbf{M}, \quad \mathbf{A} = ia_3 \sigma_2, \quad \mathbf{S} = a_1 \sigma_3 + a_2 \sigma_1, \quad a_i, i = 0, \dots, 3 \in \mathbb{R}. \quad (4.13)$$

Nous avons décomposé l'espace vectoriel de dimension quatre en la somme de deux espaces de dimensions un et un espace de dimension deux, choisissant comme base de l'espace vectoriel quatre matrices 2×2 , la matrice unité, la matrice $i\sigma_2$ qui correspond au tenseur antisymétrique, et les deux matrices σ_1, σ_3 qui forment une base des tenseurs symétriques de trace nulle :

$$\mathbf{M} = a_0 \mathbf{1} + ia_3 \sigma_2 + a_1 \sigma_3 + a_2 \sigma_1. \quad (4.14)$$

La forme bilinéaire s'écrit alors

$$\mathcal{M}(\mathbf{x}, \mathbf{y}) = a_0(x_1 y_1 + x_2 y_2) + a_2(x_1 y_2 - x_2 y_1) + a_1(x_1 y_1 - x_2 y_2) + a_2(x_1 y_2 + x_2 y_1).$$

La relation (4.11) peut se récrire

$$\mathbf{M} = \mathbf{R} \mathbf{M}' \mathbf{R}^T \Leftrightarrow \mathbf{M}' = \mathbf{R}^T \mathbf{M} \mathbf{R}.$$

(i) Si \mathbf{M} est proportionnelle à la matrice unité,

$$\mathbf{1} = \mathbf{R} \mathbf{1} \mathbf{R}^T,$$

et donc a_0 est invariant par toutes transformations, c'est un scalaire, la composante sur le produit scalaire

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2,$$

des deux vecteurs.

(ii) Si \mathbf{M} est antisymétrique, on vérifie immédiatement que \mathbf{M}' est aussi antisymétrique. De plus, par (4.6), si \mathbf{R} appartient à $SO(2)$ il commute avec σ_2 et donc la composante a_3 est invariante. Au contraire si $\mathbf{R} = \sigma_3$ (une réflexion) parce que σ_2 et σ_3 anti-commutent, la composante a_3 change de signe. On dit que a_3 est la composante pseudo-scalaire. Elle est la composante sur la composante

$$x_1 y_2 - x_2 y_1,$$

sur l'axe perpendiculaire au plan du produit vectoriel $\mathbf{x} \times \mathbf{y}$.

(iii) Si \mathbf{M} est symétrique on vérifie que \mathbf{M}' est aussi symétrique. De plus

$$\text{tr } \mathbf{M} = \text{tr}(\mathbf{R}\mathbf{M}'\mathbf{R}^T) = \text{tr}(\mathbf{M}'\mathbf{R}^T\mathbf{R}) = \text{tr } \mathbf{M}',$$

donc la trace est invariante.

Les tenseurs symétriques de trace nulle donc forment une représentation de dimension deux appelée représentation adjointe du groupe.

On peut se demander si elle isomorphe à la représentation initiale. Pour cela utilisons la paramétrisation (4.13) combinée avec la représentation (4.6),

$$\begin{aligned} a_1\sigma_3 + a_2\sigma_1 &= (\cos\theta - i\sigma_2 \sin\theta)(a'_1\sigma_3 + a'_2\sigma_1)(\cos\theta + i\sigma_2 \sin\theta) \\ &= (a'_1 \cos(2\theta) - a'_2 \sin(2\theta))\sigma_3 + (a'_1 \sin(2\theta) + a'_2 \cos(2\theta))\sigma_1. \end{aligned}$$

Nous reconnaissons une rotation d'angle 2θ , et donc une rotation d'angle π est représentée par l'identité : la représentation de $SO(2)$ n'est pas un isomorphisme.

Réflexion. Pour engendrer le groupe $O(2)$, il nous reste à calculer l'effet d'une réflexion, par exemple correspondant à σ_3 :

$$a_1\sigma_3 + a_2\sigma_1 = \sigma_3(a'_1\sigma_3 + a'_2\sigma_1)\sigma_3 = a'_1\sigma_3 - a'_2\sigma_1.$$

On trouve bien l'effet d'une réflexion par rapport à l'axe 1 dans l'espace à deux dimensions.

La transformation linéaire (4.14) a diagonalisée simultanément toutes les matrices (4.12) par blocs. Les trois sous-espaces propres correspondent maintenant à des représentations irréductibles du groupe des rotations :

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta & 0 \\ 0 & \sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a'_0 \\ a'_1 \\ a'_2 \\ a'_3 \end{pmatrix}.$$

La réflexion σ_3 est représentée ici par une matrice diagonale d'éléments diagonaux $(1, 1, -1, -1)$.

Version $U(1)$. Pour étudier les propriétés des autres formes il est commode d'introduire les nombres complexes

$$u = \mathbf{x}_1 + i\mathbf{x}_2, \quad v = \mathbf{y}_1 + i\mathbf{y}_2,$$

correspondant aux deux vecteurs et d'examiner l'effet des transformations de $U(1)$ et de la conjugaison complexe.

Le produit scalaire s'écrit $\text{Re}(u^*v)$, où u^* veut dire complexe conjugué, dont il est immédiat qu'il est invariant par $O(2)$. En effet,

$$u^*v \mapsto (e^{i\theta} u)^*(e^{i\theta} v) = u^*v, \quad u^*v \mapsto uv^* \Rightarrow \text{Re}(uv^*) = \text{Re}(u^*v).$$

Notons que la partie imaginaire de uv^* est aussi invariante par $U(1)$:

$$\text{Im}(u^*v) = \mathbf{x}_1\mathbf{y}_2 - \mathbf{x}_2\mathbf{y}_1 .$$

et donc la combinaison correspondante par $SO(2)$. Cette quantité est la composante sur l'axe perpendiculaire au plan du produit vectoriel $\mathbf{x} \times \mathbf{y}$. Par contre par conjugaison complexe elle change de signe,

$$\text{Im}(uv^*) = -\text{Im}(u^*v).$$

On dit que cette quantité se transforme comme un pseudo-scalaire.

Si maintenant nous considérons les quatre éléments $\mathbf{x}_i\mathbf{y}_j$, $i, j = 1, 2$, nous voyons qu'il n'en reste que deux combinaisons linéaires dont nous n'avons pas examiné les propriétés de transformation,

$$\{\mathbf{x}_1\mathbf{y}_1 - \mathbf{x}_2\mathbf{y}_2 = \text{Re}(uv), \mathbf{x}_2\mathbf{y}_1 + \mathbf{x}_1\mathbf{y}_2 = \text{Im}(uv)\},$$

qui sont les composantes du nombre complexe uv . Celui-ci se transforme en $uv e^{2i\theta}$ par $U(1)$ et en $u^*v^* = (uv)^*$ par réflexion. Ces deux composantes se comportent comme un vecteur, excepté que ce vecteur tourne d'un angle 2θ dans une rotation d'angle θ . Cette représentation de $SO(2)$, dite de spin deux, n'est pas un isomorphisme puisque à deux éléments de $SO(2)$ qui diffèrent d'un angle π correspondent le même élément de la représentation.

Nous notons que l'objet à deux indices $\mathbf{x}_i\mathbf{y}_j$ a été décomposé en une combinaison linéaire d'objets qui se transforment indépendamment :

$$\mathbf{x}_i\mathbf{y}_j = \frac{1}{2}\delta_{ij}\mathbf{x} \cdot \mathbf{y} + \frac{1}{2}(\mathbf{x}_i\mathbf{y}_j - \mathbf{x}_j\mathbf{y}_i) + \frac{1}{2}(\mathbf{x}_i\mathbf{y}_j + \mathbf{x}_j\mathbf{y}_i - \delta_{ij}\mathbf{x} \cdot \mathbf{y}),$$

qui ont la nature d'une trace, d'une partie antisymétrique dans les indices i, j et d'une partie symétrique de trace nulle.

Formes multilinéaires. Toute cette construction se généralise aux formes multilinéaires de ℓ vecteurs \mathbf{x}^a , $a = 1, \dots, \ell$ qui sont des combinaisons linéaires des produits de composantes $\mathbf{x}_{i_1}^1 \mathbf{x}_{i_2}^2 \dots \mathbf{x}_{i_\ell}^\ell$. Comme nous l'avons vu dans l'analyse précédente il est plus simple d'introduire les nombres complexes correspondants \mathbf{u}^a et $(\mathbf{u}^a)^*$. Les composantes $\prod_a \mathbf{u}^a$ se transforment par $e^{i\ell\theta}$ c'est à dire par la représentation dite de spin ℓ .

Dans les composantes des vecteurs réels elles correspondent à des combinaisons symétriques dans tous les indices, et dont toutes les traces partielles sont nulles.

Le produit des nombres complexes conjugués se transforme par $e^{-i\ell\theta}$, et les produits mixtes correspondent à tous les spins m tels que $m + \ell$ soit pair.

Notons que si un élément \mathbf{w} d'un espace vectoriel se transforme par $\mathbf{g}(\theta)$ comme $e^{i\ell\theta}$, alors prenant la limite $\theta \rightarrow 0$, on en déduit

$$\mathbf{g}(\theta)\mathbf{w} = e^{i\ell\theta} \mathbf{w} \Rightarrow \mathbf{T}\mathbf{w} = i\ell\mathbf{w} .$$

Cet élément est donc un vecteur propre de \mathbf{T} avec la valeur propre $i\ell$. En fait $-i\mathbf{T} = -\sigma_2$ joue le rôle de l'opérateur moment cinétique de la mécanique quantique. La représentation qui correspond aux transformations en $e^{i\ell\theta}$ correspond au moment cinétique ℓ .

Action sur des espaces de fonctions. La mécanique quantique nécessite la définition de l'action de groupes sur des fonctions (les fonctions d'onde).

Soit donc un espace vectoriel de fonctions périodiques $f(\theta)$ sur $[0, 2\pi]$ de carré sommable. Alors l'élément du groupe $SO(2)$ est un opérateur qui agit par translation de l'argument, c'est à dire,

$$\mathbf{g}(\theta)f(\theta') = f(\theta + \theta').$$

Comme extension de la définition (cf. section 3, équation (3.1)), cette représentation peut être considérée comme la représentation régulière du groupe $SO(2)$ puisque à tout vecteur est associé un élément du groupe et le transformé du vecteur correspond au produit des éléments du groupe.

Une fonction périodique régulière peut être développée sur une base. Dans la version complexe,

$$f(\theta) = \sum_{\ell=-\infty}^{+\infty} f_{\ell} e^{i\ell\theta}.$$

L'opérateur qui engendre une rotation infinitésimale et qui donc représente la matrice $-i\sigma_2$ est alors l'opérateur différentiel

$$\mathbf{T} = \frac{d}{d\theta}.$$

Dans une interprétation quantique \mathbf{T} est proportionnel à l'opérateur moment cinétique \mathbf{L} :

$$\mathbf{L} = \frac{\hbar}{i}\mathbf{T}.$$

Agissant sur une des fonctions de base, \mathbf{T} donne

$$\mathbf{T}e^{i\ell\theta} = i\ell e^{i\ell\theta} \Rightarrow \mathbf{L}e^{i\ell\theta} = \ell\hbar e^{i\ell\theta}.$$

Mesure invariante sur le groupe $SO(2)$. Considérons une fonction périodique $f(\theta)$ et l'intégrale

$$\mathcal{I}(f) = \frac{1}{2\pi} \int_0^{2\pi} d\theta f(\theta).$$

On dit que la mesure $d\theta$ est une mesure invariante par le groupe $SO(2)$: D'abord nous remarquons qu'au paramètre θ nous pouvons associer un élément $\mathbf{g}(\theta)$ du groupe $SO(2)$. Nous pouvons alors faire un changement de variables qui correspond au produit de $\mathbf{g}(\theta)$ par un élément fixe $\mathbf{g}(\theta_0)$. Cela correspond à la translation $\theta' = \theta + \theta_0$ qui laisse la mesure $d\theta$ invariante.

En conséquence l'intégrale projetée sur les fonctions invariantes par le groupe,

$$\mathcal{I}(f) = \frac{1}{2\pi} \int_0^{2\pi} d\theta f(\theta + \theta_0) \equiv \frac{1}{2\pi} \int_0^{2\pi} d\theta \mathbf{g}(\theta_0)f(\theta).$$

Exercices

Exercice 4.1

Montrer que les matrices

$$\mathbf{M}(x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$$

avec x réel quelconque, forment un groupe. Citer un groupe isomorphe.

Montrer que ce groupe admet un sous-espace invariant.

La représentation est-elle décomposable en somme de représentations irréductibles?

Exercice 4.2

Montrer qu'une transformation linéaire de l'espace à deux dimensions qui préserve la norme des vecteurs correspond à une matrice orthogonale.

Exercice 4.3

Vérifier que toute matrice de $SO(2)$ peut s'écrire sous la forme (4.10).

Exercice 4.4

Montrer que le sous-groupe des rotations d'angle $2\pi p/N$, p, N entiers (isomorphe à \mathbb{Z}_N) est un sous-groupe distingué de $O(2)$. Que peut-t'on dire de $O(2)/\mathbb{Z}_N$?

Exercice 4.5

Retour sur le groupe des permutations de trois objets. Montrer maintenant que le groupe \mathfrak{S}_3 est isomorphe à un sous-groupe de $O(2)$ (mais non de $SO(2)$). Pour cela on considère les deux matrices

$$a = \sigma_3, \quad b = \sigma_3 e^{-i(2\pi/3)\sigma_2} \equiv \sigma_3 \mathbf{g}(2\pi/3).$$

On vérifiera que les relations fondamentales (3.2, 3.3) sont satisfaites et qu'il existe une transformation qui permet de passer des matrices dans cette représentation aux matrices (3.4). Le sous-groupe de \mathfrak{S}_3 qui appartient à $SO(2)$ est le sous-groupe \mathbb{Z}_3 c'est à dire les rotations de $2\pi/3$.

Exercice 4.6

Le groupe du carré. Un autre sous-groupe utile de $O(2)$ est le groupe du carré. Il est composé de rotations de $\pi/2$ et de réflexions par rapport aux axes. Il a huit éléments. On vérifiera qu'on peut choisir comme générateurs les deux matrices σ_1 et $-\sigma_3$, que les quatre éléments de $SO(2)$ sont alors

$$\mathbf{1}, \quad -i\sigma_2, \quad -\mathbf{1}, \quad i\sigma_2,$$

et les quatre autres éléments sont obtenus en multipliant par σ_3 sont :

$$\sigma_3, \quad -\sigma_1, \quad -\sigma_3, \quad \sigma_1.$$

5 Groupes de Lie : isométries en dimension trois et spin

Nous ne revenons pas sur le groupe commutatif des translations isomorphe au groupe additif \mathbb{R}^3 , nous concentrant sur les transformations linéaires. Nous décrivons abord en un peu plus de détails les propriétés du groupe orthogonal de l'espace à trois dimensions. Nous commençons par le sous-groupe des rotations $SO(3)$. En présence de particules de spin demi-entier, le groupe d'invariance de la mécanique quantique n'est plus $SO(3)$ mais le groupe de spin $SU(2)$. Nous examinons la relation entre ces deux groupes.

5.1 Rotations et groupe orthogonal

Le groupe $SO(3)$ est le groupe des rotations de l'espace ordinaire à trois dimensions. Les matrices \mathbf{R} de $SO(3)$ satisfont donc

$$\mathbf{R}\mathbf{R}^T = \mathbf{1}, \quad \det \mathbf{R} = 1.$$

Ces relations signifient que les rotations conservent la norme des vecteurs et l'orientation d'un repère cartésien.

Les éléments du groupe voisin de l'identité satisfont

$$\mathbf{R} = \mathbf{1} + \varepsilon \mathbf{T}, \varepsilon \in \mathbb{R}, \Rightarrow \mathbf{T} + \mathbf{T}^T + \varepsilon \mathbf{T}\mathbf{T}^T = 0.$$

Dans la limite $\varepsilon \rightarrow 0$ on obtient

$$\mathbf{T} + \mathbf{T}^T = 0.$$

Les matrices \mathbf{T} sont des matrices antisymétriques 3×3 qui forment un espace vectoriel de dimension trois. Le groupe $SO(3)$ est donc un groupe à trois paramètres réels.

On peut choisir comme base les matrices

$$(\mathbf{T}_a)_{ij} = -\epsilon_{aij} \tag{5.1}$$

où de nouveau ϵ_{ijk} est le symbole complètement antisymétrique (4.4) et, donc,

$$\mathbf{T}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad \mathbf{T}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \quad \mathbf{T}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Utilisant par exemple l'identité

$$\sum_{k=1}^3 \epsilon_{abk} \epsilon_{ijk} = \delta_{ai} \delta_{bj} - \delta_{aj} \delta_{bi},$$

on vérifie que les commutateurs, ou produits de Lie, des trois matrices sont

$$[\mathbf{T}_a, \mathbf{T}_b] = \sum_c \epsilon_{abc} \mathbf{T}_c. \tag{5.2}$$

Ces matrices sont les générateurs de l'algèbre de Lie $\mathcal{L}(SO(3))$ de $SO(3)$, l'algèbre de Lie des matrices antisymétriques 3×3 .

Matrices: produit scalaire. Dans l'ensemble des matrices réelles considéré comme espace vectoriel, on peut définir un produit scalaire et donc une norme en utilisant la trace :

$$(\mathbf{X}, \mathbf{Y}) = \text{tr } \mathbf{X}^T \mathbf{Y} = \sum_{i,j} X_{ij} Y_{ij}, \quad \|\mathbf{X}\|^2 = \text{tr } \mathbf{X}^T \mathbf{X}. \quad (5.3)$$

On vérifie que cette norme satisfait

$$\|\mathbf{X}\mathbf{Y}\| \leq \|\mathbf{X}\| \|\mathbf{Y}\|. \quad (5.4)$$

Les matrices (5.1) sont orthogonales au sens de ce produit scalaire, en effet

$$\text{tr } \mathbf{T}_a \mathbf{T}_b^T = -\text{tr } \mathbf{T}_a \mathbf{T}_b = 2\delta_{ab}.$$

De plus,

$$\sum_a \mathbf{T}_a \mathbf{T}_a^T = -\sum_{a=1}^3 \mathbf{T}_a^2 = 2\mathbf{1}. \quad (5.5)$$

Toute matrice antisymétrique \mathbf{A} peut s'écrire

$$\mathbf{A} = -\sum_{a=1}^3 \mathbf{v}_a \mathbf{T}_a \Leftrightarrow \mathbf{A}_{ij} = \sum_k \epsilon_{ijk} \mathbf{v}_k \Rightarrow \mathbf{v}_i = \frac{1}{2} \sum_{jk} \epsilon_{ijk} \mathbf{A}_{jk}. \quad (5.6)$$

Exponentiation. Tout élément de $SO(3)$ peut s'écrire comme l'exponentielle d'une matrice antisymétrique :

$$\mathbf{R} = \exp(\theta \boldsymbol{\omega} \cdot \mathbf{T})$$

avec où ω_a est un vecteur réel à trois composantes de longueur unité :

$$\sum_a \omega_a^2 = 1.$$

L'exponentielle d'une matrice est définie par sa série de Taylor, qui converge au sens de la norme (5.3).

La rotation \mathbf{R} représente donc une rotation d'angle θ avec $\boldsymbol{\omega}$ comme vecteur rotation.

Vérifions le en prenant $\boldsymbol{\omega}$ dirigé le long de l'axe trois et $\omega_3 = 1$. Nous trouvons

$$e^{\theta \mathbf{T}_3} = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Agissant sur un vecteur de composantes \mathbf{x}_i c'est bien une rotation qui laisse \mathbf{x}_3 invariant, d'angle θ dans le plan $\mathbf{x}_1, \mathbf{x}_2$.

Dans le cas général un peu d'algèbre conduit à

$$\mathbf{R}_{ij} = \cos \theta \delta_{ij} + (1 - \cos \theta) \omega_i \omega_j - \sin \theta \sum_k \epsilon_{ijk} \omega_k. \quad (5.7)$$

Enfin une rotation infinitésimale agissant sur un vecteur de composantes \mathbf{x}_i s'écrit

$$(\mathbf{R}\mathbf{x})_i = \mathbf{x}_i - \theta \sum_{j,a} \epsilon_{aij} \omega_a \mathbf{x}_j + O(\theta^2),$$

ou en notation vectorielle

$$(\mathbf{R}\mathbf{x}) = \mathbf{x} + \theta \boldsymbol{\omega} \times \mathbf{x} + O(\theta^2). \quad (5.8)$$

Le groupe $O(3)$. À trois dimensions le déterminant de la matrice $-\mathbf{1}$, qui géométriquement correspond à une réflexion par rapport à l'origine, est -1 , et donc engendre par multiplication avec les éléments de $SO(3)$ tout le groupe $O(3)$. Elle commute avec tous les éléments de $SO(3)$: $O(3) \sim SO(3) \times \mathbb{Z}_2$. Les éléments de $O(3)$ ont donc comme forme générale $\pm \mathbf{R}$, avec $\mathbf{R} \in SO(3)$. Ce groupe \mathbb{Z}_2 est aussi le centre du groupe $O(3)$.

Une identité utile. Pour toute matrice 3×3 \mathbf{M} ,

$$\sum_{lmn} M_{il} M_{jm} M_{kn} \epsilon_{lmn} = \epsilon_{ijk} \det \mathbf{M}. \quad (5.9)$$

En effet le membre de gauche est antisymétrique dans les permutations de ijk et donc proportionnel à ϵ_{ijk} . Choissant $i = 1, j = 2, k = 3$, on reconnaît la formule du déterminant. Nous appliquons cette identité à une matrice de $SO(3)$ et en déduisons

$$\sum_k R_{kp} \sum_{lmn} R_{il} R_{jm} R_{kn} \epsilon_{lmn} = \det \mathbf{R} \sum_k \epsilon_{ijk} R_{kp} = \sum_{lm} R_{il} R_{jm} \epsilon_{lmp}$$

et donc,

$$\sum_{lm} R_{il} R_{jm} \epsilon_{lmp} = \sum_k \epsilon_{ijk} R_{kp}, \quad (5.10)$$

où la relation d'orthogonalité a été utilisée. Pour obtenir l'action du groupe $O(3)$, nous devons encore appliquer l'identité (5.9) à $-\mathbf{1}$. Comme $\det(-\mathbf{1}) = -1$,

$$\sum_{lm} (-\delta_{il}) (-\delta_{jm}) \epsilon_{lmp} = \sum_k \epsilon_{ijk} (+\delta_{kp}). \quad (5.11)$$

5.2 Représentation adjointe

La représentation adjointe est définie comme l'action du groupe $SO(3)$ sur des matrices appartenant à l'espace vectoriel engendré par les générateurs de l'algèbre de Lie, c'est à dire dans la représentation de définition par les matrices antisymétriques. Il agit par

$$\mathbf{A} \xrightarrow{\mathbf{R}} \mathbf{R}\mathbf{A}\mathbf{R}^{-1},$$

une application linéaire interne de l'espace des matrices antisymétriques. En termes de composantes, elle s'écrit aussi

$$A_{ij} \xrightarrow{\mathbf{R}} \sum_{k,\ell} R_{ik}R_{j\ell}A_{k\ell},$$

et correspond à l'action du produit tensoriel $\mathbf{R} \otimes \mathbf{R}$ sur les tenseurs antisymétriques.

Nous utilisons maintenant la paramétrisation (5.6) et l'identité (5.10) :

$$\sum_p \epsilon_{ijp}v_p \xrightarrow{\mathbf{R}} \sum_k \epsilon_{ijk}R_{kp}v_p,$$

ce qui implique

$$v_k \xrightarrow{\mathbf{R}} \sum_{\ell} R_{k\ell}v_{\ell}.$$

La représentation adjointe de $SO(3)$ est donc isomorphe à la représentation de définition.

Il n'en est pas de même pour $O(3)$ puisque, comme l'identité (5.11) le montre, $-\mathbf{1}$ est envoyé sur l'identité et $O(3)$ donc sur $SO(3)$.

On voit donc que, du point de vue de $SO(3)$, \mathbf{v} se transforme comme un vecteur. Par contre dans la réflexion de matrice $-\mathbf{1}$, et donc $\det \mathbf{R} = -1$, \mathbf{v} est invariant. On appelle un tel vecteur qui se transforme par $O(3)/\mathbb{Z}_2 \sim SO(3)$ un vecteur *axial*.

5.3 Les représentations matricielles de $SO(3)$

Plus généralement, comme dans l'exemple de $SO(2)$ on peut s'intéresser à l'action du groupe $SO(3)$ sur les formes multilinéaires de plusieurs vecteurs. Reprenons d'abord l'exemple des formes bilinéaires de deux vecteurs \mathbf{x}, \mathbf{y} . Comme dans le cas de $SO(2)$, elles se décomposent en

(i) une représentation de dimension un proportionnelle au produit scalaire,

$$\mathbf{x} \cdot \mathbf{y} \delta_{ij},$$

(ii) une forme antisymétrique sur laquelle agit une représentation de dimension trois, la représentation adjointe décrite en section 5.2,

$$\mathbf{x}_i\mathbf{y}_j - \mathbf{x}_j\mathbf{y}_i.$$

qu'on peut récrire en introduisant le tenseur (4.4) ϵ_{ijk} en fonction du produit vectoriel

$$\mathbf{x}_i \mathbf{y}_j - \mathbf{x}_j \mathbf{y}_i = \sum_k \epsilon_{ijk} (\mathbf{x} \times \mathbf{y})_k \Leftrightarrow \sum_{jk} \epsilon_{ijk} \mathbf{x}_j \mathbf{y}_k \equiv (\mathbf{x} \times \mathbf{y})_i.$$

(iii) un tenseur symétrique de trace nulle qui se transforme par une nouvelle représentation de dimension cinq ($4 \times 3/2 - 1$)

$$\mathbf{x}_i \mathbf{y}_j + \mathbf{x}_j \mathbf{y}_i - \frac{2}{3} \mathbf{x} \cdot \mathbf{y} \delta_{ij},$$

qui est la représentation de spin ou moment cinétique deux.

Cette décomposition s'écrit symboliquement

$$[3] \otimes [3] = [1] \oplus [3] \oplus [5].$$

La même analyse s'étend aux formes multilinéaires, mais il faut alors faire intervenir les représentations irréductibles du groupe des permutations de plus de deux indices. La représentation de spin ℓ est obtenue en particulier comme un tenseur symétrique de ℓ indices dont toutes les traces partielles sur deux indices sont nulles. La dimension de la représentation est $2\ell + 1$. Par exemple, pour trois indices on trouve la décomposition

$$[3] \otimes [3] \otimes [3] = [1] \oplus 3 \times [3] \oplus 2 \times [5] \oplus [7],$$

c'est à dire un pseudo-scalaire (dimension un) proportionnel au produit mixte de trois vecteurs, trois représentations vectorielles, correspondant à un produit scalaire de deux vecteurs par le troisième, deux représentations de spin deux correspondant à des représentations mixtes de \mathfrak{S}_3 , et la représentation de spin trois correspondant aux tenseurs symétriques de trace nulle.

Représentations de l'algèbre de Lie. Élément de Casimir. À chaque représentation $\mathcal{R}(G)$ du groupe correspond aussi une représentation de l'algèbre de Lie. Au produit tensoriel $\mathcal{R}_1(G) \otimes \mathcal{R}_2(G)$ correspondent des générateurs de la forme

$$\mathcal{L}_1(G) \otimes \mathbf{1} + \mathbf{1} \otimes \mathcal{L}_2(G),$$

où $\mathcal{L}_1(G)$ et $\mathcal{L}_2(G)$ sont les représentations correspondantes de l'algèbre de Lie.

Comme nous venons de le montrer, une telle représentation est en général réductible. Dans une représentation irréductible, les générateurs de l'algèbre de Lie prennent la forme de matrices $(2\ell + 1) \times (2\ell + 1)$ pour la représentation de spin ℓ , que nous notons $\mathbf{T}_a^{(\ell)}$ (cf. section 5.6). Dans chaque représentation on peut calculer la matrice

$$\mathcal{C} = - \sum_a \mathbf{T}_a^{(\ell)} \mathbf{T}_a^{(\ell)},$$

cette forme étant conséquence du choix de générateurs orthonormés par la trace. N'utilisant que les relations de commutation de l'algèbre de Lie (5.2) on vérifie que cette matrice, forme quadratique dans les éléments de l'algèbre de

Lie, commute avec tous les générateurs. Elle est appelée élément de Casimir. Si la représentation est irréductible, toutes les matrices qui commutent avec tous les générateurs sont proportionnelles à l'identité. En section 5.6 nous montrons que dans la représentation irréductible de spin ℓ la matrice \mathcal{C} est donnée par

$$\mathcal{C} = C_\ell \mathbf{1}^{(\ell)}, \quad C_\ell = \ell(\ell + 1), \quad (5.12)$$

et donc cette valeur est caractéristique de la représentation.

Espace de fonctions et représentation. En mécanique quantique le groupe des rotations agit sur les fonctions d'onde qui appartiennent à un espace vectoriel de dimension infini, un espace de Hilbert. En l'absence de spin ou moment cinétique intrinsèque, la représentation agit sur une fonction $\psi(\mathbf{x})$, où \mathbf{x} est un point dans l'espace à trois dimensions, par

$$(\mathbf{R}\psi)(\mathbf{x}) = \psi(\mathbf{R}\mathbf{x}).$$

La forme infinitésimale correspondant à (5.8) s'écrit

$$(\mathbf{R}\psi)(\mathbf{x}) = \psi(\mathbf{x}) + \theta(\boldsymbol{\omega} \times \mathbf{x}) \cdot \nabla_x \psi + O(\theta^2).$$

Les générateurs de l'algèbre de Lie de $SO(3)$ proportionnels aux opérateurs moment cinétique \mathbf{L} sont donc représentés par des opérateurs différentiels

$$\mathbf{L} = -i\hbar\mathbf{T} \text{ avec } \mathbf{T} = \mathbf{x} \times \nabla_x. \quad (5.13)$$

5.4 Groupe $SO(3)$ et matrices σ_i

Caractérisons maintenant les transformations linéaires qui laissent invariantes les relations (4.3) satisfaites par les matrices de Pauli σ_i ? Soit σ'_i trois matrices définies par

$$\sigma'_i = \sum_j S_{ij} \sigma_j,$$

où la matrice \mathbf{S} d'éléments S_{ij} est réelle, et qui satisfont également les relations (4.3). Calculons

$$\begin{aligned} \sigma'_i \sigma'_j &= \sum_{k,\ell} S_{ik} S_{j\ell} \sigma_k \sigma_\ell = \sum_{k,\ell} S_{ik} S_{j\ell} \left(\delta_{k\ell} \mathbf{1} + i \sum_m \epsilon_{klm} \sigma_m \right) \\ &= \delta_{ij} \mathbf{1} + i \sum_{n,m} \epsilon_{ijn} S_{nm} \sigma_m. \end{aligned}$$

Ceci implique deux relations :

$$\begin{aligned} \mathbf{S}\mathbf{S}^T &= \mathbf{1} \\ \sum_{k,\ell} S_{ik} S_{j\ell} \epsilon_{klm} &= \sum_n \epsilon_{ijn} S_{nm}. \end{aligned}$$

La première relation implique que \mathbf{S} appartient à $O(3)$. Dans la deuxième relation, nous reconnaissons l'identité (5.10) qui implique que $\det \mathbf{S} = 1$, et donc \mathbf{S} appartient à $SO(3)$. Les relations (4.3) sont donc invariantes par le groupe $SO(3)$.

5.5 Le groupe $SU(2)$

Le groupe $SU(2)$ est le groupe des matrices complexes unitaires 2×2 de déterminant 1 :

$$\mathbf{U}\mathbf{U}^\dagger = \mathbf{1}, \quad \det \mathbf{U} = 1.$$

Dans un voisinage de l'identité (ε réel),

$$\mathbf{U} = \mathbf{1} + \varepsilon \mathbf{T} \Rightarrow \mathbf{T} + \mathbf{T}^\dagger \xrightarrow{\varepsilon \rightarrow 0} 0.$$

L'algèbre de Lie est donc formée de matrices 2×2 antihermitiennes et de trace nulle. Ces matrices sont des combinaisons linéaires des trois matrices $i\sigma_i$, les matrices σ_i étant définies en (4.2). Une matrice de $SU(2)$ dépend donc de trois paramètres réels.

Exponentiation. Dans un voisinage de l'identité, nous pouvons donc écrire tout élément \mathbf{U} de $SU(2)$

$$\mathbf{U} = \exp \left[-\frac{1}{2}i\theta \sum_a \omega_a \sigma_a \right], \quad (5.14)$$

où ω est un vecteur réel à trois composantes de longueur unité. On remarque que

$$\left(\sum_a \omega_a \sigma_a \right) = \sum_a \omega_a^2 = 1.$$

Un peu d'algèbre conduit alors à

$$\mathbf{U} = \cos(\theta/2)\mathbf{1} - i \sin(\theta/2) \sum_a \omega_a \sigma_a, \quad (5.15)$$

où θ est défini mod 4π . Il est facile de vérifier que cette forme représente en effet la matrice unitaire de déterminant 1 la plus générale qui peut s'écrire aussi

$$\mathbf{U} = \begin{pmatrix} a & -b^* \\ b & a^* \end{pmatrix}.$$

où les deux nombres complexes a, b satisfont

$$|a|^2 + |b|^2 = 1.$$

Si nous décomposons ces nombres en partie réelle et imaginaire nous trouvons

$$a = a_1 + ia_2, \quad b = b_1 + ib_2, \quad a_1^2 + a_2^2 + b_1^2 + b_2^2 = 1,$$

c'est dire l'équation d'une sphère S_3 de \mathbb{R}^4 .

Conjugaison complexe. Comme les matrices $SU(2)$ ont déterminant 1,

$$\sum_{\ell, m} U_{i\ell} U_{jm} \epsilon_{\ell m} = \epsilon_{ij} \det \mathbf{U} = \epsilon_{ij},$$

où ϵ_{ij} est le tenseur antisymétrique avec $\epsilon_{12} = 1$. Agissant avec \mathbf{U}^\dagger à gauche, nous obtenons

$$\sum_{i,\ell,m} U_{in}^* U_{i\ell} U_{jm} = \sum_{\ell} U_{jm} \epsilon_{nm} = \sum_i U_{ni}^* \epsilon_{ij},$$

ce qui implique

$$U_{ij}^* = - \sum_{k\ell} \epsilon_{ik} U_{k\ell} \epsilon_{\ell j},$$

où nous avons utilisé que le carré de la matrice ϵ est $-\mathbf{1}$. Cette équation, qui peut se récrire

$$\mathbf{U}^* = \sigma_2 \mathbf{U} \sigma_2,$$

démontre que \mathbf{U} et \mathbf{U}^* sont semblables et donc que, dans le cas de $SU(2)$, la représentation de définition et sa complexe conjuguée sont équivalentes. Les éléments de $[SU(2)]^*$ s'écrivent

$$\mathbf{U}^* = \exp \left[\frac{1}{2} i \theta \sum_a \omega_a \sigma_a^* \right] = \exp \left[-\frac{1}{2} i \theta \sum_a \tilde{\omega}_a \sigma_a \right],$$

avec

$$\tilde{\omega}_1 = -\omega_1, \quad \tilde{\omega}_3 = -\omega_3, \quad \tilde{\omega}_2 = \omega_2.$$

Algèbre de Lie. Les relations de commutation des générateurs τ_a de l'algèbre de Lie du groupe $SU(2)$ se déduisent immédiatement de l'algèbre des matrices σ (équation (4.5)). En effet

$$\tau_a = \frac{1}{2} \sigma_a$$

et donc

$$[\tau_a, \tau_b] = i \sum_c \epsilon_{abc} \tau_c.$$

Nous voyons que $-i\tau_a$ satisfait les relations de commutation (5.2) de l'algèbre de Lie de $SO(3)$. Cette identité des relations de commutation renvoie à la physique quantique : les particules de spin 1/2 se transforment par le groupe $SU(2)$ quand les coordonnées subissent une rotation d'espace. Nous allons d'ailleurs rappeler plus précisément ci-dessous la relation entre les groupes $SU(2)$ et $SO(3)$. Notons par contre que

$$\sum_a \tau_a^2 = \frac{3}{4} \mathbf{1}.$$

alors que pour la représentation de $SO(3)$ nous avons trouvé 2 (équation (5.5)). Dans l'expression générale (5.12) cela correspond à la valeur $\ell = 1/2$.

Représentation adjointe de $SU(2)$ et groupe $SO(3)$. Considérons les matrices complexes 2×2 , hermitiennes et de trace nulle. Elles forment un espace vectoriel

réel de dimensions trois et, comme nous venons de le discuter, la matrice la plus générale peut s'écrire

$$\mathbf{H} = \sum_i h_i \sigma_i, \quad \mathbf{h} = (h_1, h_2, h_3) \in \mathbb{R}^3,$$

où les matrices σ_i sont les matrices de Pauli. La matrice \mathbf{H} satisfait

$$\mathbf{H}^2 = \mathbf{h}^2 \mathbf{1} \quad \Rightarrow \quad \frac{1}{2} \text{tr} \mathbf{H}^2 = \mathbf{h}^2.$$

Considérons maintenant la transformation linéaire $\mathbf{H} \mapsto \mathbf{H}_U$:

$$\mathbf{H}_U = \mathbf{U} \mathbf{H} \mathbf{U}^\dagger, \quad (5.16)$$

où \mathbf{U} est une matrice de $SU(2)$. La matrice transformée \mathbf{H}_U est toujours hermitienne et de trace nulle. En effet,

$$\text{tr} \mathbf{U} \mathbf{H} \mathbf{U}^\dagger = \text{tr} \mathbf{H} \mathbf{U}^\dagger \mathbf{U} = \text{tr} \mathbf{H} = 0,$$

où nous avons utilisé la propriété cyclique de la trace. Elle est peut donc être exprimée comme une combinaison linéaire des matrices de Pauli. Les coefficients s'expriment linéairement en fonction des coefficients initiaux h_i . À la transformation (5.16) est donc associée une transformation linéaire de \mathbb{R}^3 .

Enfin,

$$[\mathbf{U} \mathbf{H} \mathbf{U}^\dagger]^2 = \mathbf{U} \mathbf{H} \mathbf{U}^\dagger \mathbf{U} \mathbf{H} \mathbf{U}^\dagger = \mathbf{h}^2 \mathbf{1}.$$

Cette transformation linéaire conserve la norme du vecteur \mathbf{h} . C'est donc une transformation orthogonale.

De façon plus explicite, nous pouvons utiliser la forme (5.15). De plus, comme démontré en section 5.4, nous pouvons agir sur les matrices σ par une transformation de $SO(3)$ pour réduire la forme (5.15) à

$$\mathbf{U} = \cos(\theta/2) - i \sin(\theta/2) \sigma_1,$$

et \mathbf{H} à

$$\mathbf{H} = h'_1 \sigma_1 + h'_2 \sigma_2.$$

Alors

$$\begin{aligned} \mathbf{U} \mathbf{H} \mathbf{U}^\dagger &= h'_1 \sigma_1 + h'_2 [\cos(\theta/2) - i \sin(\theta/2) \sigma_1] \sigma_2 [\cos(\theta/2) + i \sin(\theta/2) \sigma_1] \\ &= h'_1 \sigma_1 + h'_2 [\cos \theta \sigma_2 + \sin \theta \sigma_3]. \end{aligned}$$

Ceci est donc une rotation d'axe $\boldsymbol{\omega}$ et d'angle θ . Ces arguments établissent donc une relation entre le groupe $SU(2)$ et le groupe $SO(3)$: $SO(3)$ est isomorphe à la représentation adjointe de $SU(2)$. On remarque cependant que le changement θ en $\theta + 2\pi$ change un élément de $SU(2)$ en son opposé mais ne change pas l'élément de $SO(3)$ correspondant. La représentation adjointe n'est pas un isomorphisme de $SU(2)$. On exprime la relation par

$$SO(3) \sim SU(2)/\mathbb{Z}_2. \quad (5.17)$$

Groupes $SU(2)$ et $O(3)$. Ecrivons la transformation (5.16) en termes de composantes :

$$[\mathbf{H}_U]_{ij} = \sum_{kl} \mathbf{U}_{ik} \mathbf{U}_{jl}^* \mathbf{H}_{kl}.$$

Cette écriture identifie aussi la représentation adjointe avec un sous-groupe de $SU(2) \times [SU(2)]^* \sim SU(2) \times SU(2)$:

$$\text{Adj}[SU(2)] = (\mathbf{U}, \mathbf{U}^* = \sigma_2 \mathbf{U} \sigma_2)$$

avec la relation d'équivalence

$$(\mathbf{U}, \mathbf{U}^*) \equiv (-\mathbf{U}, -\mathbf{U}^*).$$

Pour aussi construire un groupe qui contient $SU(2)$ et qui a $O(3)$ comme représentation, il faut ajouter à $SU(2)$ un élément qui correspond à la réflexion $-\mathbf{1}$. Cet élément doit commuter avec toutes les matrices de $SU(2)$, et donc être un multiple de l'identité. Il ne peut pas être inclus dans la représentation de dimension deux. Dans la représentation de dimension quatre que nous venons d'introduire, nous pouvons prendre l'élément $(\mathbf{1}, -\mathbf{1})$. Son carré est l'identité et il change le signe de \mathbf{H} .

Note physique. Le groupe $SU(2)$ est le groupe de transformations qui agit sur les spineurs, des vecteurs complexes à deux composantes qui sont associés aux particules de spin 1/2. Les coordonnées d'espace elles se transforment comme les particules de spin un, c'est à dire par $SO(3)$, la représentation adjointe de $SU(2)$. Ceci entraîne une propriété bien connue des particules de spin 1/2, qui ne sont pas invariantes dans une rotation de 2π mais seulement de 4π .

Les groupes $SU(2) \times SU(2)$ et $SO(4)$. De façon analogue considérons la matrice \mathbf{X} :

$$\mathbf{X} = x_4 \mathbf{1} + i \sum_{a=1}^3 x_a \sigma_a$$

où $\mathbf{x} \equiv (x_1, \dots, x_4)$ est un vecteur de \mathbb{R}^4 . On vérifie

$$\det \mathbf{X} = \sum_{a=1}^4 x_a^2,$$

et

$$\mathbf{X}^\dagger \mathbf{X} = \sum_{a=1}^4 x_a^2 \mathbf{1} \Rightarrow = \frac{1}{2} \text{tr} \mathbf{X}^\dagger \mathbf{X} = \sum_{a=1}^4 x_a^2.$$

La matrice $\mathbf{X}/|\mathbf{x}|$ est un élément arbitraire de $SU(2)$. La matrice

$$\mathbf{X}_U = \mathbf{U}_1 \mathbf{X} \mathbf{U}_2^\dagger, \tag{5.18}$$

où $\mathbf{U}_1, \mathbf{U}_2$ sont deux matrices de $SU(2)$ quelconques, est donc de nouveau une matrice de la forme $|\mathbf{x}|$ fois une matrice de $SU(2)$,

$$\mathbf{X}_U^\dagger \mathbf{X}_U = \mathbf{U}_2 \mathbf{X}^\dagger \mathbf{U}_1^\dagger \mathbf{U}_1 \mathbf{X} \mathbf{U}_2^\dagger = \sum_{a=1}^4 x_a^2,$$

et

$$\det \mathbf{X}_U = \det \mathbf{U}_1 \det \mathbf{X} (\det \mathbf{U}_2)^* = \det \mathbf{X} = \sum_{a=1}^4 x_a^2.$$

La transformation (5.18) linéaire est donc une transformation dans \mathbb{R}^4 qui laisse la longueur du vecteur \mathbf{x} invariante. Elle dépend de six paramètres comme le groupe $SO(4)$. Cette remarque permet d'établir un lien entre le produit de groupe $SU(2) \times SU(2)$ et le groupe $SO(4)$. L'algèbre de Lie de $SO(4)$ se décompose dans la somme des deux algèbres de Lie de $SU(2)$ ou $SO(3)$ et donc est semi-simple. Par ailleurs, $(\mathbf{U}_1, \mathbf{U}_2)$ et $(-\mathbf{U}_1, -\mathbf{U}_2)$ correspondent à la même rotation. Enfin, $(-\mathbf{1}, \mathbf{1})$ transforme \mathbf{x} en $-\mathbf{x}$. Nous identifions donc

$$O(4) \sim SU(2) \times SU(2) / \mathbb{Z}_2.$$

Algèbre de Lie de $SO(4)$. On peut choisir comme générateurs de l'algèbre de $SO(4)$ les trois générateurs \mathbf{t}_i de l'algèbre de Lie du sous-groupe $SO(3)$ et trois générateurs supplémentaires $\boldsymbol{\tau}_i$ qui ont comme relations de commutation

$$[\mathbf{t}_i, \mathbf{t}_j] = \sum_k \epsilon_{ijk} \mathbf{t}_k, \quad [\boldsymbol{\tau}_i, \boldsymbol{\tau}_j] = \sum_k \epsilon_{ijk} \mathbf{t}_k, \quad [\boldsymbol{\tau}_i, \mathbf{t}_j] = \sum_k \epsilon_{ijk} \boldsymbol{\tau}_k.$$

Définissant les combinaisons

$$\mathbf{T}_i^\pm = \frac{1}{\sqrt{2}} (\mathbf{t}_i \pm \boldsymbol{\tau}_i),$$

on vérifie les relations

$$[\mathbf{T}_i^+, \mathbf{T}_j^-] = 0, \quad [\mathbf{T}_i^\pm, \mathbf{T}_j^\pm] = \sum_k \epsilon_{ijk} \mathbf{T}_k^\pm,$$

ce qui démontre directement que l'algèbre de Lie de $SO(4)$ se décompose dans la somme directe de deux algèbres de $SO(3)$.

Remarque. Après prolongement analytique $x_4 = it$, on obtient une relation entre groupes pertinente pour le groupe de Lorentz $O(1, 3)$ et la physique quantique relativiste.

5.6 Représentations irréductibles de l'algèbre de Lie de $SU(2)$

Représentations irréductibles de $SU(2)$. En décomposant l'espace des formes multilinéaires de ℓ vecteurs complexes, on obtient différentes représentations matricielles de $SU(2)$. La représentation dite de spin s où s conventionnellement $2s$ est un entier, est de dimension $2s + 1$. Pour s pair on retrouve les représentations de $SO(3)$. La valeur du Casimir correspondante est $s(s + 1)$.

Les représentations irréductibles d'une algèbre de Lie sont aussi irréductibles pour les groupes de Lie associés. On peut donc commencer par construire les représentations irréductibles de l'algèbre de Lie. Nous allons donc construire explicitement les représentations irréductibles en termes de matrices hermitiennes de l'algèbre de Lie de $SU(2)$:

$$[\tau_a, \tau_b] = i \sum_k \epsilon_{abc} \tau_c, \quad \tau_a^\dagger = \tau_a,$$

où les matrices τ_α sont maintenant les générateurs dans la représentation considérée.

On vérifie immédiatement que la matrice

$$\mathcal{C} = \sum_a \tau_a^2,$$

commute avec les trois générateurs

$$[\mathcal{C}, \tau_\alpha] = 0. \tag{5.19}$$

Pour chaque représentation la matrice \mathcal{C} est une matrice hermitienne positive. Elle peut donc être diagonalisée. La relation de commutation (5.19) implique que les éléments de matrice des générateurs entre des sous-espaces propres correspondant à des valeurs propres distinctes de \mathcal{C} s'annulent. Donc ou bien \mathcal{C} est proportionnelle à la matrice unité ou la représentation par définition est réductible. Comme nous avons supposé la représentation irréductible nous concluons

$$\mathcal{C} = c\mathbf{1}, \quad c \geq 0.$$

Pour construire les représentations irréductibles la première idée est de chercher le plus grand sous-groupe abélien de $SU(2)$, qui est $U(1)$ et de classer ses représentations irréductibles, ce qui est facile. De plus un espace vectoriel irréductible pour un sous-groupe reste irréductible pour le groupe tout entier. Dans le cadre de l'algèbre de Lie cela revient à chercher la plus grande sous-algèbre dont tous les éléments commutent. Dans le cas de $SU(2)$ elle n'est composée que d'un élément, et comme tous les éléments sont équivalents nous choisissons τ_3 .

Nous avons vu qu'une représentation du groupe agit sur l'algèbre de Lie considérée comme un espace vectoriel. Si nous considérons le sous-groupe $U(1)$ cette représentation est réductible. Les générateurs de l'algèbre de Lie agissent

par commutation (équation (6.8)). Les différentes représentations irréductibles ici sont de dimension un et correspondent aux éléments τ_3, τ_+, τ_- :

$$\tau_3 : [\tau_3, \tau_3] = 0, \quad \tau_{\pm} = \tau_1 \pm i\tau_2 : [\tau_3, \tau_{\pm}] = \pm\tau_{\pm}, \quad (5.20)$$

comme le calcul le montre. Les combinaisons τ_{\pm} sont les vecteurs isotropes et sont aussi reliés aux classes $SU(2)/U(1)$.

Notons alors que

$$\tau_- = \tau_+^\dagger, \quad [\tau_+, \tau_-] = 2\tau_3, \quad \mathcal{C} = \tau_3^2 - \tau_3 + \tau_+\tau_- = \tau_3^2 + \tau_3 + \tau_-\tau_+. \quad (5.21)$$

Comme τ_3 est une matrice hermitienne, elle peut être diagonalisée. Les représentations irréductibles de $U(1)$ correspondent aux vecteurs propres de τ_3 et sont de dimension un. Soit $|\mu\rangle$ le vecteur correspondant à la valeur propre μ qui est réelle (nous utilisons la notation des bras et kets de Dirac) :

$$\tau_3 |\mu\rangle = \mu |\mu\rangle.$$

Pour obtenir une représentation irréductible de l'algèbre de Lie entière, il faut maintenant faire agir les générateurs τ_{\pm} et étudier l'espace vectoriel engendré.

Utilisant la relation de commutation de τ_3 avec τ_{\pm} , nous trouvons

$$[\tau_3, \tau_{\pm}] |\mu\rangle = \pm\tau_{\pm} |\mu\rangle = \tau_3\tau_{\pm} |\mu\rangle - \mu\tau_{\pm} |\mu\rangle$$

qui peut s'écrire

$$\tau_3\tau_{\pm} |\mu\rangle = (\mu \pm 1)\tau_{\pm} |\mu\rangle.$$

Pour chacun des deux signes cette équation a deux solutions

$$\tau_{\pm} |\mu\rangle = 0 \quad \text{ou} \quad \tau_{\pm} |\mu\rangle = a_{\pm}(\mu) |\mu \pm 1\rangle,$$

où $a_{\pm}(\mu)$ est un coefficient complexe supposé alors non-nul, et $|\mu \pm 1\rangle$ un vecteur propre de τ_3 avec valeur propre $\mu \pm 1$. Faisant agir de façon répétée τ_{\pm} nous trouvons alors des vecteurs propres de τ_3 avec des valeurs propres différant de μ par un entier positif ou négatif. Comme nous cherchons des représentations matricielles l'espace vectoriel est de dimension fini. Il existe donc une plus grande valeur propre μ_+ et le vecteur propre correspondant est tel que

$$\tau_+ |\mu_+\rangle = 0.$$

De même il existe une plus petite valeur propre μ_- et le vecteur propre correspondant est tel que

$$\tau_- |\mu_-\rangle = 0.$$

Par ailleurs de la construction il résulte que $\mu_+ - \mu_-$ est un entier. Nous le supposons non nul sinon on obtient la représentation triviale $\mathcal{R}(G) = \mathbf{1}$. Pour des raisons historiques liées à la mécanique quantique nous le notons $2s$ de sorte que s est demi-entier,

$$\mu_+ - \mu_- = 2s > 0.$$

Faisons maintenant agir l'opérateur \mathcal{C} sur ces deux vecteurs en utilisant alternativement une des deux formes (5.21) possibles

$$\begin{aligned}\mathcal{C}|\mu_+\rangle &= (\tau_3^2 + \tau_3 + \tau_-\tau_+)|\mu_+\rangle = \mu_+(\mu_+ + 1)|\mu_+\rangle \\ \mathcal{C}|\mu_-\rangle &= (\tau_3^2 - \tau_3 + \tau_+\tau_-)|\mu_-\rangle = \mu_-(\mu_- - 1)|\mu_-\rangle.\end{aligned}$$

Ces deux valeurs sont égales à puisque \mathcal{C} est proportionnelle à la matrice identité

$$\mu_+(\mu_+ + 1) = \mu_-(\mu_- - 1) \Rightarrow \mu_- = -\mu_+,$$

car $\mu_+ - \mu_- > 0$, et donc

$$\mu_{\pm} = \pm s.$$

Nous avons donc trouvé qu'il existe des représentations irréductibles pour toutes dimensions entières $2s + 1$, que le spectre d'un générateur prend toutes les valeurs $-s + m$, avec m entier, $m \leq 2s$. La valeur propre c de l'opérateur \mathcal{C} est également déterminée

$$c = s(s + 1).$$

Elle caractérise donc la représentation.

Remarque. L'expression du produit tensoriel de deux générateurs de l'algèbre de Lie entraîne que les valeurs propres extrémales de τ_3 pour le produit de représentation $[\mathcal{R}(SU(2)) \otimes]^{2s}$ sont $\pm s$, ce que nous avons retrouvé.

Représentations de $SO(3)$. Pour une représentation de $SO(3)$ l'élément du groupe $e^{2i\pi\tau_3}$ doit être identique à l'identité. Nous voyons que cela impose que s soit entier.

Ceci amène à une autre remarque. En autorisant une valeur propre μ *a priori* quelconque, nous n'avons pas vraiment utilisé l'information que le sous-groupe correspondant est un groupe $U(1)$ sous-groupe de $SU(2)$, et admis tout groupe isomorphe à un sous-groupe du groupe additif \mathbb{R} . Nous avons classé toutes les représentations de l'algèbre de Lie et non du groupe. Ce que l'analyse montre c'est que le groupe $SU(2)$ est le plus "grand" groupe qui a cette algèbre de Lie.

Par ailleurs la représentation de dimension $2\ell + 1$ avec ℓ entier correspond bien aux tenseurs symétriques de trace nulle. Vérifions le en calculant la dimension de l'espace vectoriel correspondant par récurrence. Choisissons de toujours ordonner les indices de façon non-décroissante, c'est à dire de mettre en utilisant la symétrie les indices qui ont la valeur un les plus à gauche les indices qui ont la valeur deux à leur droite et les indices qui ont la valeur trois le plus à droite. Soit N_ℓ le nombre de tenseurs symétriques de rang ℓ . On construit les tenseurs de rang $\ell + 1$ en ajoutant un indice à gauche. Si à gauche il n'y a que des uns on peut donner les trois valeurs 1, 2, 3 à l'indice supplémentaire. Cela rajoute $2 = 3 - 1$ tenseurs nouveaux. S'il a au moins un deux et pas de trois on peut ajouter à droite deux ou trois. Il y a ℓ tenseurs distincts avec au moins un deux et au plus ℓ . Cela rajoute ℓ tenseurs. Enfin s'il existe un indice qui vaut trois on doit ajouter la valeur trois et on n'obtient le même nombre de tenseurs qu'avant. D'où

$$N_{\ell+1} = N_\ell + \ell + 2 \Rightarrow N_\ell = \frac{1}{2}(\ell + 1)(\ell + 2).$$

La condition de trace pour un tenseur de rang ℓ donne $N_{\ell-2}$ conditions, le nombre de tenseurs de rang $\ell - 2$ correspondant aux indices sur lesquels on ne somme pas, d'où

$$\frac{1}{2}(\ell + 1)(\ell + 2) - \frac{1}{2}\ell(\ell - 1) = 2\ell + 1.$$

6 Groupes et algèbres de Lie plus généraux : exemples orthogonaux et unitaires

Nous faisons maintenant quelques remarques générales sans prétentions déductives destinées à souligner l'importance de la structure d'algèbre de Lie, avant de décrire les groupes orthogonaux et unitaires. Ces groupes de Lie sont dits compacts (ceci élimine par exemple le groupe de Lorentz) et sont ceux qu'on rencontre les plus fréquemment en physique. Nous rappelons leurs définitions et quelques-unes de leurs propriétés les plus simples.

6.1 Groupes et algèbres de Lie

Les groupes de Lie sont des groupes topologiques dépendant de façon continue (en fait nous supposerons ici différentiable) d'un nombre fini n de paramètres réels. Nous nous limitons dans ce qui suit aux groupes de matrices.

Norme pour les matrices complexes. Dans l'ensemble des matrices à coefficients complexes $N \times N$, considéré comme espace vectoriel on peut définir un produit scalaire et donc une norme (et donc une distance entre deux matrices) par

$$(\mathbf{X}, \mathbf{Y}) = \text{tr}(\mathbf{X}^\dagger \mathbf{Y}) = \sum_{i,j} X_{ij}^* Y_{ij}, \quad \|\mathbf{X}\|^2 = \text{tr}(\mathbf{X}^\dagger \mathbf{X}), \quad (6.1)$$

où \mathbf{X}^\dagger est la matrice hermitienne conjuguée de \mathbf{X} . Appliquée au produit de deux matrices \mathbf{X}, \mathbf{Y} , cette norme vérifie la condition

$$\|\mathbf{XY}\| \leq \|\mathbf{X}\| \|\mathbf{Y}\|.$$

Pour les matrices réelles cette norme se réduit à (équation (5.3))

$$\|\mathbf{X}\|^2 = \text{tr}(\mathbf{X}^T \mathbf{X}) = \sum_{i,j} X_{ij}^2 \quad (6.2)$$

où \mathbf{X}^T dénote la transposée de la matrice \mathbf{X} .

Groupe et algèbre de Lie. Soit \mathfrak{G} un groupe de Lie. L'existence d'une distance et la différentiabilité permettent de parler d'éléments \mathfrak{g} du groupe proche de l'identité:

$$\mathfrak{g}(\varepsilon) = \mathbf{1} + \varepsilon \mathfrak{t} + o(\varepsilon), \quad (6.3)$$

où le paramètre réel ε caractérise la distance à l'origine. La loi de groupe entraîne que les éléments \mathfrak{t} forment un espace vectoriel $\mathcal{L}(\mathfrak{G})$ sur \mathbb{R} de dimension finie n . Tout élément peut donc s'exprimer en terme d'une base de n éléments τ_α :

$$\mathfrak{t} = \sum_{\alpha=1}^n \omega^\alpha \tau_\alpha, \quad (6.4)$$

où les ω^α sont n paramètres réels.

Considérons alors le produit d'éléments du groupe

$$\mathbf{g}\mathbf{h}\mathbf{g}^{-1} = \mathbf{h}'. \quad (6.5)$$

Dans la limite où $\mathbf{h} \rightarrow \mathbf{1}$, $\mathbf{h}' \rightarrow \mathbf{1}$:

$$\mathbf{h} - \mathbf{1} \sim \mathbf{u} \Rightarrow \mathbf{h}' - \mathbf{1} \sim \mathbf{u}', \quad \mathbf{u}, \mathbf{u}' \in \mathcal{L}(\mathfrak{G}).$$

Nous déduisons de la relation (6.5) une relation entre éléments de l'espace vectoriel $\mathcal{L}(\mathfrak{G})$:

$$\mathbf{g}\mathbf{u}\mathbf{g}^{-1} = \mathbf{u}'.$$

Dans la limite (6.3), on trouve

$$[\mathbf{t}, \mathbf{u}] = \mathbf{u}'.$$

Nous en déduisons que le commutateur de deux éléments de $\mathcal{L}(\mathfrak{G})$, aussi appelée produit de Lie, appartient à $\mathcal{L}(\mathfrak{G})$. Le commutateur, considéré comme loi de composition interne, donne à l'espace vectoriel (6.4) une structure d'algèbre. Cette algèbre qui est associée à la structure de groupe de Lie est appelée algèbre de Lie de G .

Nous pouvons appliquer cette remarque aux générateurs τ_α . Nous en déduisons qu'il existe des coefficients réels $f_{\alpha\beta}^\gamma$ tels que

$$[\tau_\alpha, \tau_\beta] = \sum_{\gamma} f_{\alpha\beta}^\gamma \tau_\gamma. \quad (6.6)$$

À la dépendance dans le choix de la base des τ_α près, ces coefficients, appelés *constantes de structure*, sont des propriétés caractéristiques du groupe.

Exponentiation. De la loi de groupe on déduit alors que tous les éléments du groupe dans un voisinage fini de l'identité sont de la forme

$$\mathbf{g}(\omega) = \lim_{k \rightarrow \infty} \left(\mathbf{1} + \frac{1}{k} \sum_{\alpha=1}^n \omega^\alpha \tau_\alpha \right)^k = \exp \left[\sum_{\alpha=1}^n \omega^\alpha \tau_\alpha \right],$$

où l'exponentielle, définie par sa série de Taylor, converge au sens de la norme (6.1). Les éléments du groupe \mathfrak{G} appartiennent ainsi à une algèbre dont les générateurs sont les τ_α .

Les données de la base $\{\tau_\alpha\}$ et des coefficients $f_{\alpha\beta}^\gamma$ permettent de calculer le produit de deux éléments de la composante du groupe G dite connexe à l'identité. En effet d'après l'identité de Baker–Hausdorff, si l'on pose

$$e^{\mathbf{t}_1} e^{\mathbf{t}_2} = e^{\mathbf{t}_{12}},$$

alors \mathbf{t}_{12} appartient à l'algèbre de Lie engendrée par \mathbf{t}_1 et \mathbf{t}_2 . L'élément \mathbf{t}_{12} qui détermine le produit $e^{\mathbf{t}_1} e^{\mathbf{t}_2}$ de deux éléments du groupe peut donc être

calculé uniquement à partir des coefficients $f_{\alpha\beta}^\gamma$. Cela signifie que le groupe est déterminé dans un voisinage fini de l'identité.

Bien entendu, comme nous l'avons déjà vu un groupe de Lie peut avoir des composantes non connexes à l'identité et qui donc ne peuvent pas s'écrire comme une exponentielle d'un élément de l'algèbre de Lie. Les éléments connexes à l'identité forment alors un sous-groupe, dont un argument de continuité montre qu'il est un sous-groupe invariant ou distingué. Donc les classes correspondantes forment un groupe discret. Le groupe complet est engendré par les générateurs du groupe de Lie et les générateurs de ce groupe discret.

6.2 Représentations, représentation adjointe

Les remarques précédentes s'appliquent au groupe de Lie et à ses représentations. Ceci permet alors d'étendre par continuité la notion de représentation à l'algèbre des générateurs et à l'algèbre de Lie. On vérifie en particulier que pour toute représentation $\mathcal{R}(G)$

$$\begin{aligned}\mathcal{R}(a_1\mathbf{t}_1 + a_2\mathbf{t}_2) &= a_1\mathcal{R}(\mathbf{t}_1) + a_2\mathcal{R}(\mathbf{t}_2), \quad a_1, a_2 \in \mathbb{R} \\ [\mathcal{R}(\tau_\alpha), \mathcal{R}(\tau_\beta)] &= \sum_{\gamma} f_{\alpha\beta}^\gamma \mathcal{R}(\tau_\gamma).\end{aligned}$$

La structure d'algèbre de Lie est donc indépendante de la représentation. La classification des représentations des groupes de Lie est ainsi directement liée à la classification des représentations des algèbres de Lie.

Notons cependant que si du point de vue de l'algèbre de Lie la représentation est un isomorphisme, il n'en est pas nécessairement de même pour le groupe de Lie, comme nous le verrons sur des exemples. Les groupes correspondant ne sont que localement isomorphes, c'est à dire dans un voisinage fini de l'identité.

Enfin quelques notes de terminologie. On dit d'une algèbre de Lie qu'elle est semi-simple si elle n'a pas un élément qui commute avec tous les autres. Par exemple l'algèbre de Lie d'un groupe abélien, ou plus généralement d'un produit de groupes qui a un facteur abélien n'est pas semi-simple. Une algèbre de Lie est dite simple si elle est semi-simple et s'il est impossible de trouver deux sous-ensembles de générateurs qui commutent. Par exemple l'algèbre de Lie associée à un produit de groupes de Lie n'est pas simple.

Représentation adjointe : construction. Soit \mathbf{g} un élément fixe d'un groupe \mathfrak{G} et considérons l'application qui a tout élément \mathbf{h} du groupe associe

$$\mathbf{h} \mapsto \mathbf{g}\mathbf{h}\mathbf{g}^{-1}.$$

Cette application définit un automorphisme $A_{\mathbf{g}}$ de groupe. Par ailleurs, le produit des automorphismes $A_{g_1}A_{g_2} = A_{g_1g_2}$ définit un groupe d'automorphismes isomorphe au groupe \mathfrak{G} lui-même. En passant à l'algèbre de Lie $\mathcal{L}(\mathfrak{G})$, on définit par continuité pour tout élément \mathbf{t} de $\mathcal{L}(\mathfrak{G})$,

$$\mathbf{t} \mapsto \mathbf{g}\mathbf{t}\mathbf{g}^{-1}.$$

L'algèbre de Lie a une structure d'espace vectoriel. On peut décomposer tout élément sur une base d'éléments τ_α :

$$\mathbf{t} = \sum_{\alpha} x^{\alpha} \tau_{\alpha}. \quad (6.7)$$

Alors,

$$\mathbf{g} \tau_{\alpha} \mathbf{g}^{-1} = \mathcal{R}_{\alpha}^{\beta}(\mathbf{g}) \tau_{\beta}$$

et donc

$$x^{\alpha} \mapsto \mathcal{R}_{\beta}^{\alpha}(\mathbf{g}) x^{\beta}.$$

Les matrices $\mathcal{R}(\mathbf{g})$ définissent une représentation du groupe de Lie appelée *représentation adjointe*.

La représentation de l'algèbre de Lie s'obtient en développant \mathbf{g} au voisinage de l'identité,

$$\mathbf{g} = \mathbf{1} + \varepsilon \tau_{\gamma} + O(\varepsilon^2).$$

Ceci implique

$$\mathcal{R}_{\alpha}^{\beta} = \delta_{\alpha}^{\beta} + \varepsilon f_{\gamma\alpha}^{\beta} + O(\varepsilon^2),$$

où les $f_{\gamma\alpha}^{\beta}$ sont les constantes de structure de l'algèbre de Lie définies par l'équation (6.6)). En effet, la représentation du générateur τ_{γ} est alors

$$[\tau_{\gamma}, \tau_{\alpha}] = f_{\gamma\alpha}^{\beta} \tau_{\beta}. \quad (6.8)$$

La représentation de l'algèbre de Lie agit par commutation. Appliquant la transformation à la décomposition (6.7), on en déduit l'action de la représentation sur les coefficients x^{α} :

$$x^{\alpha} \mapsto f_{\gamma\beta}^{\alpha} x^{\beta}. \quad (6.9)$$

Les constantes de structure considérées comme matrices d'indice γ et d'éléments $\alpha\beta$ forment donc une représentation adjointe de l'algèbre de Lie. La structure d'algèbre de Lie de la représentation découle aussi de l'identité de Jacobi,

$$[\tau_{\alpha}, [\tau_{\beta}, \mathbf{t}]] + [\tau_{\beta}, [\mathbf{t}, \tau_{\alpha}]] + [\mathbf{t}, [\tau_{\alpha}, \tau_{\beta}]] = 0,$$

où les commutateurs sont exprimés en termes des constantes de structure.

Algèbres et groupes de Lie : une identité utile. Soit $\mathbf{L}(t)$ une matrice fonction dérivable d'une variable réelle t . Alors l'identité suivante

$$\frac{d}{dt} e^{\mathbf{L}(t)} = \int_0^1 d\lambda e^{(1-\lambda)\mathbf{L}(t)} \frac{d\mathbf{L}}{dt} e^{\lambda\mathbf{L}(t)},$$

peut être démontrée en développant l'intégrant en puissances de $\mathbf{L}(t)$ et en intégrant terme à terme.

Cette identité a comme conséquence simple que la matrice

$$\mathbf{J}(t) = e^{-\mathbf{L}(t)} \frac{d}{dt} e^{\mathbf{L}(t)},$$

fait partie de l'algèbre de Lie engendrée par $\{\mathbf{L}(t), d_t \mathbf{L}(t)\}$ et donc de l'algèbre de Lie de l'ensemble des matrices $\mathbf{L}(t)$. Enfin

$$e^{-\mathbf{L}(t)} e^{\mathbf{L}(t+dt)} = \mathbf{1} + dt \mathbf{J}(t) = e^{\mathbf{J}(t)dt},$$

ce qui relie la loi de groupe aux propriétés de l'algèbre.

6.3 Groupes orthogonaux

Définition. Le groupe orthogonal $O(N)$ est le groupe de matrices réelles qui conserve la norme d'un vecteur dans l'espace à N dimensions \mathbb{R}^N . Soit \mathbf{v} un vecteur à N composantes et \mathbf{R} une matrice du groupe $O(N)$. Alors pour tout vecteur \mathbf{v} :

$$|\mathbf{v}| = |\mathbf{R}\mathbf{v}| \Leftrightarrow \sum_i \mathbf{v}_i^2 = \sum_i \left(\sum_j \mathbf{R}_{ij} \mathbf{v}_j \right)^2 = \sum_{i,j,k} \mathbf{R}_{ij} \mathbf{R}_{ik} \mathbf{v}_j \mathbf{v}_k .$$

Appelons \mathbf{S} la matrice symétrique

$$\mathbf{S} = \mathbf{R}^T \mathbf{R} \quad \text{et donc} \quad \mathbf{S}_{jk} = \sum_i \mathbf{R}_{ij} \mathbf{R}_{ik} .$$

Si nous prenons un vecteur de longueur unité qui n'a qu'une composante non nulle en position ℓ nous trouvons

$$1 = \mathbf{S}_{\ell\ell} \quad \forall \ell .$$

Nous prenons maintenant un vecteur qui n'a que deux composantes non nulles, en position ℓ et $m \neq \ell$

$$\mathbf{v}_\ell = 1, \quad \mathbf{v}_m = -1 .$$

On en déduit

$$\mathbf{S}_{\ell m} = 0 .$$

Donc la matrice \mathbf{S} est la matrice identité et

$$\mathbf{R}^T \mathbf{R} = \mathbf{1} = \mathbf{R}^T \mathbf{R} .$$

Ces relations caractérisent le groupe $O(N)$. Elles ont aussi une interprétation en termes des vecteurs lignes ou colonnes de la matrice \mathbf{R} : elles expriment que les vecteurs lignes (ou colonnes) forment une base orthonormée.

Une matrice arbitraire dépend de N^2 paramètres réels. Comme la matrice $\mathbf{R}^T \mathbf{R}$ est symétrique la relation précédente ne correspond qu'à $N(N+1)/2$ équations indépendantes. Donc une matrice orthogonale dépend de $N(N-1)/2$ paramètres réels.

De nouveau, prenant le déterminant des deux membres

$$1 = \det(\mathbf{R}^T \mathbf{R}) = \det \mathbf{R}^T \det \mathbf{R} = (\det \mathbf{R})^2 .$$

Donc

$$\det \mathbf{R} = \pm 1 .$$

Les matrices qui ont déterminant un forment de nouveau un sous-groupe du groupe $O(N)$ noté $SO(N)$, et qui est le groupe des rotations du plan (mais non les matrices de déterminant -1). Pour engendrer le groupe $O(N)$ tout entier

il suffit de rajouter un élément de déterminant -1 , par exemple une réflexion qui change le signe d'une composante des vecteurs comme l'élément Π_1 :

$$\Pi_1(x_1, x_2, \dots, x_N) = (-x_1, x_2, \dots, x_N) \Rightarrow \Pi_1^T \Pi_1 = \Pi_1^2 = \mathbf{1}.$$

Tout élément de $O(N)$ soit appartient à $SO(N)$, soit peut s'écrire $\Pi_1 \mathbf{R}$, où \mathbf{R} appartient à $SO(N)$. Il découle de la définition que $SO(N)$ est un sous-groupe distingué de $O(N)$ et que l'ensemble quotient $O(N)/SO(N) \sim \mathbb{Z}_2$.

Remarquons que la matrice de réflexion par rapport à l'origine $-\mathbf{1}$ appartient ou n'appartient pas à $SO(N)$ suivant la parité de N puisque son déterminant est $(-1)^N$. Le cas $N = 2N' + 1$ impair est donc particulièrement simple puisque pour passer de $SO(N)$ à $O(N)$ il suffit dans ce cas de rajouter la matrice $-\mathbf{1}$ qui commutent avec tous les éléments de $SO(N)$, et donc $O(2N' + 1) \sim SO(2N' + 1) \times \mathbb{Z}_2$.

6.4 $SO(N)$: Algèbre de Lie

Le groupe $SO(N)$, comme le groupe $SO(2)$, est déterminé par le voisinage de l'identité. Examinons ce point de façon plus précise. Soit \mathbf{R} une matrice orthogonale proche de l'identité. Nous l'écrivons

$$\mathbf{R} = \mathbf{1} + \varepsilon \mathbf{A},$$

où ε est le paramètre réel de développement. Alors la condition d'orthogonalité devient

$$\mathbf{A}^T + \mathbf{A} = O(\varepsilon).$$

Dans la limite $\varepsilon \rightarrow 0$ la matrice \mathbf{A} devient une matrice antisymétrique. Alors qu'en dimension deux toutes les matrices antisymétriques sont proportionnelles à σ_2 , en dimension N elles forment un espace vectoriel de dimension $N(N-1)/2$. Muni du produit de Lie, elles forment l'**algèbre de Lie** de $SO(N)$.

On peut prendre comme base de l'algèbre de Lie les matrices \mathbf{L}^{ab} , $1 \leq a < b \leq N$, qui comme éléments

$$\mathbf{L}_{ij}^{ab} = \delta_{ai} \delta_{bj} - \delta_{aj} \delta_{bi}.$$

Alors toute matrice antisymétrique peut s'écrire

$$\mathbf{A} = \sum_{a < b} \alpha_{ab} \mathbf{L}^{ab}.$$

Les matrices \mathbf{L}^{ab} , qui forment un ensemble de générateurs de l'algèbre de Lie, sont orthogonales au sens de la norme (6.2) :

$$\text{tr} \mathbf{L}^{ab} (\mathbf{L}^{cd})^T = 2\delta_{ac} \delta_{cd} \quad \text{pour } a < b, c < d.$$

Les produits de Lie ou commutateurs des générateurs sont donnés par

$$[\mathbf{L}^{ab}, \mathbf{L}^{cd}] = \delta_{ad} \mathbf{L}^{bc} + \delta_{bc} \mathbf{L}^{ad} - \delta_{ac} \mathbf{L}^{bd} - \delta_{bd} \mathbf{L}^{ac}. \quad (6.10)$$

Exponentiation. Il est facile de vérifier que si \mathbf{A} est antisymétrique, toute matrice de la forme

$$\mathbf{R} = e^{\mathbf{A}} = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!},$$

est une matrice de $SO(N)$

$$\mathbf{R}^T = \sum_{k=0}^{\infty} \frac{(\mathbf{A}^k)^T}{k!} = \sum_{k=0}^{\infty} \frac{(-\mathbf{A})^k}{k!} = e^{-\mathbf{A}},$$

et

$$\ln \det \mathbf{R} = \operatorname{tr} \ln \mathbf{R} = \operatorname{tr} \mathbf{A} = 0 \Rightarrow \det \mathbf{R} = 1.$$

De plus cette expression dépend de $N(N-1)/2$ paramètres réels comme la matrice la plus générale de $SO(N)$. En fait on démontre que tout élément de $SO(N)$ peut s'écrire sous cette forme. Les matrices \mathbf{L}^{ab} sont donc bien les *générateurs* d'une *algèbre* qui contient tous les éléments de $SO(N)$.

Représentation adjointe. Faisons d'abord quelques remarques préliminaires. Considérons l'ensemble des matrices réelles qui commutent avec toute matrice antisymétrique.

En dimension deux cet ensemble correspond aux combinaisons linéaires de $\mathbf{1}$ et σ_3 . En dimensions $N > 2$ il se réduit aux multiples de l'identité.

Considérons maintenant l'espace vectoriel des matrices anti-hermitiennes \mathbf{A} . La transformation linéaire

$$\mathbf{A} \mapsto \mathbf{A}_R = \mathbf{R}\mathbf{A}\mathbf{R}^{-1} = \mathbf{R}\mathbf{A}\mathbf{R}^T,$$

où \mathbf{R} est une matrice de $O(N)$, est une représentation du groupe $O(N)$ appelée représentation adjointe. Les matrices qui laissent toutes les matrices antisymétriques invariantes sont, pour $N > 2$, multiples de l'identité. Il n'y a qu'une matrice non triviale $-\mathbf{1}$. Suivant la parité de N elle appartient à $SO(N)$ ou non. Dans le cas N impair on trouve

$$\operatorname{Adj}[O(N)] \sim O(N)/\mathbb{Z}_2 \sim SO(N) = \operatorname{Adj}[SO(N)].$$

Dans le cas N pair on trouve

$$\operatorname{Adj}[O(N)] = \operatorname{Adj}[SO(N)] \sim SO(N)/\mathbb{Z}_2.$$

6.5 Groupes unitaires

Définition. Le groupe unitaire $U(N)$ est le groupe des matrices à éléments complexes qui préservent la norme de vecteurs complexes en dimension N . Les matrices du groupe $U(N)$ satisfont

$$\mathbf{U}^\dagger \mathbf{U} = \mathbf{1} = \mathbf{U}^\dagger \mathbf{U}. \quad (6.11)$$

Ici aussi ces relations expriment que les vecteurs complexes lignes ou colonnes de \mathbf{U} forment une base orthonormée.

Une matrice complexe $N \times N$ arbitraire dépend de N^2 paramètres complexes ou $2N^2$ paramètres réels. La matrice $\mathbf{U}^\dagger \mathbf{U}$ est hermitienne et donc l'équation matricielle (6.11) ne correspond qu'à N^2 équations réelles indépendantes. Une matrice de $U(N)$ dépend donc de N^2 paramètres réels.

Conjugaison complexe. La complexe conjuguée d'une matrice unitaire est une matrice unitaire. La conjugaison complexe respecte la loi de groupe

$$(\mathbf{U}_1)^*(\mathbf{U}_2)^* = (\mathbf{U}_1 \mathbf{U}_2)^*.$$

la matrice identité est réelle et donc invariante. L'application

$$\mathbf{U} \mapsto \mathbf{U}^*.$$

est donc un isomorphisme de groupe.

Déterminant. Prenons le déterminant des deux membres de (6.11)

$$1 = \det(\mathbf{U}^\dagger \mathbf{U}) = \det \mathbf{U}^\dagger \det \mathbf{U} = |\det \mathbf{U}|^2.$$

Donc le déterminant d'une matrice unitaire est un nombre complexe de module 1. Considérons alors les matrices de déterminant 1. Elles forment un sous-groupe de $U(N)$ noté $SU(N)$ (à $N^2 - 1$ paramètres réels) :

$$\mathbf{U}^\dagger \mathbf{U} = \mathbf{1}, \quad \det \mathbf{U} = 1. \quad (6.12)$$

Réciproquement toute matrice de $U(N)$ peut être obtenue à partir d'une matrice de $SU(N)$ en la multipliant par un nombre complexe de module 1, et donc un élément de $U(1)$. De plus ce produit est commutatif, à la différence de ce qui se passe dans la relation entre $SO(N)$ et $O(N)$. On peut décrire tout élément de $U(N)$ comme un couple (\mathbf{U}, v) avec $\mathbf{U} \in SU(N)$ et $v \in U(1)$, et le produit satisfait

$$(\mathbf{U}_1, v_1)(\mathbf{U}_2, v_2) = (\mathbf{U}_1 \mathbf{U}_2, v_1 v_2).$$

Cette structure se résume dans l'identité

$$U(N) \sim SU(N) \times U(1).$$

L'étude du groupe $U(N)$ se ramène donc à l'étude du groupe $U(1)$, déjà faite, et l'étude du groupe $SU(N)$.

Notons enfin que la condition $\det \mathbf{U} = 1$ étant réelle, la conjugaison complexe transforme aussi une représentation de $SU(N)$ en une autre représentation de $SU(N)$. Notons que pour $N > 2$, la représentation fondamentale (ou de définition) et sa complexe conjuguée ne sont pas équivalentes.

Matrices unitaires et orthogonales. Le groupe orthogonal $O(N)$ est le sous-groupe des matrices réelles du groupe $U(N)$.

Réciproquement toute matrice unitaire peut aussi être écrite comme une matrice orthogonale de dimension double. Il suffit de décomposer les complexes en partie réelle et imaginaire. La norme du vecteur complexe à N composantes est aussi la norme du vecteur réel à $2N$ composantes. La transformation préserve la norme du vecteur complexe et donc du vecteur réel. Nous concluons que cette décomposition réelle de la matrice est orthogonale. Nous en concluons que le groupe $U(N)$ est un sous-groupe du groupe $O(2N)$. Les deux groupes ne peuvent être isomorphes que pour $N = 1$ car pour $N > 1$ $U(N)$ dépend d'un nombre plus petit de paramètres que $O(2N)$.

6.6 Algèbre de Lie de $SU(N)$

Le groupe $SU(N)$ est déterminé par le voisinage de l'identité. Nous posons donc de nouveau

$$\mathbf{U} = \mathbf{1} + i\varepsilon\mathbf{H},$$

où ε est un paramètre réel que nous allons faire tendre vers zéro. Au premier ordre en ε on trouve

$$\mathbf{H} - \mathbf{H}^\dagger = 0.$$

La matrice \mathbf{H} est donc hermitienne et donc la matrice $i\mathbf{H}$ anti-hermitienne.

Nous devons encore exprimer que la matrice \mathbf{U} a un déterminant unité. Nous utilisons l'identité $\ln \det = \text{tr} \ln$ valable pour toute matrice et donc

$$\det(\mathbf{1} + i\varepsilon\mathbf{H}) = 1 + i\varepsilon \text{tr} \mathbf{H} + O(\varepsilon^2).$$

La condition d'ordre ε est donc

$$\text{tr} \mathbf{H} = 0.$$

Les matrices \mathbf{H} sont les matrices hermitiennes de trace nulle. Elles forment un espace vectoriel de dimension $N^2 - 1$. On peut choisir une base \mathbf{T}_α orthogonale au sens de la trace des matrices complexes, c'est à dire satisfaisant

$$\text{tr}(\mathbf{T}_\alpha \mathbf{T}_\beta^\dagger) = \text{tr}(\mathbf{T}_\alpha \mathbf{T}_\beta) = N\delta_{\alpha\beta}. \quad (6.13)$$

Toute matrice hermitienne de trace nulle \mathbf{H} peut alors s'écrire

$$\mathbf{H} = \sum_{\alpha} h_{\alpha} \mathbf{T}_{\alpha}. \quad (6.14)$$

Comme nous l'avons déjà mentionné dans le cas des matrices orthogonales, la structure de groupe est liée à la structure d'algèbre de Lie des matrices \mathbf{T}_α . Les matrices hermitiennes \mathbf{T}_α base de l'espace sont les générateurs de l'algèbre de Lie de $SU(N)$. Si on ajoute la matrice unité on obtient l'ensemble des générateurs de l'algèbre de Lie de $U(N)$. Le produit de Lie est le commutateur des matrices :

$$[\mathbf{T}_\alpha, \mathbf{T}_\beta] = i \sum_{\gamma} f_{\alpha\beta}^{\gamma} \mathbf{T}_{\gamma}, \quad (6.15)$$

où les $f_{\alpha\beta}^\gamma$ sont des coefficients réels, constantes de structure de l'algèbre de Lie. Ces coefficients sont antisymétriques en $\alpha \leftrightarrow \beta$ et satisfont une identité, appelée identité de Jacobi, conséquence de la relation (2.3) appliquée à $\mathbf{T}_\alpha, \mathbf{T}_\beta, \mathbf{T}_\gamma$,

$$\sum_{\delta} f_{\alpha\beta}^{\delta} f_{\delta\gamma}^{\epsilon} + f_{\beta\gamma}^{\delta} f_{\delta\alpha}^{\epsilon} + f_{\gamma\alpha}^{\delta} f_{\delta\beta}^{\epsilon} = 0.$$

Introduisant les matrices \mathbf{F}_α

$$\mathbf{F}_{\alpha\gamma}^{\beta} = f_{\beta\alpha\gamma} = -\mathbf{F}_{\alpha\gamma}^{\beta}, \quad (6.16)$$

nous pouvons récrire l'identité de Jacobi

$$[\mathbf{F}^\alpha, \mathbf{F}^\beta] = f_{\alpha\beta\delta} \mathbf{F}^\delta.$$

Les matrices $-i\mathbf{F}^\alpha$ forment une représentation de l'algèbre de Lie isomorphe à la représentation adjointe.

Pour un choix de matrices satisfaisant les relations d'orthogonalité (6.13), on peut calculer les constantes de structure en multipliant les deux membres de (6.13) par \mathbf{T}_δ et en prenant la trace. On en déduit

$$f_{\alpha\beta\gamma} = -\frac{i}{N} [\text{tr}(\mathbf{T}_\alpha \mathbf{T}_\beta \mathbf{T}_\gamma) - \text{tr}(\mathbf{T}_\alpha \mathbf{T}_\gamma \mathbf{T}_\beta)].$$

Utilisant la propriété cyclique de la trace, on en déduit que $f_{\alpha\beta\gamma}$ est antisymétrique dans ses trois indices. Le même résultat peut être obtenu dans le cas orthogonal. Interprété en termes des matrices (6.16) ce résultat signifie que les matrices sont antisymétriques et correspondent à une représentation orthogonale du groupe unitaire.

Élément de Casimir. On vérifie que comme conséquence de l'antisymétrie des constantes de structure, l'élément

$$\mathcal{C} = \sum_{\alpha} \mathbf{T}_\alpha^2, \quad (6.17)$$

commute avec tous les générateurs de l'algèbre de Lie :

$$[\mathcal{C}, \mathbf{T}_\alpha] = 0.$$

Exponentiation. On démontre ensuite que toute matrice de $SU(N)$ peut s'écrire comme une exponentielle d'une matrice anti-hermitienne $i\mathbf{H}$ de trace nulle :

$$\mathbf{U} = e^{i\mathbf{H}} = \sum_{k=0}^{\infty} \frac{i^k}{k!} \mathbf{H}^k.$$

On vérifie qu'une telle expression est bien une matrice unitaire

$$\mathbf{U}^\dagger = \sum_{k=0}^{\infty} \frac{(-i)^k}{k!} \mathbf{H}^k = e^{-i\mathbf{H}},$$

de déterminant égal à un

$$\ln \det \mathbf{U} = \operatorname{tr} \ln \mathbf{U} = i \operatorname{tr} \mathbf{H} = 0.$$

L'introduction de la représentation (6.14) permet de paramétrer les éléments de $SU(N)$ en termes des paramètres réels \mathbf{h}_α .

Quelques remarques utiles.

(i) Une matrice qui commute avec toutes les matrices hermitiennes de trace nulle est proportionnelle à l'identité.

En effet toute matrice complexe \mathbf{M} peut s'écrire

$$\mathbf{M} = (\mathbf{M} + \mathbf{M}^\dagger)/2 + i [(\mathbf{M} - \mathbf{M}^\dagger)/2i],$$

et est donc une combinaison linéaire à coefficients complexes de matrices hermitiennes. Une base est fournie par les générateurs de l'algèbre de Lie de $U(N)$, c'est à dire de l'algèbre de Lie de $SU(N)$ auxquels on ajoute l'identité. Toute matrice commute avec la matrice identité et donc une telle matrice commute avec toute matrice hermitienne, et à cause de l'argument précédent donc avec toute matrice. Une telle matrice est proportionnelle à l'identité.

Une conséquence simple : toute matrice qui commute avec tous les éléments de $SU(N)$ commute avec les générateurs de l'algèbre de Lie et est proportionnelle à la matrice identité.

(ii) Les matrices de $SU(N)$ qui sont proportionnelles à l'identité ont la forme

$$\mathbf{U} = u\mathbf{1} \Rightarrow u = e^{i\theta}, \quad \det \mathbf{U} = 1 \Rightarrow e^{iN\theta} = 1.$$

Les matrices de $SU(N)$ qui commutent avec toute matrice complexe et tous les éléments du groupe sont les racines $N^{\text{ième}}$ de l'unité, et forment le sous-groupe \mathbb{Z}_N de $SU(N)$. Un sous-groupe qui a la propriété de commuter avec tous les autres éléments est appelé *centre* du groupe.

6.7 Représentations: représentation adjointe

Représentation adjointe. Considérons l'espace vectoriel \mathcal{H} des matrices hermitiennes de trace nulle, espace vectoriel isomorphe à l'espace vectoriel des générateurs de l'algèbre de Lie. Si \mathbf{H} appartient à cet espace et \mathbf{U} est une matrice unitaire

$$\mathbf{H}_U = \mathbf{U}\mathbf{H}\mathbf{U}^\dagger \tag{6.18}$$

appartient à l'espace

$$\mathbf{H}_U^\dagger = \mathbf{H}_U, \quad \operatorname{tr} \mathbf{H}_U = \operatorname{tr}(\mathbf{U}\mathbf{H}\mathbf{U}^\dagger) = \operatorname{tr} \mathbf{H} = 0,$$

où nous avons utilisé la propriété cyclique de la trace. La transformation (6.18) est donc une transformation linéaire dans l'espace \mathcal{H} . Notons que la transformation laisse le déterminant de \mathbf{H} invariant.

Par ailleurs, si nous décomposons la matrice unitaire \mathbf{U} dans le produit d'une matrice de $SU(N)$ par élément de $U(1)$ nous voyons que \mathbf{H}_U ne dépend que

de l'élément de $SU(N)$. Nous pouvons donc nous restreindre aux éléments de $SU(N)$.

Récrivons l'équation (6.18) en termes de composantes,

$$[\mathbf{H}_U]_{ij} = \sum_{kl} \mathbf{U}_{ik} \mathbf{U}_{jl}^* \mathbf{H}_{kl}.$$

Les transformations (6.18) de l'espace \mathcal{H} forment donc un groupe, représentation du groupe $SU(N)$:

$$\mathbf{U} \in SU(N) \longmapsto (\mathbf{U}, \mathbf{U}^*),$$

sous-groupe de $SU(N) \times SU(N)^*$.

Les éléments de cette représentation de $SU(N)$, appelée représentation adjointe, sont des matrices $N^2 - 1 \times N^2 - 1$ agissant sur les $N^2 - 1$ coefficients du développement de \mathbf{H} sur une base.

Caractérisons maintenant les éléments de $SU(N)$ qui laissent tout élément de \mathcal{H} invariant

$$\forall \mathbf{H} : \mathbf{U} \mathbf{H} \mathbf{U}^\dagger = \mathbf{H} \Leftrightarrow [\mathbf{U}, \mathbf{H}] = 0.$$

Or nous avons vu qu'une matrice qui a cette propriété est un multiple de l'identité, ce qui dans le cas de $SU(N)$ signifie un élément du centre \mathbb{Z}_N de $SU(N)$. Cette propriété s'exprime en écrivant

$$\text{Adj}[SU(N)] \sim SU(N)/\mathbb{Z}_N.$$

Autres représentations. Nous discutons dans un contexte plus général la décomposition des produits de représentations en représentations irréductibles. Cependant, quelques remarques élémentaires peuvent être faites. Il est utile de considérer un sous-groupe abélien maximal et les générateurs de l'algèbre de Lie correspondants. Dans le cas de $SU(N)$, $(N-1)$ matrices diagonales linéairement indépendantes de trace nulle forment un tel ensemble de générateurs. Il est alors commode de bâtir les représentations de l'algèbre de Lie avec comme vecteurs de base les vecteurs propres communs aux générateurs de ce sous-groupe abélien. Par ailleurs, comme nous l'avons déjà noté, au produit tensoriel de deux représentations $\mathcal{R}_1(G) \otimes \mathcal{R}_2(G)$ d'un groupe de Lie G correspondent des générateurs de la forme

$$\mathcal{L}_1(G) \otimes \mathbf{1} + \mathbf{1} \otimes \mathcal{L}_2(G),$$

où $\mathcal{L}_1(G)$ et $\mathcal{L}_2(G)$ sont les représentations correspondantes de l'algèbre de Lie. Ainsi, si $\{\mathbf{v}_\alpha^1\}$ et $\{\mathbf{v}_\alpha^2\}$ sont les vecteurs propres dans les représentations 1 et 2 respectivement, $\{\mathbf{v}_\alpha^1 \otimes \mathbf{v}_\beta^2\}$ forment un ensemble de vecteurs propres dans le produit des représentations.

Note. Les groupes unitaires jouent un rôle important dans la théorie des interactions fondamentales. Les propriétés très voisines, en l'absence d'interactions électromagnétiques, du proton et du neutron, conduisent à une symétrie approchée $SU(2)$ dite d'isospin.

Le Modèle Standard, bien établi expérimentalement, est basé le produit de groupes $U(1) \times SU(2) \times SU(3)$. Il a été spéculé que ce produit est le sous-groupe résiduel à basse énergie d'un groupe plus vaste comme $SU(5)$ ou $SO(10)$.

6.8 Un exemple : le groupe $SU(3)$.

Nous appliquons ces remarques au groupe $SU(3)$ des matrices complexes unitaires 3×3 de déterminant 1. Par contraste avec le groupe $SU(2)$, la représentation fondamentale et sa conjuguée complexe ne sont pas équivalentes.

Remarque. $SU(3)$ a comme sous-groupe $SU(2) \times U(1)$. En fait la variété quotient $SU(3)/SU(2) \times U(1)$ est un espace symétrique ce qui se reflète dans la structure de l'algèbre de Lie comme nous allons l'expliquer. De plus, nous pouvons partir des représentations de l'algèbre de Lie de $SU(2) \times U(1)$ pour construire les représentations de $SU(3)$.

L'algèbre de Lie a huit générateurs de la forme $i\mathcal{T}_\alpha$, où les \mathcal{T}_α sont des matrices hermitiennes de trace nulle que nous choisissons orthonormées par la trace :

$$\text{tr } \mathcal{T}_\alpha \mathcal{T}_\beta = \frac{1}{2} \delta_{\alpha\beta}.$$

La représentation adjointe est donc de dimension huit. Le sous-groupe abélien maximal comprend cette fois deux générateurs que nous pouvons choisir de la forme

$$T_3 = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T_4 = \frac{\sqrt{3}}{2} \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & -\frac{2}{3} \end{pmatrix}.$$

Avec la même normalisation, nous définissons encore

$$T_1 = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T_2 = \begin{pmatrix} 0 & -\frac{i}{2} & 0 \\ \frac{i}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

T_1, T_2, T_3 sont les générateurs du sous-groupe $SU(2)$ et T_4 de $U(1)$. Enfin pour les quatre générateurs restants nous pouvons choisir

$$\begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & -\frac{i}{2} \\ 0 & 0 & 0 \\ \frac{i}{2} & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\frac{i}{2} \\ 0 & \frac{i}{2} & 0 \end{pmatrix}.$$

En fait, de même que nous avons l'avons fait pour τ_1 et τ_2 dans le but de construire les représentations de $SU(2)$ (équations (5.20)), il est utile d'introduire les combinaisons complexes S_1, S_2 des matrices précédentes,

$$[S_1]_{kl} = \delta_{k1}\delta_{l3}, \quad [S_2]_{kl} = \delta_{k2}\delta_{l3}$$

ainsi que S_1^\dagger et S_2^\dagger .

Représentation de $SU(2) \times U(1)$ et relations de commutation. Alors l'action de l'algèbre de Lie de $SU(2)$ est

$$\begin{aligned} [T_3, S_1] &= \frac{1}{2}S_1, & [T_3, S_2] &= -\frac{1}{2}S_2 \\ [T_1, S_1] &= \frac{1}{2}S_2, & [T_1, S_2] &= \frac{1}{2}S_1 \\ [T_2, S_1] &= -\frac{1}{2}iS_2, & [T_2, S_2] &= \frac{1}{2}iS_1, \end{aligned} \tag{6.19}$$

c'est à dire S_1 et S_2 se transforme suivant la représentation fondamentale de $SU(2)$. Les matrices S_1^\dagger et S_2^\dagger se transforment par la représentation conjuguée.

De plus,

$$\begin{aligned} [T_4, S_1] &= \frac{1}{2}\sqrt{3}S_1, & [T_4, S_2] &= \frac{1}{2}\sqrt{3}S_2 \\ [T_4, T_1] &= [T_4, T_2] = [T_4, T_3] = 0. \end{aligned} \quad (6.20)$$

Il reste à déterminer les relations de commutation entre S_1 et S_2 ainsi que les matrices conjuguées:

$$[S_1, S_2] = 0, \quad [S_1, S_2^\dagger] = T_1 + iT_2, \quad [S_2, S_1^\dagger] = T_1 - iT_2, \quad [S_1^\dagger, S_2^\dagger] = 0,$$

et

$$[S_1, S_1^\dagger] = T_3 + \sqrt{3}T_4, \quad [S_2, S_2^\dagger] = -T_3 + \sqrt{3}T_4.$$

Notons les relations de commutation sont compatible avec une notion de parité pour les générateurs de l'algèbre de Lie telle que les générateurs de $SU(2) \times U(1)$ soient pairs et les générateurs $S_1, S_2, S_1^\dagger, S_2^\dagger$ soient impairs. Cette propriété est caractéristique des espaces quotients, ici $SU(3)/SU(2) \times U(1)$, symétriques.

6.9 Représentations de $SU(3)$

Remarque préliminaire. La condition $\det \mathbf{U} = 1$ implique

$$\sum_{a,b,c} U_{ia}U_{jb}U_{kc}\epsilon_{abc} = \det \mathbf{U} \epsilon_{ijk} = \epsilon_{ijk}.$$

Comme première conséquence, on trouve la représentation identité dans la réduction du produit tensoriel $(SU(3) \otimes)^3$. De plus, multipliant les deux membres par U_{kl}^* , sommant sur k et utilisant la relation d'unitarité, on trouve

$$\sum_{a,b} U_{ia}U_{jb}\epsilon_{abl} = \sum_k \epsilon_{ijk}U_{kl}^*.$$

Ceci implique que dans la réduction du produit tensoriel $SU(3) \otimes SU(3)$, on trouve la représentation $SU(3)^*$.

Dans ce qui suit nous notons 3 et 3^* la représentation fondamentale de $SU(3)$.

Pour construire les représentations de $SU(3)$, on peut partir des représentations de $SU(2) \times U(1)$ qui sont caractérisées par un demi-entier s pour $SU(2)$ et un entier ν pour $U(1)$. Par ailleurs, les générateurs T_3 et T_4 forment une sous-algèbre abélienne maximale. Nous choisissons donc comme base des vecteurs tels que

$$T_3 |\mu, \nu\rangle = \mu |\mu, \nu\rangle, \quad T_4 |\mu, \nu\rangle = \frac{1}{2}\sqrt{3}\nu |\mu, \nu\rangle.$$

Élément de Casimir. L'élément de Casimir peut s'écrire

$$\mathcal{C} = T_1^2 + T_2^2 + T_3^2 + T_4^2 + \frac{1}{2} \left(S_1 S_1^\dagger + S_1^\dagger S_1 + S_2 S_2^\dagger + S_2^\dagger S_2 \right).$$

Pour la représentation fondamentale (ou de définition), on trouve

$$\mathcal{C} = \frac{4}{3}\mathbf{1}.$$

Exemples: représentations fondamentale et adjointe. Par exemple, pour la représentation fondamentale, les trois vecteurs sont

$$\left|\frac{1}{2}, \frac{1}{3}\right\rangle, \left|-\frac{1}{2}, \frac{1}{3}\right\rangle, \left|0, -\frac{2}{3}\right\rangle$$

et du point de vue de $SU(3)$ correspondent à la somme des représentations $s = 1/2$ et $s = 0$.

Pour la représentation adjointe, nous lisons directement les valeurs propres de T_3 et T_4 se lisent sur les relations de commutation. Nous trouvons la somme de la représentation $s = 1$ avec les vecteurs

$$|1, 0\rangle, |0, 0\rangle, |-1, 0\rangle,$$

de la représentation $s = 1/2$ et sa conjuguée,

$$\left|\frac{1}{2}, 1\right\rangle, \left|-\frac{1}{2}, 1\right\rangle \text{ et } \left|\frac{1}{2}, -1\right\rangle, \left|-\frac{1}{2}, -1\right\rangle.$$

Enfin une représentation identité $s = 0$ correspondant au générateur T_4 et au vecteur $|0, 0\rangle$.

Pour calculer l'élément de Casimir, on peut agir sur le vecteur $|1, 0\rangle$ et utiliser les relations de commutation qui conduisent à

$$\mathcal{C} |1, 0\rangle = (T_1^2 + T_2^2 + T_3^2 + T_3 + T_4^2) |1, 0\rangle = 3 |1, 0\rangle$$

et donc

$$\mathcal{C} = 3\mathbf{1}.$$

Construction générale. Les relations de commutation (6.19) et (6.20) impliquent

$$\begin{aligned} S_1 |\mu, \nu\rangle &\propto \left|\mu + \frac{1}{2}, \nu + 1\right\rangle \\ S_2 |\mu, \nu\rangle &\propto \left|\mu - \frac{1}{2}, \nu + 1\right\rangle. \end{aligned}$$

De façon corrélée,

$$\begin{aligned} S_1^\dagger |\mu, \nu\rangle &\propto \left|\mu - \frac{1}{2}, \nu - 1\right\rangle \\ S_2^\dagger |\mu, \nu\rangle &\propto \left|\mu + \frac{1}{2}, \nu - 1\right\rangle. \end{aligned}$$

Le produit tensoriel $3 \otimes 3$. Appliquons ces résultats au produit tensoriel $3 \otimes 3$. Notons que puisque les générateurs du produit tensoriel sont de la forme $T \otimes \mathbf{1} + \mathbf{1} \otimes T$, le produit tensoriel des vecteurs propres communs à T_3 et T_4 sont des vecteurs propres de la représentation de T_3 et T_4 , les nouvelles valeurs propres étant toutes les sommes deux à deux des valeurs propres.

Du point de vue du groupe $SU(2)$, nous trouvons une représentation $s = 1$, deux représentations $s = \frac{1}{2}$ et deux représentations $s = 0$ correspondant aux vecteurs

$$\left|1, \frac{2}{3}\right\rangle, \left|0, \frac{2}{3}\right\rangle \times 2, \left|-1, \frac{2}{3}\right\rangle, \left|\frac{1}{2}, -\frac{1}{3}\right\rangle \times 2, \left|-\frac{1}{2}, -\frac{1}{3}\right\rangle \times 2, \left|0, -\frac{4}{3}\right\rangle.$$

Nous avons déjà vu que dans la réduction apparaît la représentation 3^* associée aux tenseurs antisymétriques et aux vecteurs de base

$$\left|-\frac{1}{2}, -\frac{1}{3}\right\rangle, \left|\frac{1}{2}, -\frac{1}{3}\right\rangle, \left|0, \frac{2}{3}\right\rangle.$$

Les tenseurs symétriques forment eux un espace vectoriel de dimension six correspondant donc à trois représentations $s = 1$, $s = \frac{1}{2}$ et $s = 0$:

$$\begin{aligned} &\left|1, \frac{2}{3}\right\rangle, \left|0, \frac{2}{3}\right\rangle, \left|-1, \frac{2}{3}\right\rangle, \\ &\left|\frac{1}{2}, -\frac{1}{3}\right\rangle, \left|-\frac{1}{2}, -\frac{1}{3}\right\rangle, \\ &\left|0, -\frac{4}{3}\right\rangle, \end{aligned}$$

S_1^\dagger faisant passer d'une ligne à la suivante. Alors,

$$\mathcal{C} = \frac{10}{3} \mathbf{1}.$$

Le produit tensoriel $3 \otimes 3 \otimes 3$. Dans le produit tensoriel, nous trouvons d'abord le produit $3 \otimes 3^*$ qui contient la représentation adjointe et la représentation identité. De plus, il faut réduire le produit $3 \otimes 6$. Du point de vue de $SU(2)$, nous trouvons maintenant une représentation $s = \frac{3}{2}$, deux représentations $s = 1$, trois représentations $s = \frac{1}{2}$ et deux représentations $s = 0$ correspondant aux vecteurs

$$\begin{aligned} &\left|\frac{3}{2}, 1\right\rangle, \left|\frac{1}{2}, 1\right\rangle, \left|-\frac{1}{2}, 1\right\rangle, \left|-\frac{3}{2}, 1\right\rangle, \\ &[|1, 0\rangle, |0, 0\rangle, |-1, 0\rangle] \times 2, \\ &\left|\frac{1}{2}, 1\right\rangle, \left|\frac{1}{2}, 1\right\rangle \text{ et } \left[\left|\frac{1}{2}, -1\right\rangle, \left|\frac{1}{2}, -1\right\rangle\right] \times 2, \\ &|0, 0\rangle \text{ et } |0, -2\rangle. \end{aligned}$$

Les représentations irréductibles se composent alors d'une représentation isomorphe à la représentation adjointe et d'une représentation dix :

$$\begin{aligned} &\left|\frac{3}{2}, 1\right\rangle, \left|\frac{1}{2}, 1\right\rangle, \left|-\frac{1}{2}, 1\right\rangle, \left|-\frac{3}{2}, 1\right\rangle, \\ &|1, 0\rangle, |0, 0\rangle, |-1, 0\rangle, \\ &\left|\frac{1}{2}, -1\right\rangle, \left|\frac{1}{2}, -1\right\rangle, \\ &|0, -2\rangle. \end{aligned}$$

La représentation de dimension dix correspond à l'action du produit tensoriel $3 \otimes 3 \otimes 3$ sur les tenseurs symétriques, les deux représentation de dimension huit sur les tenseurs mixtes (au sens des représentations irréductibles du groupe symétrique \mathfrak{S}_3 , cf. section 7.1) et la représentation identité sur les tenseurs complètement antisymétriques.

L'élément de Casimir est

$$\mathcal{C} = 6\mathbf{1}.$$

Plus généralement la représentation correspondant aux tenseurs complètement symétriques de rang m a dimension $\frac{1}{2}(m+1)(m+2)$ et comme Casimir $\mathcal{C} = \frac{1}{3}m(m+3)\mathbf{1}$.

7 Produit tensoriel. Tenseurs.

Nous avons déjà introduit la notion de *tenseur* et de *produit tensoriel* en discutant les propriétés de transformation de formes multilinéaires de plusieurs vecteurs sur lesquelles agissent des transformations linéaires d'un groupe G . De plus, nous avons montré qu'en général l'espace vectoriel de ces formes se décomposait en différentes représentations du groupe G . On dit que la représentation initiale a été réduite. Dans le cas des groupes $SO(2)$ et $SO(3)$ la décomposition a engendré une somme de représentations dites *irréductibles*.

Rappelons que si une représentation est irréductible pour un sous-groupe elle est irréductible pour le groupe. Réciproquement si une représentation est réductible pour un groupe, elle est réductible pour tout sous-groupe. Une dernière remarque : pour les groupes orthogonaux ou unitaires on peut toujours décomposer toute représentation en une somme de représentations irréductibles.

Produit tensoriel. Nous rappelons la définition du produit tensoriel $\mathcal{R}_1(G) \otimes \mathcal{R}_2(G)$ de deux représentations matricielles $\mathcal{R}_1(G)$ de dimension n_1 et $\mathcal{R}_2(G)$ de dimension n_2 d'un groupe G . On considère alors un espace vectoriel \mathcal{V} de dimension $n_1 \times n_2$ et les vecteurs de cet espace que nous notons $T^{i_1 i_2}$. Si à l'élément g de G correspondent les matrices $\mathbf{R}_1(g)$ et $\mathbf{R}_2(g)$ dans les deux représentations, nous définissons l'action de $\mathcal{R}_1(G) \otimes \mathcal{R}_2(G)$ sur \mathcal{V} par

$$[\mathcal{R}_1(G) \otimes \mathcal{R}_2(G)T]^{i_1 i_2} \equiv \sum_{j_1, j_2} [\mathbf{R}_1]_{j_1}^{i_1} [\mathbf{R}_2]_{j_2}^{i_2} T^{j_1 j_2} .$$

Cette définition s'étend de façon naturelle au produit tensoriel d'un nombre arbitraire de représentations.

Nous rappelons aussi que G est un groupe de Lie, cette définition implique une relation entre les algèbres de Lie correspondantes. Si \mathbf{L}_1 et \mathbf{L}_2 représentent un élément de l'algèbre de Lie dans les deux représentations,

$$\mathbf{L}_1 \otimes \mathbf{1} + \mathbf{1} \otimes \mathbf{L}_2$$

est l'élément de l'algèbre de Lie dans la représentation produit.

Dans la suite, nous nous intéressons aux produits tensoriels n'impliquant que deux représentations $\mathcal{R}(G)$ et la représentation isomorphe $\tilde{\mathcal{R}}(G)$ où la représentation $\tilde{\mathcal{R}}(G)$ est obtenue à partir de la représentation $\mathcal{R}(G)$ par l'application qui à toute matrice \mathbf{R} de $\mathcal{R}(G)$ associe la matrice $[\mathbf{R}^{-1}]^T$:

$$\mathbf{R} \in \mathcal{R}(G) \Rightarrow [\mathbf{R}^{-1}]^T \in \tilde{\mathcal{R}}(G).$$

En général les deux représentations $\mathcal{R}(G)$ et $\tilde{\mathcal{R}}(G)$ ne sont pas équivalentes. Une exception notable est évidemment fournie par les groupes orthogonaux où les deux représentations sont identiques. Dans le cas des groupes unitaires, $\tilde{\mathcal{R}}(G) \equiv [\mathcal{R}(G)]^*$

Formes multilinéaires, polynômes et produit tensoriel. Dans la théorie des transitions de phase nous voudrions construire des polynômes dans des composantes de vecteurs sur lesquels agit un groupe de transformations linéaires.

Comme nous l'avons déjà indiqué, à cause de la linéarité les polynômes de degré différent se transforment indépendamment. Il suffit d'étudier les transformations agissant sur des polynômes homogènes. Ces polynômes homogènes sont des restrictions de formes multilinéaires dans un ensemble de vecteurs, et c'est le problème qu'il suffit de résoudre.

Formes multilinéaires et transformations linéaires. Soit un espace vectoriel \mathcal{V} réel ou complexe de dimension N , isomorphe à \mathbb{R}^N ou \mathbb{C}^N , sur lequel agit linéairement une représentation matricielle d'un groupe G . Considérons alors un ensemble de p vecteurs \mathbf{v}_a où l'indice $a = 1, \dots, p$ indique le numéro du vecteur et non une composante, de composantes v_a^i . Une forme multilinéaire est un objet de la forme

$$\mathbf{F}(\mathbf{v}) = \sum_{i_1, i_2, \dots, i_p} F_{i_1 i_2 \dots i_p} \prod_{a=1}^p v_a^{i_a}. \quad (7.1)$$

Une transformation linéaire agissant sur les vecteurs \mathbf{v}_a est équivalente à une transformation linéaire sur les coefficients $F_{i_1 i_2 \dots i_p}$, coefficients qui peuvent être considérés comme les composantes de vecteurs d'un espace de dimension Np .

Par exemple, dans le cas d'une forme linéaire,

$$F(\mathbf{v}) = \sum_i F_i v^i, \quad (7.2)$$

soit \mathbf{R} une matrice d'éléments $R_{ij}^{\hat{}}$ appartenant à la représentation $\mathcal{R}(G)$ de G agissant sur le vecteur \mathbf{v} ,

$$\mathbf{v}^i \mapsto u^i = \sum_j R_j^i v^j.$$

Substituant dans l'expression (7.2), nous voyons que dans le vecteur transformé la forme s'écrit

$$\mathbf{F}(\mathbf{u}) = \sum_i F_i \sum_j R_j^i u^j = \sum_i F_i^R u^i$$

avec

$$F_i^R = \sum_j F_j R_i^j.$$

La forme se transforme par la matrice transposée. Il est plus commode d'inverser la transformation. Alors,

$$F_i = \sum_j [(\mathbf{R}^{-1})^T]_i^j F_j^R.$$

Sous cette forme apparaissent les matrices $(\mathbf{R}^{-1})^T$ qui appartiennent à la représentation $\tilde{\mathcal{R}}(G)$ du groupe G .

Plus généralement, supposons que la matrice \mathbf{R} de la représentation $\mathcal{R}(G)$ de G agit sur un ensemble de p vecteurs \mathbf{v}_a de composantes v_a^i ,

$$v_a^i \mapsto u_a^i = \sum_j \mathbf{R}_{j_a}^i v_a^j.$$

Substituant dans l'expression (7.1), nous voyons que dans les vecteurs transformés la forme s'écrit

$$\mathbf{F}(\mathbf{u}) = \sum_{i_1, i_2, \dots, i_p} F_{i_1 i_2 \dots i_p} \prod_{a=1}^p \sum_{j_a} \mathbf{R}_{j_a}^{i_a} \mathbf{u}_a^{j_a} = \sum_{i_1, i_2, \dots, i_p} F_{i_1 i_2 \dots i_p}^R \prod_{a=1}^p \mathbf{u}_a^{i_a}$$

avec

$$F_{i_1 i_2 \dots i_p}^R = \sum_{j_1, j_2, \dots, j_p} F_{j_1, j_2, \dots, j_p} \prod_a \mathbf{R}_{j_a}^{i_a}.$$

La forme se transforme par les matrices transposées qui dans le langage de la théorie générale des groupes appartiennent à une représentation du groupe opposé. Il est plus commode de l'écrire sous la forme inverse

$$F_{i_1, i_2, \dots, i_p} = \sum_{j_1, j_2, \dots, j_p} \left(\prod_a [(\mathbf{R}^{-1})^T]_{i_a}^{j_a} \right) F_{j_1, j_2, \dots, j_p}^R.$$

Nous trouvons de nouveau une représentation du groupe agissant des tenseurs de rang p , qui est le produit tensoriel $\tilde{\mathcal{R}}(G) \otimes \tilde{\mathcal{R}}(G) \dots \otimes \tilde{\mathcal{R}}(G)$. Si nous voulons connaître les ensembles de formes qui se transforment les unes dans les autres, nous devons décomposer cette représentation en représentations irréductibles.

7.1 Groupe symétrique et tenseurs : réduction des représentations

Nous avons déjà vu que les propriétés des tenseurs dans une permutation des indices conduisaient à la réduction de la représentation. Ce phénomène est tout à fait général. En fait la décomposition du produit tensoriel d'ordre n s'appuie sur la décomposition en représentations irréductibles du groupe symétrique \mathfrak{S}_n ou groupe des permutations de n objets.

Considérons une représentation $[\mathcal{R}(G) \otimes]^p$ agissant sur les tenseurs de rang p :

$$[\mathbf{R}T]^{i_1, i_2, \dots, i_p} \equiv \sum_{j_1, j_2, \dots, j_p} \left(\prod_a \mathbf{R}_{j_a}^{i_a} \right) T^{j_1, j_2, \dots, j_p}.$$

Il est commode d'introduire la notation

$$T^{i_1, i_2, \dots, i_p} \equiv T(\{i\})$$

$$\left(\prod_a \mathbf{R}_{j_a}^{i_a} \right) \equiv R(\{i\}, \{j\}),$$

où $\{i\}$, resp. $\{j\}$ indiquent l'ensemble des indices i_1, i_2, \dots, i_p , resp. j_1, \dots, j_p . La transformation s'écrit alors

$$[\mathbf{RT}]({i}) = \sum_{\{j\}} R(\{i\}, \{j\})T(\{j\}).$$

Nous définissons une représentation du groupe symétrique \mathfrak{S}_p par des transformations linéaires agissant sur les tenseurs par l'action d'une permutation \mathbf{P} sur T :

$$[\mathbf{PT}]^{i_1, i_2, \dots, i_p} = T^{i_{P_1}, i_{P_2}, \dots, i_{P_p}} \equiv T(\mathbf{P}\{i\}).$$

Avec cette définition

$$[\mathbf{RPT}]({i}) = \sum_{\{j\}} R(\{i\}, \{j\})T(\mathbf{P}\{j\}).$$

Mais il résulte de la forme factorisée de $R(\{i\}, \{j\})$ que

$$R(\{i\}, \{j\}) = R(\mathbf{P}\{i\}, \mathbf{P}\{j\}).$$

Substituant, nous trouvons

$$[\mathbf{RPT}]({i}) = \sum_{\{j\}} R(\mathbf{P}\{i\}, \mathbf{P}\{j\})T(\mathbf{P}\{j\}).$$

Mais l'ensemble des indices $\mathbf{P}\{j\}$ est un ensemble d'indices muets que nous pouvons rebaptiser $\{j\}$. Donc

$$[\mathbf{RPT}]({i}) = \sum_{\{j\}} R(\mathbf{P}\{i\}, \{j\})T(\{j\}) = [\mathbf{PRT}]({i}).$$

Le groupe symétrique ou des permutations agissant sur les indices des tenseurs commute avec les transformations du groupe.

Dans le cas de deux matrices si elles commutent on peut les diagonaliser simultanément. De même dans le cas de groupes qui commutent on peut réduire les représentations simultanément. Si l'on décompose la représentation régulière du groupe symétrique en représentations irréductibles, on réduit du même coup la représentation ci-dessus. L'argument s'applique au groupe linéaire général et ne garantit pas que pour des sous-groupes du groupe linéaire les représentations ainsi obtenues soient irréductibles. D'autre part cette analyse n'est pas générale puisque nous n'avons considéré que les produits d'une seule représentation initiale. Si on considère des produits de la représentation $\mathcal{R}(G)$ et $\tilde{\mathcal{R}}(G)$ alors l'opération trace réduit des représentations puisque

$$\sum_i R_i^{j_1} [\mathbf{R}^{-1}]_{j_2}^i = \delta_{j_2}^{j_1} \equiv \delta_{j_1 j_2}.$$

En particulier, dans le cas des groupes orthogonaux puisque $(\mathbf{R}^{-1})^T = \mathbf{R}$, prendre les traces sur deux indices réduit aussi la représentation.

Enfin notons que pour les groupes qui peuvent être entièrement engendrés par leur algèbre de Lie, pour classer les représentations irréductibles du groupe il suffit de classer les représentations irréductibles de l'algèbre de Lie.

Groupe symétrique : exemple de réduction de représentation. La représentation régulière de \mathfrak{S}_3 est de dimension six. Elle agit linéairement sur les vecteurs de base

$$[123], [132], \dots$$

Cette représentation est réductible et se décompose en

(i) une représentation complètement symétrique de dimension un qui laisse invariant la combinaison

$$[123] + [231] + [312] + [213] + [321] + [132],$$

(ii) une représentation complètement antisymétrique de dimension un qui laisse invariant la combinaison

$$[123] + [231] + [312] - [213] - [321] - [132],$$

(ii) deux représentations mixtes de dimension deux, obtenues en antisymétrisant par rapport à une des transpositions et en symétrisant par rapport à l'autre, de vecteurs de bases

$$[123] + [132] - [213] - [231], \quad [213] + [312] - [123] - [321],$$

et

$$[123] + [213] - [132] - [312], \quad [132] + [231] - [123] - [321],$$

respectivement.

Cette décomposition appliquée à un tenseur de rang trois permet de réduire la représentation.

7.2 Groupe linéaire général et analyse tensorielle

À titre de complément nous donnons brièvement quelques détails supplémentaires sur la notion de tenseur et de produit tensoriel de représentations dans le cadre du groupe linéaire général des matrices de déterminant non-nul dont tous les groupes de matrices sont sous-groupes. Ceci nous permet d'introduire des notations usuelles dans ce cadre que nous avons évité jusqu'à présent.

Le groupe linéaire général. On note $GL(N, \mathbb{R})$, resp. $GL(N, \mathbb{C})$, le groupe engendré par l'ensemble des matrices réelles, resp. complexes, de déterminant non-nul. Tous les groupes de matrices réelles ou complexes sont donc des sous-groupe de $GL(N)$. Toute matrice \mathbf{g} de déterminant non-nul peut s'écrire

$$\mathbf{g} = \mathbf{s} \det \mathbf{g} \Rightarrow \det \mathbf{s} = 1.$$

Le groupe $GL(N)$ est donc le produit d'un groupe abélien isomorphe au groupe additif des réels ou complexes par le groupe $SL(N)$ des matrices de déterminant unité :

$$GL(N, \mathbb{R}) = \mathbb{R} \times SL(N, \mathbb{R}), \quad GL(N, \mathbb{C}) = \mathbb{C} \times SL(N, \mathbb{C}).$$

L'étude des représentations de $GL(N)$ se réduit donc à l'étude des représentations de $SL(N)$.

De façon générale, si $\mathbf{g} \in GL(N)$, alors la matrice inverse transposée \mathbf{g}^{-1T} appartient à un groupe $\tilde{GL}(N)$ isomorphe à $GL(N)$, les deux représentations n'étant pas équivalentes. Le sous-groupe des matrices de déterminant unité forment alors une représentation $\tilde{SL}(N)$ de $SL(N)$.

Dans le cas des matrices complexes, la conjugaison complexe engendre encore une autre représentation isomorphe mais inéquivalente.

Algèbre de Lie de $SL(N)$ et représentation adjointe. L'algèbre de Lie de $SL(N)$ est composé de l'ensemble des matrices de trace nulle et correspondant à un espace vectoriel de dimension réelle ou complexe $(N^2 - 1)$. La représentation adjointe du groupe agit sur cet espace vectoriel. La représentation adjointe est aussi obtenue dans la réduction du produit tensoriel $SL(N) \otimes \tilde{SL}(N)$:

$$SL(N) \otimes \tilde{SL}(N) = \text{Adj}[SL(N)] \oplus \mathbf{1}.$$

Notation. Nous considérons dans ce qui suit les vecteurs \mathbf{v} et $\tilde{\mathbf{v}}$ resp., appartenant à deux espaces vectoriels \mathcal{V} et $\tilde{\mathcal{V}}$ resp., isomorphes à \mathbb{R}^N ou \mathbb{C}^N que nous appelons duaux et dont nous notons les composantes $\{v^i\}$ et $\{v_i\}$ resp., $i = 1, \dots, N$ c'est à dire avec des indices en position haute ou basse pour les distinguer (on parle parfois de vecteurs contravariants et covariants, resp.). Par ailleurs, la convention de sommation implicite sur des indices haut et bas répétés sera maintenant utilisée :

$$u^i v_i \equiv \sum_{i=1}^N u^i v_i.$$

Sur les vecteurs $\{v^i\}$ agit le groupe linéaire de matrices \mathbf{g} d'éléments g_j^i dont nous notons l'action

$$(\mathbf{g}\mathbf{v})^i = g_j^i v^j.$$

Sur les vecteurs duaux agit la représentation isomorphe $\tilde{GL}(N)$ des matrices inverses transposées $\tilde{\mathbf{g}} = [\mathbf{g}^{-1}]^T$:

$$(\mathbf{g}\tilde{\mathbf{v}})_i = \tilde{g}_i^j v_j$$

avec

$$g_k^i \tilde{g}_j^k = \tilde{g}_k^i g_j^k = \delta_j^i,$$

où δ_j^i représente la matrice unité.

Nous notons qu'avec ces définitions le 'produit scalaire'

$$(\mathbf{gu})^i(\mathbf{gv})_i = u^j g_j^i \tilde{g}_i^k v_k = u^i v_i,$$

est invariant par le groupe. En terme de représentations, la représentation $GL(N) \otimes \tilde{GL}(N)$ est réductible et contient la représentation identité dans sa décomposition.

Tenseurs. On appelle tenseurs de rang n des éléments \mathbf{t} des espaces vectoriels $\mathcal{T}_{n-p}^p \sim \mathcal{V}^p \times \tilde{\mathcal{V}}^{n-p}$, $0 \leq p \leq n$, de dimensions Nn , dont on note les composantes avec n indices, et sur lesquels le groupe agit par le produit

$$(\mathbf{gt})_{i_{p+1} \dots i_n}^{i_1 i_2 \dots i_p} = \prod_{a=1}^p \mathbf{g}_{j_a}^{i_a} \prod_{a=p+1}^n \tilde{\mathbf{g}}_{i_a}^{j_a} \mathbf{t}_{j_{p+1} \dots j_n}^{j_1 j_2 \dots j_p}.$$

Ceci définit la représentation en général réductible du groupe linéaire notée $[GL(N) \otimes]^p \otimes [\tilde{GL}(N)]^{n-p}$.

On appelle trace partielle d'un tenseur toute somme sur une paire d'indices haut bas, par exemple

$$\mathbf{t}_{i_{p+1} \dots i_{n-1} k}^{i_1 i_2 \dots i_{p-1} k},$$

qui généralise le produit scalaire. Une telle trace se transforme comme un tenseur de rang $n - 2$. En effet, reprenant le même exemple,

$$\begin{aligned} (\mathbf{gt})_{i_{p+1} \dots i_{n-1} k}^{i_1 i_2 \dots i_{p-1} k} &= \mathbf{g}_{j_p}^k \tilde{\mathbf{g}}_k^{j_{n-p}} \prod_{a=1}^{p-1} \mathbf{g}_{j_a}^{i_a} \prod_{a=p+1}^{n-1} \tilde{\mathbf{g}}_{i_a}^{j_a} \mathbf{t}_{j_{p+1} \dots j_n}^{j_1 j_2 \dots j_p} \\ &= \prod_{a=1}^{p-1} \mathbf{g}_{j_a}^{i_a} \prod_{a=p+1}^{n-1} \tilde{\mathbf{g}}_{i_a}^{j_a} \mathbf{t}_{j_{p+1} \dots j_{n-1} k}^{j_1 j_2 \dots j_{p-1} k}. \end{aligned}$$

Cette représentation de $GL(N)$ peut donc se réduire par les traces partielles et la commutation avec les deux groupes des permutations agissant sur les p indices du haut et les $n - p$ indices du bas.

Notons enfin que dans le cas du groupe orthogonal les deux représentations duales sont identiques, et dans le cas du groupe unitaire elles sont complexes conjuguées.

8 Symétries en mécanique lagrangienne : théorème de Noether

Rappelons quelques éléments de mécanique lagrangienne. Dans ce cadre les équations du mouvement classique peuvent se déduire d'un principe d'action, où l'action est l'intégrale du lagrangien,

$$\mathcal{A}(q) = \int_{t'}^{t''} dt \mathcal{L}(q, \dot{q}; t), \quad (8.1)$$

le point initial $q(t')$ et final $q(t'')$ étant fixés.

Les équations classiques expriment que les trajectoires classiques $q(t)$ rendent l'action stationnaire : si $q(t)$ est une trajectoire classique, la variation $\delta\mathcal{A}$ de l'action quand on change $q(t)$ en $q(t) + \delta q(t)$, $\delta q(t') = \delta q(t'') = 0$, s'annule au premier ordre en δq . Cette condition s'écrit

$$\begin{aligned} \delta\mathcal{A} &= \int_{t'}^{t''} dt \left(\frac{\mathcal{L}}{\partial \dot{q}} \delta \dot{q} + \frac{\partial \mathcal{L}}{\partial q} \delta q \right) = 0 \quad \forall \delta q(t) \\ &= \delta q(t'') \left. \frac{\partial \mathcal{L}}{\partial \dot{q}} \right|_{t=t''} - \delta q(t') \left. \frac{\partial \mathcal{L}}{\partial \dot{q}} \right|_{t=t'} + \int_{t'}^{t''} dt \delta q(t) \left(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right). \end{aligned}$$

Le terme intégré dans l'intégration par parties s'annule. L'intégrale restante ne peut s'annuler pour toute fonction $\delta q(t)$ que si le coefficient de $\delta q(t)$ est nul ce qui conduit à la forme d'Euler-Lagrange des équations du mouvement :

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} = \frac{\partial \mathcal{L}}{\partial q}. \quad (8.2)$$

8.1 Symétries. Lois de conservation

Montrons maintenant qu'aux symétries continues de l'action correspondent des quantités conservées dans le mouvement classique. Comme les symétries continues correspondent à des groupes de Lie, il suffit de d'exprimer l'invariance dans des transformations infinitésimales qui s'expriment en termes des générateurs de l'algèbre de Lie.

La stratégie générale pour obtenir ces lois de conservation est de faire une transformation infinitésimale dépendante du temps, et d'exprimer la stationnarité de l'action. Comme le lagrangien ne contient que des dérivées premières, la variation du lagrangien au premier ordre est une combinaison linéaire des paramètres de la transformation et de leurs dérivées par rapport au temps. Si le lagrangien est symétrique la variation s'annule quand ces paramètres sont constants. Donc la variation du lagrangien est proportionnelle aux dérivées. Intégrant par parties et exprimant la stationnarité de l'action on obtient que certaines quantités ne dépendent pas du temps. Donnons maintenant quelques exemples pour illustrer cette démarche.

Invariance par translation du temps et conservation de l'énergie. Ceci constitue l'exemple le plus général et aussi le plus compliqué.

On dit que le lagrangien est invariant par translation du temps s'il ne dépend du temps qu'à travers la trajectoire $q(t)$ mais pas explicitement

$$\frac{\partial \mathcal{L}(q, \dot{q}, t)}{\partial t} = 0.$$

Nous faisons un changement de variables

$$t = t' + \varepsilon(t'),$$

dans l'action (8.1), dont nous évaluons l'effet au premier ordre en ε ,

$$\begin{aligned} dt &= dt'(1 + \dot{\varepsilon}(t')) \\ q(t) &= q(t') + \varepsilon(t')\dot{q}(t') \\ \frac{dq(t)}{dt} &= \dot{q}(t') + \varepsilon(t')\ddot{q}(t'). \end{aligned}$$

Nous en déduisons

$$\mathcal{A}(q) = \int dt (1 + \dot{\varepsilon}(t)) \mathcal{L}(q + \varepsilon\dot{q}, \dot{q} + \varepsilon\ddot{q}) + O(\varepsilon^2). \quad (8.3)$$

Exprimons maintenant que l'action est stationnaire par rapport à une variation de la trajectoire de la forme

$$\delta q(t) = \varepsilon(t)q(t) \Rightarrow \delta \dot{q}(t) = \dot{\varepsilon}(t)q(t) + \varepsilon(t)\dot{q}(t).$$

Si nous utilisons cette propriété dans le membre de droite de l'équation (8.3) développé au premier ordre en ε , nous voyons que seuls les termes proportionnels à $\dot{\varepsilon}$ survivent et donc

$$\delta \mathcal{A} = \int dt \dot{\varepsilon} \left(\dot{q} \frac{\partial \mathcal{L}}{\partial \dot{q}} - \mathcal{L} \right) = 0.$$

Intégrant par parties cette identité valable pour toute fonction $\varepsilon(t)$ nous trouvons

$$\frac{d}{dt} \left(\dot{q} \frac{\partial \mathcal{L}}{\partial \dot{q}} - \mathcal{L} \right) = 0,$$

ou, en introduisant le moment conjugué et l'hamiltonien,

$$p(t) = \frac{\partial \mathcal{L}}{\partial \dot{q}}, \quad H(p, q) = p(t)\dot{q}(t) - \mathcal{L}, \quad (8.4)$$

l'équation

$$\frac{d}{dt} H(p, q) = 0 \Rightarrow H(p, q) = E,$$

où E est l'énergie conservée du système.

Si cette démonstration est un peu lourde, elle met clairement en évidence la relation entre conservation de l'énergie et invariance par translation dans le temps. Notons que la transformation (8.4) qui relie hamiltonien H et lagrangien \mathcal{L} est appelée transformation de Legendre.

Translations spatiales et conservation de l'impulsion ou quantité de mouvement. Supposons maintenant que le lagrangien soit invariant dans le changement $q(t) \mapsto q(t) + a$, a étant une constante arbitraire. Ceci signifie évidemment que le lagrangien ne dépend que de \dot{q} , par exemple

$$\mathcal{L} = \frac{1}{2}m\dot{q}^2.$$

Une variation qui est une translation dépendant du temps est en fait une variation quelconque de $q(t)$

$$\delta q(t) = a(t).$$

Comme le lagrangien ne dépend que de \dot{q} , la variation correspondante de l'action est

$$\delta \mathcal{A} = \int dt \dot{a} \frac{\partial \mathcal{L}(\dot{q}; t)}{\partial \dot{q}} = 0,$$

En fonction du moment conjugué $p(t)$ défini en (8.4), on voit après intégration par parties que cette équation entraîne

$$\frac{dp(t)}{dt} = 0 \Rightarrow p(t) = p_0.$$

Invariance par translation entraîne donc conservation de l'impulsion.

Rotations. Soit \mathbf{q} un point dans l'espace à un, deux, trois, ..., N dimensions, dont nous notons les coordonnées q_i , et supposons que le lagrangien est invariant par rotation, c'est à dire par des transformations linéaires du groupe $SO(N)$,

$$q_i(t) \mapsto \sum_j R_{ij} q_j(t),$$

où \mathbf{R} d'éléments R_{ij} est une matrice orthogonale $\mathbf{R}^T \mathbf{R} = \mathbf{1}$,

$$\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}; t) = \mathcal{L}(\mathbf{R}\mathbf{q}, \mathbf{R}\dot{\mathbf{q}}; t).$$

On exprime de nouveau que si $q_i(t)$ est une solution classique, l'action est stationnaire par rapport à toute variation de $q_i(t)$ et donc, en particulier, par rapport à une variation qui a la forme d'une rotation infinitésimale dépendante du temps,

$$R_{ij}(t) = \delta_{ij} + \tau_{ij}(t).$$

où τ_{ij} est une matrice quelconque antisymétrique $\tau_{ji} = -\tau_{ij}$, associée à un élément de l'algèbre de Lie de $SO(N)$. Dans ces conditions,

$$\delta q_i(t) = \sum_j \tau_{ij}(t) q_j(t), \quad \delta \dot{q}_i(t) = \sum_j (\tau_{ij}(t) \dot{q}_j(t) + \dot{\tau}_{ij}(t) q_j(t)).$$

Si τ_{ij} est indépendant du temps la variation de l'action s'annule à cause de la symétrie. La variation de l'action au premier ordre en τ ne provient donc que de la dérivée de τ , c'est à dire de \dot{q} ,

$$\delta\mathcal{A}(q) = \int dt \sum_{ij} \dot{\tau}_{ij}(t) q_j(t) \frac{\partial\mathcal{L}}{\partial\dot{q}_i(t)}.$$

Annuler la variation de l'action, c'est exprimer que la dérivée du coefficient de $\dot{\tau}$ s'annule, et donc que le coefficient de $\dot{\tau}$ est constant :

$$\sum_{ij} \tau_{ij}(t) \frac{d}{dt} \left(q_j(t) \frac{\partial\mathcal{L}}{\partial\dot{q}_i(t)} \right) = 0.$$

Comme τ_{ij} est une matrice antisymétrique arbitraire cette équation entraîne que son coefficient antisymétrisé en ij est nul :

$$L_{ij} = q_i \frac{\partial\mathcal{L}}{\partial\dot{q}_j} - q_j \frac{\partial\mathcal{L}}{\partial\dot{q}_i}, \quad \dot{L}_{ij} = 0. \quad (8.5)$$

À l'invariance par rotation sont associées des constantes du mouvement L_{ij} , dont l'ensemble est appelé moment cinétique, en correspondance biunivoque avec les générateurs du groupe des rotations.

En dimension trois, la représentation antisymétrique est isomorphe à la représentation vectorielle. Les constantes du mouvement correspondent au vecteur moment cinétique

$$\mathbf{L} = \mathbf{q} \times \frac{\partial\mathcal{L}}{\partial\mathbf{q}}.$$

Un exemple en est fourni par une particule soumise à un potentiel central comme les potentiels de Coulomb ou newtonien

$$\mathcal{L} = \frac{1}{2} m \dot{\mathbf{q}}^2 - V(|\mathbf{q}|).$$

Alors,

$$\frac{\partial\mathcal{L}}{\partial\dot{q}_i} = m\dot{q}_i,$$

et donc

$$\mathbf{L} = m\mathbf{q} \times \dot{\mathbf{q}}.$$

On reconnaît la forme la plus fréquente du moment cinétique.

8.2 Mécanique hamiltonienne

La transformation qui relie hamiltonien et lagrangien,

$$\mathcal{L}(q, \dot{q}; t) + H(p, q; t) - p(t)\dot{q}(t) = 0, \quad (8.6)$$

où $p(t)$, la quantité $p(t)$ conjuguée à \dot{q} , est obtenue en exprimant que l'équation est stationnaire par rapport à \dot{q} à p, q fixés :

$$p(t) = \frac{\partial \mathcal{L}}{\partial \dot{q}}, \quad (8.7)$$

est une transformation de Legendre. La fonction $H(p, q)$ est alors la transformée de Legendre du lagrangien. La transformation de Legendre est involutive, c'est à dire que \mathcal{L} et \dot{q} jouent un rôle symétrique à H et p . En effet en variant l'équation par rapport à p (à q fixé), \dot{q} étant considéré comme une fonction de p par (8.7), on trouve

$$\left. \frac{\partial H}{\partial p} \right|_q = \left. \frac{\partial}{\partial p} \right|_q [p\dot{q} - \mathcal{L}(q, \dot{q})] = \dot{q} + \left. \frac{\partial \dot{q}}{\partial p} \right|_q \left. \frac{\partial}{\partial \dot{q}} \right|_{q,p} [p\dot{q} - \mathcal{L}(q, \dot{q})],$$

et donc

$$\dot{q}(t) = \frac{\partial H}{\partial p}.$$

Enfin la condition de stationnarité entraîne la relation

$$\left. \frac{\partial H}{\partial q} \right|_p = \left. \frac{\partial}{\partial q} \right|_p [p\dot{q} - \mathcal{L}(q, \dot{q})] = - \left. \frac{\partial \mathcal{L}(q, \dot{q})}{\partial q} \right|_{\dot{q}} + \left. \frac{\partial \dot{q}}{\partial q} \right|_p \left. \frac{\partial}{\partial \dot{q}} \right|_{q,p} [p\dot{q} - \mathcal{L}(q, \dot{q})],$$

et donc

$$\frac{\partial H}{\partial q} + \frac{\partial \mathcal{L}}{\partial q} = 0.$$

Dans ces conditions on vérifie que les équations du mouvement dans l'espace de phase (p, q) s'écrivent

$$\dot{q}(t) = \frac{\partial H}{\partial p}, \quad \dot{p}(t) = -\frac{\partial H}{\partial q}, \quad (8.8)$$

et s'obtiennent à partir d'un principe d'action

$$\mathcal{A}(p, q) = \int_{t'}^{t''} dt [p(t)\dot{q}(t) - H(p(t), q(t); t)], \quad (8.9)$$

en exprimant que l'action est stationnaire à la fois par rapport à $p(t)$ et $q(t)$, les valeurs initiales et finales de $q(t)$ étant fixés. Alternativement on peut intégrer par parties et remplacer $\int p\dot{q}$ par $-\int q\dot{p}$, mais les conditions aux limites portent alors sur $p(t)$.

Le terme

$$\int dt p(t)\dot{q}(t) = \oint pdq, \quad (8.10)$$

ne dépend que de la trajectoire dans l'espace de phase mais pas de la vitesse sur la trajectoire. En fait il est égal à l'aire comprise entre la trajectoire et l'axe $p = 0$ (à plusieurs variables à la somme des aires). On peut l'antisymétriser

en p et q . Dans le langage mathématique des formes, il est l'intégrale d'une deux-forme, la forme symplectique $\omega = \sum_i dp_i \wedge dq_i$. Ce point de vue est particulièrement utile quand l'espace de phase a une topologie non-triviale.

Symétries. Sous la forme (8.9), la conservation de l'énergie quand H ne dépend pas explicitement du temps est très facile à prouver, de même que la conservation de l'impulsion comme conséquence de l'invariance par translation.

Dans le cas du groupe $SO(N)$, les quantités conservées (8.5) qui sont associées aux générateurs du groupe des rotations s'écrivent dans le formalisme hamiltonien

$$L_{ij} = q_i p_j - q_j p_i. \quad (8.11)$$

La vérification de cette propriété à partir de l'action (8.9) est très simple. Les variables p et q se transforment de la même manière par une rotation

$$q_i \mapsto \sum_j R_{ij} q_j, \quad p_i \mapsto \sum_j R_{ij} p_j.$$

Si le hamiltonien est invariant l'action est invariante car la quantité

$$\int dt \sum_i p_i(t) \dot{q}_i(t)$$

est aussi invariante par rotation. Dans une rotation dépendante du temps seul \dot{q}_i donne un terme additionnel, et on retrouve simplement la forme (8.11).

Dans le cas de l'espace à trois dimensions nous avons déjà vu qu'une matrice antisymétrique était équivalente à un vecteur et les quantités conservées peuvent s'écrire comme les trois composantes du vecteur moment cinétique

$$\mathbf{L} = \mathbf{q} \times \mathbf{p}.$$

Transformations canoniques. Les transformations canoniques sont les transformations de l'espace de phase $\{q_i, p_i\} \mapsto \{Q_i, P_i\}$ qui laissent la forme symplectique

$$\int dt p(t) \dot{q}(t) = \int dp \wedge dq,$$

invariante. Elles ont une structure de groupe. Un ensemble de transformations triviales correspondent à ajouter à p_i un gradient $\partial_i F(q)$. Cette transformation est induite par l'ajout au lagrangien d'un terme de dérivée totale par rapport au temps. Il est facile ensuite de caractériser les transformations infinitésimales qui ne sont pas de cette forme. On trouve

$$Q_i = q_i + \varepsilon \frac{\partial T(p, q)}{\partial p_i} \quad P_i = p_i - \varepsilon \frac{\partial T(p, q)}{\partial q_i}, \quad (8.12)$$

où $T(q, p)$ est arbitraire. On reconnaît les équations du mouvement si T est l'hamiltonien. Donc l'ensemble des transformations canoniques est associé à

l'ensemble des hamiltoniens : l'application qui associe la position dans l'espace de phase au temps t à la position au temps initial est une transformation canonique.

On en déduit la forme intégrée de ces transformations. On se donne une fonction génératrice $\mathcal{S}(q, Q)$ (l'action classique de la trajectoire qui relie q à Q) et on pose

$$p_i = \frac{\partial \mathcal{S}}{\partial q_i} \quad P_i = -\frac{\partial \mathcal{S}}{\partial Q_i}. \quad (8.13)$$

On vérifie directement l'invariance de la forme en faisant le changement en deux étapes $\{q_i, p_i\} \mapsto \{q_i, Q_i\}$ puis $\{q_i, Q_i\} \mapsto \{Q_i, P_i\}$. Par la même méthode on vérifie que ce changement laisse aussi invariante la mesure de Liouville $\prod_i dq_i dp_i$. Ceci justifie en particulier la cohérence du choix de la mesure thermique

$$\rho(p, q) = dpdq e^{-\beta H(p, q)}$$

en mécanique statistique classique.

Crochets de Poisson. Il est commode d'introduire les crochets de Poisson,

$$\{A(p, q), B(p, q)\} = \frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial B}{\partial q_i} \frac{\partial A}{\partial p_i}, \quad (8.14)$$

limite semi-classique des commutateurs de la mécanique quantique

$$\{A(p, q), B(p, q)\} = \lim_{\hbar \rightarrow 0} \frac{1}{i\hbar} [A(\hat{p}, \hat{q}), B(\hat{p}, \hat{q})] \Big|_{\hat{p} \rightarrow p, \hat{q} \rightarrow q},$$

où \hat{p}, \hat{q} sont les opérateurs quantiques d'impulsion et de position correspondant et l'ordre des opérateurs dans A et B correspond à un choix hermitien.

Avec cette définition $\{q_i, p_j\} = \delta_{ij}$, et une transformation canonique infinitésimale (8.12) agissant sur toute fonction de A de p, q s'écrit

$$A(P, Q) = A(p, q) + \varepsilon \{A, T\}. \quad (8.15)$$

Comme les transformations canoniques forment un groupe, les crochets de Poisson ont une structure de produit de Lie. Cela signifie qu'ils satisfont à la propriété cyclique

$$\{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0. \quad (8.16)$$

Les combinaisons linéaires à coefficients réels constants des fonctions de p, q munies du produit crochet de Poisson forment une algèbre de Lie.

La dérivée par rapport au temps de toute fonction de p, q devient

$$d_t F(p, q) = \{F, H\}. \quad (8.17)$$

On en déduit l'invariance canonique des crochets de Poisson $\{q_i(t), p_j(t)\}$.

Par ailleurs les quantités conservées forment une algèbre de Lie. En effet on vérifie que si $A(p, q)$ et $B(p, q)$ sont deux constantes du mouvement

$$\{A, H\} = \{B, H\} = 0, \Rightarrow \{\{A, B\}, H\} = 0,$$

et donc $\{A, B\}$ l'est aussi. Notons d'ailleurs que les quantités (8.11) satisfont

$$\{L_{ab}, L_{cd}\} = -\delta_{ad} L_{bc} - \delta_{bc} L_{ad} + \delta_{ac} L_{bd} + \delta_{bd} L_{ac},$$

où nous reconnaissons les relations de commutation (6.10) de l'algèbre de Lie de $SO(N)$.

8.3 Théorie classique des champs

Les symétries continues en théorie classique des champs conduisent aussi à des identités remarquables et des lois de conservation. La trajectoire classique est remplacée par des champs classiques $\varphi(x)$, eux-mêmes fonctions du temps $x^0 \equiv t$, et de variables d'espace $\{x_1, \dots, x_d\}$. La densité de lagrangien \mathcal{L} est maintenant une fonction des champs et de leurs dérivées partielles premières, et l'action \mathcal{A} l'intégrale sur le temps et l'espace de \mathcal{L} :

$$\mathcal{A} = \int dx_0 dx_1 \dots dx_d \mathcal{L}(\varphi(x), \partial_\mu \varphi(x)),$$

où l'indice μ prend les valeurs $\mu = 0, 1, \dots, d$.

Les équations de champs sont obtenues en exprimant que l'action est stationnaire. On obtient les équations d'Euler–Lagrange généralisées

$$\frac{\partial \mathcal{L}}{\partial \varphi(x)} - \partial_\mu \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi(x)} = 0,$$

où la convention de sommation sur indices-exposants μ répétées est ici adoptée.

Nous supposons maintenant que le lagrangien est invariant par un groupe continu agissant linéairement sur les champs φ_i

$$\varphi_i(x) \mapsto \sum R_{ij} \varphi_j(x),$$

Nous pouvons faire la même analyse en utilisant des transformations R_{ij} qui sont maintenant fonction du temps et de l'espace. La variation de l'action est linéaire dans les dérivées et donc maintenant s'écrit

$$\delta \mathcal{A}(\varphi) = \int d^{d+1}x \sum_{ij} \partial_\mu \tau_{ij}(x) \varphi_i(x) \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi_j(x)}.$$

Après intégration par parties et développement des transformations sur une base de l'algèbre de Lie, on trouve une équation qu'on peut écrire

$$\sum_{\mu} \partial_\mu J_\alpha^\mu(x) = 0,$$

où les fonctions $J_\alpha^\mu(x)$ sont les courants associés aux générateurs de l'algèbre de Lie

$$J_\alpha^\mu(x) = \sum_{ij} \ell_{ij}^\alpha \varphi_j(x) \frac{\partial \mathcal{L}}{\partial \partial_\mu \varphi_i(x)}.$$

L'équation qui relie la variation du lagrangien (supposée ici nulle) à la divergence du courant est appelée théorème de Noether. On vérifie que les quantités conservées dans le temps sont les charges Q_α associées, intégrale sur tout l'espace de la composante de temps J_0 du courant :

$$Q_\alpha = \int dx_1 \dots dx_d J_\alpha^0(x).$$

Exemple. Considérons une théorie des champs avec un champ à N composantes $\phi_i(x)$ et une densité de lagrangien

$$\mathcal{L}(\phi) = \frac{1}{2} \sum_i \left[(\partial_t \phi_i)^2 - (\nabla_x \phi_i)^2 - m^2 \right] - \frac{1}{4} g \left[\sum_i (\phi_i)^2 \right].$$

Ce lagrangien a une symétrie orthogonale $O(N)$. Les courants conservés sont

$$J_\mu^{ij} = \phi_i \partial_\mu \phi_j - \phi_j \partial_\mu \phi_i.$$

Dans le cas du groupe $O(3)$ la matrice antisymétrique peut être paramétrée en termes d'un vecteur, et le courant devient

$$\mathbf{J}_\mu = \phi \times \partial_\mu \phi.$$

8.4 Tenseurs : un exemple, le moment d'inertie

Nous donnons maintenant un exemple de tenseurs intervenant en mécanique classique, le moment d'inertie. En mécanique classique, l'énergie E d'un solide libre en rotation s'exprime en terme du vecteur rotation ω de composantes ω_i et du moment d'inertie \mathbf{I}

$$E = \frac{1}{2} \omega \mathbf{I} \omega \equiv \sum_{ij} \omega_i \mathbf{I}_{ij} \omega_j,$$

où \mathbf{I}_{ij} a la forme d'une matrice symétrique. Si nous faisons maintenant un changement de coordonnées sous forme d'une rotation \mathbf{R} du système d'axes, les composantes du vecteur de rotation ω deviennent

$$\omega'_i = \sum_j \mathbf{R}_{ij} \omega_j \Rightarrow \omega_i = \sum_j \mathbf{R}_{ji} \omega'_j, \quad (8.18)$$

où nous avons utilisé que pour une matrice de rotation $\mathbf{R}^T \mathbf{R} = \mathbf{1}$. L'énergie de rotation du solide est indépendante du système de coordonnées, et donc en fonction des nouvelles composantes ω'_i s'écrit

$$E = \sum_{ijkl} \omega'_i \mathbf{R}_{ik} \mathbf{I}_{kl} \mathbf{R}_{jl} \omega'_j,$$

En conséquence à la transformation par rotation (8.18) du vecteur ω correspond la transformation des composantes du moment d'inertie

$$\mathbf{I}'_{ij} = \sum_{kl} \mathbf{R}_{ik} \mathbf{R}_{jl} \mathbf{I}_{kl}. \quad (8.19)$$

Il est facile de vérifier que cette transformation est une nouvelle représentation du groupe des rotations. Les éléments sur lesquels agit le groupe des rotations

par la loi (8.19), et qui forment un espace vectoriel, sont des tenseurs. On peut évidemment généraliser cette définition à des objets qui ont plus de deux indices, et c'est même nécessaire en mécanique quantique.

Décomposition en représentations irréductibles. Reprenons alors l'exemple des tenseurs de rang deux comme le moment d'inertie. Si dans la transformation (8.19) nous supposons, comme c'est toujours le cas pour le moment d'inertie, que le tenseur \mathbf{I}_{ij} est symétrique, alors \mathbf{I}'_{ij} est automatiquement symétrique. Si nous supposons que le tenseur \mathbf{I}_{ij} est antisymétrique, alors \mathbf{I}'_{ij} est automatiquement antisymétrique. Enfin si nous supposons que le solide est invariant par rotation alors le moment d'inertie prend la forme d'un multiple de la matrice identité

$$\mathbf{I}_{ij} = I\delta_{ij}.$$

Alors, parce que les matrices de rotations sont orthogonales, les composantes du moment d'inertie sont invariantes.

Prenant l'exemple de l'espace de dimension trois. La dimension de l'espace vectoriel des tenseurs à deux indices est $3^2 = 9$. Cet espace est réductible pour la représentation du groupe des rotations. Les tenseurs proportionnels à δ_{ij} forment un sous-espace de dimension un qui est invariant. Cela signifie que dans ce sous-espace tous les éléments du groupe sont envoyés sur l'identité.

Les tenseurs antisymétriques $\mathbf{I}_{ij} = -\mathbf{I}_{ji}$ forment un autre sous-espace vectoriel invariant par le groupe de dimension trois, et on vérifie que la représentation du groupe dans ce sous-espace est isomorphe à $SO(3)$, le groupe initial.

Enfin, les tenseurs symétriques de trace nulle $\sum_i \mathbf{I}_{ii} = 0$, forment un sous-espace vectoriel invariant de dimension cinq, dont on montre qu'il est irréductible, dans lequel le groupe $SO(3)$ est donc représenté par des matrices 5×5 .

En termes de mécanique quantique nous avons décomposé les états à deux particules de spin un, en une somme directe d'une composante scalaire, une composante de spin un et une composante de spin deux.

9 Symétries en mécanique quantique

Rappelons brièvement comment ces propriétés se généralisent en mécanique quantique.

Toute symétrie est représentée dans l'espace de Hilbert des états de la mécanique quantique par un opérateur unitaire ou anti-unitaire (unitaire plus conjugaison complexe) car la conservation des probabilités entraîne l'invariance de la norme de tout vecteur. Vérifions le cas unitaire :

$$|\psi_U\rangle = \mathbf{U}|\psi\rangle \Rightarrow \langle\psi_U| = \langle\psi|\mathbf{U}^\dagger \quad (9.1)$$

et donc

$$\mathbf{U}^\dagger\mathbf{U} = \mathbf{1} \Rightarrow \langle\psi_U|\psi_U\rangle = \langle\psi|\psi\rangle,$$

où nous avons utilisé la notation des bras et kets pour représenter le vecteur associé à l'état ψ et son hermitien conjugué.

Vecteurs et observables. Les quantités physiques (mesurables) ou observables \mathcal{O} sont des valeurs moyennes dans des états d'opérateurs (hermitiens par réalité) de l'espace de Hilbert,

$$\mathcal{O} = \langle\psi|\mathbf{O}|\psi\rangle \quad \text{avec} \quad \langle\psi|\psi\rangle = 1,$$

où \mathbf{O} est un opérateur de l'espace de Hilbert. Dans ces conditions une transformation unitaire (9.1) sur les états est équivalente du point de vue des observables à une transformation des opérateurs avec des états fixes

$$\langle\psi_U|\mathbf{O}|\psi_U\rangle = \langle\psi|\mathbf{O}_U|\psi\rangle \quad \text{avec} \quad \mathbf{O}_U = \mathbf{U}^\dagger\mathbf{O}\mathbf{U}.$$

Dans ce qui suit nous discuterons la transformation des opérateurs plutôt que celle des états, ce qui correspond au formalisme de Heisenberg plutôt qu'au formalisme de Schrödinger.

9.1 Evolution dans le temps. Quantités conservées

À l'évolution temporelle correspond donc un opérateur unitaire, l'opérateur d'évolution $U(t, t')$ qui décrit l'évolution du temps t' au temps t . On suppose que pour un système isolé l'opérateur d'évolution satisfait la loi de composition

$$U(t_3, t_2)U(t_2, t_1) = U(t_3, t_1), \quad (9.2)$$

c'est à dire que l'évolution dans le temps est markovienne : l'évolution est sans mémoire. Il agit sur les états ψ de l'espace de Hilbert par

$$|\psi(t)\rangle = U(t, t')|\psi(t')\rangle.$$

Comme conséquence de l'unitarité et de la loi de groupe

$$U^\dagger(t_2, t_1) = U(t_1, t_2).$$

Comme l'opérateur d'évolution satisfait l'identité (9.2), son action peut être obtenue à partir d'évolutions infinitésimales. On pose alors

$$U(t + \varepsilon, t) = 1 - \frac{i\varepsilon}{\hbar} H(t) + O(\varepsilon^2).$$

L'opérateur $H(t)$ qui représente le *générateur* des translations dans le temps, est l'opérateur hamiltonien comme on le voit dans la limite classique où les commutateurs deviennent des crochets de Poisson.

Dans ce formalisme l'équation de Schrödinger qui décrit l'évolution des états dans le temps s'écrit

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H(t) |\psi(t)\rangle.$$

Si le hamiltonien est indépendant du temps, l'opérateur d'évolution prend la forme

$$U(t, t') = e^{-i(t-t')H/\hbar}.$$

La valeur moyenne de l'hamiltonien à un temps donné est une observable, l'énergie E . Si l'hamiltonien est indépendant du temps,

$$H_U = e^{i(t-t')H/\hbar} H e^{-i(t-t')H/\hbar} = H,$$

et donc l'énergie moyenne E est conservée,

$$E = \langle \psi(t) | H | \psi(t) \rangle = \langle \psi(0) | H_U | \psi(0) \rangle = \langle \psi(0) | H | \psi(0) \rangle.$$

Plus généralement, l'évolution dans le temps d'une observable est associée à l'évolution de l'opérateur $O(t)$ de l'espace de Hilbert correspondant

$$O(t) = e^{i(t-t')H/\hbar} O e^{-i(t-t')H/\hbar}.$$

L'observable a une valeur moyenne indépendante du temps si $O(t)$ ne dépend pas du temps. De nouveau prenant la limite des translations infinitésimales, et gardant le terme d'ordre t , on trouve la relation de commutation

$$[H, O] = 0.$$

Les observables conservées dans le temps correspondent à des opérateurs qui commutent avec le hamiltonien.

Notons que les opérateurs correspondant à des quantités conservées ont une structure d'algèbre de Lie. En effet si

$$[H, O_1] = 0 \quad \text{et} \quad [H, O_2] = 0,$$

alors

$$[H, [O_1, O_2]] = -[O_1, [O_2, H]] - [O_2, [H, O_1]] = 0.$$

Donc l'observable associée au commutateur $[O_1, O_2]$ est aussi conservée.

Un groupe de symétrie a une représentation en terme d'un groupe d'opérateurs unitaires. D'après l'argument précédent l'évolution dans le temps est compatible avec la symétrie si les opérateurs correspondants commutent avec le hamiltonien. Dans le cas d'un groupe de Lie ceci entraîne que le hamiltonien commute avec les opérateurs représentant les générateurs de l'algèbre de Lie du groupe.

9.2 Transformations et symétries

Pour discuter des symétries d'espace il faut introduire l'observable correspondante, l'opérateur position $\hat{\mathbf{q}}$. Les composantes des vecteurs d'état sur les états propres $|\mathbf{q}\rangle$ de $\hat{\mathbf{q}}$ sont les fonctions d'onde $\psi(\mathbf{q})$

$$\hat{\mathbf{q}}|\mathbf{q}\rangle = \mathbf{q}|\mathbf{q}\rangle, \quad \psi(\mathbf{q}) = \langle \mathbf{q}|\psi\rangle.$$

Dans ce qui suit l'espace a une, deux, trois dimensions... et donc $\hat{\mathbf{q}}$ a une ou plusieurs composantes $\hat{\mathbf{q}}_i$.

Translations d'espace. Nous introduisons donc d'abord la représentation par opérateur unitaire du groupe de translation d'espace. Une translation de vecteur \mathbf{a} est représentée par une transformation unitaire $T(\mathbf{a})$:

$$T^\dagger(\mathbf{a})T(\mathbf{a}) = \mathbf{1}.$$

Le groupe des translations est abélien :

$$T(\mathbf{a})T(\mathbf{b}) = T(\mathbf{a} + \mathbf{b}) = T(\mathbf{b})T(\mathbf{a}).$$

Prenant la limite $|\mathbf{a}| \rightarrow 0$, nous définissons la représentation des générateurs du groupe des translations

$$T(\mathbf{a}) = \mathbf{1} + i\mathbf{a} \cdot \hat{\mathbf{p}}/\hbar + O(|\mathbf{a}|^2),$$

où les opérateurs $\hat{\mathbf{p}}$ sont hermitiens. Leurs composantes commutent :

$$[\hat{p}_\alpha, \hat{p}_\beta] = 0.$$

Les éléments de la représentation du groupe peuvent être écrits sous forme exponentielle en termes des générateurs

$$T(\mathbf{a}) = e^{i\mathbf{a} \cdot \hat{\mathbf{p}}/\hbar}.$$

Cherchons maintenant les vecteurs invariants par translation. Comme la multiplication d'un vecteur par une phase ne change pas l'état physique, cela revient à chercher les vecteurs propres de $T(\mathbf{a})$ qui alors correspondent à des valeurs propres de module 1 :

$$T(\mathbf{a})|\widetilde{\mathbf{k}}\rangle = e^{i\mathbf{k} \cdot \mathbf{a}/\hbar}|\widetilde{\mathbf{k}}\rangle,$$

où \mathbf{k} est un vecteur quelconque. Les opérateurs de translation ont un spectre continu, et en conséquence leurs vecteurs propres n'appartiennent pas à l'espace de Hilbert, puisque la condition d'orthogonalité s'écrit

$$|\widetilde{\mathbf{k}}\rangle|\widetilde{\mathbf{k}}'\rangle = (2\pi)^d \delta(\mathbf{k} - \mathbf{k}').$$

Ils forment néanmoins une base complète.

Enfin l'interprétation physique du vecteur \mathbf{k} est l'impulsion associée à l'état invariant par translation.

Translations d'impulsion. Introduisons maintenant les opérateurs unitaires $V(\mathbf{k})$ qui ajoutent à l'impulsion des états un vecteur \mathbf{k} : Ils peuvent être définis par leur action sur les vecteurs $|\tilde{\psi}(\mathbf{k})\rangle$ qui forment une base complète :

$$V(\mathbf{k})|\widetilde{\mathbf{k}'}\rangle = |(\mathbf{k}' + \mathbf{k})\widetilde{\rangle}.$$

Ils forment un groupe abélien et peuvent être exprimés en fonction de générateurs hermitiens $\hat{\mathbf{q}}$:

$$V(\mathbf{k}) = e^{-i\mathbf{k}\cdot\hat{\mathbf{q}}/\hbar}.$$

Les générateurs composantes de $\hat{\mathbf{q}}$ commutent.

De nouveau ces opérateurs ont un spectre continu, et l'analyse est la même que dans le cas précédent.

Relations de commutations. Calculons

$$T(\mathbf{a})V(\mathbf{k})|\widetilde{\mathbf{k}'}\rangle = e^{i\mathbf{a}\cdot(\mathbf{k}+\mathbf{k}')/\hbar}|\widetilde{\mathbf{k}'}\rangle.$$

De même

$$V(\mathbf{k})T(\mathbf{a})|\widetilde{\mathbf{k}'}\rangle = e^{i\mathbf{a}\cdot\mathbf{k}'/\hbar}|\widetilde{\mathbf{k}'}\rangle.$$

Comme ces relations sont vraies pour tous les vecteurs de base, nous en déduisons les relations de commutation

$$T(\mathbf{a})V(\mathbf{k}) = e^{i\mathbf{a}\cdot\mathbf{k}/\hbar} V(\mathbf{k})T(\mathbf{a}). \quad (9.3)$$

Développant pour $\mathbf{a}, \mathbf{k} \rightarrow 0$, on en déduit les relations de commutation entre générateurs

$$[q_\alpha, p_\beta] = i\hbar\delta_{\alpha\beta}.$$

On en déduit également

$$\hat{\mathbf{q}} + \mathbf{a} = T(\mathbf{a})\hat{\mathbf{q}}T^\dagger(\mathbf{a}). \quad (9.4)$$

Fonctions d'onde. Considérons maintenant les vecteurs propres de $V(\mathbf{k})$. Ils dépendent d'un vecteur \mathbf{q} et nous les notons simplement $|\mathbf{q}\rangle$.

$$V(\mathbf{k})|\mathbf{q}\rangle = e^{-i\mathbf{k}\cdot\mathbf{q}}|\mathbf{q}\rangle.$$

Nous utilisons de nouveau les relations de commutation (9.3), agissant sur le vecteur $|\mathbf{q}\rangle$. Nous en déduisons

$$T(\mathbf{a})|\mathbf{q}\rangle = |\mathbf{q} + \mathbf{a}\rangle.$$

Le vecteur \mathbf{q} caractérise une position attachée au vecteur $|\mathbf{q}\rangle$. Par ailleurs

$$\langle\mathbf{q}'|\mathbf{q}\rangle = \delta(\mathbf{q} - \mathbf{q}').$$

L'opérateur $\hat{\mathbf{q}}$ doit donc être identifié avec un opérateur de position. Les vecteurs propres $|\mathbf{q}\rangle$ correspondent donc à des états strictement localisés au point d'espace \mathbf{q} .

Représentation des fonctions d'onde. Tout vecteur de l'espace de Hilbert $|\psi\rangle$ peut être développé sur la base $|\mathbf{q}\rangle$. On définit la fonction d'onde

$$\psi(\mathbf{q}) = \langle \mathbf{q} | \psi \rangle,$$

dont le module carré caractérise la probabilité de trouver une particule au point \mathbf{q} .

De même on peut développer les vecteurs sur la base des vecteurs propres de $\hat{\mathbf{p}}$:

$$\tilde{\psi}(\mathbf{p}) = \langle \psi | \tilde{\mathbf{p}} \rangle.$$

Les deux représentations sont reliées par transformation de Fourier

$$\int d^d q e^{i\mathbf{q}\cdot\mathbf{p}/\hbar} |\mathbf{q}\rangle = |\tilde{\mathbf{p}}\rangle \Rightarrow \psi(\mathbf{q}) = \langle \mathbf{q} | \psi \rangle = \frac{1}{(2\pi\hbar)^d} \int d^d p e^{i\mathbf{q}\cdot\mathbf{p}/\hbar} \tilde{\psi}(\mathbf{p}).$$

Dans la représentation des fonctions d'onde, les fonctions d'onde à impulsion fixé sont les ondes planes.

Invariance par translation. Dire que l'évolution dans le temps est invariante par translation, signifie que faire une translation puis laisser évoluer ou laisser évoluer puis faire une translation conduit au même résultat :

$$[U(t', t), T(\mathbf{a})] = 0.$$

Passant aux générateurs, c'est à dire développant au premier ordre en t et \mathbf{a} on trouve les relations de commutations :

$$[H, \hat{\mathbf{p}}] = 0.$$

Dans ces conditions ils existent des valeurs moyennes conservées, les valeurs moyennes de $\hat{\mathbf{p}}$ qui sont les impulsions ou quantités de mouvement. Un exemple d'un tel hamiltonien est fourni par le hamiltonien de la particule libre

$$H = (\hat{\mathbf{p}})^2 / 2m.$$

Rotations. Les mêmes arguments montrent que les rotations de l'espace à trois dimensions sont représentées par des transformations unitaires $U(\mathbf{R})$ telles que

$$U(\mathbf{R}) \hat{\mathbf{q}} U^\dagger(\mathbf{R}) = \sum_j \mathbf{R}_{ij} \hat{\mathbf{q}}_j,$$

où \mathbf{R} est une matrice de $SO(3)$. Comme les opérateurs $U(\mathbf{R})$ appartiennent à une représentation de $SO(3)$ on peut les exprimer en termes de trois opérateurs \mathbf{L} générateurs du groupe unitaire et d'un vecteur réel $\boldsymbol{\Omega}$

$$U(\mathbf{R}) = \exp[-i\boldsymbol{\Omega} \cdot \mathbf{L}/\hbar].$$

Faisant une rotation infinitésimale et introduisant les générateurs du groupe $SO(3)$ (5.1) nous trouvons

$$\frac{i}{\hbar}[\hat{\mathbf{q}}_i, \mathbf{L}_a] = \sum_j (\mathbf{T}_a)_{ij} \hat{\mathbf{q}}_j = - \sum_j \epsilon_{aij} \hat{\mathbf{q}}_j. \quad (9.5)$$

Les relations de commutation de l'algèbre de Lie de $SO(3)$ entraînent alors

$$[\mathbf{L}_i, \mathbf{L}_j] = i\hbar \sum_k \epsilon_{ijk} \mathbf{L}_k.$$

Par ailleurs on vérifie que les opérateurs

$$\mathbf{L} = \hat{\mathbf{q}} \times \hat{\mathbf{p}},$$

satisfont aux relations de commutation (9.5).

L'évolution dans le temps est invariante par rotation si

$$U^\dagger(\mathbf{R}) H U(\mathbf{R}) = H.$$

Ceci entraîne que le hamiltonien commute avec les trois générateurs \mathbf{L}

$$[H, \mathbf{L}] = 0,$$

et par conséquent les valeurs moyennes de ces trois opérateurs sont conservées dans le temps.

Remarque. L'opérateur position se transforme par la représentation adjointe du groupe unitaire $U(\mathbf{R})$. Dans ces conditions si l'on remplace le groupe des matrices \mathbf{R} de $SO(3)$ par les matrices du groupe $SU(2)$ rien n'est changé. C'est exactement ce qui se passe dans le cas de particules de spin 1/2 ou plus généralement demi-entier où l'état se transforme par une représentation de $SU(2)$.

Spin et tenseurs. Nous considérons ici des particules caractérisées uniquement par leur spin. Ceci signifie qu'un état à une particule est associé à un vecteur : l'espace de Hilbert est un espace vectoriel de dimension $2s+1$ pour une particule de spin s . Un état à deux particules de spin s_1 et s_2 est alors associé au *produit tensoriel* de deux espaces vectoriels qu'on peut décrire par un objet à deux indices vectoriels correspondant aux deux particules, un espace de dimension $(2s_1 + 1)(2s_2 + 1)$.

De façon générale une fonction d'onde à n particules par exemple de spin 1/2 sera décrite par un objet à n indices $\psi_{i_1 i_2 \dots i_n}$ et appartiendra à un espace vectoriel de dimension complexe 2^n . Une fonction d'onde à n particules de spin 1 sera décrite par un objet à n indices $\psi_{i_1 i_2 \dots i_n}$ et appartiendra à un espace vectoriel de dimension réelle 3^n . Dans ce dernier cas dans une rotation d'espace la fonction d'onde à une particule se transforme comme un vecteur de l'espace

à trois dimensions. La fonction d'onde à n particules se transforme comme un tenseur de rang n

$$\psi_{R, i_1 i_2 \dots i_n} = \sum_{j_1 j_2 \dots j_n} \left(\prod_{k=1}^n \mathbf{R}_{i_k j_k} \right) \psi_{j_1 j_2 \dots j_n} .$$

On vérifie facilement que la loi de transformation précédente définit une représentation du groupe des rotations. Les particules de spins demi-entiers se transforment par des représentations du groupe $SU(2)$.

10 Mouvement brownien ou marche au hasard : Limite continue

Nous considérons un processus de marche au hasard sur le réseau à N dimensions, $N = 1, 2, 3, \dots$ des points de coordonnées entières $\mathbf{q} \equiv (q_1, \dots, q_N)$. Nous supposons le processus markovien (sans mémoire), ce qui signifie que le déplacement au temps n ne dépend que de la position \mathbf{q}_n au temps n mais pas de l'histoire qui a conduit à ce point. Il est donc entièrement spécifié par une probabilité de transition $p(\mathbf{q}, \mathbf{q}')$ du point \mathbf{q}' vers le point \mathbf{q} .

Soit alors $P_n(\mathbf{q})$ la probabilité au temps n d'être au point \mathbf{q} . La probabilité $P_n(\mathbf{q})$ satisfait à l'équation de récurrence (parfois appelée équation maîtresse)

$$P_{n+1}(\mathbf{q}) = \sum_{\mathbf{q}' \in \mathbb{Z}^N} p(\mathbf{q}, \mathbf{q}') P_n(\mathbf{q}'). \quad (10.1)$$

Les quantités P_n et p étant des probabilités sont positives par définition. La conservation des probabilités implique les conditions

$$\sum_{\mathbf{q} \in \mathbb{Z}^N} P_n(\mathbf{q}) = 1, \quad \sum_{\mathbf{q} \in \mathbb{Z}^N} p(\mathbf{q}, \mathbf{q}') = 1. \quad (10.2)$$

Systèmes invariants par translation. Nous considérons maintenant l'exemple d'une probabilité de transition invariante par translation

$$p(\mathbf{q}, \mathbf{q}') \equiv p(\mathbf{q} - \mathbf{q}'). \quad (10.3)$$

L'équation (10.1) peut être interprétée comme l'action itérative d'une matrice \mathbf{p} d'éléments $p(\mathbf{q}, \mathbf{q}')$ (ou d'un opérateur pour un réseau infini) sur un vecteur $P_n(\mathbf{q})$. On peut alors introduire l'opérateur translation $\mathbf{T}(\mathbf{a})$ d'un vecteur spatial \mathbf{a}

$$\mathbf{T}(\mathbf{a})P_n(\mathbf{q}) = P_n(\mathbf{q} + \mathbf{a}).$$

L'invariance par translation est équivalente à la commutation des opérateurs \mathbf{p} et \mathbf{T}

$$[\mathbf{p}, \mathbf{T}(\mathbf{a})] = 0,$$

qui peuvent donc être diagonalisés simultanément.

Le groupe des translations est isomorphes au groupe additif \mathbb{Z}^N . C'est un groupe commutatif qui a comme générateurs les translations \mathbf{T}_i d'une unité dans chacune des N directions possibles

$$\mathbf{T}(\mathbf{a}) = \prod_{i=1}^N (\mathbf{T}_i)^{a_i}.$$

Un vecteur propre de $\mathbf{T}(\mathbf{a})$ satisfait

$$\mathbf{T}(\mathbf{a})\psi(\mathbf{q}) = \psi(\mathbf{q} + \mathbf{a}) = \tau(\mathbf{a})\psi(\mathbf{q}).$$

La loi de groupe permet d'écrire les valeurs propres

$$\tau(\mathbf{a}) = \prod_{i=1}^N (\tau_i)^{a_i}.$$

Les opérateurs \mathbf{T}_i agissent sur des distributions qui sont des fonctions bornées. Les valeurs propres de \mathbf{T}_i sont des nombres complexes. Si l'on veut qu'une translation quelconque soit une opération interne aux fonctions bornées en module, il faut que ce nombre complexe soit de module un

$$\tau(\mathbf{a}) = e^{-i\mathbf{k}\cdot\mathbf{a}},$$

où \mathbf{k} est un vecteur réel. Comme le vecteur \mathbf{a} n'a que des composantes entières les fonctions $e^{-i\mathbf{k}\cdot\mathbf{a}}$ sont des fonctions périodiques des composantes k_i du vecteur \mathbf{k} de période 2π , et ces composantes peuvent donc être restreintes à $-\pi \leq k_i < \pi$ (une zone de Brillouin).

Les vecteurs propres de l'opérateur translation $\mathbf{T}(\mathbf{a})$ correspondants sont des exponentielles :

$$\mathbf{T}(\mathbf{a}) e^{-i\mathbf{k}\cdot\mathbf{q}} = e^{-i\mathbf{k}\cdot(\mathbf{q}+\mathbf{a})} = e^{-i\mathbf{k}\cdot\mathbf{a}} e^{-i\mathbf{k}\cdot\mathbf{q}}.$$

On est donc amené à décomposer le vecteur $P_n(\mathbf{q})$ sur la base des ondes planes, c'est à dire à l'écrire comme une transformée de Fourier

$$P_n(\mathbf{q}) = \frac{1}{(2\pi)^N} \int d^N k e^{-i\mathbf{k}\cdot\mathbf{q}} \tilde{P}_n(\mathbf{k}), \quad (10.4)$$

et réciproquement

$$\tilde{P}_n(\mathbf{k}) = \sum_{\mathbf{q} \in \mathbb{Z}^N} e^{i\mathbf{k}\cdot\mathbf{q}} P_n(\mathbf{q}) \Rightarrow \tilde{P}_n^*(\mathbf{k}) = \tilde{P}_n(-\mathbf{k}), \quad \tilde{P}(\mathbf{k} = 0) = 1, \quad (10.5)$$

où la propriété (10.2) a été utilisée.

Calculons maintenant $\tilde{P}_{n+1}(\mathbf{k})$ à partir de l'équation (10.1), utilisant la forme (10.3) :

$$\begin{aligned} \tilde{P}_{n+1}(\mathbf{k}) &= \sum_{\mathbf{q}, \mathbf{q}' \in \mathbb{Z}^N} e^{i\mathbf{k}\cdot\mathbf{q}} p(\mathbf{q} - \mathbf{q}') P_n(\mathbf{q}') \\ &= \sum_{\mathbf{q}, \mathbf{q}' \in \mathbb{Z}^N} e^{i\mathbf{k}\cdot\mathbf{q}} p(\mathbf{q} - \mathbf{q}') \frac{1}{(2\pi)^N} \int d^N k' e^{-i\mathbf{k}'\cdot\mathbf{q}'} \tilde{P}_n(\mathbf{k}'). \end{aligned}$$

Mais

$$\begin{aligned} \sum_{\mathbf{q}, \mathbf{q}' \in \mathbb{Z}^N} e^{i\mathbf{k}\cdot\mathbf{q}} p(\mathbf{q} - \mathbf{q}') e^{-i\mathbf{k}'\cdot\mathbf{q}'} &= \sum_{\mathbf{q}, \mathbf{q}' \in \mathbb{Z}^N} e^{i\mathbf{k}\cdot\mathbf{q}} p(\mathbf{q}) e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{q}'} \\ &= (2\pi)^N \delta^{(N)}(\mathbf{k} - \mathbf{k}') \sum_{\mathbf{q} \in \mathbb{Z}^N} e^{i\mathbf{k}\cdot\mathbf{q}} p(\mathbf{q}). \end{aligned}$$

Nous posons alors

$$\tilde{p}(\mathbf{k}) = \sum_{\mathbf{q} \in \mathbb{Z}^N} e^{i\mathbf{k} \cdot \mathbf{q}} p(\mathbf{q}) \Rightarrow \tilde{p}^*(\mathbf{k}) = \tilde{p}(-\mathbf{k}), \quad \tilde{p}(0) = 1. \quad (10.6)$$

La fonction \tilde{p} est aussi une fonction périodique des composantes k_i du vecteur \mathbf{k} . Enfin prenant le module de l'équation de définition nous trouvons la borne

$$|\tilde{p}(\mathbf{k})| \leq \sum_{\mathbf{q} \in \mathbb{Z}^N} |e^{i\mathbf{k} \cdot \mathbf{q}} p(\mathbf{q})| = 1. \quad (10.7)$$

De plus dans la zone de Brillouin la borne n'est atteinte qu'à l'origine $\mathbf{k} = 0$. Avec cette définition l'équation (10.1) s'écrit donc

$$\tilde{P}_{n+1}(\mathbf{k}) = \tilde{p}(\mathbf{k}) \tilde{P}_n(\mathbf{k}) \Rightarrow \tilde{P}_n(\mathbf{k}) = \tilde{p}^n(\mathbf{k}) \tilde{P}_0(\mathbf{k}).$$

Par conséquent

$$P_n(\mathbf{q}) = \frac{1}{(2\pi)^N} \int d^N k e^{-i\mathbf{k} \cdot \mathbf{q}} \tilde{p}^n(\mathbf{k}) \tilde{P}_0(\mathbf{k}). \quad (10.8)$$

Symétrie cubique. Nous supposons maintenant que la probabilité de transition a la symétrie cubique du réseau. Nous avons vu que le groupe cubique \mathfrak{C}_N admet comme générateurs les générateurs du groupe des permutations et une réflexion. Il suffit donc d'imposer

$$\begin{cases} p(q_1, q_2, \dots, q_N) = p(-q_1, q_2, \dots, q_N), \\ p(q_1, \dots, q_i, q_{i+1}, \dots, q_N) = p(q_1, \dots, q_{i+1}, q_i, \dots, q_N). \end{cases} \quad (10.9)$$

Exprimons maintenant les conséquences de la symétrie cubique. Notons \mathbf{t} un des générateurs du groupe de symétrie explicités en (10.9), qui tous satisfont $\mathbf{t}^2 = \mathbf{1}$. On vérifie alors

$$\mathbf{k} \cdot \mathbf{t}\mathbf{q} = \mathbf{t}\mathbf{k} \cdot \mathbf{q},$$

et donc

$$\begin{aligned} \tilde{p}(\mathbf{k}) &= \sum_{\mathbf{q} \in \mathbb{Z}^N} e^{i\mathbf{k} \cdot \mathbf{q}} p(\mathbf{q}) = \sum_{\mathbf{q} \in \mathbb{Z}^N} e^{i\mathbf{k} \cdot \mathbf{q}} p(\mathbf{t}\mathbf{q}) = \sum_{\mathbf{t}\mathbf{q} \in \mathbb{Z}^N} e^{i\mathbf{k} \cdot \mathbf{t}\mathbf{q}} p(\mathbf{q}) \\ &= \sum_{\mathbf{q} \in \mathbb{Z}^N} e^{i\mathbf{k} \cdot \mathbf{t}\mathbf{q}} p(\mathbf{q}) = \tilde{p}(\mathbf{t}\mathbf{k}). \end{aligned}$$

La fonction $\tilde{p}(\mathbf{k})$ satisfait donc des conditions analogues à (10.9). Elle a une symétrie cubique, c'est à dire qu'elle est symétrique et paire dans toutes composantes du vecteur \mathbf{k} ce qui entraîne aussi qu'elle est réelle.

Déplacements à courte distance. Nous nous restreignons maintenant aux probabilités de transition $p(\mathbf{q})$ qui décroissent au moins exponentiellement avec

la distance $|\mathbf{q}|$, et ceci comprend l'exemple classique où les seuls déplacements possibles correspondent aux voisins sur le réseau.

$$p(\mathbf{q}) \leq M e^{-\mu|\mathbf{q}|}, \quad \mu > 0.$$

Dans ce cas la fonction $\tilde{p}(\mathbf{k})$, somme uniformément convergente de fonctions entières

$$|\tilde{p}(\mathbf{k})| \leq M \sum_{\mathbf{q}} e^{-(\mu - |\operatorname{Im} \mathbf{k}|)|\mathbf{q}|}$$

pour $|\operatorname{Im} \mathbf{k}| < \mu$, est une fonction analytique sur l'axe réel. Posons

$$\tilde{p}(\mathbf{k}) = e^{-w(\mathbf{k})} \Rightarrow w^*(\mathbf{k}) = w(-\mathbf{k}), \quad w(\mathbf{0}) = 0. \quad (10.10)$$

La régularité de \tilde{p} et la condition $\tilde{p}(\mathbf{0}) = 1$ impliquent que $w(\mathbf{k})$ a un développement régulier pour $|\mathbf{k}| \rightarrow 0$.

La symétrie de réflexion $k_i \mapsto -k_i$ interdit tout terme impair en \mathbf{k} . Examinons les conséquences de la symétrie cubique sur les termes quadratiques dans les composantes du vecteur \mathbf{k} :

$$\sum_i a_i k_i^2 + \sum_{i \neq j} b_{ij} k_i k_j.$$

Imposant l'invariance dans le changement de signe d'une composante $k_\ell \mapsto -k_\ell$ nous voyons que le premier terme est invariant et que le coefficient de k_ℓ dans le second terme change de signe. Donc les coefficients b_{ij} doivent s'annuler. Dans le premier nous échangeons k_i et k_{i+1} et la forme doit être invariante, ce qui implique $a_i = a_{i+1}$. On en déduit qu'un seul terme quadratique est symétrique, le carré scalaire du vecteur $\sum_i k_i^2$. Remarquons alors que ce terme a une symétrie plus grande que la symétrie cubique puisqu'il est invariant par le groupe $O(N)$ des rotations-réflexions.

Faisons le même exercice pour les termes quartiques. À cause de l'invariance par réflexion il est nécessairement de la forme

$$\sum_i a_i k_i^4 + \sum_{i \neq j} b_{ij} k_i^2 k_j^2.$$

Si nous exprimons maintenant l'invariance dans l'échange k_i et k_{i+1} nous trouvons $a_i = a_{i+1} = a$, $b_{ij} = b_{i+1,j} = b$, d'où la forme générale

$$a \sum_i k_i^4 + b \sum_{i \neq j} k_i^2 k_j^2 = b \left(\sum_i k_i^2 \right)^2 + (a - b) \sum_i k_i^4.$$

Si $a \neq b$ le deuxième terme brise l'invariance par rotation et ne laisse que l'invariance cubique.

À $|\mathbf{k}|$ petit la fonction $w(\mathbf{k})$ peut donc être paramétrée comme

$$w(\mathbf{k}) = w_2 \mathbf{k}^2 / 2 + O(k^4), \quad w_2 > 0, \quad (10.11)$$

la positivité de w_2 résultant de l'inégalité (10.7).

Comportement asymptotique aux temps longs. Pour $n \rightarrow \infty$, l'intégrale (10.8) est dominée par les valeurs maximales de $|\tilde{p}(\mathbf{k})|$, et donc le voisinage de $\mathbf{k} = \mathbf{0}$. À cause de la forme (10.11) de $w(\mathbf{k})$, pour $\mathbf{k} \rightarrow \mathbf{0}$ seules les valeurs de \mathbf{k} d'ordre $1/\sqrt{n}$ contribuent à l'intégrale (10.8). À l'ordre dominant nous pouvons donc négliger les termes d'ordre k^4 dans $w(\mathbf{k})$.

Nous supposons aussi que la distribution initiale $P_0(\mathbf{q})$ décroît assez rapidement avec $|\mathbf{q}|$ pour que $\tilde{P}_n(\mathbf{k})$ soit régulier

$$\tilde{P}_n(\mathbf{k}) = \exp [i\mathbf{k} \cdot \mathbf{q}_0 + O(k^2)],$$

où \mathbf{q}_0 est la position initiale moyenne. Nous pouvons alors négliger les termes d'ordre k^2 ou plus. Pour les mêmes raisons que dans la méthode du col, dans la limite $n \rightarrow \infty$ nous pouvons intégrer dans (10.8) sur toutes les valeurs de \mathbf{k} réelles sans restriction à la zone de Brillouin. Nous en déduisons la forme asymptotique de $P_n(\mathbf{q})$,

$$P_n(\mathbf{q}) \sim \frac{1}{(2\pi w_2 n)^{N/2}} \exp \left[-\frac{(\mathbf{q} - \mathbf{q}_0)^2}{2w_2 n} \right]. \quad (10.12)$$

Nous sommes ramenés à une situation très semblable au théorème de la limite centrale. Pour $n \rightarrow \infty$ la probabilité $P_n(\mathbf{q})$ prend une forme gaussienne universelle qui de nouveau ne dépend que de propriétés générales de $p(\mathbf{q} - \mathbf{q}')$. Nous notons

(i) Cette forme gaussienne asymptotique est valable à temps grands et pour des distances $|\mathbf{q}| \ll n$.

(ii) La distribution asymptotique a une symétrie $O(N)$ de rotations-réflexion, indépendamment de la distribution initiale, et admet un groupe de symétrie plus grand que la probabilité de transition qui n'a que la symétrie cubique du réseau.

(iii) Faisons un changement d'échelles de temps et de distance

$$t = n\varepsilon, \quad \mathbf{x} = (\mathbf{q} - \mathbf{q}_0)\sqrt{\varepsilon}, \quad (10.13)$$

où nous prenons une limite $n \rightarrow \infty$ à t fixé, et donc $\varepsilon = O(1/n)$. En fonction de ces variables *macroscopiques* la distribution asymptotique devient (le changement de variables sur \mathbf{q} entraîne un changement de normalisation)

$$P(t, \mathbf{x}) = \frac{1}{(2\pi w_2 t)^{N/2}} e^{-\mathbf{x}^2/2w_2 t}. \quad (10.14)$$

Le temps t et les coordonnées \mathbf{x} sont les variables qui décrivent à l'échelle macroscopique le mouvement brownien engendré par la dynamique microscopique (10.1). L'universalité a permis de définir une *limite continue*, dans la mesure où la structure de réseau et les détails du processus élémentaire ont disparu.

(iv) Il est simple d'étudier comment des perturbations à la distribution gaussienne limite décroissent avec n . Développons la fonction régulière $w(\mathbf{k})$ définie en (10.10) en puissances de \mathbf{k} jusqu'à l'ordre quatre (10.11) :

$$w(\mathbf{k}) = w_2 \mathbf{k}^2 / 2 + \frac{1}{4!} w_4 (\mathbf{k}^2)^2 + \frac{1}{4!} w'_4 \sum_i k_i^4 + O(\mathbf{k}^6).$$

À l'ordre k^4 les corrections brisent l'invariance par rotation. Nous faisons alors les changements d'échelle (10.13), ce qui sur les variables de Fourier correspond à $\mathbf{k} = \kappa \sqrt{\varepsilon}$. Dans la limite $\varepsilon \rightarrow 0$, κ fixé, la zone de Brillouin devient toute la droite réelle. Dans le cas d'une distribution initiale localisée exactement à $\mathbf{q} = \mathbf{q}_0$ la distribution dans la limite continue prend la forme

$$P(t, \mathbf{x}) = \frac{1}{(2\pi)^N} \int d^N \kappa e^{-i\kappa \cdot \mathbf{x}} e^{-t\omega(\kappa)},$$

où l'intégrale sur κ n'est plus restreinte, avec

$$\begin{aligned} n\omega(\mathbf{k}) &= t\omega(\kappa), \\ \omega(\kappa) &= \frac{w_2}{2!} \kappa^2 + \frac{1}{4!} \varepsilon w_4 (\kappa^2)^2 + \frac{1}{4!} \varepsilon w'_4 \sum_i \kappa_i^4 + O(\kappa^6). \end{aligned}$$

Développant la transformée de Fourier en puissances de $\varepsilon = t/n$ on voit qu'à l'ordre ε des contributions apparaissent proportionnelles à $(\mathbf{x}^2)^2$ qui modifie la forme gaussienne, et $\sum_i x_i^4$ qui brise la symétrie de rotation.

Exemple. On pourra vérifier les propriétés démontrées dans cette section sur l'exemple de la marche au hasard avec mouvements limités aux proches voisins. Dans cet exemple $p(\mathbf{q})$ s'annule sauf si \mathbf{q} est un des vecteurs de coordonnées $(0, \dots, 0, \pm 1, 0, \dots, 0)$. Dans ce dernier cas $p(\mathbf{q})$ vaut $1/2N$. Donc

$$\tilde{p}(\mathbf{k}) = \frac{1}{N} \sum_{i=1}^N \cos k_i, \quad w(\mathbf{k}) = \mathbf{k}^2 / 2N + O(k^4),$$

où k_i est une composante du vecteur \mathbf{k} . On prendra aussi comme condition initiale un point de départ \mathbf{q}_0

$$P_0(\mathbf{q}) = \delta_{\mathbf{q}, \mathbf{q}_0} \Rightarrow \tilde{P}_0(\mathbf{k}) = e^{i\mathbf{k} \cdot \mathbf{q}_0}.$$

11 Transitions de phase et brisure spontanée de symétrie

Nous abordons maintenant l'étude des propriétés des transitions de phase continues ou du second ordre dans des systèmes classiques (non quantiques) du point de vue des symétries.

Nous supposons ici que la phénoménologie des transitions de phase dans les systèmes simples comme les transitions liquide–vapeur ou les transitions magnétiques est connue. Nous voulons seulement rappeler quelques subtilités de la notion de transition de phase du point de vue de la mécanique statistique.

Une première remarque importante est la suivante : les transitions de phase ne sont possibles que dans des systèmes de volume infini, c'est à dire des systèmes ayant un nombre infini de degrés de liberté. Ceci montre déjà que l'existence de transitions de phase n'est en soi pas quelque chose d'évident.

Une caractérisation simple et assez générale d'une transition de phase est dynamique. On appelle ici espace de phase l'ensemble des configurations d'un système. Par exemple si nous considérons le modèle d'Ising, un modèle de spins classiques sur réseau où chaque spin ne peut prendre que deux valeurs, alors l'espace de phase pour N spins contient 2^N éléments. Les transitions que nous allons considérer ont alors le caractère suivant : aussi longtemps que le volume du système est fini, tout élément de l'espace de phase est atteint au cours de l'évolution temporelle et ceci quelle que soit la température (si le système n'est pas discret comme le système de spins d'Ising tout élément doit être remplacé par tout élément de volume de l'espace des phases aussi petit soit-il). On dit que le système est *ergodique*. Du point de vue de dynamiques stochastiques cela signifie qu'il existe une probabilité finie de relier deux éléments quelconques. Si le système a alors un état d'équilibre thermodynamique, les moyennes temporelles sont identiques aux moyennes statiques calculées avec un poids de Boltzmann en sommant sur tout l'espace de phase.

Par contre dans la limite du volume infini (à densité fixée pour un système de particules), suivant les valeurs de la température, le système reste ergodique ou au contraire subit une brisure d'ergodicité. Dans ce dernier cas l'espace de phase se décompose en des sous-ensembles disjoints. Quand le système est préparé initialement dans un de ces sous-ensembles il y reste. Par exemple pour un système de type Ising en dessous de la température critique, les deux sous-ensembles correspondent aux deux valeurs possibles de l'aimantation spontanée.

Du point de vue statique les régions à une ou plusieurs phases se distinguent aussi par des sensibilités différentes aux conditions aux limites. La région à une seule phase ne garde pas trace de la manière spécifique dont la limite thermodynamique, c'est à dire de volume infini, est atteinte. Il n'en est pas de même dans la région à plusieurs phases où la phase atteinte en dépend.

Pour les systèmes simples que nous voulons étudier, il est possible de trouver des observables locales dont les valeurs discriminent entre les phases. Nous appelons une de ces observables *paramètre d'ordre*. C'est par exemple le spin dans le cas de systèmes ferromagnétiques.

De plus dans ces systèmes la transition de phase correspond aussi à une

brisure spontanée de symétrie. Pour expliquer ce concept nous allons utiliser l'exemple du modèle d'Ising.

L'énergie de configuration du modèle d'Ising ne change pas quand on change le signe de tous les spins (symétrie \mathbb{Z}_2). On s'attend donc à ce que la valeur moyenne du spin soit toujours nulle.

Ajoutons maintenant à l'énergie de configuration un terme qui brise explicitement la symétrie du système (un terme de champ magnétique pour un système ferromagnétique), prenons la limite de volume infini, et faisons tendre ensuite l'amplitude du terme de brisure vers zéro. Suivant la température deux cas peuvent se présenter. Dans la phase désordonnée la symétrie est restaurée en ce sens que toutes les fonctions de corrélation ont la symétrie du système. Dans la région où plusieurs phases sont possibles (phases dites ordonnées) ce n'est pas le cas. On parle alors de brisure spontanée de symétrie. Dans cette région de paramètre la limite thermodynamique et la limite de brisure nulle ne commutent pas, à la différence de la phase symétrique. Dans le cas des spins on trouve une valeur du spin non-nulle, c'est à dire une aimantation spontanée. Le signe de l'aimantation spontanée dépend du signe du champ magnétique. Du point de vue du poids de Boltzmann, pour obtenir une de ces limites, il ne faut plus sommer sur toutes les configurations, mais seulement sur celles dont la valeur moyenne du spin est égale à l'aimantation spontanée. Comme le système n'est plus ergodique, c'est bien cette somme qui reproduit les moyennes temporelles.

Propriété d'amas. Dans la région à phase unique, pour des systèmes avec interactions de courte portée (par exemple avec décroissance exponentielle) on peut définir des fonctions de corrélation connexes (l'équivalent des cumulants d'une distribution) qui ont la propriété suivante. Considérons la corrélation entre deux observables non triviales localisées dans deux régions spatiales disjointes. Alors cette corrélation décroît exponentiellement quand la séparation entre les deux régions tend vers l'infini : cette propriété est aussi appelée propriété d'*amas*. On peut relier la notion de paramètre d'ordre à la propriété d'amas dans la phase de symétrie brisée à basse température. On appelle *longueur de corrélation* l'inverse du plus petit taux de décroissance.

11.1 Brisure spontanée de symétrie

Nous allons essayer maintenant d'illustrer cette notion de brisure spontanée de symétrie en examinant d'autres situations physiques.

En mécanique classique la notion de brisure spontanée de symétrie est banale. Il suffit de prendre l'exemple d'une particule dans un potentiel $V(x)$ régulier (c'est à dire autant de fois dérivable que nécessaire) qui a un axe de symétrie

$$V(x) = V(-x) \quad \Rightarrow \quad V'(0) = 0.$$

Deux cas peuvent se présenter, ou bien $x = 0$ est un minimum du potentiel et $x = 0$ est une position d'équilibre stable, ou au contraire $x = 0$ est un maximum relatif. Dans cette situation si le potentiel a un autre minimum $x = a \neq 0$ il en a deux $x = \pm a$ qui sont des positions d'équilibre stable. Dans le premier cas non seulement le potentiel est symétrique, mais la position de la particule au repos

est symétrique. Dans le deuxième cas, bien que le potentiel soit symétrique, la position ne l'est pas. On appelle cette situation une *brisure spontanée* de symétrie. Bien entendu il est également possible que le potentiel ait à la fois des minima symétriques et non symétriques et dans ces conditions les deux situations peuvent être réalisées.

Cette analyse est valable pour tout groupe de symétrie. Par exemple si \mathbf{x} est un point de l'espace à trois dimensions et V un potentiel radial, c'est à dire ne dépendant que de $|\mathbf{x}|$ dans la situation de symétrie brisée il existe une sphère de minima $|\mathbf{x}| = a$.

Exprimons ces propriétés en langage plus formel. Soit G le groupe de symétrie du potentiel, et supposons que $\mathbf{x} = 0$ est le centre de symétrie. Alors

$$V(\mathbf{g}\mathbf{x}) = V(\mathbf{x}) \quad \forall \mathbf{g} \in G.$$

Si une position d'équilibre stable \mathbf{x}_m satisfait

$$\mathbf{g}\mathbf{x}_m = \mathbf{x}_m \quad \forall \mathbf{g} \in G,$$

c'est à dire est invariante par le groupe, la symétrie n'est pas brisée. Si elle n'est pas invariante l'ensemble des points

$$\mathbf{g}\mathbf{x}_m \quad \forall \mathbf{g} \in G,$$

sont des positions d'équilibre stable et la symétrie est brisée spontanément.

Dans cette situation considérons l'ensemble H des éléments de G tels que

$$\mathbf{h}\mathbf{x}_m = \mathbf{x}_m \quad \forall \mathbf{h} \in H.$$

Les éléments de H forment un groupe, sous-groupe de G , et il existe une bijection entre l'ensemble des points $\mathbf{g}\mathbf{x}_m$ appelé orbite du groupe et l'ensemble quotient G/H .

Groupes discrets et continus. Notons ici une différence très importante entre les groupes discrets et continus. Dans le cas de la symétrie de réflexion \mathbb{Z}_2 , il faut fournir à la particule une énergie finie, au moins égale à la hauteur de la barrière de potentiel qui les sépare, pour qu'elle puisse passer par les deux minima. Dans le cas d'une symétrie continue au contraire on voit intuitivement qu'une énergie infinitésimale suffit. Vérifions le de façon plus précise sur un exemple.

Considérons de nouveau l'exemple d'un potentiel radial dans le plan. On peut décrire le plan par des coordonnées polaires r, θ . L'action classique s'écrit alors

$$\mathcal{A} = \int dt \left[\frac{1}{2}m \left(\dot{r}^2 + r^2\dot{\theta}^2 \right) - V(r) \right].$$

Soit $r = a$ le cercle de minima. Posons $r(t) = a + \varepsilon(t)$, $\varepsilon(0) \ll 1$, et développons en $\varepsilon(t)$. On vérifie que $\dot{\theta}$ est d'ordre ε et donc :

$$\mathcal{A} = \int dt \left[\frac{1}{2}m \left(\dot{\varepsilon}^2 + a^2\dot{\theta}^2 \right) - \frac{1}{2}\varepsilon^2 V''(a) \right] + O(\varepsilon^3).$$

La variable angulaire θ satisfait l'équation du mouvement libre et donc

$$\theta(t) = \omega t.$$

Malgré la condition que la vitesse angulaire ω est d'ordre ε , $\theta(t)$ augmente indéfiniment. Au contraire le mouvement de ε correspond à une force de rappel harmonique qui conduit à un mouvement oscillatoire autour de $\varepsilon = 0$:

$$\varepsilon(t) \propto \varepsilon(0) \cos \left(t \sqrt{V''(a)/m} \right).$$

Mécanique quantique. En mécanique quantique la situation est très différente à cause de l'effet tunnel. Pour un nombre fini de degrés de liberté il n'y a pas de brisure spontanée de symétrie. La fonction d'onde du fondamental est symétrique, et donc la particule a une probabilité de présence égale dans tous les minima dégénérés. On voit donc que la notion de brisure spontanée de symétrie a des aspects non triviaux. Seuls des systèmes macroscopiques, parce qu'ils ont un très grand nombre de degrés de liberté, se comportent de façon classique parce que l'effet tunnel est négligeable.

Mécanique statistique. En mécanique statistique aussi la situation est subtile à cause non plus des fluctuations quantiques mais thermiques. Examinons l'exemple le plus simple, le modèle d'Ising ferromagnétique avec interaction de proche voisin sur un réseau. Pour se convaincre de l'existence d'une transition on peut se placer à basse température, car l'aimantation spontanée ne peut que décroître à plus haute température. Dans ces conditions l'analyse est purement énergétique. Partant d'une configuration où les spins prennent presque tous la valeur $+1$, il faut déterminer la probabilité pour des fluctuations thermiques engendrent une configuration avec les spins retournés.

La probabilité qu'une fluctuation thermique crée une goutte de spins -1 dans un environnement de spins $+1$ est de la forme $e^{-\beta \Delta E}$, où ΔE est la différence des énergies entre la configuration initiale et la configuration avec la goutte. La différence ΔE est entièrement due à la surface de la goutte où les spins voisins sont opposés, et donc proportionnelle à l'aire de la surface. Si R est la dimension linéaire de la goutte, en dimension d'espace d l'aire est proportionnelle à R^{d-1} . Pour que tous les spins soient retournés il faut qu'une goutte de taille infinie puisse être créée et donc que la probabilité reste finie quand $R \rightarrow \infty$. On en conclut qu'à une dimension cela est possible et il ne peut pas donc y avoir de transition de phase. Au contraire une transition de phase est possible en dimension plus grande que un. Et en effet la solution exacte du modèle d'Ising à deux dimensions exhibe une transition de phase.

L'exemple de la mécanique classique suggère qu'il devrait être plus difficile de briser une symétrie continue qu'une symétrie discrète. En effet au lieu de faire tourner brutalement tous les spins d'un angle α correspondant à deux directions de l'aimantation spontanée, on peut les faire tourner continûment pour diminuer la variation de l'énergie. Prenant le cas de spins classiques à N composantes de longueur unité $\mathbf{S}_i^2 = 1$, i dénotant un site du réseau, avec une

interaction ferromagnétique de proche voisin sous forme d'un produit scalaire $\mathbf{S}_i \mathbf{S}_j = \cos(\theta_{ij})$, où θ_{ij} est l'angle entre les spins. Faisons varier linéairement l'angle θ définissant l'orientation des spins entre une direction initiale d'angle $\theta = 0$ au centre d'une sphère, cercle... et une direction au bord d'angle α . Nous appelons R le rayon de la sphère et r la distance au centre :

$$\theta(r) = \alpha r / R.$$

Comme la variation d'angle d'un site au voisin est faible pour R grand, la variation de l'énergie par site est proportionnelle à

$$1 - \cos(\theta_i - \theta_j) \sim \frac{1}{2}(\theta_i - \theta_j)^2 \sim \frac{1}{2}(n_{ij} \cdot \nabla \theta)^2,$$

où n_{ij} est le vecteur joignant les deux sites i, j . Sommant sur toutes les directions d'un réseau cubique on reconstitue le carré du vecteur gradient. La variation ΔE de l'énergie totale en dimension d est alors proportionnelle à :

$$\int_{|r| \leq R} d^d r \left(\frac{\partial \theta}{\partial r} \right)^2 \propto \alpha^2 R^{d-2}.$$

On voit donc qu'une transition de phase avec aimantation spontanée c'est à dire brisure spontanée de symétrie n'est possible qu'en dimension plus grande que deux, un résultat qui peut être démontré plus rigoureusement pour tout système avec interaction de courte portée.

11.2 Potentiel thermodynamique. Transformation de Legendre

L'étude de l'approximation de champ moyen pour de nombreux systèmes montre qu'il est plus simple de discuter les transitions de phase en terme de potentiel thermodynamique que l'énergie libre.

À partir de maintenant nous nous restreignons aux transitions de phase de systèmes ferromagnétiques. Cette restriction est partiellement une restriction de langage. En effet, comme conséquence de l'*universalité* des phénomènes critiques, une propriété que nous n'allons ici qu'évoquer, les résultats universels obtenus dans le cadre ferromagnétique s'appliquent à de nombreux autres systèmes physiques qui n'ont rien de magnétiques, comme la transition liquide-vapeur, les transitions de démixtion des mélanges binaires, la transition superfluide de l'Hélium *etc...* À cette liste il convient d'ajouter, un problème qui ne semble pas à première vue relever des phénomènes critiques, les propriétés statistiques des polymères, ou du point de vue théorique des chemins aléatoires sans intersection sur réseau.

Dans l'exemple des systèmes ferromagnétiques on comprend facilement qu'il est préférable de travailler à aimantation fixée plutôt qu'à champ magnétique fixé. En effet dans la phase de symétrie brisée (ou phase ordonnée) en champ nul l'aimantation peut avoir plusieurs valeurs, et l'état du système est donc déterminé par l'aimantation et non le champ.

Nous rappelons donc les définitions correspondantes.

Nous notons S_i un spin classique sur le site i d'un réseau. Nous appelons $\mathcal{E}(S)$ l'énergie de configuration des spins en champ nul, β l'inverse de la température et $d\rho(S)$ la distribution de spin (normalisée) qui pondère les configurations de spin à chaque site. La fonction de partition dans un champ magnétique H s'écrit alors

$$Z(H) = \int \prod_i d\rho(S_i) \exp \left[-\beta \mathcal{E}(S) + \sum_i H S_i \right], \quad (11.1)$$

où l'échelle de champ magnétique inclut un facteur de température ($\beta H \mapsto H$).

L'énergie libre associée $W(H)$ par unité de volume (notre définition de l'énergie libre diffère ici et plus tard par un facteur de température, pour nous sans importance, des définitions usuelles) est

$$W(H) = \Omega^{-1} \ln Z(H),$$

où Ω est le volume. Pour travailler à aimantation M fixée on passe de l'énergie libre au potentiel thermodynamique $G(M)$. Les deux fonctions $W(H)$ et $G(M)$ sont reliées par une transformation de Legendre.

Transformation de Legendre. Le concept de transformée de Legendre apparaît dans plusieurs domaines de la physique, comme la mécanique analytique (reliant lagrangien et hamiltonien) et la mécanique statistique.

Le potentiel thermodynamique $G(M)$ fonction de l'aimantation M est relié à l'énergie libre $W(H)$ par la transformation

$$W(H) + G(M) = HM \quad (11.2a)$$

$$M = \frac{\partial W(H)}{\partial H}. \quad (11.2b)$$

La transformation de Legendre est localement inversible aussi longtemps que

$$\frac{\partial^2 W}{(\partial H)^2} \neq 0.$$

Appelant S le spin total

$$S = \sum_i S_i,$$

on trouve que la condition est satisfaite puisque

$$\frac{\partial^2 W}{(\partial H)^2} = \left\langle (S - \langle S \rangle)^2 \right\rangle > 0,$$

où valeur moyenne $\langle \bullet \rangle$ signifie valeur moyenne en présence du champ H .

On vérifie alors que la relation est symétrique. En effet la première relation entraîne

$$\frac{\partial G(M)}{\partial M} = H + \frac{\partial H}{\partial M} \frac{\partial}{\partial H} [HM - W(H)].$$

Utilisant (11.2b) on trouve que le deuxième terme s'annule. Donc

$$H = \frac{\partial G(M)}{\partial M}. \quad (11.3)$$

L'équation (11.3) est en général la forme la plus simple de l'équation d'état.

Dérivées secondes. La susceptibilité magnétique χ est définie par

$$\chi = \frac{\partial M}{\partial H} = \frac{\partial^2 W}{\partial H \partial H}.$$

De la transformation de Legendre on déduit une expression de χ impliquant le potentiel thermodynamique. Elle se déduit de la relation entre dérivées secondes obtenue en dérivant l'équation (11.2b) par rapport à M et utilisant l'équation (11.3)

$$\frac{\partial^2 W}{\partial H \partial H} \frac{\partial^2 G}{\partial M \partial M} = 1. \quad (11.4)$$

Donc $G''(M)$ est aussi positif en général. C'est en particulier le cas sur un réseau fini. Dans la limite thermodynamique au contraire $W''(H)$, qui est la susceptibilité magnétique, peut diverger en champ nul, et dans ce cas $G''(M)$ s'annule.

En conséquence sur un réseau fini la transformation de Legendre est toujours localement bijective. Dans la limite thermodynamique la question devient plus subtile. Si $W''(H)$ qui est la susceptibilité magnétique, diverge, $G''(M)$ s'annule et il peut y avoir une bifurcation entre une situation où à H donné il n'y a qu'une valeur de l'aimantation et une situation où il y a plusieurs solutions. On trouve alors une transition de phase.

Paramètre d'ordre à plusieurs composantes. Quand le paramètre d'ordre a N composantes (cas par exemple du spin classique à trois composantes) le formalisme précédent se généralise immédiatement :

$$W(H) + G(M) = \sum_{\alpha} H_{\alpha} M_{\alpha}, \quad M_{\alpha} = \frac{\partial W(H)}{\partial H_{\alpha}}. \quad (11.5)$$

On vérifie de nouveau que cette relation est symétrique et permet d'exprimer l'équation d'état en fonction du potentiel thermodynamique :

$$H_{\alpha} = \frac{\partial G(M)}{\partial M_{\alpha}}. \quad (11.6)$$

Il existe alors une matrice $\chi_{\alpha\beta}$ de susceptibilités magnétiques

$$\chi_{\alpha\beta} = \frac{\partial M_{\alpha}}{\partial H_{\beta}} = \frac{\partial^2 W(H)}{\partial H_{\alpha} \partial H_{\beta}}.$$

Cette matrice symétrique est positive (toutes les valeurs propres sont positives). L'inverse de la matrice $\chi_{\alpha\beta}$ est la matrice des dérivées partielles secondes de $G(M)$

$$\sum_{\gamma} \chi_{\alpha\gamma} \frac{\partial^2 G(M)}{\partial M_{\gamma} \partial M_{\beta}} = \delta_{\alpha\beta}.$$

11.3 Symétries continues et modes de Goldstone

Si la variable initiale de spin S_i est un vecteur à N composantes, et si à la fois l'interaction et la distribution de spin ont une symétrie continue, à toute température dans la phase de symétrie brisée certaines susceptibilités divergent. Les modes correspondant sont appelés modes de Goldstone.

Nous illustrons ces remarques par l'exemple de la symétrie $O(N)$.

Modes de Goldstone. Rappelons d'abord brièvement dans ce contexte la généralité du concept de modes de Goldstone, dans la phase ordonnée, en champ nul.

La symétrie orthogonale $O(N)$ implique que le potentiel thermodynamique $G(M)$ n'est qu'une fonction du carré du vecteur aimantation \mathbf{M} . Alors la relation entre le champ magnétique de composantes H_α et aimantation de composantes M_α prend la forme :

$$G(M) = \phi(M^2/2), \Rightarrow H_\alpha = M_\alpha \phi'(M^2/2). \quad (11.7)$$

L'inverse de la matrice des susceptibilités magnétiques est donné par :

$$G_{\alpha\beta}^{(2)} = \frac{\partial^2 G}{\partial M_\alpha \partial M_\beta} = M_\alpha M_\beta \phi''(M^2/2) + \delta_{\alpha\beta} \phi'(M^2/2). \quad (11.8)$$

La matrice $G_{\alpha\beta}^{(2)}$ a deux sous-espaces propres correspondant au vecteur M_α (donc de dimension un) et aux vecteurs X_α orthogonaux à M_α (donc de dimension $N - 1$), avec les valeurs propres $G_L^{(2)}$, $G_T^{(2)}$ respectivement :

$$\begin{aligned} \sum_{\beta} G_{\alpha\beta}^{(2)} M_\beta &= M_\alpha [M^2 \phi''(M^2/2) + \phi'(M^2/2)] \\ \sum_{\beta} G_{\alpha\beta}^{(2)} X_\beta &= X_\alpha \phi'(M^2/2). \end{aligned}$$

Les deux valeurs propres correspondant à ces sous-espaces sont respectivement

$$G_L^{(2)} = M^2 \phi''(M^2/2) + \phi'(M^2/2), \quad (11.9)$$

$$G_T^{(2)} = \phi'(M^2/2). \quad (11.10)$$

Les valeurs propres de la matrice des susceptibilités sont les inverses de celles de $G^{(2)}$. Utilisant alors l'équation (11.7), nous obtenons pour l'inverse de la susceptibilité transverse χ_T

$$G_T^{(2)} = \chi_T^{-1} = H/M. \quad (11.11)$$

Dans la phase ordonnée quand H tend vers zéro M tend vers l'aimantation spontanée qui est non nulle, et donc la susceptibilité transverse diverge. Ce comportement signale la présence de $N - 1$ modes de Goldstone (de masse nulle au sens de la physique des particules).

Une autre façon de comprendre ce résultat est la suivante. En champ nul au-dessous de T_c l'aimantation est non nulle, alors que le potentiel est symétrique. Cela signifie que le potentiel a un minimum pour $\mathbf{M} \neq 0$, et donc par symétrie une sphère de minima $|\mathbf{M}| = M$. Si nous prenons un minimum et faisons une rotation (une transformation orthogonale) nous obtenons un autre minimum. Une rotation infinitésimale correspond à l'addition à \mathbf{M} d'un vecteur tangent à la sphère et donc orthogonal à \mathbf{M} . Appelons \mathbf{X} un tel vecteur, $|\mathbf{X}| \ll 1$. Alors, développant la condition d'extremum au premier ordre en \mathbf{X} , nous obtenons la condition

$$0 = \frac{\partial G(\mathbf{M} + \mathbf{X})}{\partial M_\alpha} = \sum_\beta \frac{\partial^2 G(\mathbf{M})}{\partial M_\alpha \partial M_\beta} X_\beta.$$

Nous retrouvons les $N - 1$ vecteurs propres de valeur propre nulle, correspondant aux modes de Goldstone.

11.4 Modèle quasi-gaussien ou théorie de Landau : universalité

Dans ce qui suit nous ne nous intéressons qu'à la classe spéciale des transitions de phase continues ou du second ordre, au voisinage de la température de transition T_c . Ces transitions sont caractérisées par la propriété que le paramètre d'ordre s'annule continûment à la température de transition, et donc près de T_c et en champ magnétique faible l'aimantation est petite. De plus dans ces conditions une échelle de distance, grande par rapport à l'échelle des longueurs microscopiques (portée des forces, maille de réseau), est engendrée dynamiquement. Cette échelle, qui peut être caractérisée par la longueur de corrélation, diverge au point de transition. L'existence de cette nouvelle échelle entraîne l'apparition d'une physique de longue distance ou macroscopique non triviale, dont on peut s'attendre qu'elle soit peu sensible aux détails des modèles microscopiques et qu'on s'efforce de caractériser.

La description la plus ancienne et la plus simple des transitions de phase est basée sur des approximations de type champ moyen (TCM). Les résultats obtenus par l'approximation de champ moyen et les principes sous-jacents ont été formalisés par la théorie de Landau, qui repose en particulier sur le concept de découplage des différentes échelles de physique : on suppose qu'un petit nombre de degrés de liberté macroscopiques peuvent remplacer l'infinité de degrés de liberté microscopiques, les effets résiduels pouvant être traités de façon perturbative. La théorie de Landau postule que le potentiel thermodynamique général est une fonction régulière du paramètre d'ordre et des autres variables thermodynamiques, dont la forme générale est déterminée par les propriétés de symétrie. La théorie résultante a des propriétés de longue distance remarquablement *universelles*, c'est à dire indépendantes des détails des interactions microscopiques et, dans une large mesure, de la géométrie, de la dimension d'espace. En particulier le comportement des quantités thermodynamiques dans le voisinage d'une transition, en particulier leurs singularités comme fonction de la température à la température critique, est *universel*.

L'approximation de champ moyen ou la théorie de Landau peuvent être qualifiées de modèles quasi-gaussiens. En réalité on montre qu'elles ne peuvent être

exactes qu'en dimension d'espace plus grande que quatre, et donc pas pour des systèmes réalistes. Ce ne sont que des approximations car les différentes échelles ne se découplent pas, un phénomène dont l'étude nécessite un outil entièrement nouveau : le *groupe de renormalisation*. Dans ce qui suit nous discutons néanmoins les transitions de phase des systèmes ferromagnétiques dans le cadre de la théorie de Landau, car elle constitue une approche qualitative très raisonnable.

Le potentiel thermodynamique près de T_c . Comme l'aimantation M est faible et qu'on suppose le potentiel $G(M)$ régulier, on peut le développer en puissances de l'aimantation. On vérifie que $G(M)$ a les mêmes symétries que l'énergie de configuration.

Symétrie \mathbb{Z}_2 . Nous allons d'abord discuter le potentiel thermodynamique pour des systèmes ayant une symétrie de réflexion \mathbb{Z}_2 comme le modèle d'Ising, où la symétrie correspond au renversement de tous les spins. Dans ce cas $G(M)$ est une fonction paire de l'aimantation

$$G(M) = G(-M) \Rightarrow G(M) = \frac{1}{2}g_2M^2 + \frac{1}{4}g_4M^4 + O(M^6).$$

Dans la phase de haute température $M = 0$ est le minimum absolu de $G(M)$. Ceci implique $g_2 > 0$. Au-dessous de T_c mais près de T_c on veut que le potentiel ait un minimum absolu proche de $M = 0$. Ceci impose $g_2 < 0$ et $g_4 > 0$. En effet $g_2 > 0$ conduit à une transition discontinue, M passant de la valeur nulle à une valeur finie à T_c . Il en est de même pour $g_4 < 0$.

Puisque g_2 change de signe à T_c et qu'on a supposé la régularité dans toutes les variables thermodynamiques, on en déduit pour $|T - T_c| \ll 1$,

$$g_2(T) \sim a_2(T - T_c), \quad a_2 > 0.$$

L'équation d'état $H = G'(M)$ prend alors la forme

$$H/M = a_2(T - T_c) + g_4M^2, \quad (11.12)$$

pour $|M| \ll 1$, $|T - T_c| \ll 1$. Par ailleurs la dépendance de g_4 en T peut être négligée.

Nous avons donc construit un modèle très simple d'équation d'état qui est tout à fait universel dans la classe des modèles avec symétrie \mathbb{Z}_2 et interactions à courte portée.

On peut en déduire quelques résultats simples.

Aimantation spontanée. Pour $T < T_c$, en champ nul, apparaît une aimantation spontanée M :

$$a_2(T - T_c) + g_4M^2 = 0 \Rightarrow M \sim \pm [a_2/g_4]^{1/2} (T_c - T)^{1/2} \quad \text{pour } |T - T_c| \ll 1. \quad (11.13)$$

Près de la température critique T_c , l'aimantation a donc un comportement en loi de puissances avec un exposant magnétique β :

$$M \propto (T_c - T)^\beta, \quad \beta = 1/2. \quad (11.14)$$

La valeur $\beta = 1/2$ est aussi appelée valeur “classique” de l’exposant.

Susceptibilité magnétique. L’inverse de la susceptibilité magnétique χ est donnée par

$$\chi^{-1} = \left(\frac{\partial M}{\partial H} \right)^{-1} = \frac{\partial H}{\partial M} = a_2(T - T_c) + 3g_4 M^2.$$

En champ nul nous trouvons :

$$\begin{aligned} \chi_+^{-1} &= a_2(T - T_c) & T > T_c, \\ \chi_-^{-1} &= 2a_2(T_c - T) & T < T_c, \end{aligned} \quad (11.15)$$

où l’équation (11.13) a été utilisée en dessous de T_c . La susceptibilité magnétique diverge donc à T_c avec des exposants de susceptibilité γ, γ' :

$$\begin{aligned} \chi_+ &\sim C_+ (T - T_c)^{-\gamma}, & \gamma &= 1, \\ \chi_- &\sim C_- (T_c - T)^{-\gamma'}, & \gamma' &= 1, \end{aligned} \quad (11.16)$$

et :

$$C_+ / C_- = 2. \quad (11.17)$$

Equation d’état. Revenons à l’équation générale (11.12). À T_c on trouve

$$H \propto M^3. \quad (11.18)$$

On définit généralement $H \propto M^\delta$ et l’exposant critique δ a donc la valeur “classique”

$$\delta = 3. \quad (11.19)$$

Pour $H, T - T_c, M$ faibles, Widom a conjecturé l’existence d’une forme d’échelle universelle :

$$H = M^\delta f\left((T - T_c)M^{-1/\beta}\right), \quad (11.20)$$

ce qui signifie que le rapport H/M^δ n’est pas fonction des variables T et M indépendamment, mais seulement de la combinaison $(T - T_c)/M^{1/\beta}$.

Dans le cadre de la théorie de Landau on trouve bien cette forme d’échelle, et après un changement d’échelle de champ, température et aimantation, la fonction $f(x)$ peut s’écrire :

$$f(x) = 1 + x. \quad (11.21)$$

On vérifie facilement que les termes omis d’ordre M^6 et $M^4(T - T_c)$ n’induisent bien que des corrections à cette forme d’échelle.

La valeur $x = -1$, où H s’annule avec M non nul, correspond à la courbe de coexistence dans la région à deux phases, et redonne l’aimantation spontanée.

Les résultats précédents sont universels. En particulier ils ne dépendent pas de la dimension de l’espace. En réalité ils ne peuvent pas être vrais en général car

ils contredisent par exemple les résultats exacts du modèle d'Ising en dimension deux.

Chaleur spécifique. En champ nul la dérivée de l'énergie libre par unité de volume par rapport à β , l'inverse de la température, donne l'opposé de l'énergie moyenne. La dérivée par rapport à T rajoute un facteur sans importance à T_c . Comme conséquence de la stationnarité on trouve pour tout paramètre, et donc cela s'applique à T , la relation

$$\left. \frac{\partial \Gamma(M)}{\partial T} \right|_{M \text{ fixé}} + \left. \frac{\partial \mathcal{W}(H)}{\partial T} \right|_{H \text{ fixé}} = 0.$$

Nous en déduisons

$$\left. \frac{\partial \mathcal{W}(H)}{\partial T} \right|_{H=0} = -g'_0(T) - a_2 M^2 (H=0).$$

On trouve qu'au dessus de T_c l'énergie moyenne s'annule et au dessous de T_c elle est proportionnelle au carré de l'aimantation spontanée. La dérivée de l'énergie moyenne par rapport à la température donne la chaleur spécifique. Calculant la dérivée par rapport à T à T_c , nous obtenons un résultat proportionnel à la chaleur spécifique \mathcal{C}

$$\mathcal{C}(T \rightarrow T_{c+}) = -g''_0(T_c), \quad \mathcal{C}(T \rightarrow T_{c-}) = -g''_0(T_c) + a_2^2/g_4. \quad (11.22)$$

Nous trouvons que dans la théorie de Landau la chaleur spécifique a un saut non universel.

Symétrie $O(N)$. Nous considérons maintenant un spin à N composantes et un système qui a une symétrie de rotation-réflexion $O(N)$. Le champ magnétique et l'aimantation sont aussi des vecteurs à N composantes, et on vérifie facilement que $G(M)$ ne dépend que du module de l'aimantation, une propriété que nous avons déjà exploitée en section 11.3. Les arguments donnés précédemment se généralisent simplement, et l'on peut écrire

$$G(\mathbf{M}) = \frac{1}{2}g_2 M^2 + \frac{1}{4}g_4 M^4 + O(M^6),$$

où M est la longueur du vecteur \mathbf{M} . Nous négligeons de nouveau les termes d'ordre M^6 ou supérieur et la dépendance de g_4 en T . L'équation d'état s'en déduit

$$\mathbf{H} = \mathbf{M} (a_2(T - T_c) + g_4 M^2).$$

Si nous prenons le module des deux membres nous trouvons une équation tout à fait analogue au cas \mathbb{Z}_2 , avec la même loi d'échelle et les mêmes exposants. Les exposants critiques de la théorie de Landau sont non seulement indépendants de la dimension d'espace mais aussi du groupe de symétrie.

Simplement dans la phase ordonnée l'aimantation a une longueur fixée mais une direction arbitraire

$$M_\alpha = [a_2/g_4]^{1/2} (T_c - T)^{1/2} u_\alpha, \quad \sum_\alpha u_\alpha^2 = 1.$$

Une différence apparaît pour les susceptibilités. L'inverse de la matrice des susceptibilité est donnée par

$$(\chi)_{\alpha\beta}^{-1} = \frac{\partial H_\alpha}{\partial M_\beta} = \delta_{\alpha\beta} (a_2(T - T_c) + g_4 M^2) + 2g_4 M_\alpha M_\beta.$$

Dans le coefficient de $\delta_{\alpha\beta}$ nous reconnaissons bien H/M . La matrice $(\chi)_{\alpha\beta}^{-1}$ a deux valeurs propres correspondant aux vecteurs M_α et

$$\sum_{\beta} (\chi)_{\alpha\beta}^{-1} M_\beta = (H/M + 2g_4 M^2) M_\alpha$$

$$\chi_L = 1/(H/M + 2g_4 M^2),$$

orthogonaux à M_α : $\mathbf{X} \cdot \mathbf{M} = 0$,

$$\sum_{\beta} (\chi)_{\alpha\beta}^{-1} X_\beta = (H/M) X_\alpha$$

$$\chi_T = M/H,$$

en accord avec le résultat général.

La susceptibilité transverse diverge en champ nul pour tout $T < T_c$. La susceptibilité longitudinale ne diverge qu'à T_c avec la même loi que dans le cas \mathbb{Z}_2

$$\chi_L = \frac{1}{2a_2(T_c - T)}.$$

Autre exemple : la symétrie cubique. Nous supposons maintenant que le modèle n'a qu'une symétrie cubique. Ceci peut être la conséquence d'une brisure de la symétrie $O(N)$ par l'interaction entre les spins et le réseau. Dans l'étude de la marche au hasard nous avons déjà déterminé la forme des invariants quadratiques et quartiques :

$$G(\mathbf{M}) = \frac{1}{2} a_2 (T - T_c) \sum_{\alpha} M_\alpha^2 + \frac{1}{4} g_4 \left(\sum_{\alpha} M_\alpha^2 \right)^2 + \frac{1}{4} h_4 \sum_{\alpha} M_\alpha^4.$$

Pour que la transition soit du second ordre il faut que le potentiel quartique soit borné inférieurement. Cela donne les deux conditions obtenues la première en choisissant $M_\alpha = 0$ pour $\alpha \neq 1$, la deuxième $M_\alpha = M$ quel que soit α :

$$g_4 + h_4 > 0, \quad N g_4 + h_4 > 0.$$

Pour $T > T_c$ fixé et champ \mathbf{H} faible l'aimantation est faible et le terme quartique est négligeable et la symétrie apparente est une symétrie de rotation $O(N)$

$$H_\alpha \sim a_2 (T - T_c) M_\alpha.$$

De façon générale

$$H_\alpha = a_2(T - T_c)M_\alpha + g_4M_\alpha\mathbf{M}^2 + h_4M_\alpha^3.$$

En dessous de T_c l'aimantation spontanée est donnée par l'équation

$$M_\alpha = 0 \quad \text{ou} \quad a_2(T - T_c) + g_4 \sum_{\beta} M_\beta^2 + h_4M_\alpha^2 = 0.$$

Soit $1 \leq m \leq N$ le nombre de composantes non nulles. Elles sont telles que

$$M_\alpha = \mu\epsilon_\alpha, \quad \epsilon = \pm 1.$$

Reportant dans l'équation on trouve

$$a_2(T - T_c) + (mg_4 + h_4)\mu^2 = 0,$$

et donc

$$M_\alpha = \epsilon_\alpha \sqrt{a_2/(mg_4 + h_4)} \sqrt{T_c - T}.$$

Notons que $mg_4 + h_4$ est bien positif car

$$mg_4 + h_4 = \frac{N - m}{N - 1}(g_4 + h_4) + \frac{m - 1}{N - 1}(Ng_4 + h_4) > 0.$$

Il reste à trouver les solutions qui minimisent le potentiel thermodynamique. On trouve

$$G(M) = -\frac{1}{4}a_2^2(T - T_c)^2 \frac{m}{mg_4 + h_4}.$$

La dérivée par rapport à m est proportionnelle à $-h_4$. Si $h_4 > 0$ le minimum est $m = N$. Toutes les composantes de \mathbf{M} sont égales au signe près. Il ya 2^N minima équivalents. On peut choisir par exemple $M_\alpha = \mu$ pour tout α et on voit que le sous-groupe non brisé est le groupe des permutations qui laisse ce vecteur invariant.

Si $h_4 < 0$ (et alors $|h_4| < g_4$), c'est le contraire et $m = 1$. Il y a $2N$ minima équivalents. Le sous-groupe non brisé est le groupe cubique à $N - 1$ dimensions. On peut choisir la première composante non-nulle.

Ces résultats sont bien une réflexion de la symétrie cubique, toutes les valeurs possibles de l'aimantation spontanée sont déduites de l'une d'entre elles par l'action du groupe cubique.

Matrice des susceptibilités. Au-dessus de T_c la matrice des susceptibilités est proportionnelle à la matrice unité

$$\chi_{\alpha\beta} \sim \frac{\delta_{\alpha\beta}}{a_2(T - T_c)},$$

avec donc la valeur propre unique $1/a_2(T - T_c)$.

En dessous de T_c on trouve

$$\chi_{\alpha\beta}^{-1} = \delta_{\alpha\beta} (a_2(T - T_c) + g_4\mathbf{M}^2 + 3h_4M_\alpha^2) + 2g_4M_\alpha M_\beta.$$

Nous devons maintenant distinguer les deux cas.

$h_4 > 0$:

Alors

$$M_\alpha^2 = \mu^2 = \frac{a_2(T_c - T)}{Ng_4 + h_4},$$

et utilisant l'équation donnant μ ,

$$\chi_{\alpha\beta}^{-1} = 2h_4\mu^2\delta_{\alpha\beta} + 2g_4M_\alpha M_\beta.$$

De nouveau on trouve la somme d'une matrice proportionnelle à la matrice unité et d'un projecteur sur \mathbf{M} , comme dans le cas invariant par rotation. Cette matrice a deux valeurs propres correspondant au vecteur aimantation et aux vecteurs perpendiculaires. Appelant de nouveau χ_L et χ_T les valeurs de la matrice des susceptibilités nous trouvons

$$\chi_L = \frac{1}{2a_2(T_c - T)},$$

et

$$\chi_T = \frac{Ng_4 + h_4}{2a_2h_4(T_c - T)}.$$

$h_4 < 0$:

Dans ce cas la matrice $\chi_{\alpha\beta}^{-1}$ est diagonale. On trouve aussi deux valeurs propres. Si $M_1 \neq 0$:

$$\chi_{11} = \frac{1}{2a_2(T_c - T)}, \quad \chi_{\alpha\alpha} = \frac{g_4 + h_4}{h_4a_2(T - T_c)} \text{ pour } \alpha > 1.$$

Nous notons que pour $h_4 \neq 0$, c'est à dire en présence de brisure cubique de la symétrie de rotation, la susceptibilité transverse ne diverge plus qu'à T_c et de plus le rapport d'amplitude correspondant n'est pas universel à la différence de la susceptibilité longitudinale.

Remarque. Dans tous les cas que nous avons examinés la symétrie implique que le terme quadratique est proportionnel à $\sum_\alpha M_\alpha^2$.

Si la symétrie n'est pas suffisante les coefficients de tous les termes M_α^2 génériquement ne sont pas identiques. Il n'y a alors aucune raison pour que les coefficients de tous les termes M_α^2 s'annulent à la même température. On s'attend ainsi à trouver plusieurs températures de transition qui chacune n'implique qu'un sous-ensemble des composantes du paramètre d'ordre. À chacune des transitions les composantes qui ne sont pas impliquées peuvent être négligées et on se trouve ramené à la situation précédente.

11.5 Fonction de corrélation spin–spin

La fonction de partition dans un champ magnétique H_i , dont l'amplitude varie de site en site, s'écrit

$$\mathcal{Z}(H) = \int \prod_i d\rho(S_i) \exp \left[-\mathcal{E}(S)/T + \sum_i H_i S_i \right], \quad (11.23)$$

où $d\rho(S_i)$ est la mesure en chaque site, \mathcal{E} l'énergie de configuration, et l'échelle de champ magnétique inclut un facteur de température ($H/T \mapsto H$). Alors

$$\langle S_{i_1} S_{i_2} \dots S_{i_n} \rangle = \mathcal{Z}^{-1}(H) \frac{\partial^n \mathcal{Z}(H)}{\partial H_{i_1} \partial H_{i_2} \dots \partial H_{i_n}}.$$

En prenant la limite $H_i = H$ nous obtenons les fonctions de corrélation en champ uniforme, et pour $H_i = 0$ en champ nul.

L'énergie libre associée $\mathcal{W}(H)$

$$\mathcal{W}(H) = \ln \mathcal{Z}(H),$$

est alors la fonction génératrice des fonctions de corrélation connexes $W^{(n)}$. Par exemple

$$W_{ij}^{(2)} = \frac{\partial^2 \mathcal{W}(H)}{\partial H_i \partial H_j} = \langle (S_i - \langle S_i \rangle) (S_j - \langle S_j \rangle) \rangle.$$

La transformée de Legendre $\Gamma(M)$ de $\mathcal{W}(H)$ est alors une fonction de aimantation locale M_i

$$\mathcal{W}(H) + \Gamma(M) = \sum_i H_i M_i,$$

avec

$$M_i = \frac{\partial \mathcal{W}(H)}{\partial H_i}.$$

Comme nous l'avons montré la dérivée seconde $\Gamma_{ij}^{(2)}$

$$\Gamma_{ij}^{(2)} = \frac{\partial^2 \Gamma(M)}{\partial M_i \partial M_j},$$

est l'inverse au sens des matrices de $W_{ij}^{(2)}$

$$\sum_k W_{ik}^{(2)} \Gamma_{kj}^{(2)} = \delta_{ij}. \quad (11.24)$$

Dans un système invariant par translation la fonction à deux points ne dépend que de $\mathbf{r}_i - \mathbf{r}_j$, où $\mathbf{r}_i, \mathbf{r}_j$ sont les positions des sites i et j . L'équation (11.24)

précédente est donc une équation de convolution qui devient simple en transformée de Fourier. Posons

$$\tilde{W}^{(2)}(\mathbf{p}) = \sum_{\mathbf{r}} e^{i\mathbf{p}\cdot\mathbf{r}} W^{(2)}(\mathbf{r}),$$

et de même pour $\Gamma^{(2)}$

$$\tilde{\Gamma}^{(2)}(\mathbf{p}) = \sum_{\mathbf{r}} e^{i\mathbf{p}\cdot\mathbf{r}} \Gamma^{(2)}(\mathbf{r}).$$

L'équation (11.24) peut se récrire

$$\sum_k W^{(2)}(\mathbf{r}_i - \mathbf{r}_k) \Gamma^{(2)}(\mathbf{r}_k - \mathbf{r}_j) = \delta(\mathbf{r}_i - \mathbf{r}_j).$$

Multipliant par $e^{i\mathbf{p}\cdot\mathbf{r}_i}$ et sommant sur i nous trouvons

$$\sum_{ik} e^{i\mathbf{p}\cdot\mathbf{r}_i} W^{(2)}(\mathbf{r}_i - \mathbf{r}_k) \Gamma^{(2)}(\mathbf{r}_k - \mathbf{r}_j) = e^{i\mathbf{p}\cdot\mathbf{r}_j}.$$

Nous utilisons alors successivement :

$$\begin{aligned} \sum_i e^{i\mathbf{p}\cdot\mathbf{r}_i} W^{(2)}(\mathbf{r}_i - \mathbf{r}_k) &= \sum_{\mathbf{r}} e^{i\mathbf{p}\cdot(\mathbf{r}_k + \mathbf{r})} W^{(2)}(\mathbf{r}) = e^{i\mathbf{p}\cdot\mathbf{r}_k} \tilde{W}^{(2)}(\mathbf{p}) \\ \sum_k e^{i\mathbf{p}\cdot\mathbf{r}_k} \Gamma^{(2)}(\mathbf{r}_k - \mathbf{r}_j) &= e^{i\mathbf{p}\cdot\mathbf{r}_j} \tilde{\Gamma}^{(2)}(\mathbf{p}). \end{aligned}$$

Nous en déduisons la relation

$$\tilde{W}^{(2)}(\mathbf{p}) = 1/\tilde{\Gamma}^{(2)}(\mathbf{p}).$$

Notons que $\tilde{W}^{(2)}(0)$ est la susceptibilité magnétique. La théorie de Landau fait l'hypothèse que $\Gamma_{kj}^{(2)}$ décroît exponentiellement quand la distance entre i et j tend vers l'infini. Dans ces conditions sa transformée de Fourier est une fonction régulière du vecteur k . Le comportement à grande distance de $W_{ij}^{(2)}$ est dominée par le développement à $|p| \rightarrow 0$ de $\tilde{W}^{(2)}$ et donc $\tilde{\Gamma}^{(2)}$.

Si le système a la symétrie cubique du réseau, ce développement prend la forme

$$\tilde{\Gamma}^{(2)}(\mathbf{p}) = a_2(T - T_c) + \ell^2 \mathbf{p}^2 + O(p^4),$$

où ℓ^2 est une constante, et son inverse a une forme d'Ornstein-Zernicke. À l'ordre p^2 la fonction $\tilde{W}^{(2)}(\mathbf{p})$ a une symétrie de rotation $O(d)$, plus étendue que les symétries discrètes du réseau. On en déduit que pour $T > T_c$ $W^{(2)}(\mathbf{r})$ décroît exponentiellement

$$\begin{aligned} W^{(2)}(\mathbf{r}) &= \frac{1}{(2\pi)^d} \int d^d p e^{-i\mathbf{p}\cdot\mathbf{r}} \tilde{W}^{(2)}(\mathbf{p}) \sim \frac{1}{(2\pi)^d} \int d^d p \frac{e^{-i\mathbf{p}\cdot\mathbf{r}}}{a_2(T - T_c) + \ell^2 \mathbf{p}^2} \\ &\underset{r \rightarrow \infty}{\propto} \frac{1}{r^{(d-1)/2}} e^{-r/\xi}, \end{aligned}$$

avec

$$\xi_{T \rightarrow T_c} \propto (T - T_c)^{-1/2}.$$

ce qui donne la loi de divergence universelle de la longueur de corrélation pour $T \rightarrow T_{c+}$. En général on appelle ν l'exposant et donc $\nu = 1/2$. Enfin à T_c

$$W^{(2)}(\mathbf{r}) \sim \frac{1}{(2\pi)^{d/2}} \int d^d p \frac{e^{-i\mathbf{p}\cdot\mathbf{r}}}{p^2} \underset{r \rightarrow \infty}{\propto} \propto 1/|\mathbf{r}|^{d-2}.$$

Conclusion. Nous avons donc montré sur quelques exemples simples comment la théorie de Landau combinée avec des considérations de groupe de symétrie permettait de mettre en évidence un certain nombre de propriétés universelles des transitions de phase continues au voisinage de la température de transition.

Même si les hypothèses qui sont à la base de la théorie de Landau sont incorrectes, cette théorie fournit une première approche qualitative utile des transitions de phase. De façon assez remarquable si les exposants critiques et certaines prédictions universelles sont incorrectes, la théorie correcte basée sur le groupe de renormalisation prédit la survivance d'une universalité quoique plus restreinte. Les notions d'exposants critiques et autres quantités universelles restent valables, mais si ces valeurs ne dépendent toujours pas des détails des interactions microscopiques, ils dépendent cette fois d'autres caractéristiques des systèmes comme par exemple la dimension d'espace et le groupe de symétrie.

POINTS DIVERS

A1 Solide rigide libre

Nous considérons n points matériels avec positions \mathbf{q}^a et masses m_a rigidement liés. Nous introduisons le centre de masse de ces points dont nous notons la position $\mathbf{Q}(t)$, et la masse totale $M = \sum_a m_a$

$$M\mathbf{Q}(t) = \sum_a m_a \mathbf{q}^a(t).$$

Introduisant une matrice de rotation $\mathbf{R}(t)$, nous pouvons exprimer que le système est rigide en paramétrant $\mathbf{q}^a(t)$ par

$$\mathbf{q}^a(t) = \mathbf{Q}(t) + \mathbf{R}(t)\mathbf{x}^a,$$

où les vecteurs \mathbf{x}^a sont indépendants du temps. En effet on trouve

$$|\mathbf{q}^a - \mathbf{q}^b|^2 = |\mathbf{x}^a - \mathbf{x}^b|^2.$$

De plus par définition du centre de masse

$$\sum_a m_a \mathbf{x}^a = 0. \quad (\text{A1.1})$$

L'énergie E du système libre est alors

$$E = \frac{1}{2} \sum_a m_a (\dot{\mathbf{q}}^a)^2 = \frac{1}{2} M (\dot{\mathbf{Q}}(t))^2 + \frac{1}{2} \sum_a m_a |\dot{\mathbf{R}}(t)\mathbf{x}^a|^2,$$

où l'équation (A1.1) a été utilisée. Nous allons omettre dans ce qui suit la contribution du centre de masse qui se découple. Nous pouvons transformer la contribution de la particule a

$$|\dot{\mathbf{R}}(t)\mathbf{x}^a|^2 = \mathbf{x}^a \cdot \dot{\mathbf{R}}^T \dot{\mathbf{R}} \mathbf{x}^a = \mathbf{x}^a \cdot \dot{\mathbf{R}}^T \mathbf{R} \mathbf{R}^T \dot{\mathbf{R}} \mathbf{x}^a.$$

L'expression peut se récrire en terme de la matrice $\mathbf{L}(t)$

$$\mathbf{L}(t) = \mathbf{R}^T \dot{\mathbf{R}},$$

qui est antisymétrique

$$\mathbf{R}^T \mathbf{R} = \mathbf{1} \Rightarrow \dot{\mathbf{R}}^T \mathbf{R} + \mathbf{R}^T \dot{\mathbf{R}} = 0 \Leftrightarrow \mathbf{L}^T(t) + \mathbf{L}(t) = 0.$$

La matrice $\mathbf{L}(t)$ est donc associée à l'algèbre de Lie du groupe, et l'énergie devient

$$E = \frac{1}{2} \sum_a m_a (\dot{\mathbf{q}}^a)^2 = \frac{1}{2} \sum_a m_a \mathbf{x}^a \cdot \mathbf{L}^T(t) \mathbf{L}(t) \mathbf{x}^a.$$

Pour un système en rotation uniforme la matrice \mathbf{L} est constante et dans l'espace à trois dimensions les éléments de la matrice peuvent s'écrire

$$\mathbf{L}_{ij} = \sum_k \epsilon_{ijk} \omega_k .$$

On trouve alors

$$E = \frac{1}{2} \sum_{ij} \omega_i I_{ij} \omega_j ,$$

avec

$$I_{ij} = \sum_a m_a (\mathbf{x}^a \cdot \mathbf{x}^a \delta_{ij} - \mathbf{x}_i^a \mathbf{x}_j^a) .$$

Rien ne change dans l'argument si la distribution de matière est continue, si ce n'est que la somme sur a est remplacée par une intégrale.

A2 Oscillateur harmonique quantique et algèbre de Lie

Nous donnons ici un calcul du spectre de l'oscillateur harmonique moins standard que la méthode utilisant les opérateurs de création et d'annihilation, mais basée sur des considérations de représentation d'algèbre de Lie.

Nous pouvons toujours ramener le hamiltonien de l'oscillateur harmonique à un multiple d'une forme standard H

$$H = \frac{1}{2}(\hat{p}^2 + \hat{q}^2), \quad [\hat{q}, \hat{p}] = i .$$

Considérons alors les deux autres opérateurs hermitiens

$$\begin{aligned} L &= \frac{1}{2}(\hat{p}^2 - \hat{q}^2) \\ M &= \frac{1}{2}(\hat{p}\hat{q} + \hat{q}\hat{p}) . \end{aligned}$$

Un calcul simple permet d'établir les relations de commutation

$$[M, L] = 2iH, \quad [H, L] = 2iM, \quad [M, H] = 2iL .$$

Elles sont analogues aux relations de commutation de l'algèbre de Lie de $SO(3)$, mais le signe de la première relation est différent. Par ailleurs on trouve aussi

$$C = H^2 - M^2 - L^2 = -\frac{3}{4}\mathbf{1} .$$

Comparant avec l'expression de l'opérateur de Casimir de $SO(3)$ on observe de nouveau une différence de signe. En fait cette algèbre correspond à un groupe de transformations qui préserve la pseudo-métrique $x_0^2 - x_1^2 - x_2^2$. Ce groupe est le groupe de Lorentz de la Relativité Spéciale en deux dimensions d'espace.

Pour trouver le spectre de H on procède en analogie avec la classification des représentations de l'algèbre de Lie de $SO(3)$. Définissons les opérateurs

$$J_{\pm} = L \pm iM \Rightarrow [H, J_{\pm}] = \pm 2J_{\pm} .$$

Alors

$$\frac{3}{4} = J_+ J_- + 2H - H^2 = J_- J_+ - 2H - H^2.$$

Soit $|\mu\rangle$ un vecteur propre de H avec valeur propre μ . Les relations de commutation entraînent

$$H J_{\pm} |\mu\rangle = (\mu \pm 2) J_{\pm} |\mu\rangle.$$

Ceci implique que soit $J_{\pm} |\mu\rangle$ s'annule soit

$$J_{\pm} |\mu\rangle \propto |\mu \pm 2\rangle.$$

C'est la condition de positivité de l'opérateur H qui arrête l'itération. Donc si $0 < \mu < 2$

$$J_- |\mu\rangle = 0.$$

Agissant alors avec \mathcal{C} sur $|\mu\rangle$ on trouve

$$\mathcal{C} |\mu\rangle = -\frac{3}{4} |\mu\rangle = \mu(\mu - 2) |\mu\rangle,$$

et donc

$$\mu = \frac{1}{2} \quad \text{et} \quad \mu = \frac{3}{2}.$$

Il y a donc deux séries de valeurs propres $1/2 + 2n$ et $3/2 + 2n$ ce qui donne bien le spectre usuel en $n + 1/2$ de l'oscillateur harmonique. Notons que parce que le groupe $SO(1, 2)$ n'est pas compact, la représentation est de dimension infinie.

Cette méthode pour déterminer le spectre de l'oscillateur harmonique est un peu compliquée mais elle s'applique aussi à un hamiltonien plus général de la forme

$$H = \frac{1}{2}(\hat{p}^2 + \hat{q}^2 + \alpha/\hat{q}^2),$$

où α est une constante positive. On retrouve alors que H et

$$L = \frac{1}{2}(\hat{p}^2 - \hat{q}^2 + \alpha/\hat{q}^2), \quad M = \frac{1}{2}(\hat{p}\hat{q} + \hat{q}\hat{p}),$$

forment la même algèbre de Lie que précédemment. Le seul changement provient de l'opérateur \mathcal{C} qui vaut maintenant

$$\mathcal{C} = \left(-\frac{3}{4} + \alpha\right) \mathbf{1}.$$

En conséquence les valeurs propres partent des solutions de l'équation

$$\mu(\mu - 2) + \frac{3}{4} - \alpha = 0 \Rightarrow \mu = 1 \pm \frac{1}{2}\sqrt{1 + 4\alpha}.$$

Nous voyons que le spectre reste en fait réel pour tout $\alpha > -1/4$. Ceci entraîne qu'il est possible de paramétrer α en termes d'un paramètre réel $\ell \geq -1/2$ de la façon suivante

$$\alpha = \ell(\ell + 1) \Rightarrow \alpha > -1/4 \quad \text{et} \quad \mu = 1 \pm (\ell + 1/2).$$

La valeur du Casimir est alors

$$\mathcal{C} = \left(\ell + \frac{3}{2}\right)\left(\ell - \frac{1}{2}\right).$$

Pour $\alpha \geq 0$ et donc $\ell \geq 0$, le hamiltonien est positif et donc la deuxième série n'existe que si

$$\mu > 0 \Rightarrow \ell \leq \frac{1}{2} \quad \text{ou} \quad \alpha \leq \frac{3}{4}.$$

Enfin, notons que pour ℓ entier le hamiltonien H est aussi le hamiltonien d'un oscillateur harmonique à plusieurs dimensions, invariant par rotation et exprimé en terme de variables radiales après diagonalisation du moment cinétique.