

D. BERNARD, Y. LASZLO et D. RENARD

ELÉMENTS DE THÉORIE DES GROUPES
ET SYMÉTRIES QUANTIQUES

D. BERNARD, Y. LASZLO et D. RENARD

10 septembre 2012

**ELÉMENTS DE THÉORIE DES GROUPES ET SYMÉTRIES
QUANTIQUES**

D. BERNARD, Y. LASZLO et D. RENARD

TABLE DES MATIÈRES

Introduction.....	1
Auteurs.....	2
I. Groupes, actions de groupes et représentations.....	3
I.1. Un exemple fondamental et quelques définitions.....	3
I.2. Exemples de groupes et d'actions de groupes.....	7
I.3. Le groupe symétrique.....	9
I.4. Produit semi-direct.....	12
I.5. Représentations.....	13
I.6. Opérations sur les représentations : sommes directes et produits, produits tensoriels, représentation contragrédiente.....	18
I.7. Décomposition canonique et règles de sélection.....	23
I.8. Représentations des groupes topologiques.....	24
I.9. Mesure de Haar.....	26
II. Groupes linéaires et leurs algèbres de Lie.....	29
II.1. Le groupe $\mathbf{GL}(n, \mathbb{K})$	29
II.2. L'application exponentielle.....	30
II.3. Groupes linéaires.....	31
II.4. Sous-groupes de Lie. Connexité.....	33
II.5. Homomorphismes de groupes linéaires. Revêtements.....	36
II.6. Représentations de dimension finie des groupes linéaires connexes.....	39
II.7. Représentation adjointe.....	41
II.8. Représentations projectives.....	41
III. Représentations de $\mathfrak{sl}(2, \mathbb{C})$, $\mathbf{SU}(2)$, et $\mathbf{SO}(3)$.....	43
III.1. Le revêtement $\mathbf{SU}(2) \rightarrow \mathbf{SO}(3)$	43

III.2. Représentations de $\mathfrak{sl}(2, \mathbb{C})$	45
III.3. Les harmoniques sphériques.....	49
IV. Représentations du groupe de Heisenberg.....	57
V. Symétries en mécanique quantique.....	67
V.1. Actions sur les états quantiques.....	67
V.2. Symétries, dégénérescences et règles de sélection.....	73
V.3. Observables.....	78
V.4. Permutations et tresses.....	80
V.5. Les classes de Wigner.....	86
V.6. Espaces de Fock.....	88
VI. Du modèle des quarks au modèle standard.....	93
VI.1. L'isospin et l'algèbre $su(2)$	93
VI.2. Symétrie de saveur et l'algèbre $su(3)$	95
VI.3. Le modèle des quarks.....	100
VI.4. Eléments du modèle standard.....	104
VII. Groupes de Lorentz, de Poincaré et équation de Dirac.....	107
VII.1. Groupes de Lorentz et de Poincaré.....	107
VII.2. Représentations du groupe de Lorentz.....	111
VII.3. Equation de Dirac.....	113
VII.4. Représentations du groupe de Poincaré.....	115
VIII. Invariance conforme en physique.....	119
VIII.1. Transformations conformes en physique classique.....	119
VIII.2. Le mouvement brownien bidimensionnel.....	123
VIII.3. Systèmes électroniques mésoscopiques 1D.....	128
VIII.4. Champs quantiques et algèbre de Virasoro.....	135
IX. Oscillateurs harmoniques et représentations du groupe de Heisenberg.....	139
IX.1. Introduction.....	140
IX.2. Groupe et algèbre d'Heisenberg.....	141
IX.3. Le théorème de Stone - von Neumann.....	147
IX.4. Algèbre d'Heisenberg de dimension infinie.....	152
IX.5. Algèbre de Virasoro.....	158
IX.6. Appendice.....	162

Bibliographie..... 165

Introduction

Ces notes de cours traitent de la théorie des groupes et de ses applications à la physique, en particulier à la physique des hautes énergies. Elles constituent une simple introduction à ce sujet, sans chercher à en couvrir tous les aspects. Elles sont donc dédiées à un public d'amateurs (au deux sens du mot) mathématiciens ou physiciens. Elles ont été écrites conjointement par deux mathématiciens et un physicien théoricien et illustrent les différents points de vue, mathématiques ou physiques. Ces différences de perspective et de langage sont, nous le croyons, une richesse qu'il est intéressant d'exploiter, avec aussi l'espoir de renforcer les liens entre mathématiciens et physiciens.

La partie mathématique de ces notes présente de manière succincte quelques éléments de base de la théorie des groupes et de leurs représentations, en vue de leurs applications en physique. Le style adopté est celui d'une exposition mathématique classique de la forme « définition-lemme-théorème » et autant que faire se peut, rigoureuse, même si beaucoup de démonstrations sont omises. En revanche, les nombreux exemples illustrant ces notes ont toujours une application physique.

La partie physique a pour but d'illustrer l'apport de la théorie des groupes en physique, et en particulier en physique quantique. Bien qu'une grande part des exemples soit issue de la physique des hautes énergies, les notions introduites se recyclent et émergent dans d'autres domaines de la physique. Certains des sujets traités sont des classiques (par exemple, les groupes de Lorentz et de Poincaré, l'équation de Dirac, le modèle des quarks), d'autres font appel à des notions de physiques plus récentes (par exemple, le modèle standard ou l'invariance conforme).

Auteurs

Les chapitres I à IV ont été écrits par David Renard, les chapitres V à IX par Denis Bernard et le chapitre X par Yves Laszlo.

Une grande partie du texte des chapitres I à IV est tiré de l'ouvrage « Groupes et représentations » publié par D. Renard aux éditions de l'Ecole Polytechnique ([8]), auquel on pourra se reporter pour certains compléments, en particulier les démonstrations.

CHAPITRE I

GROUPES, ACTIONS DE GROUPES ET REPRÉSENTATIONS

« Un principe directeur des mathématiques modernes tient en cette leçon : lorsque vous avez affaire à une entité S munie d'une certaine structure, essayez de déterminer son groupe d'automorphismes, le groupe des transformations de ses éléments qui préservent les relations structurales. Vous pouvez espérer gagner une profonde compréhension de la constitution de S de cette manière » Hermann Weyl⁽¹⁾.

Le but de ce chapitre est de rappeler les définitions et les résultats de base de la théorie des groupes, supposées déjà plus ou moins connues du lecteur, en en profitant pour introduire la terminologie et les notations employées par la suite.

I.1. Un exemple fondamental et quelques définitions

Soit X un ensemble. Notons $\mathbf{Aut}(X)$ l'ensemble des bijections de X dans lui-même. Cet ensemble est muni de la loi de composition des applications :

$$(I.1.1) \quad \mu : \mathbf{Aut}(X) \times \mathbf{Aut}(X) \rightarrow \mathbf{Aut}(X), \quad (\phi, \psi) \mapsto \phi \circ \psi.$$

La loi de composition est associative, c'est-à-dire que quels que soient ϕ_1 , ϕ_2 et ϕ_3 dans $\mathbf{Aut}(X)$,

$$(I.1.2) \quad \mu(\mu(\phi_1, \phi_2), \phi_3) = \mu(\phi_1, \mu(\phi_2, \phi_3)),$$

ou plus simplement $(\phi_1 \circ \phi_2) \circ \phi_3 = \phi_1 \circ (\phi_2 \circ \phi_3)$.

D'autre part, cette loi admet un élément neutre, l'identité de X , notée Id_X :

$$(I.1.3) \quad (\forall \phi \in \mathbf{Aut}(X)), \quad \text{Id}_X \circ \phi = \phi \circ \text{Id}_X = \phi.$$

Enfin, tout élément ϕ de $\mathbf{Aut}(X)$ admet un inverse, c'est-à-dire un élément de $\mathbf{Aut}(X)$, noté ϕ^{-1} , vérifiant

$$(I.1.4) \quad \phi \circ \phi^{-1} = \phi^{-1} \circ \phi = \text{Id}_X.$$

1. Traduit librement d'une traduction de l'allemand en anglais... j'espère que le sens général se sera conservé.

Le lecteur instruit reconnaît là le fait que $\mathbf{Aut}(X)$ est muni d'une structure de groupe. Pour les autres, nous rappelons la définition d'un groupe ci-dessous, qui consiste à prendre comme axiomes ces propriétés de $\mathbf{Aut}(X)$, de μ et de Id_X .

Remarquons que nous disposons aussi d'une application canonique

$$(I.1.5) \quad a : \mathbf{Aut}(X) \times X \longrightarrow X, \quad (\phi, x) \mapsto \phi(x).$$

L'application a vérifie les propriétés suivantes : quels que soient ϕ_1, ϕ_2 dans $\mathbf{Aut}(X)$ et x dans X ,

$$(I.1.6) \quad a(\mu(\phi_1, \phi_2), x) = a(\phi_1, a(\phi_2, x)),$$

et de plus

$$(I.1.7) \quad a(\text{Id}_X, x) = x.$$

Autrement dit l'application a définit une action du groupe $\mathbf{Aut}(X)$ sur X . Donnons maintenant les définitions générales.

Définition I.1.1. — Un groupe est un ensemble G , muni d'une loi

$$(I.1.8) \quad \mu : G \times G \rightarrow G, \quad (g, h) \mapsto gh := \mu(g, h),$$

appelée produit du groupe, et vérifiant :

(i) (associativité) quels que soient g, h, k dans G ,

$$\mu(\mu(g, h), k) = \mu(g, \mu(h, k)),$$

(ou encore, $(gh)k = g(hk)$),

(ii) (élément neutre) il existe un élément $e = e_G$ de G , appelé l'élément neutre, tel que pour tout $g \in G$, $\mu(g, e) = \mu(e, g) = g$ (ou encore $ge = eg = g$),

(iii) (inverse) quelque soit $g \in G$, il existe un élément de G , noté g^{-1} , tel que $\mu(g, g^{-1}) = \mu(g^{-1}, g) = e$ (ou encore $gg^{-1} = g^{-1}g = e$).

Remarque I.1.2. — On déduit facilement de ces axiomes l'unicité de l'élément neutre et de l'inverse d'un élément donné.

Définition I.1.3. — Soit G un groupe, et X un ensemble. Une action (à gauche) de G sur X est la donnée d'une application

$$(I.1.9) \quad a : G \times X \rightarrow X, \quad (g, x) \mapsto g \cdot x$$

vérifiant :

$$(I.1.10) \quad (\forall g, h \in G), (\forall x \in X), \quad a(\mu(g, h), x) = a(g, a(h, x)).$$

Lorsque l'action est notée par un “ \cdot ”, ceci s'écrit $(gh) \cdot x = g \cdot (h \cdot x)$, et de plus,

$$(I.1.11) \quad (\forall x \in X), \quad a(e, x) = e \cdot x = x.$$

Définition I.1.4. — On appelle G -ensemble un ensemble X muni d'une action de G . Un G -morphisme du G ensemble X vers le G -ensemble Y est une application $f : X \rightarrow Y$ compatible avec les actions de G , c'est-à-dire

$$f(g \cdot x) = g \cdot f(x), \quad (x \in X), (g \in G).$$

On peut penser aux G -ensembles X, Y , comme à des ensembles munis de « symétries », et aux G -morphisme comme à des applications préservant ces symétries.

Un morphisme de groupes est une application d'un groupe vers un autre qui préserve la structure de groupe :

Définition I.1.5. — Soient G et H deux groupes. Une application de $f : G \rightarrow H$ est un morphisme de groupes si quels que soient g, h dans G ,

$$f(gh) = f(g)f(h).$$

Dans ce cas, l'ensemble des $g \in G$ tels que $f(g) = e_H$ est appelé noyau du morphisme f . C'est un sous-groupe de G . On le note $\ker f$. L'image du morphisme f est un sous-groupe de H que l'on note $\text{Im } f$.

Remarque I.1.6. — La donnée d'une action a d'un groupe G sur un ensemble X est équivalente à la donnée d'un morphisme de groupes

$$A : G \rightarrow \mathbf{Aut}(X).$$

On passe de a à A et réciproquement par

$$A(g)(x) = a(g, x), \quad (x \in X), (g \in G).$$

Définition I.1.7. — Un sous-ensemble H d'un groupe G est un sous-groupe s'il contient l'élément neutre e et est stable par produits et passage aux inverses.

On obtient de nombreux exemples de groupes et d'actions de groupes à partir de l'exemple fondamental $(X, \mathbf{Aut}(X))$ ci-dessus, et en supposant que l'ensemble X est muni d'une structure supplémentaire, clairement spécifiée par le contexte. On redéfinit alors $\mathbf{Aut}(X)$ comme l'ensemble des bijections de X dans lui-même qui préservent, ainsi que leurs inverses, la structure de X . Les applications μ et a définies comme en (I.1.1) et (I.1.5) vérifient encore (I.1.2), (I.1.3), (I.1.4), (I.1.6), (I.1.7). Lorsque X est muni d'une structure supplémentaire, on supposera, souvent de manière implicite, qu'une action d'un groupe G sur X préserve cette structure. Remarquons aussi que si l'on part d'un ensemble X muni d'une certaine structure et de son groupe d'automorphismes $\mathbf{Aut}(X)$, et que l'on rajoute une structure supplémentaire, l'ensemble des éléments de $\mathbf{Aut}(X)$ préservant de plus cette nouvelle structure est un sous-groupe de $\mathbf{Aut}(X)$. Ces considérations un peu abstraites seront illustrées par des exemples dans la section suivante.

Une autre manière d'obtenir des exemples de groupes à partir d'une action d'un groupe G sur un ensemble X est de considérer, pour toute partie Y de X

$$\mathbf{Fix}_G(Y) = \{g \in G \mid (\forall y \in Y), g \cdot y = y\},$$

$$\mathbf{Stab}_G(Y) = \{g \in G \mid (\forall y \in Y), g \cdot y \in Y\}.$$

(Pour $\mathbf{Stab}_G(Y)$, on doit supposer que la partie Y est finie pour assurer la stabilité par passage à l'inverse, où bien l'imposer : $(\forall y \in Y), g^{-1} \cdot y \in Y$). On vérifie facilement que l'on obtient ainsi des sous-groupes du groupe G . Il est intéressant de remarquer que tout sous-groupe d'un groupe G peut s'obtenir ainsi.

Définition I.1.8. — Soit G un groupe agissant sur un ensemble X . On appelle orbite d'un point x de X sous l'action de G l'ensemble des points de la forme $g \cdot x$, g décrivant le groupe G . Notons $G \cdot x$ l'orbite d'un point x de X . On dit que l'action de G sur X est transitive s'il n'y a qu'une seule orbite, et qu'elle est fidèle si le morphisme

$$A : G \rightarrow \mathbf{Aut}(X)$$

défini par l'action est injectif. On dit que l'action est libre si tout élément différent de l'élément neutre agit sans point fixe. Une action libre est fidèle.

Proposition I.1.9. — Soit G un groupe agissant sur un ensemble X . La relation binaire

$$x \sim y \quad \text{si} \quad G \cdot x = G \cdot y$$

est une relation d'équivalence sur X . Les orbites de l'action de G forment donc une partition de l'ensemble X .

Nous laissons la vérification de ce fait au lecteur.

On note $G \backslash X$ l'ensemble des orbites de l'action de G sur X . On appelle système de représentants des orbites de G dans X un ensemble $\{x_i\}$ d'éléments de X tel que

$$\{x_i\} \rightarrow G \backslash X, \quad x_i \mapsto G \cdot x_i$$

soit une bijection.

Remarque I.1.10. — Soit G un groupe agissant sur un ensemble X , l'action étant fidèle et transitive (donc libre). Alors il découle facilement des définitions que tout choix d'un point de base $x \in X$ donne une bijection $X \simeq G$. On dit alors que X est un espace homogène principal sur G . Par exemple, si G est un espace vectoriel sur un corps k , un espace homogène principal sur G est un espace affine.

L'action d'un groupe G sur un ensemble X muni d'une certaine structure, est un moyen puissant d'obtenir des informations sur la structure du groupe G , ou sur celle de l'espace X , selon la nature du problème considéré.

I.2. Exemples de groupes et d'actions de groupes

Exemple I.2.1. — Le groupe des bijections (on dit aussi permutations dans ce contexte) de l'ensemble $\{1, \dots, n\}$ est noté \mathfrak{S}_n .

Exemple I.2.2. — Soit V un espace vectoriel sur un corps k . L'ensemble des applications linéaires bijectives de V dans lui-même est souvent noté $\mathbf{GL}(V)$ plutôt que $\mathbf{Aut}(V)$. On appelle ce groupe le groupe général linéaire.

Une action d'un groupe G dans un espace vectoriel V qui préserve la structure d'espace vectoriel (on parle aussi d'action linéaire) est donc équivalent à la donnée d'un morphisme de groupes :

$$A : G \rightarrow \mathbf{GL}(V).$$

De telles actions apparaissent dans de nombreux contextes en mathématique (et en physique) et l'importance de ce concept justifie une terminologie spécifique. On dit que l'espace vectoriel V , muni d'une action linéaire d'un groupe G , est une **représentation** du groupe G . Lorsque $V = k^n$, on utilise la notation $\mathbf{GL}_n(k)$ pour $\mathbf{GL}(V)$.

Exemple I.2.3. — Soit G un groupe agissant sur un ensemble X , et soit $\mathcal{F}(X)$ l'espace vectoriel des fonctions sur X à valeurs complexes. Alors $\mathcal{F}(X)$ est lui aussi muni d'une action linéaire de G , donnée par

$$(g \cdot f)(x) = f(g^{-1} \cdot x), \quad (g \in G), (f \in \mathcal{F}(X)), (x \in X).$$

Cette nouvelle action, d'une certaine manière, contient autant d'information que l'ancienne, mais présente l'avantage de pouvoir utiliser les techniques d'algèbre linéaire. C'est pourquoi l'on s'attache plus particulièrement à l'étude des actions linéaires des groupes, c'est-à-dire, de leurs représentations.

Exemple I.2.4. — Soit V un espace vectoriel sur \mathbb{C} , muni d'une structure hilbertienne, c'est-à-dire d'un produit scalaire hermitien défini positif. Le sous-groupe de $\mathbf{GL}(V)$ préservant ce produit hermitien est noté $\mathbf{U}(V)$ et est appelé groupe unitaire. Lorsque $V = \mathbb{C}^n$, muni du produit hermitien canonique, on le note $\mathbf{U}(n)$. Une action d'un groupe G dans V préservant la structure hilbertienne est équivalente à la donnée d'un morphisme de groupes :

$$A : G \rightarrow \mathbf{U}(V).$$

On dit alors que la représentation de G dans V est unitaire.

Exemple I.2.5. — Soit V un espace vectoriel sur \mathbb{R} muni d'un produit scalaire. Le sous-groupe de $\mathbf{GL}(V)$ préservant ce produit scalaire est noté $\mathbf{O}(V)$ et est appelé groupe orthogonal. Lorsque $V = \mathbb{R}^n$, muni du produit scalaire canonique, on le note $\mathbf{O}(n)$.

Si $p + q = n$, munissons \mathbb{R}^n de la forme bilinéaire symétrique :

$$(x, y)_{p,q} = x_1 y_1 + x_2 y_2 + \dots + x_p y_p - x_{p+1} y_{p+1} - \dots - x_n y_n,$$

où $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ sont des vecteurs de \mathbb{R}^n . Cette forme est non dégénérée, de signature (p, q) . Le sous-groupe de $\mathbf{GL}_n(\mathbb{R})$ préservant la forme $(\cdot, \cdot)_{p,q}$ est noté $\mathbf{O}(p, q)$. Si l'on

note J_{pq} la matrice diagonale formée de 1 puis de -1 avec pour multiplicités respectives p et q , on a

$$\mathbf{O}(p, q) = \{A \in \mathbf{GL}_n(\mathbb{R}) \mid AJ_{pq} {}^t A = J_{pq}\}.$$

Le groupe $\mathbf{O}(3, 1)$ joue un rôle important comme groupe de symétrie en électromagnétisme et en théorie de la relativité. Il s'appelle le groupe de Lorentz.

De même, si $p + q = n$, on définit les groupes $\mathbf{U}(p, q)$, sous-groupes de $\mathbf{GL}(n, \mathbb{C})$.

Exemple I.2.6. — Si V est un espace vectoriel de dimension finie sur le corps k , on note $\mathbf{SL}(V)$ le sous-groupe de $\mathbf{GL}(V)$ des éléments de déterminant 1. Ce groupe s'appelle le groupe spécial linéaire. Remarquons que

$$\det : \mathbf{GL}(V) \rightarrow k^*$$

est un morphisme de groupes, et donc $\mathbf{SL}(V)$ est son noyau. L'intersection d'un sous-groupe H de $\mathbf{GL}(V)$ avec $\mathbf{SL}(V)$ sera notée \mathbf{SH} . En reprenant les exemples ci-dessus, on obtient $\mathbf{SU}(V)$, $\mathbf{SO}(V)$, $\mathbf{SO}(p, q)$...

Exemple I.2.7. — Considérons l'action naturelle de $\mathbf{O}(2)$ dans \mathbb{R}^2 , et soit $Y \subset \mathbb{R}^2$ un polygone régulier à n cotés ($n \leq 3$), centré en 0. Le sous-groupe de $\mathbf{O}(2)$ laissant invariant Y est le groupe diédral D_n . Son ordre est $2n$. Son intersection avec $\mathbf{SO}(2)$ est isomorphe à $\mathbb{Z}/n\mathbb{Z}$.

Exemple I.2.8. — Une forme symplectique sur un espace vectoriel V (défini sur un corps k) est une forme $\omega : V \times V \rightarrow k$ bilinéaire, antisymétrique, et non dégénérée (c'est-à-dire que si $\omega(X, Y) = 0$ pour tout $X \in V$, alors $Y = 0$). En dimension finie, ceci impose à la dimension de l'espace d'être paire.

Sur k^{2n} la forme symplectique canonique est donnée par :

$$X = {}^t(x_1, \dots, x_n, y_1, \dots, y_n), \quad X' = {}^t(x'_1, \dots, x'_n, y'_1, \dots, y'_n)$$

$$\omega(X, X') = x_1 y'_1 + \dots + x_n y'_n - y_1 x'_1 - \dots - y_n x'_n.$$

ou encore, en notant J_{2n} la matrice

$$J_{2n} = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix},$$

$$\omega(X, X') = {}^t X J_{2n} X'$$

Soit (V, ω) un espace symplectique. Le groupe symplectique $\mathbf{Sp}(V, \omega)$ est alors le sous-groupe de $\mathbf{GL}(V)$ des isomorphismes préservant la forme symplectique, c'est-à-dire que

$$\mathbf{Sp}(V, \omega) = \{A \in \mathbf{GL}(V) \mid \omega(AX, AX') = \omega(X, X'), (\forall X, X' \in V)\}$$

Dans le cas où $V = k^{2n}$ muni de la forme symplectique canonique,

$$\mathbf{Sp}(2n, k) = \{A \in \mathbf{GL}_{2n}(k) \mid {}^t A J_{2n} A = J_{2n}\}.$$

Exemple I.2.9. — **Le groupe de Heisenberg** Soit (V, ω) un espace symplectique. Pour fixer les idées, nous supposons ici que le corps de base est \mathbb{R} , mais les constructions qui vont suivre peuvent se faire sur un corps quelconque. Notons $H(V, \omega) = V \oplus \mathbb{R}$, que nous munissons du produit suivant :

$$(X, z)(X', z') = (X + X', z + z' + \omega(X, X'))$$

On vérifie facilement que $H(V, \omega)$ est un groupe.

Dans le cas où $V = \mathbb{R}^{2n}$ muni de la forme symplectique canonique, on obtient une réalisation matricielle du groupe de Heisenberg de la manière suivante

$$(X = {}^t(x_1, \dots, x_n, y_1, \dots, y_n), z) \mapsto \begin{pmatrix} 1 & x_1 & \dots & x_n & y_1 & \dots & y_n & z \\ 0 & 1 & \dots & \dots & \dots & \dots & \dots & y_1 \\ 0 & 0 & 1 & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & y_n \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & -x_1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & 1 & -x_n \\ 0 & 0 & \dots & \dots & \dots & \dots & 0 & 1 \end{pmatrix}$$

Exemple I.2.10. — Si X est un espace métrique, où plus généralement topologique, $\mathbf{Aut}(X)$ est l'ensemble des homéomorphismes de X dans lui-même. Si a est une action d'un groupe G sur X , on demande que les applications $a(g, \cdot) : X \rightarrow X$ soient continues. Dans le cas où X est une variété différentiable, $\mathbf{Aut}(X)$ est l'ensemble des difféomorphismes de X dans lui-même. Les actions sur X sont alors supposées différentiables.

Un groupe important apparaissant en physique est celui des difféomorphismes du cercle, noté $\mathbf{Diff}(S^1)$. Le sous-groupe des difféomorphismes du cercle préservant l'orientation est noté $\mathbf{Diff}^+(S^1)$.

I.3. Le groupe symétrique

Nous rappelons rapidement dans cette section les notations et les principaux résultats concernant le groupe symétrique \mathfrak{S}_n , le groupe des permutations de l'ensemble $\{1, \dots, n\}$.

Nous adoptons les notations usuelles pour les éléments de \mathfrak{S}_n . Ainsi, par exemple

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 1 & 3 \end{pmatrix}$$

est la bijection de l'ensemble $\{1, 2, 3, 4, 5\}$ envoyant 1 sur 4, 2 sur 2, 3 sur 5, 4 sur 1 et 5 sur 3.

Ceci permet d'effectuer facilement les calculs de produits, si l'on n'oublie pas que dans une composition de fonctions $f \circ g$, c'est la fonction g qui agit avant f :

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 2 & 5 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 1 & 3 & 2 \end{pmatrix}$$

Soit $\sigma \in \mathfrak{S}_n$ et $k \in \{1, \dots, n\}$. On appelle par abus de langage orbite de k sous σ l'orbite de k sous l'action du sous-groupe $\langle \sigma \rangle$ de \mathfrak{S}_n ($\langle \sigma \rangle$ est le sous-groupe engendré par σ).

Voyons maintenant certaines permutations particulières. Si i et j sont deux éléments différents de $\{1, \dots, n\}$, on appelle **transposition** de i et de j la permutation de \mathfrak{S}_n (notée τ_{ij}) qui échange i et j et laisse tous les autres éléments fixes.

On appelle **cycle** une permutation dont toutes les orbites sauf au plus une sont des singletons. On appelle longueur du cycle le cardinal de cette orbite particulière, la longueur de l'identité étant 1. Ainsi une transposition est un cycle de longueur 2.

On appelle **permutation circulaire** de \mathfrak{S}_n une permutation n'ayant qu'une seule orbite. Les permutations circulaires sont donc les cycles de longueur n .

On peut aussi noter les permutations selon leur décomposition en cycles, par exemple,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 7 & 4 & 6 & 9 & 2 & 5 & 8 & 1 & 3 \end{pmatrix},$$

est aussi notée

$$\sigma = (178)(249365).$$

De même, on note aussi (i, j) la transposition τ_{ij} .

Remarquons qu'il n'y a pas d'unicité d'une telle écriture :

$$\sigma = (178)(249365) = (817)(365249) \in \mathfrak{S}_9,$$

mais, en dehors de ces ambiguïtés évidentes, la décomposition en cycles est essentiellement déterminée.

On omet généralement les cycles de longueur 1 (les points fixes) d'une telle écriture :

$$\sigma = (178)(2)(4536)(9) = (178)(4536).$$

Remarquons que dans cette dernière écriture, il n'est plus apparent que σ soit un élément de \mathfrak{S}_9 .

La décomposition de $\{1, \dots, n\}$ en orbites sous σ est apparente dans une écriture en cycles de σ .

Théorème I.3.1. — *Les transpositions $\tau_{i,i+1}$, $i = 1, \dots, n-1$, engendrent \mathfrak{S}_n .*

Démonstration. (Esquisse). Par récurrence sur n , on montre que \mathfrak{S}_n est engendré par les transpositions. On montre ensuite qu'une transposition quelconque est produit de transpositions de la forme $\tau_{i,i+1}$. \square

Ecrivons une permutation σ comme produit de transpositions. Bien sûr, il n'y a pas unicité de cette écriture, ni même unicité du nombre de transpositions intervenant dans cette écriture. En revanche, la parité de ce nombre de transpositions est déterminée par σ , comme l'affirme le théorème suivant :

Théorème I.3.2. — *Soit $\sigma \in \mathfrak{S}_n$. Il y a égalité entre les nombres suivants :*

(i) $(-1)^T$ où T est le nombre de transpositions dans une écriture de σ comme produit de transpositions.

(ii) $(-1)^D$, $D = n - m$ où m est le nombre d'orbites de σ dans $\{1, \dots, n\}$.

(iii) $(-1)^S$ où S est le cardinal de l'ensemble des couples $(i, j) \in \{1, \dots, n\}^2$ tels que $i < j$ et $\sigma(i) > \sigma(j)$.

(iv) $\prod_{i < j} \frac{\sigma(i) - \sigma(j)}{i - j}$.

On appelle ce nombre la **signature** de σ et on le note $\text{sgn}(\sigma)$. Si $\text{sgn}(\sigma) = 1$, on dit que σ est **paire**, et **impaire** si $\text{sgn}(\sigma) = -1$

Démonstration. l'égalité entre (iii) et (iv) est évidente car tous les facteurs $(i-j)$, au signe près, apparaissent une et une seule fois au numérateur et au dénominateur. La valeur absolue de (iv) est donc 1, et son signe est donné par le nombre de couples (i, j) tels que $i < j$ et $\sigma(i) > \sigma(j)$.

Montrons l'égalité entre (i) et (ii), ce qui montre au passage que (i) est bien défini, c'est-à-dire ne dépend pas de l'écriture de σ en un produit de transpositions. Notons $\epsilon(\sigma)$ la quantité définie en (ii). On montre d'abord que si $\tau = \tau_{ij}$ est une transposition $\epsilon(\sigma\tau) = -\epsilon(\sigma)$, en distinguant deux cas :

- si i et j sont dans la même orbite sous σ , alors les orbites sous $\sigma\tau$ sont les mêmes que celles sous σ , sauf l'orbite contenant i et j qui se scinde en deux.

- si i et j ne sont pas dans la même orbite sous σ , alors les orbites sous $\sigma\tau$ sont les mêmes que celles sous σ , sauf celles contenant i et j qui n'en forment plus qu'une.

On raisonne alors par récurrence sur le nombre de transpositions dans une écriture de σ comme produit de celles-ci. Ceci montre au passage que la signature d'un cycle de longueur k est $(-1)^{k-1}$.

Montrons maintenant que (i) = (iv). On a, quels que soient $\sigma, \tau \in \mathfrak{S}_n$,

$$\begin{aligned} \prod_{i < j} \frac{\sigma\tau(i) - \sigma\tau(j)}{i - j} &= \prod_{i < j} \frac{\sigma\tau(i) - \sigma\tau(j)}{\tau(i) - \tau(j)} \prod_{i < j} \frac{\tau(i) - \tau(j)}{i - j} \\ &= \prod_{i < j} \frac{\sigma(i) - \sigma(j)}{i - j} \prod_{i < j} \frac{\tau(i) - \tau(j)}{i - j} \end{aligned}$$

par un changement de variables dans le premier produit. Si l'on note $\epsilon'(\sigma)$ la quantité définie en (iv), on a donc $\epsilon'(\sigma\tau) = \epsilon'(\sigma)\epsilon'(\tau)$. On conclut alors encore par récurrence sur le nombre de transpositions dans une écriture de σ comme produit de celles-ci, en remarquant que $\epsilon'(\tau) = -1$ si τ est une transposition. \square

Corollaire I.3.3. — *L'application*

$$\mathfrak{S}_n \rightarrow \{\pm 1\}, \quad \sigma \mapsto \text{sgn}(\sigma)$$

est un morphisme de groupes.

Le noyau du morphisme sgn , c'est-à-dire l'ensemble des permutations paires, est appelé le **groupe alterné** et noté \mathfrak{A}_n .

Exercice I.3.4. — Montrer que le groupe alterné \mathfrak{A}_n est engendré par les 3-cycles. Montrer que si $n \geq 5$, tous les 3-cycles sont conjugués dans \mathfrak{A}_n .

Nous allons décrire les classes de conjugaison dans le groupe symétrique \mathfrak{S}_n . Rappelons qu'une **partition** de l'entier n est une collection d'entiers ≥ 1 (avec répétitions) $\{n_1, \dots, n_k\}$ tel que

$$n = n_1 + \dots + n_k.$$

On note souvent une partition en ordonnant les n_i dans l'ordre décroissant :

$$\lambda = (n_1, \dots, n_k)$$

avec $n_1 \geq n_2 \geq \dots \geq n_k$, $n = n_1 + \dots + n_k$. Une autre notation souvent utilisée pour une partition est d'indiquer, pour chaque entier $1, 2, \dots$ la multiplicité avec lequel celui-ci intervient dans la partition par un exposant (en omettant les entiers n'intervenant pas) . Par exemple, la notation

$$\lambda = (1^2, 2^2, 4^1)$$

désigne la partition

$$10 = 4 + 2 + 2 + 1 + 1.$$

A chaque élément $\sigma \in \mathfrak{S}_n$, on associe une partition de n donnée par les longueurs des cycles dans la décomposition en cycles de σ . Par exemple, $\sigma = (178)(2)(4536)(9)$ donne la partition $9 = 4 + 3 + 1 + 1$.

Théorème I.3.5. — Deux permutations σ et τ de \mathfrak{S}_n sont conjuguées si et seulement si les partitions de n données par leur décomposition en cycles sont les mêmes.

La démonstration est laissée en exercice.

Exercice I.3.6. — Calculer le cardinal de la classe de conjugaison de \mathfrak{S}_n correspondant à la partition $\lambda = (\lambda_1^{\alpha_1}, \dots, \lambda_r^{\alpha_r})$ de n .

I.4. Produit semi-direct

Cette section est donnée sous forme d'exercice.

— 1. *Le point de vue interne.* Soient G un groupe et H et N deux sous-groupes de G vérifiant :

- (a) N est distingué dans G ,
- (b) H et N engendrent G , $H \cap N = \{e\}$.

Montrer que tout élément $g \in G$ se décompose de manière unique sous la forme $g = nh$, $n \in N$, $h \in H$. En déduire que l'on a une bijection

$$N \times H \rightarrow G, \quad (n, h) \mapsto nh.$$

Montrer que la loi de groupe sur $N \times H$ induite de celle de G par transport de structure est

$$(n, h)(n', h') = (n(hn'h^{-1}), hh').$$

— 2. *Le point de vue externe.* Soient H et N deux groupes, et supposons que H agisse sur N par automorphismes de groupe, c'est-à-dire que l'on dispose d'un morphisme de groupes

$$\phi : H \rightarrow \mathbf{Aut}(N),$$

et l'on pose $h \cdot n = \phi(h)(n)$. On définit sur $N \times H$ le produit

$$(n, h)(n', h') = (n(h \cdot n'), hh').$$

Montrer que $N \times H$ muni de ce produit est un groupe, que l'on appelle le produit semi-direct de N et H , et que l'on note $N \rtimes H$. Vérifier que les parties $N \times \{e_H\}$ et $\{e_N\} \times H$ sont deux sous-groupes de $N \rtimes H$, respectivement isomorphes à N et H . On identifie ainsi N et H à deux sous-groupes de $N \rtimes H$. Montrer qu'ils vérifient les hypothèses du 1.

— 3. *Extensions.* Une suite exacte de groupes est une suite de groupes G_i , et de morphismes $\phi_i : G_i \rightarrow G_{i+1}$,

$$\cdots G_{i-1} \xrightarrow{\phi_{i-1}} G_i \xrightarrow{\phi_i} G_{i+1} \xrightarrow{\phi_{i+1}} G_{i+2} \xrightarrow{\phi_{i+2}} \cdots$$

telle que pour tout i , $\ker \phi_{i+1} = \text{Im } \phi_i$. Une suite exacte courte est une suite exacte de la forme

$$\{e\} \rightarrow N \xrightarrow{\phi} G \xrightarrow{\psi} H \rightarrow \{e\}.$$

Le morphisme ϕ est injectif, et ψ est surjectif.

Supposons que soit donnée une suite exacte courte comme ci-dessus. Une section de cette suite exacte est un morphisme de groupes $s : H \rightarrow G$ tel que $\psi \circ s = \text{Id}_H$.

Montrer que s est injective. Montrer que $\phi(N)$ et $s(H)$ sont deux sous-groupes vérifiant les hypothèses du 1. Faire le lien avec le point de vue externe.

— 4. *Exemples.* Montrer que le groupe diédral D_n est isomorphe au produit semi-direct $\mathbb{Z}/n\mathbb{Z} \rtimes \mathbb{Z}/2\mathbb{Z}$. Déterminer les classes de conjugaison de D_n . Montrer que le groupe $E(2)$ des isométries affines du plan est le produit semi-direct $\mathbb{R}^2 \rtimes \mathbf{O}(2)$. Chercher dans la littérature ou sur internet la définition du groupe de Poincaré.

I.5. Représentations

Dans ce chapitre, les espaces vectoriels sont définis sur le corps des nombres complexes.

Rappelons que si V est un espace vectoriel, $\mathbf{GL}(V)$ désigne le groupe des isomorphismes linéaires de V dans lui-même. Si V est de plus un espace de Hilbert pour le produit hermitien $(\cdot | \cdot)_V$, $\mathbf{U}(V)$ désigne le sous-groupe de $\mathbf{GL}(V)$ des applications linéaires u préservant le produit hermitien, c'est-à-dire

$$(u(v) | u(w))_V = (v | w)_V, \quad (v, w \in V).$$

I.5.1. Premières définitions. Représentations unitaires. — Soit G un groupe.

Définition I.5.1. — Une représentation (ρ, V) du groupe G est la donnée d'un espace vectoriel V , appelé espace de la représentation, et d'un morphisme de groupes

$$\rho : G \rightarrow \mathbf{GL}(V).$$

Si V est un espace de Hilbert pour le produit hermitien $(\cdot | \cdot)_V$, la représentation (ρ, V) est dite **unitaire** si ρ est à valeurs dans $\mathbf{U}(V)$, c'est-à-dire si pour tout $g \in G$, pour tous $v, w \in V$,

$$(\rho(g) \cdot v | \rho(g) \cdot w)_V = (v | w)_V.$$

La dimension de la représentation (ρ, V) est la dimension de V . On la note d_ρ .

La représentation triviale de G est celle où $V = \mathbb{C}$ et tout $g \in G$ agit comme l'identité de \mathbb{C} .

L'espace vectoriel $\{0\}$ est aussi un espace de représentation pour tout groupe G (de manière unique, puisque $\mathbf{GL}(\{0\})$ est le groupe à un élément). Nous l'appellerons **représentation nulle** de G .

Théorème I.5.2. — Soit (ρ, V) une représentation de dimension finie de G d'un groupe fini. On peut munir V d'un produit hermitien $(\cdot | \cdot)_V$ qui rend la représentation (ρ, V) unitaire.

Démonstration. Munissons V d'un produit hermitien $(\cdot | \cdot)_0$ quelconque. Définissons un nouveau produit hermitien $(\cdot | \cdot)_1$ par

$$(v, w)_1 = \frac{1}{|G|} \sum_{g \in G} (\rho(g) \cdot v | \rho(g) \cdot w)_0, \quad (v, w \in V).$$

Ce nouveau produit vérifie les propriétés de sesquilinearité requises et est positif, comme on peut le voir immédiatement. Il est défini car si

$$(v|v)_1 = \frac{1}{|G|} \sum_{g \in G} (\rho(g) \cdot v | \rho(g) \cdot v)_0 = 0$$

alors tous les termes de la somme étant positifs, ils sont nuls. Pour $g = e$, ceci donne $(v|v)_0 = 0$, et donc $v = 0$.

Vérifions que ce nouveau produit hermitien est invariant par ρ . Pour tout $h \in H$:

$$\begin{aligned} (\rho(h) \cdot v | \rho(h) \cdot w)_1 &= \frac{1}{|G|} \sum_{g \in G} (\rho(g) \cdot \rho(h) \cdot v | \rho(g) \cdot \rho(h) \cdot w)_0 \\ &= \frac{1}{|G|} \sum_{g \in G} (\rho(gh) \cdot v | \rho(gh) \cdot w)_0 \\ &= \frac{1}{|G|} \sum_{g \in G} (\rho(g) \cdot v | \rho(g) \cdot w)_0 \\ &= (v|w)_1 \end{aligned}$$

Le point crucial du calcul est donc juste un changement de variable dans la somme.

Remarquons que l'hypothèse de la dimension finie ne sert qu'à s'assurer que V est bien un espace de Hilbert. Si l'on suppose au départ que $(V, (\cdot | \cdot)_0)$ est un espace de Hilbert de dimension infinie, le même procédé de moyenne donne un nouveau produit hermitien $(\cdot | \cdot)_1$ invariant par G . Il est facile de voir que la topologie définie par ce nouveau produit hermitien est la même que l'ancienne (les normes induites sont équivalentes), et donc que V est encore un espace de Hilbert pour $(\cdot | \cdot)_1$. \square

Remarque I.5.3. — Pour des groupes plus généraux que les groupes finis, il nous faudra remplacer les arguments basés sur ce procédé de moyenne par quelque chose de plus général, à savoir l'existence de mesure de Haar sur les groupes (topologiques localement compacts). Nous

ne définissons pas la notion de mesure de Haar pour l'instant, mais nous remarquons simplement que l'on peut munir l'ensemble fini G de sa **mesure de comptage normalisée** μ_G . Plus explicitement, pour toute fonction f sur G

$$\int_G f(g) d\mu_G(g) = \frac{1}{|G|} \sum_{g \in G} f(g).$$

La propriété fondamentale de cette mesure est que quels que soient x, y dans G ,

$$\int_G f(x^{-1}gy) d\mu_G(g) = \int_G f(g) d\mu_G(g),$$

c'est-à-dire que μ_G est invariante par translation à gauche et à droite.

Dans la suite de ce chapitre, les groupes finis sont toujours munis de leurs mesures de comptage normalisée.

I.5.2. Sous-représentations, représentations irréductibles. — Soit (ρ, V) une représentation du groupe G . Un sous-espace W de V est dit invariant par ρ si pour tout $g \in G$, $\rho(g) \cdot W \subset W$. On peut alors parler de la restriction de ρ à W , que l'on note $(\rho|_W, W)$. Une telle représentation restreinte à un sous-espace invariant s'appelle une **sous-représentation** de G .

Définition I.5.4. — Une représentation (ρ, V) du groupe G est dite **irréductible** si elle est non nulle n'admet aucun sous-espace autre que $\{0\}$ et V invariant par ρ .

Proposition I.5.5. — Une représentation irréductible d'un groupe fini est de dimension finie.

Démonstration. Soit (ρ, V) une représentation irréductible du groupe fini G . Soit $v \in V$, non nul, et soit W le sous-espace engendré par les vecteurs de la forme $\rho(g) \cdot v$, $g \in G$. Ce sous-espace est donc de dimension finie, et il est immédiat de vérifier qu'il est invariant par ρ . On a donc $W = V$ et V est de dimension finie. \square

Soit (ρ, V) une représentation du groupe G et supposons que l'espace V soit somme directe de sous-espaces W_i (non nuls), $i = 1, \dots, r$:

$$V = \bigoplus_{i=1, \dots, r} W_i$$

et que ces espaces W_i soient invariants par ρ . On dit alors que la représentation (ρ, V) se décompose en somme directe des représentations $(\rho|_{W_i}, W_i)$ et l'on écrit

$$(\rho, V) = \bigoplus_{i=1, \dots, r} (\rho|_{W_i}, W_i).$$

L'étude de la représentation (ρ, V) se ramène alors à celle des $(\rho|_{W_i}, W_i)$. Il paraît raisonnable d'espérer pouvoir décomposer toute représentation en somme directe de représentations, jusqu'à ce que toutes celles-ci soient irréductibles. Ceci n'est pourtant pas totalement évident, même pour des représentations de dimension finie, le problème étant le suivant : si (ρ, V) est une représentation qui n'est pas irréductible, alors il existe un sous-espace W invariant par ρ . Pour pouvoir décomposer (ρ, V) , il faudrait pouvoir exhiber un supplémentaire de W dans V qui

soit lui aussi invariant par ρ . Le théorème ci-dessous affirme que pour les représentations d'un groupe fini, ceci est toujours possible. Pour des représentations plus générales, ce n'est pas le cas. Il est donc utile d'introduire la terminologie **représentation indécomposable** pour une représentation qui ne peut pas s'écrire comme somme directe non triviale. Une représentation irréductible est toujours indécomposable, l'inverse n'étant pas vrai en général (mais l'est pour les représentations des groupes finis).

Théorème I.5.6. — Soient G un groupe fini et (ρ, V) une représentation de dimension finie de G . Soit W un sous-espace de V invariant par ρ . Alors W admet un supplémentaire invariant W' , de sorte que l'on peut décomposer (ρ, V) en somme directe de $(\rho|_W, W)$ et $(\rho|_{W'}, W')$.

Démonstration. D'après le théorème I.5.2, on peut munir V d'un produit hermitien invariant $(\cdot | \cdot)_V$. Il est alors immédiat de voir que l'orthogonal W^\perp de W dans V pour ce produit hermitien est invariant par ρ . Ceci fournit une décomposition

$$V = W \oplus W^\perp$$

en somme directe de sous-espaces invariants. □

Corollaire I.5.7. — Toute représentation de dimension finie (ρ, V) d'un groupe fini G se décompose en somme directe de représentations irréductibles.

Démonstration. Ceci est facile à établir par récurrence sur la dimension de la représentation. Remarquons que le fait que le groupe soit fini permet de montrer l'existence d'un supplémentaire stable, et le fait que la représentation soit de dimension finie permet la récurrence. □

Une représentation est dite **complètement réductible**, ou **semi-simple** si elle peut s'écrire comme somme directe de représentations irréductibles. Le corollaire affirme que toute représentation de dimension finie d'un groupe fini est complètement réductible. Ceci permet de réduire dans une certaine mesure l'étude des représentations de dimension finie du groupe G à celle des représentations irréductibles.

Exemple I.5.8. — La théorie des représentations des groupes est une généralisation de l'algèbre linéaire. Soit $g \in \mathbf{GL}(n, \mathbb{C})$, et considérons le sous-groupe G de $\mathbf{GL}(n, \mathbb{C})$ engendré par g ,

$$G = \{g^n, n \in \mathbb{Z}\}.$$

L'espace $E = \mathbb{C}^n$ est naturellement une représentation de G . Supposons que g soit diagonalisable, c'est-à-dire que E est somme directe de sous-espace de dimension 1, stabilisés par g , et donc par tous les éléments de G . On obtient donc ainsi une décomposition de E en somme directe de représentations irréductibles (de dimension 1) de G .

En revanche, si G n'est pas diagonalisable, chaque sous-espace caractéristique de g est stable sous l'action de G , mais si cet espace n'est pas un sous-espace propre, il n'est pas complètement réductible.

I.5.3. Opérateurs d'entrelacement. Lemme de Schur. —

Définition I.5.9. — Soient (ρ, V) et (τ, W) deux représentations du groupe G . Un opérateur d'entrelacement $T : V \rightarrow W$ est une application linéaire de V dans W vérifiant

$$T(\rho(g) \cdot v) = \tau(g) \cdot T(v), \quad (g \in G), (v \in V)$$

Autrement dit, un opérateur d'entrelacement est un G -morphisme linéaire (cf. Définition I.1.4). On dit aussi que T est G -**équivariant**.

On note $\text{Hom}_G(V, W)$ ou parfois $\text{Hom}_G(\rho, \tau)$ l'ensemble des opérateurs d'entrelacement entre (ρ, V) et (τ, W) . Il est clair que c'est un sous-espace vectoriel de l'espace des applications linéaires de V vers W .

Il devient maintenant possible de définir la notion de **représentations équivalentes** ou **isomorphes**.

Définition I.5.10. — Soient (ρ, V) et (τ, W) deux représentations du groupe G . Elles sont équivalentes (ou isomorphes) s'il existe un opérateur d'entrelacement inversible $T : V \rightarrow W$.

Si T est un tel opérateur d'entrelacement inversible, T^{-1} est bien sûr aussi un opérateur d'entrelacement et

$$\tau(g) = T \circ \rho(g) \circ T^{-1}, \quad (g \in G)$$

L'équivalence dans le sens défini ci-dessus est une relation d'équivalence sur l'ensemble des représentations du groupe G . Dans la pratique, comme souvent en mathématique, on a tendance à confondre équivalence et égalité, c'est-à-dire à confondre une représentation et sa classe d'équivalence, ou dans le sens contraire, une classe d'équivalence et l'un de ses représentants. Il s'agit là d'abus de langage la plupart du temps inoffensifs.

Lemme I.5.11. — (i) Soient (ρ, V) et (τ, W) deux représentations d'un groupe G et $T : V \rightarrow W$ un opérateur d'entrelacement. Alors $\ker T$ est un sous-espace de V invariant par ρ , et $\text{Im } T$ est un sous-espace de W invariant par τ .

(ii) Soit (ρ, V) une représentation du groupe G , et T un opérateur d'entrelacement de (ρ, V) avec elle-même. Alors tout sous-espace propre de T est invariant par ρ .

Démonstration. (i) Si $v \in \ker T$, alors, pour tout $g \in G$,

$$T(\rho(g) \cdot v) = \tau(g) \cdot T(v) = 0$$

donc $\rho(g) \cdot v \in \ker T$. Si $w \in \text{Im } T$, il existe $v \in V$ tel que $T(v) = w$, et pour tout $g \in G$,

$$\tau(g) \cdot w = \tau(g) \cdot T(v) = T(\rho(g) \cdot v)$$

donc $\tau(g) \cdot w \in \text{Im } T$. □

(ii) Soit λ une valeur propre de T , et V_λ le sous-espace propre correspondant. Alors pour tout $g \in G$, pour tout $v \in V_\lambda$,

$$T(\rho(g) \cdot v) = \rho(g) \cdot T(v) = \lambda \rho(g) \cdot v$$

et donc $\rho(g) \cdot v \in V_\lambda$. □

Théorème I.5.12 (Lemme de Schur). — Soit T un opérateur d'entrelacement entre deux représentations irréductibles (ρ_1, V_1) et (ρ_2, V_2) d'un groupe G . Alors

- si (ρ_1, V_1) et (ρ_2, V_2) ne sont pas équivalentes, $T = 0$,
- si (ρ_1, V_1) et (ρ_2, V_2) sont équivalentes et de dimension finie, $\text{Hom}_G(V_1, V_2)$ est de dimension 1. De manière équivalente, $\text{Hom}_G(V_1, V_1)$ est l'ensemble des multiples scalaires de l'identité de V_1 .

Démonstration. Ceci découle facilement du lemme précédent. En effet, si (ρ_1, V_1) et (ρ_2, V_2) ne sont pas équivalentes, T n'est pas inversible. S'il n'est pas injectif, son noyau est non trivial. Mais (ρ_1, V_1) étant irréductible, ceci donne $\ker T = V_1$, et donc $T = 0$. De même, s'il n'est pas surjectif, son image $\text{Im } T$ est un sous-espace invariant de V_2 , et donc V_2 étant irréductible, $\text{Im } T = \{0\}$, donc $T = 0$.

Pour le second point, soit $T \in \text{Hom}_G(V_1, V_1)$, considérons une valeur propre λ de T , et soit V_λ le sous-espace de V_1 correspondant (c'est ici qu'intervient l'hypothèse de dimension finie, il faut pouvoir assurer l'existence d'un sous-espace propre non trivial). Il est non nul par hypothèse, et donc par irréductibilité de (ρ_1, V_1) , c'est V_1 tout entier. Ceci montre que $T = \lambda \text{Id}_{V_1}$. L'équivalence entre les deux formulations du second point vient du fait que si $S : V_1 \rightarrow V_2$ est un opérateur d'entrelacement inversible réalisant l'équivalence entre (ρ_1, V_1) et (ρ_2, V_2) , il est clair que

$$\begin{aligned} \text{Hom}_G(V_1, V_1) &\rightarrow \text{Hom}_G(V_1, V_2) \\ T &\mapsto S \circ T \end{aligned}$$

est un isomorphisme linéaire d'inverse donné par $T \mapsto S^{-1} \circ T$. □

I.6. Opérations sur les représentations : sommes directes et produits, produits tensoriels, représentation contragrédiente

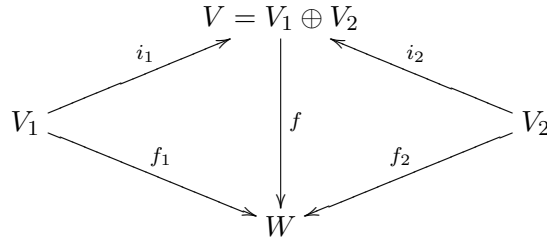
Dans ce qui suit, on ne fait pas d'hypothèses sur G , qui est un groupe quelconque. Les représentations de ce groupe ne sont pas non plus supposées de dimension finie.

I.6.1. Sommes directes et produits. — Nous avons vu dans la section I.5.2 comment une représentation (ρ, V) d'un groupe G pouvait parfois se décomposer en somme directe de sous-représentations (point de vue interne). Voyons maintenant comment former la somme directe de deux représentations de G n'ayant a priori rien à voir l'une avec l'autre (point de vue externe). Soient donc (ρ_1, V_1) et (ρ_2, V_2) deux représentations de G . La somme directe de V_1 et V_2 est un espace vectoriel V , muni de deux G -morphisms

$$i_1 : V_1 \rightarrow V, \quad i_2 : V_2 \rightarrow V,$$

vérifiant la propriété universelle suivante : pour toute représentation (τ, W) de G et toute paire de G -morphisms $f_1 : V_1 \rightarrow W$, $f_2 : V_2 \rightarrow W$, il existe un unique G -morphisme $f : V \rightarrow W$ tel que $f_1 = f \circ i_1$ et $f_2 = f \circ i_2$. On note $V = V_1 \oplus V_2$.

(I.6.1)



Cette définition, un peu abstraite, appelle plusieurs remarques :

— Il y a unicité de la somme directe, à *unique isomorphisme près*. En effet, soit (V, i_1, i_2) et (V', i'_1, i'_2) deux sommes directes de V_1 et V_2 . La propriété universelle de V appliquée à $W = V'$ et $f_1 = i'_1, f_2 = i'_2$ donne un unique G -morphisme $f : V \rightarrow V'$ tel que $i'_1 = f \circ i_1$ et $i'_2 = f \circ i_2$. En renversant les rôles de V et V' , on obtient de même un unique G -morphisme $f' : V' \rightarrow V$ tel que $i_1 = f' \circ i'_1$ et $i_2 = f' \circ i'_2$.

Considérons maintenant $f' \circ f : V \rightarrow V$. Il vérifie $f' \circ f \circ i_1 = f' \circ i'_1 = i_1$ et $f' \circ f \circ i_2 = f' \circ i'_2 = i_2$. Or un autre G -morphisme de V dans V satisfait aux mêmes propriétés, il s'agit de l'identité de V . La condition d'unicité de la propriété universelle de V appliquée à $W = V$ et $f_1 = i_1, f_2 = i_2$ nous donne alors $f' \circ f = \text{Id}_V$. En renversant les rôles de V et V' , on obtient de même $f \circ f' = \text{Id}_{V'}$. Ceci montre que V et V' sont isomorphes, et que l'isomorphisme entre eux est unique, et justifie l'abus de langage courant consistant à parler de « la » somme directe de V_1 et V_2 .

— La somme directe existe : on prend le produit ensembliste usuel $V_1 \times V_2$ de V_1 et V_2 , muni de la structure d'espace vectoriel produit et des injections

$$i_1 : V_1 \rightarrow V_1 \times V_2, \quad v_1 \mapsto (v_1, 0)$$

$$i_2 : V_2 \rightarrow V_1 \times V_2, \quad v_2 \mapsto (0, v_2).$$

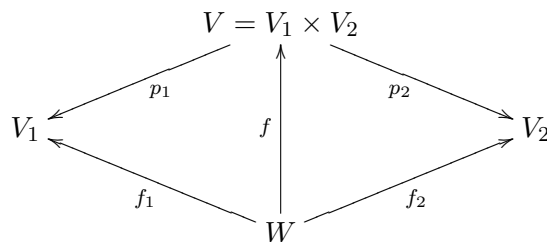
L'action ρ de G sur $V_1 \times V_2$ étant donné par

$$\rho(g) \cdot (v_1, v_2) = (\rho_1(g) \cdot v_1, \rho_2(g) \cdot v_2).$$

On vérifie aisément que cette construction donne bien un objet vérifiant la propriété universelle voulue.

Le produit ensembliste usuel $V = V_1 \times V_2$ muni de la structure vectorielle produit, de la structure de représentation de G donnée ci-dessus et des projections canoniques : $p_1 : V_1 \times V_2 \rightarrow V_1, p_2 : V_1 \times V_2 \rightarrow V_2$ vérifie aussi une propriété universelle, à savoir que pour toute représentation (τ, W) de G et toute paire de G -morphisms $f_1 : W \rightarrow V_1, f_2 : W \rightarrow V_2$, il existe un unique G -morphisme $f : W \rightarrow V$ tel que $f_1 = p_1 \circ f$ et $f_2 = p_2 \circ f$.

(I.6.2)



Une représentation (π, V) de G vérifiant cette propriété universelle est appelé **produit direct** des représentations (π_1, V_1) et (π_2, V_2) . De même que pour la somme directe, deux espaces V et V' vérifiant cette propriété universelle sont isomorphes, l'isomorphisme étant unique. Ceci justifie l'abus de langage consistant à parler « du » produit de V_1 et V_2 .

Résumons la discussion ci-dessus : nous disposons de deux notions bien distinctes, celle de somme directe et celle de produit direct de deux représentations (π_1, V_1) et (π_2, V_2) de G . Ce sont des représentations de G , vérifiant chacune une propriété universelle différente (bien que proche : la seconde est obtenue de la première en inversant le sens des flèches dans le diagramme I.6.1). Le fait que l'on puisse construire la somme directe ou le produit de la même façon (à partir du produit ensembliste usuel, comme expliqué ci-dessus) ne doit pas masquer cette différence, et d'ailleurs, lorsqu'on généralise à une famille infinie de représentations, les représentations obtenues comme somme directe et produit direct ne sont plus isomorphes.

Généralisons tout ceci à une famille quelconque de représentations (ρ_i, V_i) de G , i variant dans un ensemble d'indices I . La somme directe des $(\rho_i, V_i)_{i \in I}$ est une représentation notée $(\rho, V) = \bigoplus_{i \in I} (\rho_i, V_i)$ munie de G -morphisms $i_i : V_i \rightarrow V$ vérifiant la propriété universelle suivante :

pour toute représentation (τ, W) de G et toute famille de G -morphisms $f_i : V_i \rightarrow W$, $i \in I$, il existe un unique G -morphisme $f : V \rightarrow W$ tel que $f_i = f \circ i_i$ pour tout $i \in I$.

Le produit direct des $(\rho_i, V_i)_{i \in I}$ est une représentation notée $(\rho, V) = \prod_{i \in I} (\rho_i, V_i)$ munie de G -morphisms $p_i : V \rightarrow V_i$ vérifiant la propriété universelle suivante :

pour toute représentation (τ, W) de G et toute famille de G -morphisms $f_i : W \rightarrow V_i$, $i \in I$, il existe un unique G -morphisme $f : W \rightarrow V$ tel que $f_i = i_i \circ f$ pour tout $i \in I$.

Comme précédemment, somme directe et produit direct sont uniques à unique isomorphisme près, et on peut les construire ensemblistement de la manière suivante : pour le produit direct, on prend pour V le produit direct ensembliste des V_i , $i \in I$, muni de sa structure d'espace vectoriel canonique et des projections canoniques p_i . L'action de G sur un élément de V est donné par l'action de G sur chaque facteur :

$$\rho(g) \cdot (v_i)_{i \in I} = (\rho_i(g) \cdot v_i)_{i \in I}.$$

Pour la somme directe, on prend pour V le sous-espace du produit ensembliste des V_i constitués des familles $(v_i)_{i \in I}$ où seul un nombre fini de v_i sont non nuls. C'est un sous-espace vectoriel du produit ensembliste des V_i , et même une sous-représentation de G , comme on le vérifie facilement. Les morphismes i_i sont les inclusions canoniques $i_i : V_i \rightarrow V$.

Les propriétés universelles de la somme et du produit direct d'une famille de représentations peuvent se réécrire de la manière suivante.

Théorème I.6.1. — Soit $(\rho_i, V_i)_{i \in I}$ une famille de représentations du groupe G . Pour toute représentation (τ, W) de G , on a

$$(I.6.3) \quad \text{Hom}_G(\bigoplus_{i \in I} V_i, W) \simeq \prod_{i \in I} \text{Hom}_G(V_i, W),$$

$$(I.6.4) \quad \text{Hom}_G(W, \prod_{i \in I} V_i) \simeq \prod_{i \in I} \text{Hom}_G(W, V_i).$$

I.6.2. Produits tensoriels. — Dans le paragraphe précédent, nous avons muni l'ensemble des représentations d'un groupe G d'une somme :

$$((\rho_1, V_1), (\rho_2, V_2)) \mapsto (\rho_1 \oplus \rho_2, V_1 \oplus V_2),$$

La terminologie et la notation « additive » se justifient par le fait que $(\rho_1 \oplus \rho_2, V_1 \oplus V_2)$ est toujours isomorphe à $(\rho_2 \oplus \rho_1, V_2 \oplus V_1)$ et que

$$\dim(V_1 \oplus V_2) = \dim V_1 + \dim V_2,$$

lorsque V_1 et V_2 sont de dimension finie. La représentation de G dans l'espace nul $\{0\}$ est un « élément neutre » pour cette somme. Mais remarquons qu'une représentation (ρ, V) non nulle n'admet pas d'inverse.

Nous voudrions maintenant construire une opération analogue à un produit :

$$((\rho_1, V_1), (\rho_2, V_2)) \mapsto (\rho_1 \otimes \rho_2, V_1 \otimes V_2)$$

ayant de bonnes propriétés de distributivité par rapport à la somme définie précédemment, et vérifiant

$$(I.6.5) \quad \dim(V_1 \otimes V_2) = \dim V_1 \times \dim V_2,$$

lorsque V_1 et V_2 sont de dimension finie.

Définition I.6.2. — Soient V_1 et V_2 deux espaces vectoriels. Le **produit tensoriel** $V_1 \otimes V_2$ est un espace vectoriel muni d'une application

$$\iota : V_1 \times V_2 \rightarrow V_1 \otimes V_2, \quad (v_1, v_2) \mapsto v_1 \otimes v_2$$

vérifiant :

(i) ι est bilinéaire,

(ii) si $(e_i)_{i \in I}$ est une base de V_1 et si $(f_j)_{j \in J}$ est une base de V_2 ,

$$(e_i \otimes f_j)_{i \in I, j \in J}$$

est une base de $V_1 \otimes V_2$.

Remarques I.6.3. — Un tel espace existe et est déterminé à isomorphisme près. La propriété (ii) entraîne la formule (I.6.5) lorsque les espaces sont de dimension finie. L'espace $V_1 \otimes V_2$ vérifie la propriété universelle suivante :

(ii') soit $\phi : V_1 \times V_2 \rightarrow W$ une application bilinéaire quelconque. Alors il existe une unique application linéaire

$$\tilde{\phi} : V_1 \otimes V_2 \rightarrow W$$

tel que $\tilde{\phi} \circ \iota = \phi$.

La propriété (ii') est équivalente à (ii). Comme dans le paragraphe précédent pour les sommes et produits directs la propriété universelle garantit l'unicité du produit tensoriel à unique isomorphisme près, et justifie l'abus de langage consistant à parler « du » produit tensoriel. Nous

ne donnons pas de détails concernant la construction du produit tensoriel, mais celle-ci est, nous semble-t-il, claire si l'on considère la propriété (ii), et que l'on se débarrasse de ses scrupules à utiliser l'axiome du choix.

Soient (ρ_1, V_1) une représentation d'un groupe G_1 , et (ρ_2, V_2) une représentation d'un groupe G_2 . On peut munir l'espace $V_1 \otimes V_2$ d'une représentation notée $\rho_1 \boxtimes \rho_2$ de $G_1 \times G_2$. Une définition évidente est de poser

$$(I.6.6) \quad \begin{aligned} (\rho_1 \boxtimes \rho_2)(g_1, g_2) \cdot (v_1 \otimes v_2) &= \rho_1(g_1) \cdot v_1 \otimes \rho_2(g_2) \cdot v_2 \\ (v_1 \in V_1), (v_2 \in V_2), (g_1 \in G_1), (g_2 \in G_2). \end{aligned}$$

Comme $V_1 \otimes V_2$ est engendré par les $v_1 \otimes v_2$, $v_1 \in V_1$, $v_2 \in V_2$, par linéarité, ces formules suffisent à définir l'opérateur $(\rho_1 \boxtimes \rho_2)(g_1, g_2)$ sur $V_1 \otimes V_2$, pour peu que l'on ait montré que si un vecteur de $V_1 \otimes V_2$ se décompose de deux manières en tenseurs élémentaires

$$\sum_i v_i \otimes w_i = \sum_j v'_j \otimes w'_j$$

alors

$$\sum_i \rho_1(g_1) \cdot v_i \otimes \rho_2(g_2) \cdot w_i = \sum_j \rho_1(g_1) \cdot v'_j \otimes \rho_2(g_2) \cdot w'_j.$$

C'est en fait une conséquence de la propriété universelle du produit tensoriel. En effet, considérons l'application bilinéaire :

$$V_1 \times V_2, \quad (v_1, v_2) \mapsto \rho_1(g_1) \cdot v_1 \otimes \rho_2(g_2) \cdot v_2.$$

D'après la propriété universelle, il existe un endomorphisme (unique) de $V_1 \otimes V_2$ (c'est le $(\rho_1 \boxtimes \rho_2)(g_1, g_2)$ que l'on cherche et c'est donc ainsi qu'on le note) vérifiant (I.6.6).

On vérifie ensuite facilement que $(\rho_1 \boxtimes \rho_2)$ est une représentation de $G_1 \times G_2$ dans $V_1 \otimes V_2$.

Lorsque $G_1 = G_2$ on obtient une représentation de G , notée $\rho_1 \otimes \rho_2$ définie par

$$\begin{aligned} (\rho_1 \otimes \rho_2)(g) \cdot (v_1 \otimes v_2) &= \rho_1(g) \cdot v_1 \otimes \rho_2(g) \cdot v_2 \\ (v_1 \in V_1), (v_2 \in V_2), (g \in G). \end{aligned}$$

I.6.3. Représentation contragrédiente. — Si V est un espace vectoriel, notons V^* son dual, c'est-à-dire l'espace des formes linéaires sur V . Il est bien connu que

$$\mathbf{ev} : V \rightarrow (V^*)^*, \quad \mathbf{ev}(v) : \lambda \in V^* \mapsto \lambda(v)$$

est une application linéaire, injective. Si V est de dimension finie, par égalité des dimensions, cette application est un isomorphisme.

Si (π, V) est une représentation d'un groupe G , on définit une représentation $\tilde{\pi}$ de G dans V^* , appelée **représentation contragrédiente**, par la formule suivante :

$$(\tilde{\pi}(g) \cdot \lambda)(v) = \lambda(\pi(g)^{-1} \cdot v), \quad (\lambda \in V^*), (v \in V), (g \in G).$$

Il est clair que $(\tilde{\tilde{\pi}}, (V^*)^*) = (\pi, V)$ lorsque V est de dimension finie et que $(V^*)^*$ est identifié à V par la remarque ci-dessus.

Proposition I.6.4. — Soit (π, V) une représentation d'un groupe fini G . Alors (π, V) est irréductible si et seulement si $(\tilde{\pi}, V^*)$ est irréductible.

Démonstration. D'après la remarque que $(\tilde{\pi}, (V^*)^*) = (\pi, V)$ lorsque V est de dimension finie, il suffit de montrer une seule implication pour obtenir l'équivalence. Supposons (π, V) irréductible, et soit W un sous-espace invariant de V^* . Alors l'orthogonal dans V de W est aussi invariant, et donc ne peut-être que $\{0\}$ ou V . Ceci montre que $W = \{0\}$ ou V^* . \square

I.7. Décomposition canonique et règles de sélection.

Soit (ρ, V) une représentation de dimension finie du groupe fini G . Nous avons vu que (ρ, V) est semi-simple (ou complètement réductible) ce qui signifie que l'on peut écrire (ρ, V) comme la somme directe de représentations irréductibles (ρ_i, W_i) . Cette décomposition en représentation irréductible n'est en général pas unique. Nous allons expliquer comment regrouper les (ρ_i, W_i) par « paquets », de manière à obtenir une décomposition moins fine, mais canonique. L'idée est de mettre dans le même paquet toutes les représentations (ρ_i, W_i) isomorphes.

Pour tout élément $\delta \in \widehat{G}$, soit $V'(\delta)$ le sous-espace de V obtenu comme la somme de toutes les sous-représentations irréductibles de G dans la classe δ . Cet espace s'appelle la composante isotypique de type δ de V . D'autre part, fixons une décomposition

$$V = \bigoplus_i W_i$$

de V en somme directe de sous-représentations irréductibles. Soit $V(\delta)$ la somme (directe donc) de tous les W_j dans la classe δ . Il est donc clair que $V(\delta) \subset V'(\delta)$. Fixons un représentant (π_δ, V_δ) de la classe d'équivalence $\delta \in \widehat{G}$. Notons

$$H_\delta = \text{Hom}_G(\pi_\delta, \rho).$$

Lemme I.7.1. — $\dim H_\delta$ est la multiplicité m_δ de δ dans la décomposition $V = \bigoplus_i W_i$ (c'est-à-dire le nombre de W_i dans la classe δ). En particulier, cette multiplicité ne dépend pas de la décomposition choisie.

Démonstration. On a, d'après (I.6.4), en utilisant le fait qu'une somme directe finie de représentations est isomorphe au produit direct de ces représentations :

$$H_\delta = \text{Hom}_G(\pi_\delta, \rho) = \text{Hom}_G(\pi_\delta, \bigoplus_i \rho_i) \simeq \bigoplus_i \text{Hom}_G(\pi_\delta, \rho_i)$$

et d'après le lemme de Schur, la dimension de chaque $\text{Hom}_G(\pi_\delta, \rho_i)$ est 1 si ρ_i est dans la classe δ , 0 sinon. \square

Lemme I.7.2. — Le morphisme

$$\Phi_\delta : H_\delta \otimes V_\delta \rightarrow V, \quad \phi \otimes v \mapsto \phi(v)$$

réalise un isomorphisme G -équivariant entre $H_\delta \otimes V_\delta$ et $V'(\delta)$.

De plus $V(\delta) = V'(\delta)$ et $V(\delta)$ ne dépend pas de la décomposition choisie.

Démonstration. l'action de G sur $H_\delta \otimes V_\delta$ est le produit tensoriel de l'action triviale de G sur H_δ et de l'action π_δ sur V_δ .

On vérifie sans difficulté que Φ_δ est un opérateur d'entrelacement. La surjectivité de Φ_δ est une conséquence des définitions : soit $w \in V'(\delta)$. Alors w est dans une certaine sous-représentation irréductible (ρ_W, W) de (ρ, V) dans la classe δ . Soit ϕ un isomorphisme entre π_δ et (ρ_W, W) que l'on voit comme un élément de H_δ . Il est alors clair que w est dans l'image de Φ_δ .

Puisque Φ_δ est surjective, on a

$$\dim(H_\delta \otimes V_\delta) = \dim(H_\delta) \times \dim(V_\delta) \geq \dim(V'(\delta)).$$

D'autre part

$$\dim V'(\delta) \geq \dim V(\delta) = m_\delta \dim(V_\delta) = \dim(H_\delta) \times \dim(V_\delta)$$

d'après le lemme précédent. Toute les inégalités sont donc des égalités, et l'inclusion de $V(\delta)$ dans $V'(\delta)$ est en fait aussi une égalité. Comme $V'(\delta)$ ne dépend pas de la décomposition choisie, $V(\delta)$ non plus.

On déduit de ceci le théorème de décomposition suivant :

Théorème I.7.3. — *Le morphisme*

$$\Phi : \bigoplus_{\delta \in \hat{G}} H_\delta \otimes V_\delta \rightarrow V, \quad \Phi = \sum_{\delta} \Phi_\delta$$

est un isomorphisme G -équivariant, qui envoie chaque $H_\delta \otimes V_\delta$ sur la composante isotypique $V(\delta) = V'(\delta)$ de type δ de (ρ, V) .

Théorème I.7.4 (Règles de sélection). — *Soit (ρ, V) et (τ, W) deux représentations de dimension finie du groupe G , et soit T un opérateur d'entrelacement entre ces deux représentations. Alors T envoie la composante isotypique $V(\delta)$ de V dans la composante isotypique $W(\delta)$ de W .*

Démonstration. C'est une conséquence directe de (I.6.3) et (I.6.4) et du lemme de Schur.

Remarque I.7.5. — Lorsqu'on choisit des bases de V et W adaptées aux décompositions canoniques en composantes isotypiques, la matrice de T dans ces bases est diagonale par blocs.

I.8. Représentations des groupes topologiques

La plupart des groupes de symétries apparaissant en physique ne sont pas des groupes finis. Beaucoup de ces groupes sont des groupes de matrices, munis naturellement d'une structure topologique (ce sont des variétés différentiables). Pour développer une théorie des représentations convenable pour de tels groupes, il faut tenir compte des structures topologiques.

I.8.1. Représentations des groupes topologiques. —

Définition I.8.1. — Un **groupe topologique** est à la fois un groupe et un espace topologique séparé, où l'on exige que ces deux structures soient compatibles, c'est-à-dire que les applications

$$m : G \times G \rightarrow G, \quad (g, h) \mapsto gh$$

et

$$\iota : G \rightarrow G, \quad g \mapsto g^{-1}$$

soient continues.

Il convient d'adapter le vocabulaire des représentations lorsque l'on a affaire à des groupes topologiques :

Une **représentation** (π, V) d'un groupe topologique dans un espace vectoriel topologique V est la donnée d'un morphisme

$$\pi : G \rightarrow \mathbf{GL}(V)$$

tel que l'application

$$G \times V \rightarrow V, \quad (g, v) \mapsto \pi(g) \cdot v$$

soit **continu**.

Si V est un espace de Hilbert, muni d'un produit scalaire invariant, on dit que (π, V) est unitaire.

Lorsque l'espace de représentation V est de dimension finie, on le suppose toujours muni de la topologie transcendante.

Une **sous-représentation** de (π, V) est un sous-espace **fermé** de V invariant sous l'action de G .

Une représentation (π, V) est **irréductible** si elle n'admet aucune autre sous-représentation qu'elle-même et $\{0\}$.

Soient (π_1, V_1) et (π_2, V_2) deux représentations d'un groupe topologique G . Un **opérateur d'entrelacement** T entre ces représentations est un G -morphisme linéaire **continu** $T : V_1 \rightarrow V_2$. Ceci permet de définir la notion de représentations isomorphes ou équivalentes.

Dans ce contexte, nous avons l'analogie du théorème I.5.6 :

Théorème I.8.2. — *Soit (π, V) une représentation unitaire du groupe topologique G dans un espace de Hilbert V . Soit W un sous-espace invariant fermé de V . Alors W^\perp est stable sous l'action de G et la représentation V se décompose en somme directe*

$$V = W \oplus W^\perp$$

Démonstration. La démonstration est la même que pour le théorème I.5.6, en remarquant de plus que si W est fermé, alors W^\perp aussi et que l'on a $V = W \oplus W^\perp$. \square

Il n'est pas possible d'en dire plus à ce niveau de généralité. En particulier, il n'est pas vrai :

- qu'une représentation irréductible soit toujours de dimension finie ;

- qu'une représentation (même de dimension finie) soit toujours unitaire ;
- qu'une représentation de dimension finie soit toujours complètement réductible.

Un espace topologique est dit **localement compact** si tout point admet une base de voisinages compacts. Un groupe topologique est dit localement compact (resp. compact) s'il est localement compact en tant qu'espace topologique (resp. compact).

I.9. Mesure de Haar

Dans le chapitre sur les représentations des groupes finis, nous avons fait un usage immédiat (théorème I.5.2) de la mesure de comptage normalisée sur un groupe fini, et de ses propriétés d'invariance par translation à gauche ou à droite (voir la remarque I.5.3). Définissons l'analogie pour un groupe topologique localement compact quelconque.

Rappelons qu'une mesure de Radon λ sur un espace topologique localement compact X est une forme linéaire continue sur $\mathcal{C}_c(X, \mathbb{R})$, positive (c'est-à-dire que si $f \geq 0$, $\lambda(f) \geq 0$). Un théorème d'analyse nous dit qu'une telle forme linéaire est donnée par intégration par rapport à une mesure borélienne μ

$$\lambda(f) = \int_X f d\mu$$

possédant certaines propriétés (elle est localement finie et régulière intérieurement).

Définition I.9.1. — Soit G un groupe topologique localement compact. Une mesure de Radon μ_G sur G sera dite mesure de Haar à gauche (resp. à droite) sur G , si pour toute fonction intégrable f de G à valeurs dans \mathbb{C} , et pour tout $t \in G$,

$$\int_G f(t^{-1}g) d\mu_G(g) = \int_G f(g) d\mu_G(g)$$

(resp.

$$\int_G f(gt) d\mu_G(g) = \int_G f(g) d\mu_G(g). \quad)$$

Nous admettrons le théorème suivant.

Théorème I.9.2. — Soit G un groupe localement compact. Alors il existe une mesure de Haar à gauche μ_G sur G . Une telle mesure est unique à un facteur scalaire réel positif près.

Remarque I.9.3. — Dans l'énoncé du théorème précédent, on peut remplacer « mesure de Haar à gauche » par « mesure de Haar à droite », mais les deux notions sont distinctes, une mesure de Haar à gauche n'est pas nécessairement une mesure de Haar à droite.

Théorème I.9.4. — Soit G un groupe compact. Toute mesure de Haar à gauche est aussi une mesure de Haar à droite. On peut normaliser la mesure de Haar μ_G de sorte que $\mu_G(G) = 1$.

Exemples I.9.5. — si G est un groupe fini, il est compact pour la topologie discrète, et la mesure de comptage normalisée est une mesure de Haar.

— Dans un groupe topologique localement compact abélien, toute mesure de Haar à gauche est aussi une mesure de Haar à droite.

— Une mesure de Haar sur $(\mathbb{R}, +)$ est un multiple scalaire de la mesure de Lebesgue.

— Si $G = \mathbf{U}(1) = \{z \in \mathbb{C} \mid |z| = 1\}$, une mesure de Haar est un multiple scalaire de $\frac{dz}{2i\pi z}$. Si l'on identifie $U(1)$ et $\mathbb{R}/2\pi\mathbb{Z}$ par $\theta \mapsto e^{i\theta}$, cette mesure est donnée par $\frac{d\theta}{2\pi}$.

I.9.1. Représentations des groupes compacts : unitarité et complète réductibilité.

— Nous pouvons ensuite généraliser le théorème I.5.2 aux groupes compacts.

Théorème I.9.6. — *Toute représentation d'un groupe compact G de dimension finie peut être munie d'un produit hermitien invariant, qui rend la représentation unitaire.*

Démonstration. La démonstration est la même que pour les groupes finis, grâce à l'existence de la mesure de Haar sur G . □

Corollaire I.9.7. — *Toute représentation d'un groupe compact G de dimension finie est complètement réductible.*

Démonstration. Ceci découle du théorème précédent et du théorème I.8.2 de la même façon que pour les groupes finis. □

Corollaire I.9.8. — *Si (π, V) est une représentation de dimension finie d'un groupe compact G , pour tout $g \in G$, l'opérateur $\pi(g)$ est diagonalisable.*

Tout espace vectoriel complexe de dimension finie peut être muni d'un produit hermitien qui en fait un espace de Hilbert. Le théorème I.9.6 ci-dessus montre que l'on peut, par un argument de moyenne, rendre ce produit hermitien invariant sous l'action d'un groupe compact G . Que se passe-t-il si l'on part d'une représentation (non nécessairement unitaire) de G dans un espace de Hilbert ? Peut-on la rendre unitaire ? L'argument de moyenne marche encore, et l'on obtient un produit hermitien invariant. Le problème est de vérifier que la nouvelle topologie définie par ce produit hermitien invariant coïncide avec l'ancienne, en particulier, que l'espace reste complet. La réponse à cette question est positive :

Proposition I.9.9. — *Toute représentation d'un groupe compact dans un espace de Hilbert H peut être munie d'un produit hermitien invariant, qui rend la représentation unitaire et définit la même structure topologique sur H .*

CHAPITRE II

GROUPES LINÉAIRES ET LEURS ALGÈBRES DE LIE

La plupart des groupes apparaissant comme groupes de symétries en physique sont des groupes de matrices. Ce sont des groupes topologiques, mais leur structure est en fait bien plus riche puisque ce sont des variétés différentiables. La compatibilité de la structure différentiable et de la structure de groupe donne des propriétés de rigidité très fortes. De plus, les outils du calculs différentiels seront fondamentaux dans l'étude des représentations de ces groupes.

La classe de groupes que nous allons étudier est celle des groupes de Lie linéaires. Nous allons voir comment leur associer une algèbre de Lie, et comment utiliser celle-ci pour l'étude de leurs représentations.

II.1. Le groupe $\mathbf{GL}(n, \mathbb{K})$

La notation \mathbb{K} désigne un corps pouvant être soit le corps des nombres réels \mathbb{R} , soit le corps des nombres complexes \mathbb{C} .

Soit $\mathcal{M}_n(\mathbb{K})$ l'algèbre sur \mathbb{K} des matrices carrés $n \times n$ à coefficients dans \mathbb{K} . C'est un espace de dimension n^2 , et toutes les normes sur cet espace sont donc équivalentes. On suppose \mathbb{K}^n muni de la norme usuelle provenant du produit scalaire canonique sur \mathbb{R}^n ou du produit hermitien canonique sur \mathbb{C}^n . Munissons $\mathcal{M}_n(\mathbb{K})$ de la norme d'opérateur

$$\|X\| = \sup_{x \in \mathbb{K}^n \setminus \{0\}} \frac{\|X(x)\|}{\|x\|}.$$

Cette norme vérifie

$$(II.1.1) \quad \|XY\| \leq \|X\| \|Y\|, \quad (X, Y \in \mathcal{M}_n(\mathbb{K})).$$

L'algèbre $\mathcal{M}_n(\mathbb{K})$ est donc munie d'une norme qui en fait un espace vectoriel normé complet (un espace de Banach).

Rappelons maintenant quelques propriétés du groupe des matrices inversibles $\mathbf{GL}(n, \mathbb{K})$. On note Id_n la matrice identité de $\mathcal{M}_n(\mathbb{K})$.

Proposition II.1.1. — Le groupe des matrices inversibles $\mathbf{GL}(n, \mathbb{K})$ est un groupe topologique pour la topologie induite de celle de $\mathcal{M}_n(\mathbb{K})$. De plus, $\mathbf{GL}(n, \mathbb{K})$ est un ouvert dense de $\mathcal{M}_n(\mathbb{K})$.

II.2. L'application exponentielle

Définition II.2.1. — Soit X un élément de $\mathcal{M}_n(\mathbb{K})$. l'exponentielle de X , notée $\exp X$, désigne la somme de la série (normalement convergente dans l'espace de Banach $\mathcal{M}_n(\mathbb{K})$)

$$\sum_{n=0}^{+\infty} \frac{X^n}{n!}.$$

Donnons quelques propriétés de l'exponentielle.

Proposition II.2.2. — Quels que soient X et Y dans $\mathcal{M}_n(\mathbb{K})$:

(i) Si X et Y commutent $\exp X \exp Y = \exp(X + Y)$.

(ii) L'exponentielle est à valeurs dans $\mathbf{GL}(n, \mathbb{K})$ et

$$(\exp X)^{-1} = \exp(-X).$$

(iii) Quels que soient t, s dans \mathbb{K} ,

$$\exp(sX) \exp(tX) = \exp((s + t)X)$$

(iv) L'application $\mathbb{R} \rightarrow \mathbf{GL}(n, \mathbb{K})$, $t \mapsto \exp(tX)$ est l'unique solution différentiable de l'équation différentielle du premier ordre

$$a'(t) = X a(t)$$

avec la condition initiale $a(0) = \text{Id}_n$.

(v) L'application $\mathbb{R} \rightarrow \mathbf{GL}(n, \mathbb{K})$, $t \mapsto \exp(tX)$ est l'unique solution différentiable de l'équation fonctionnelle

$$a(s)a(t) = a(s + t), \quad a(0) = \text{Id}_n, \quad a'(0) = X.$$

(vi) Pour tout $g \in \mathbf{GL}(n, \mathbb{K})$, $g \exp X g^{-1} = \exp(gXg^{-1})$.

Remarque II.2.3. — On peut reformuler (iii) en disant que

$$t \mapsto \exp tX$$

est un morphisme de groupes (continu) de \mathbb{R} dans $\mathbf{GL}(n, \mathbb{K})$. On appelle un tel morphisme un **sous-groupe à un paramètre** de $\mathbf{GL}(n, \mathbb{K})$.

Exercice II.2.4. — Montrer qu'un morphisme de groupe continu $\phi : \mathbb{R} \rightarrow \mathbf{GL}(n, \mathbb{K})$ est différentiable. En déduire que tout sous-groupe à un paramètre de $\mathbf{GL}(n, \mathbb{K})$ est de la forme $t \mapsto \exp tX$ pour un certain X dans $\mathcal{M}_n(\mathbb{K})$.

Proposition II.2.5. — L'application exponentielle

$$\exp : \mathcal{M}_n(\mathbb{K}) \longrightarrow \mathbf{GL}(n, \mathbb{K})$$

est de classe \mathcal{C}^∞ ; sa différentielle à l'origine est l'application identique de $\mathcal{M}_n(\mathbb{K})$.

Corollaire II.2.6. — Il existe un voisinage \mathcal{U} ouvert de 0 dans $\mathcal{M}_n(\mathbb{K})$ et un voisinage ouvert \mathcal{V} de l'identité dans $\mathbf{GL}(n, \mathbb{K})$ tels que l'application exponentielle réalise un difféomorphisme de \mathcal{U} sur \mathcal{V} .

Démonstration. C'est une conséquence immédiate du théorème d'inversion locale. □

II.3. Groupes linéaires

Définition II.3.1. — On appelle groupe linéaire un sous-groupe d'un groupe général linéaire $\mathbf{GL}(n, \mathbb{R})$.

Remarques II.3.2. — Tout sous-groupe d'un groupe linéaire est linéaire. Le groupe $\mathbf{GL}(n, \mathbb{C})$ est linéaire, puisqu'on peut le plonger naturellement dans $\mathbf{GL}(2n, \mathbb{R})$.

Exemples II.3.3. — Les groupes $\mathbf{SL}(n, \mathbb{K})$, $\mathbf{O}(n)$, $\mathbf{O}(p, q)$, $\mathbf{SO}(n)$, $\mathbf{SO}(p, q)$, $\mathbf{U}(n)$, $\mathbf{U}(p, q)$, $\mathbf{SU}(n)$, $\mathbf{SU}(p, q)$, $\mathbf{Sp}(2n, k)$ des exemples I.2.4, I.2.5 I.2.6 et I.2.8.

— Le groupe $\mathbf{B}(n, \mathbb{K})$ des matrices triangulaires supérieures inversibles.

— Le groupe $\mathbf{N}(n, \mathbb{K})$ des matrices des matrices triangulaires supérieures avec des 1 sur la diagonale.

— Les groupes de Heisenberg de l'exemple I.2.9.

Définition II.3.4. — Soit G un groupe linéaire, disons $G \subset \mathbf{GL}(n, \mathbb{R})$. L'espace tangent \mathfrak{g} à G en Id_n est l'espace des matrices $X \in \mathfrak{gl}(n, \mathbb{R})$ telles qu'il existe une courbe $a(t)$, définie sur un intervalle ouvert autour de 0 dans \mathbb{R} , de classe \mathcal{C}^1 , à valeurs dans G , et vérifiant $a(0) = \text{Id}_n$ et $a'(0) = X$.

Exercice II.3.5. — Montrer que cet espace tangent en l'identité pour un groupe linéaire fini est $\{0\}$.

Proposition II.3.6. — Soit G un groupe linéaire. L'espace tangent en l'identité \mathfrak{g} est un sous-espace vectoriel de $\mathfrak{gl}(n, \mathbb{R})$.

Démonstration. Soient $X, Y \in \mathfrak{g}$ et $\alpha, \beta \in \mathbb{R}$. Soient $a(t)$ et $b(t)$ des courbes de classe \mathcal{C}^1 dans G vérifiant $a(0) = b(0) = \text{Id}_n$, $a'(0) = X$, $b'(0) = Y$. Posons $c(t) = a(\alpha t)b(\beta t)$. On a bien $c(0) = \text{Id}_n$ et $c'(0) = \alpha X + \beta Y$. Ceci montre que \mathfrak{g} est un sous-espace vectoriel de $\mathfrak{gl}(n, \mathbb{R})$. \square

Donnons la propriété fondamentale de cet espace tangent. Posons, pour s suffisamment proche de 0,

$$c_s(t) = a(s)b(t)a(s)^{-1}.$$

C'est une courbe dans G vérifiant $c_s(0) = \text{Id}_n$ et $c'_s(0) = a(s)Ya(s)^{-1}$. Ceci montre que $a(s)Ya(s)^{-1} \in \mathfrak{g}$ pour tout s suffisamment proche de 0. Maintenant, $s \mapsto a(s)Ya(s)^{-1}$ est une courbe dans \mathfrak{g} , et son vecteur tangent en 0 est dans \mathfrak{g} , donc

$$\frac{d}{ds}(a(s)Ya(s)^{-1})|_{s=0} = XY - YX = [X, Y] \in \mathfrak{g}.$$

L'espace \mathfrak{g} est donc stable sous l'opération (appelée crochet de Lie) $(X, Y) \mapsto [X, Y] = XY - YX$. Rappelons les propriétés de cette opération : le crochet de Lie est bilinéaire, antisymétrique et vérifie l'**identité de Jacobi** :

$$(II.3.1) \quad [X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0.$$

Ceci nous conduit à définir la notion suivante.

Définition II.3.7. — Une algèbre de Lie L sur \mathbb{K} est un \mathbb{K} -espace vectoriel muni d'une application (appelée "crochet") :

$$[\cdot, \cdot] : L \times L \rightarrow L$$

bilinéaire antisymétrique vérifiant l'identité de Jacobi (II.3.1).

Une sous-algèbre de Lie de L est un sous-espace vectoriel de L stable par crochets.

Un idéal de l'algèbre de Lie L est un sous-espace vectoriel I de L tel que quels que soient $X \in I$ et $Y \in L$, $[X, Y] \in I$.

L'algèbre des matrices $\mathcal{M}_n(\mathbb{K})$ muni du crochet de Lie est une algèbre de Lie. On la note $\mathfrak{gl}(n, \mathbb{K})$. Plus généralement, si V est un espace vectoriel sur \mathbb{K} , $\text{End}(V)$ est une algèbre de Lie, notée $\mathfrak{gl}(V)$.

Nous pouvons maintenant être plus précis que la proposition précédente.

Proposition II.3.8. — Soit G un groupe linéaire. L'espace tangent en l'identité \mathfrak{g} est une sous-algèbre de Lie de $\mathfrak{gl}(n, \mathbb{R})$.

Exemples II.3.9. — Nous avons vu que $\mathcal{M}_n(\mathbb{K})$ est une algèbre de Lie. On la note $\mathfrak{gl}(n, \mathbb{K})$. Plus généralement, si V est un espace vectoriel sur \mathbb{K} , $\text{End}(V)$ est une algèbre de Lie, notée $\mathfrak{gl}(V)$.

— $\mathfrak{sl}(n, \mathbb{K})$ est la sous-algèbre de Lie de $\mathfrak{gl}(n, \mathbb{K})$ des matrices de trace nulle.

— $\mathfrak{b}(n, \mathbb{K})$ est la sous-algèbre de Lie de $\mathfrak{gl}(n, \mathbb{K})$ des matrices triangulaires supérieures.

- $\mathfrak{n}(n, \mathbb{K})$ est la sous-algèbre de Lie de $\mathfrak{gl}(n, \mathbb{K})$ des matrices triangulaires supérieures strictes. C'est un idéal de la précédente.
- $\mathfrak{so}(n)$ est la sous-algèbre de Lie de $\mathfrak{gl}(n, \mathbb{R})$ des matrices antisymétriques.
- $\mathfrak{su}(n)$ est la sous-algèbre de Lie de $\mathfrak{gl}(n, \mathbb{C})$ des matrices antihermitiennes.
- Soit M une variété différentiable et $\mathcal{X}(M)$ l'espace des champs de vecteurs sur M . Alors $\mathcal{X}(M)$, muni du crochet de Lie des champs de vecteurs, est une algèbre de Lie sur \mathbb{R} .

Exercice II.3.10. — Déterminer les algèbres de Lie des groupes II.3.3

Définition II.3.11. — Un morphisme d'algèbres de Lie est une application linéaire entre algèbres de Lie respectant le crochet. Un morphisme d'algèbres de Lie d'une algèbre de Lie L dans $\mathfrak{gl}(V)$ est appelée représentation de L dans l'espace vectoriel V .

Soit G un groupe linéaire. On appelle l'espace tangent à G en l'identité \mathfrak{g} l'algèbre de Lie du groupe linéaire G . Un résultat crucial dans la théorie des groupes linéaires est que l'application exponentielle renvoie l'algèbre de Lie \mathfrak{g} dans le groupe G .

Théorème II.3.12. — Soit G un groupe linéaire et soit \mathfrak{g} son algèbre de Lie. Alors l'application \exp envoie \mathfrak{g} dans G .

La démonstration se trouve par exemple dans [8].

Corollaire II.3.13. — L'algèbre de Lie \mathfrak{g} d'un groupe linéaire G dans $\mathbf{GL}(n, \mathbb{R})$ est l'ensemble des éléments X dans $\mathcal{M}_n(\mathbb{R})$ tels que $\exp tX \in G$ pour tout $t \in \mathbb{R}$.

Démonstration. Une inclusion provient du théorème, et l'autre de la définition de \mathfrak{g} en considérant les courbes $t \mapsto \exp tX$. □

Remarque II.3.14. — Une terminologie ancienne, encore parfois utilisée en physique, est celle de « générateurs infinitésimaux » du groupe G pour les éléments $X \in \mathfrak{g}$. Chaque « générateur infinitésimal » X engendre le sous-groupe à un paramètre $t \mapsto \exp tX$ de G .

II.4. Sous-groupes de Lie. Connexité

Certains groupes linéaires ont une topologie (induite de celle de $\mathbf{GL}(n, \mathbb{R})$ dans lequel ils sont plongés) peu agréable. Par exemple \mathbb{Q}^\times est un groupe linéaire, puisqu'il est inclu dans $\mathbb{R}^\times \simeq \mathbf{GL}(1, \mathbb{R})$. Il est dense dans \mathbb{R}^\times mais son algèbre de Lie est $\{0\}$. Sa topologie ne peut donc pas être décrite de manière simple.

Pour une bonne classe de groupes linéaires, comprenant tous les groupes qui nous intéressent, la situation topologique est très bonne.

Théorème II.4.1. — Soit G un sous-groupe fermé de $\mathbf{GL}(n, \mathbb{R})$. Alors une base de voisinage d'un point $g \in G$ est donnée par les parties $g \exp \mathcal{U}$, où \mathcal{U} est un voisinage de 0 dans \mathfrak{g} . Si \mathcal{U} est un ouvert contenant 0 suffisamment petit dans \mathfrak{g} , $X \mapsto g \exp X$ est un difféomorphisme de \mathcal{U} sur l'ouvert $g \exp \mathcal{U}$ de G .

Les voisinages d'un élément de G ressemblent donc à ceux d'un espace vectoriel. En termes techniques, G est muni d'une structure de **variété différentielle**. Ceci permet de définir sur G la notion de fonction différentiable, par exemple, et plus généralement, d'étendre toute les notions du calcul différentiel sur \mathbb{R}^n . Le produit et le passage à l'inverse sont \mathcal{C}^∞ . On dit que G est un **groupe de Lie**.

Nous avons vu que les groupes finis sont linéaires. Or, il est clair d'après la définition que l'algèbre de Lie d'un groupe fini est l'espace vectoriel nul $\{0\}$. En fait, et on le comprend bien en regardant la définition, l'algèbre de Lie d'un groupe linéaire ne dépend que de la composante connexe du groupe contenant l'identité (pour un groupe fini, cette composante connexe est $\{\text{Id}\}$). Il existe plusieurs notions de connexité en topologie (connexité, connexité par arcs...). Pour les groupes linéaires, elles coïncident toutes, comme le montre le résultat suivant

Proposition II.4.2. — Soit G un sous-groupe fermé de $\mathbf{GL}(n, \mathbb{R})$. Les conditions suivantes sont équivalentes :

- (i) Deux éléments quelconques de G peuvent être reliés par un chemin continu.
- (ii) G n'est pas l'union de deux ouverts disjoints non vides.
- (iii) G est engendré par un voisinage quelconque de Id .
- (iv) G est engendré par $\exp U$, pour tout voisinage U de 0 dans \mathfrak{g} .

Si ces conditions sont vérifiées, on dit que G est connexe.

Démonstration. (i) \Rightarrow (ii) Supposons (i) et soient U et V deux ouverts disjoints non vides dont la réunion est G . Alors un chemin continu reliant un élément de U à un élément de V donnerait une partition de l'intervalle de définition du chemin en deux ouverts non vides disjoints, contradiction.

(ii) \Rightarrow (iii) Supposons (ii) et soit G_0 le sous-groupe de G engendré par un voisinage de Id dans G . Alors G_0 est ouvert dans G , car il contient un voisinage de chacun de ses points (c'est clair pour Id , et pour tout autre point par translation). De même, chaque partie de la forme gG_0 , $g \in G$, est ouverte dans G . Comme G peut être décomposé en une union disjointe de telles classes à gauche, l'hypothèse entraîne que $G = G_0$.

(iii) \Rightarrow (iv) est clair car $\exp U$ est un voisinage de Id .

(iv) \Rightarrow (i) Si l'on suppose (iv), alors tout élément g de G peut se mettre sous la forme

$$g = \exp X_1 \exp X_2 \dots \exp X_k,$$

les X_i étant dans U . En particulier, étant donnés deux éléments a_0 et a_1 dans G , on peut écrire

$$a_1 = a_0 \exp X_1 \exp X_2 \dots \exp X_k.$$

On prend alors $a(t) = a_0 \exp tX_1 \exp tX_2 \dots \exp tX_k$. C'est un chemin continu dans G vérifiant $a(0) = a_0, a(1) = a_1$. \square

Corollaire II.4.3. — Soit G un sous-groupe fermé de $\mathbf{GL}(n, \mathbb{R})$ d'algèbre de Lie \mathfrak{g} . Alors la composante neutre G_0 de G (c'est-à-dire la composante connexe de G contenant Id) est le sous-groupe de G engendré par $\exp \mathfrak{g}$. C'est un sous-groupe ouvert, fermé, distingué de G et c'est le seul sous-groupe ouvert connexe de G . La composante connexe d'un élément g de G est la classe à droite gG_0 .

Démonstration. La première assertion est claire d'après la démonstration de la proposition. Le reste se vérifie facilement.

Définition II.4.4. — Le groupe quotient G/G_0 est appelé groupe des composantes connexes de G .

Exercice II.4.5. — Composantes connexes de $\mathbf{O}(2, 1)$

Le groupe $\mathbf{O}(2, 1)$ est-il connexe ?

Rappelons que $\mathbf{O}(2, 1)$ est le sous-groupe de $\mathbf{GL}(3, \mathbb{R})$ préservant la forme bilinéaire symétrique non-dégénérée de \mathbb{R}^3

$$(x, y) = x_1y_1 + x_2y_2 - x_3y_3.$$

Soit \mathcal{C} le cône $\{x \in \mathbb{R}^3 \mid (x, x) = 0\}$. Montrer que \mathbb{R}^3 est l'union disjointe de

- \mathcal{C}
- $A = \{x \in \mathbb{R}^3 \mid (x, x) > 0\}$
- $B^+ = \{x \in \mathbb{R}^3 \mid (x, x) < 0, (x, e_3) > 0\}$
- $B^- = \{x \in \mathbb{R}^3 \mid (x, x) < 0, (x, e_3) < 0\}$

Soit $a \in \mathbf{SO}(2, 1)$. Montrer que les conditions suivantes sont équivalentes :

- (i) a stabilise chacun des ensembles A, B^+, B^- .
- (ii) a stabilise B^+ .
- (iii) $(ae_3, e_3) < 0$.

Le groupe $\mathbf{SO}(2, 1)$ est-il connexe ?

Soit $\mathbf{SO}_0(2, 1)$ le sous-groupe des éléments de $\mathbf{SO}(2, 1)$ vérifiant ces conditions équivalentes. On veut maintenant montrer que $\mathbf{SO}_0(2, 1)$ est la composante neutre de $\mathbf{SO}(2, 1)$ et de $\mathbf{O}(2, 1)$. Pour cela, introduisons les transformations linéaires suivantes de \mathbb{R}^3 :

- les rotations euclidiennes. On fixe deux vecteurs u et v de \mathbb{R}^3 vérifiant

$$(u, u) = 1, \quad (v, v) = 1, \quad (u, v) = 0$$

et $\alpha \in \mathbb{R}$. On pose

$$\begin{aligned} a(u) &= \cos \alpha u - \sin \alpha v \\ a(v) &= \sin \alpha u + \cos \alpha v \\ a(x) &= x \text{ si } (x, u) = 0, (x, v) = 0 \end{aligned}$$

- les *rotations hyperboliques*. On fixe deux vecteurs u et v de \mathbb{R}^3 vérifiant

$$(u, u) = 1, \quad (v, v) = -1, \quad (u, v) = 0$$

et $\alpha \in \mathbb{R}$. On pose

$$\begin{aligned} a(u) &= \cosh \alpha u + \sinh \alpha v \\ a(v) &= \sinh \alpha u + \cosh \alpha v \\ a(x) &= x \text{ si } (x, u) = 0, (x, v) = 0 \end{aligned}$$

Montrer que les rotations euclidiennes et hyperboliques sont dans le groupe $\mathbf{SO}_e(2, 1)$, puis que tout élément de $\mathbf{SO}_e(2, 1)$ s'écrit $a = kh$ où k est une rotation euclidienne d'axe e_3 et h une rotation hyperbolique dans un plan contenant e_3 .

Exercice II.4.6. — Etudier s'inspirant de l'exercice précédent les composantes connexes du groupe de Lorentz $\mathbf{O}(3, 1)$.

II.5. Homomorphismes de groupes linéaires. Revêtements

On suppose que tous les groupes linéaires considérés sont fermés dans le groupe linéaire dans lequel on les considère et qu'ils sont ainsi munis d'une structure de groupe de Lie. Les morphismes entre groupes linéaires sont toujours supposés différentiables. Le résultat qui suit montre qu'un morphisme de groupe linéaire connexe est entièrement déterminé par sa différentielle en l'identité. C'est ce résultat remarquable qui nous permettra de ramener l'étude des représentations de dimension finie d'un tel groupe à celles de son algèbre de Lie.

Théorème II.5.1. — Soit $f : G \rightarrow H$ un morphisme (différentiable) entre groupes linéaires. Soit $\phi = df_{\text{Id}}$ la différentielle de f en Id . Alors

$$\phi : \mathfrak{g} \rightarrow \mathfrak{h}$$

est un morphisme d'algèbres de Lie et

$$f(\exp X) = \exp(\phi(X)), \quad (\forall X \in \mathfrak{g})$$

Démonstration. C'est encore et toujours la même méthode. Montrons que $f(\exp X) = \exp(\phi(X))$. Posons $a(s) = f(\exp sX)$. On a

$$\begin{aligned} \frac{d}{ds} a(s) &= \frac{d}{dt} f(\exp((s+t)X))|_{t=0} = \frac{d}{dt} f(\exp sX \exp tX)|_{t=0} \\ &= f(\exp sX) \frac{d}{dt} f(\exp tX)|_{t=0} = a(s) \phi(X) \end{aligned}$$

Ceci montre que $a(s)$ vérifie l'équation différentielle $a'(s) = a(s)\phi(X)$, avec condition initiale $a(0) = \text{Id}$. Ceci montre que $a(s) = \exp(s\phi(X))$, et en particulier $f(\exp X) = \exp(\phi(X))$.

Montrons que ϕ est un morphisme d'algèbres de Lie, c'est-à-dire que quels que soient X, Y dans \mathfrak{g} ,

$$\phi([X, Y]) = [\phi(X), \phi(Y)].$$

On part de l'équation

$$f(\exp tX \exp sY \exp -tX) = \exp t\phi(X) \exp s\phi(Y) \exp -t\phi(X)$$

que l'on dérive par rapport à s et que l'on évalue en $s = 0$:

$$\phi(\exp tX Y \exp -tX) = \exp t\phi(X) \phi(Y) \exp -t\phi(X).$$

On dérive par rapport à t et l'on évalue en $t = 0$:

$$\phi([X, Y]) = [\phi(X), \phi(Y)].$$

□

Un morphisme différentiable $f : G \rightarrow H$ est localement bijectif si $\phi = df_{\text{Id}}$ est un isomorphisme entre \mathfrak{g} et \mathfrak{h} . En effet, d'après le théorème d'inversion locale, on a dans ce cas :

— pour tout $g \in G$, il existe un voisinage ouvert \mathcal{U} de 0 dans \mathfrak{g} tel que f réalise un difféomorphisme de $g \exp \mathcal{U}$ vers $f(g \exp \mathcal{U})$.

Un morphisme localement bijectif $p : \tilde{G} \rightarrow G$ entre groupes linéaires connexes est appelé un **revêtement**⁽¹⁾ de G . Le noyau d'un tel morphisme est discret dans \tilde{G} et p est surjectif. Réciproquement, un morphisme différentiable surjectif à noyau discret entre deux groupes connexes est localement bijectif. Le noyau de p est nécessairement contenu dans le centre de \tilde{G} . En effet tout élément g de \tilde{G} peut être relié à Id par un chemin continu $a(t)$. Pour tout $z \in \ker p$, on a $a(t)za(t)^{-1} \in \ker p$ et en $t = 0$, $a(0)za(0)^{-1} = z$. La fonction $t \mapsto a(t)za(t)^{-1}$ est continue, à valeurs dans un sous-groupe discret de \tilde{G} , donc constante, ce qui prouve l'assertion. Réciproquement, pour tout sous-groupe central discret Z de \tilde{G} , la suite exacte

$$1 \longrightarrow Z \longrightarrow \tilde{G} \longrightarrow \tilde{G}/Z \rightarrow 1.$$

fait apparaître \tilde{G} comme revêtement de \tilde{G}/Z . (Il faut munir \tilde{G}/Z de la topologie quotient et montrer que c'est bien un groupe linéaire).

Soient G et H deux groupes linéaires, G étant connexe, respectivement d'algèbres de Lie \mathfrak{g} et \mathfrak{h} , et supposons donné un morphisme d'algèbres de Lie

$$\phi : \mathfrak{g} \rightarrow \mathfrak{h}.$$

Pouvons nous relever ϕ en un morphisme de groupes linéaires

$$f : G \rightarrow H$$

de telle sorte que $df_{\text{Id}} = \phi$?

L'exemple de $G = \mathbb{C}^\times$, $H = \mathbb{C}$, $\mathfrak{g} = \mathfrak{h} = \mathbb{C}$ et $\phi = \text{Id}_{\mathbb{C}}$ montre que tel n'est pas le cas. En effet, supposons que f existe. Alors $f(e^z) = \exp(\phi(z)) = \exp z$ pour tout $z \in \mathbb{C}$. Mais il

1. C'est alors un revêtement au sens de la topologie.

faut prendre garde au fait que dans le membre de droite de cette équation, l'exponentielle n'est pas l'exponentielle usuelle des nombres complexes. Pour bien comprendre ceci, il faut revenir aux définitions, qui demandent de réaliser $(\mathbb{C}^\times, \times)$ et $(\mathbb{C}, +)$ comme des groupes linéaires. Or, si $(\mathbb{C}^\times, \times) = \mathbf{GL}(1, \mathbb{C})$ est trivialement linéaire, il faut être plus subtil pour $(\mathbb{C}, +)$. Ce groupe se plonge dans $\mathbf{GL}(2, \mathbb{C})$ par

$$z \mapsto \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix},$$

et donc son algèbre de Lie \mathbb{C} s'identifie à une sous-algèbre de Lie de $\mathfrak{gl}(2, \mathbb{R})$

$$x \mapsto \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix},$$

On a alors

$$\exp \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

Donc, il faut interpréter la formule ci-dessus comme

$$f(e^z) = \exp z = z.$$

Or il est bien connu que la fonction $z \mapsto e^z$ n'admet pas d'inverse global sur \mathbb{C}^\times . Si G et H sont donnés comme ci-dessus, un morphisme d'algèbres de Lie $\phi : \mathfrak{g} \rightarrow \mathfrak{h}$ ne se relève donc pas forcément en un morphisme de groupes linéaires. Mais il existe toujours un revêtement de G qui va permettre ce relèvement :

Théorème II.5.2. — Soit G et H deux groupes linéaires connexes et $\phi : \mathfrak{g} \rightarrow \mathfrak{h}$ un morphisme d'algèbres de Lie. Alors il existe un revêtement $p : \tilde{G} \rightarrow G$ tel que ϕ se relève en un morphisme de groupes

$$f : \tilde{G} \rightarrow H.$$

Démonstration. Dans l'énoncé ci-dessus, il faut comprendre que les algèbres de Lie de G et \tilde{G} sont identifiées par l'isomorphisme dp_{Id} . Posons :

$$\tilde{\mathfrak{g}} = \{(X, Y) \in \mathfrak{g} \times \mathfrak{h} \mid Y = \phi(X)\}.$$

C'est le graphe de ϕ , et c'est une algèbre de Lie linéaire. Soit \tilde{G} le groupe lui correspondant par la correspondance de Lie : c'est un sous-groupe de $G \times H$. Soit $p : \tilde{G} \rightarrow G$ la restriction à \tilde{G} de la projection de $G \times H$ sur G et $f : \tilde{G} \rightarrow H$ la restriction à \tilde{G} de la projection de $G \times H$ sur H . On a

$$dp_{\text{Id}} : \tilde{\mathfrak{g}} \rightarrow \mathfrak{g}, \quad (X, \phi(X)) \mapsto X$$

qui est un isomorphisme, ce qui montre que $p : \tilde{G} \rightarrow G$ est un revêtement. D'autre part

$$df_{\text{Id}} : \tilde{\mathfrak{g}} \rightarrow \mathfrak{h}, \quad (X, \phi(X)) \mapsto \phi(X),$$

ce qui montre que f est un relèvement de ϕ . □

Exemple II.5.3. — Dans la discussion précédente avec $G = \mathbb{C}^\times$, $H = \mathbb{C}$ le morphisme d'algèbres de Lie

$$\phi : \mathbb{C} \rightarrow \mathbb{C}, \quad z \mapsto z$$

se relève en un morphisme de $\tilde{G} \rightarrow \mathbb{C}$, où

$$\tilde{G} = \left\{ M(z) = \begin{pmatrix} e^z & 0 & 0 \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix}, (z \in \mathbb{C}) \right\}.$$

En effet $p : \tilde{G} \rightarrow \mathbb{C}^\times$, $M(z) \mapsto e^z$ est un isomorphisme local et $f : \tilde{G} \rightarrow \mathbb{C}$, $M(z) \mapsto z$ relève ϕ .

On dit qu'un groupe linéaire connexe G est **simplement connexe** lorsqu'il vérifie la condition suivante. Tout morphisme d'algèbres de Lie entre son algèbre de Lie \mathfrak{g} et une algèbre de Lie linéaire \mathfrak{h} se relève en un morphisme de groupes linéaires entre G et le groupe linéaire H donné par la correspondance de Lie. Un groupe est simplement connexe en ce sens s'il l'est au sens de la topologie (admis).

Exemples II.5.4. — Les groupes $\mathbf{SL}(n, \mathbb{C})$ sont simplement connexes, ainsi que les groupes compacts $\mathbf{SU}(n)$. Les groupes $\mathbf{SO}(n, \mathbb{C})$ ne le sont pas, ni les groupes compacts $\mathbf{SO}(n)$. Il existe un revêtement d'ordre 2 de $\mathbf{SO}(n)$ simplement connexe que l'on note $\mathbf{Spin}(n)$. Lorsque $n = 3$, $\mathbf{Spin}(3) \simeq \mathbf{SU}(2)$. Cet exemple sera détaillé dans le chapitre suivant.

Une question naturelle est de savoir si tout groupe linéaire connexe G admet un revêtement simplement connexe \tilde{G} . En topologie, tout espace connexe « raisonnable » (une variété différentiable par exemple), admet un revêtement simplement connexe (appelé revêtement universel). Un **groupe de Lie** est une variété différentiable, munie d'une structure de groupe telle que le produit et le passage à l'inverse soient \mathcal{C}^∞ . Le revêtement universel d'un groupe de Lie connexe peut être muni d'une structure de groupe de Lie. Un groupe linéaire connexe est un groupe de Lie, mais son revêtement universel n'est pas nécessairement linéaire. Ainsi, par exemple, le groupe $\mathbf{SL}(2, \mathbb{R})$ admet un revêtement universel $\widetilde{\mathbf{SL}}(2, \mathbb{R})$ qui est un groupe de Lie non linéaire. Le noyau de $p : \widetilde{\mathbf{SL}}(2, \mathbb{R}) \rightarrow \mathbf{SL}(2, \mathbb{R})$ est isomorphe à \mathbb{Z} .

II.6. Représentations de dimension finie des groupes linéaires connexes

Soit G un groupe linéaire connexe. Une représentation de dimension finie (π, V) de G est donnée par un morphisme de groupes linéaires :

$$\pi : G \rightarrow \mathbf{GL}(V).$$

La différentielle de ce morphisme est un morphisme d'algèbre de Lie :

$$d\pi_{\text{Id}} : \mathfrak{g} \rightarrow \mathfrak{gl}(V).$$

Le résultat suivant est immédiat, et montre que l'étude des représentations de G se ramène dans une large mesure à l'étude des représentations de son algèbre de Lie \mathfrak{g} .

Proposition II.6.1. — *Un sous-espace W de V est stable sous l'action de G si et seulement s'il est stable sous l'action de \mathfrak{g} .*

Démonstration. Dans un sens, on différencie, de l'autre, on utilise le fait que G étant connexe, il est engendré par l'image de l'exponentielle. \square

Ainsi une représentation irréductible de G donne une représentation irréductible de \mathfrak{g} . Si l'on cherche par exemple à déterminer toutes les représentations irréductibles de dimension finie de G , on commencera par le problème, en pratique plus simple, de déterminer toutes les représentations irréductibles de dimension finie de \mathfrak{g} . Bien sûr, si G n'est pas simplement connexe, une représentation de \mathfrak{g} ne se remonte pas forcément en une représentation de G . La stratégie sera alors la suivante. On essaie d'identifier le revêtement universel \tilde{G} de G . Toute représentation de \mathfrak{g} se remonte en une représentation de \tilde{G} . Une telle représentation définit aussi une représentation de G si et seulement si elle est triviale sur le noyau de $p : \tilde{G} \rightarrow G$.

Dans le chapitre suivant, nous allons utiliser cette stratégie pour étudier les représentations du groupe compact $\mathbf{SO}(3)$.

Exercice II.6.2. — Soient (π, V) et (τ, W) deux représentations de dimension finie d'un groupe linéaire G . Quelle est l'action de \mathfrak{g} dans $V \otimes W$?

Une représentation de dimension finie de l'algèbre de Lie \mathfrak{g} est un morphisme d'algèbres de Lie :

$$\phi : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$$

où V est un espace vectoriel **complexe** de dimension finie. Or \mathfrak{g} est un espace vectoriel **réel** et ϕ est \mathbb{R} -linéaire. Il est avantageux, comme nous le verrons dans l'exemple du chapitre suivant, de remplacer \mathfrak{g} par sa **complexification**.

Définition II.6.3. — Soit E un espace vectoriel réel. La complexification de E , notée $E_{\mathbb{C}}$, est l'espace vectoriel complexe

$$E_{\mathbb{C}} = E \otimes_{\mathbb{R}} \mathbb{C}.$$

La multiplication par un scalaire $\lambda \in \mathbb{C}$ d'un élément $v \otimes z$ est donnée par

$$\lambda \cdot (v \otimes z) = v \otimes \lambda z.$$

En pratique, il suffit de comprendre que si $(e_i)_{i \in I}$ est une base de E (sur \mathbb{R}), alors c'est aussi une base de $E_{\mathbb{C}}$ (sur \mathbb{C}). Si $\dim_{\mathbb{R}} E = n$, alors $\dim_{\mathbb{C}} E_{\mathbb{C}} = n$. Remarquons que tout espace vectoriel sur \mathbb{C} est naturellement un espace vectoriel sur \mathbb{R} , mais que la complexification n'est pas l'inverse de cette opération. Ainsi, la complexification de \mathbb{R} est \mathbb{C} , et $\mathbb{C} \simeq \mathbb{R}^2$ en tant qu'espace vectoriel réel.

Proposition II.6.4. — Toute application \mathbb{R} -linéaire

$$\phi : E \rightarrow V$$

où V est un espace vectoriel complexe se prolonge naturellement en une application \mathbb{C} -linéaire (encore notée ϕ pour ne pas alourdir)

$$\phi : E_{\mathbb{C}} \rightarrow V.$$

Réciproquement, toute application \mathbb{C} -linéaire $\phi : E_{\mathbb{C}} \rightarrow V$ est entièrement déterminée par sa restriction à $E \simeq E \otimes \mathbb{R} \subset E \otimes_{\mathbb{R}} \mathbb{C} = E_{\mathbb{C}}$ qui est \mathbb{R} -linéaire.

Démonstration. C'est tautologique, les définitions sont faites pour. \square

En résumé, il est équivalent d'étudier les représentations \mathbb{R} -linéaires de \mathfrak{g} dans des espaces vectoriels complexes ou les représentations \mathbb{C} -linéaires de sa complexification $\mathfrak{g}_{\mathbb{C}}$.

II.7. Représentation adjointe

Soit G un groupe linéaire. Il agit sur son algèbre de Lie \mathfrak{g} (ou sa complexification $\mathfrak{g}_{\mathbb{C}}$) par

$$g \cdot X = gXg^{-1}, \quad (X \in \mathfrak{g}), (g \in G).$$

En effet, comme pour tout $t \in \mathbb{R}$, $g \exp tXg^{-1} = \exp(gXg^{-1})$, on voit que $gXg^{-1} \in \mathfrak{g}$. Cette action est linéaire, comme on le vérifie facilement, et la représentation de G dans \mathfrak{g} (ou $\mathfrak{g}_{\mathbb{C}}$) ainsi définie s'appelle la **représentation adjointe** de G . On la note Ad .

Sa différentielle en l'identité définit une représentation de \mathfrak{g} dans elle-même, appelée de même représentation adjointe, et notée ad . Quels que soient $X, Y \in \mathfrak{g}$, on a

$$\text{ad}(X) \cdot Y = [X, Y].$$

II.8. Représentations projectives

Nous introduisons une notion plus faible que celle de représentation : les représentations projectives. Nous verrons que les représentations projectives d'un groupe G se relèvent en de véritables représentations d'une extension centrale \tilde{G} de G .

Définition II.8.1. — Soit G un groupe. Une représentation projective de G est la donnée d'un espace vectoriel (complexe) V , et d'une application

$$\rho : G \rightarrow \text{GL}(V)$$

vérifiant :

$$(II.8.1) \quad (\forall g, h \in G), \exists c(g, h) \in \mathbb{C}^{\times} \text{ tel que } \rho(g)\rho(h) = c(g, h)\rho(gh).$$

Remarques II.8.2. — Si $c(g, h)$ est identiquement égal à 1, on retrouve la définition d'une représentation. Une représentation projective est une représentation « à multiplication scalaire » près.

— Comme la loi de groupe est associative, on a d'une part

$$\rho(g_1)\rho(g_2)\rho(g_3) = c(g_1, g_2)\rho(g_1g_2)\rho(g_3) = c(g_1, g_2)c(g_1g_2, g_3)\rho(g_1g_2g_3),$$

et d'autre part

$$\rho(g_1)\rho(g_2)\rho(g_3) = \rho(g_1)c(g_2, g_3)\rho(g_2g_3) = c(g_2, g_3)c(g_1, g_2g_3)\rho(g_1g_2g_3).$$

Il s'ensuit la relation suivante, appelée **relation de cocycle**

$$(II.8.2) \quad c(g_1, g_2)c(g_1g_2, g_3) = c(g_1, g_2g_3)c(g_2, g_3).$$

On appelle cocycle (plus précisément 2-cocycle, à valeurs dans \mathbb{C}^\times , ou plus généralement dans un sous-groupe A de \mathbb{C}^\times , par exemple le groupe μ_n des racines n -ièmes de l'unité) une application de $G \times G \rightarrow A$ vérifiant (II.8.2).

— On fabrique certains cocycles particulier de la manière suivante : soit $b : G \rightarrow \mathbb{C}^\times$ une application quelconque. Alors $\tilde{b} : (g, h) \mapsto b(g)b(h)b(gh)^{-1}$ est un cocycle. Les cocycles obtenus de cette manière s'appellent les **cobords**.

— Soit (ρ, V) une représentation projective de G , et soit c le cocycle associé. Soit $b : G \rightarrow \mathbb{C}^\times$ une application et \tilde{b} le cobord qui s'en déduit. Alors

$$\rho_1 : G \rightarrow \mathrm{GL}(V), \quad g \mapsto b(g)\rho(g)$$

est encore une représentation projective de G , dont le cocycle associé est

$$(g, h) \mapsto c_1(g, h) = \tilde{b}(g, h)c(g, h).$$

Etant donné un cocycle c à valeurs dans un sous-groupe A de \mathbb{C}^\times , on fabrique une extension centrale \tilde{G} de G de la manière suivante. En tant qu'ensemble, \tilde{G} est le produit cartésien $G \times A$, la multiplication étant donnée par

$$(g, z)(g', z') = (gg', zz'c(g, g')), \quad (g, g' \in G), (z, z' \in A).$$

La relation de cocycle rend cette multiplication associative, comme il siéd à une loi de groupe. Il est clair que $(e_G, 1)$ est l'élément neutre de \tilde{G} , l'inverse de (g, z) étant $(g^{-1}, z^{-1}c(g, g^{-1})^{-1})$. Les homomorphismes

$$\begin{aligned} A &\rightarrow \tilde{G}, & z &\mapsto (e_G, z) \\ \tilde{G} &\rightarrow G, & (g, z) &\mapsto g \end{aligned}$$

donnent une suite exacte

$$\{1\} \longrightarrow A \longrightarrow \tilde{G} \longrightarrow G \longrightarrow \{1\}.$$

Proposition II.8.3. — Soit (ρ, V) une représentation projective de G , et soit c le cocycle associé, à valeurs dans un sous-groupe A de \mathbb{C}^\times . Soit \tilde{G} l'extension centrale de G par A fabriquée ci-dessus. Alors

$$\tilde{\rho} : \tilde{G} \rightarrow \mathrm{GL}(V), \quad \tilde{\rho}(g, z) = z\rho(g)$$

définit une représentation du groupe \tilde{G} .

Démonstration. On calcule :

$$\begin{aligned} \tilde{\rho}(g, z)\tilde{\rho}(g', z') &= zz'\rho(g)\rho(g') = zz'c(g, g')\rho(gg') \\ \tilde{\rho}((g, z)(g', z')) &= \tilde{\rho}(gg', zz'c(g, g')) = zz'c(g, g')\rho(gg') \end{aligned}$$

□

CHAPITRE III

REPRÉSENTATIONS DE $\mathfrak{sl}(2, \mathbb{C})$, $\mathbf{SU}(2)$, ET $\mathbf{SO}(3)$

Notre but est d'étudier les représentations du groupe compact $\mathbf{SO}(3)$ selon les principes dégagés dans la dernière section du chapitre précédent. Nous commençons par identifier le revêtement universel de $\mathbf{SO}(3)$.

III.1. Le revêtement $\mathbf{SU}(2) \rightarrow \mathbf{SO}(3)$

Rappelons que $\mathbf{SU}(2)$ est le groupe de matrices :

$$\mathbf{SU}(2) = \{a \in \mathbf{GL}(2, \mathbb{C}) \mid {}^t \bar{a} a = \text{Id}, \det a = 1\},$$

et que son algèbre de Lie est :

$$\mathfrak{su}(2) = \{X \in \mathfrak{gl}(2, \mathbb{C}) \mid {}^t \bar{X} + X = 0, \text{Tr } X = 0\}.$$

Plus explicitement, on a donc :

$$\mathbf{SU}(2) = \left\{ \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix}, \alpha, \beta \in \mathbb{C}, |\alpha|^2 + |\beta|^2 = 1 \right\},$$

$$\mathfrak{su}(2) = \left\{ \begin{pmatrix} iz & -y + ix \\ y + ix & -iz \end{pmatrix}, x, y, z \in \mathbb{R} \right\}.$$

On voit donc que, topologiquement, $\mathbf{SU}(2)$ est la sphère unité dans $\mathbb{C}^2 \simeq \mathbb{R}^4$. En particulier :

Proposition III.1.1. — $\mathbf{SU}(2)$ est un groupe connexe, simplement connexe.

Passons maintenant à $\mathbf{SO}(3)$. C'est le groupe de matrices :

$$\mathbf{SO}(3) = \{a \in \mathbf{GL}(3, \mathbb{R}) \mid {}^t a a = \text{Id}, \det a = 1\},$$

dont l'algèbre de Lie est

$$\begin{aligned} \mathfrak{so}(3) &= \{X \in \mathfrak{gl}(3, \mathbb{R}) \mid {}^t X + X = 0\} \\ &= \left\{ \begin{pmatrix} 0 & -z & y \\ z & 0 & x \\ -y & -x & 0 \end{pmatrix}, x, y, z \in \mathbb{R} \right\} \end{aligned}$$

Une base de $\mathfrak{su}(2)$ est donnée par :

$$I = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad K = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

et l'on a les relations de commutation suivantes :

$$[I, J] = 2K, \quad [J, K] = 2I, \quad [K, I] = 2J.$$

Munissons $\mathfrak{su}(2)$ de la forme bilinéaire κ symétrique définie par

$$\kappa(X, Y) = \text{Tr}(\text{ad}(X)\text{ad}(Y)), \quad (X, Y \in \mathfrak{su}(2)).$$

Les relations de commutation ci-dessus donnent les matrices de $\text{ad}(I)$, $\text{ad}(J)$, $\text{ad}(K)$ dans la base I, J, K : On calcule alors facilement que

$$\kappa(I, I) = \kappa(J, J) = \kappa(K, K) = -8, \quad \kappa(I, J) = \kappa(J, K) = \kappa(K, I) = 0.$$

La base $\{I, J, K\}$ de $\mathfrak{su}(2)$ est donc orthogonale, et l'on constate que la forme κ est définie négative. Si l'on préfère travailler avec un produit scalaire (défini positif) et une base orthonormale, on peut remplacer κ par $-\frac{1}{8}\kappa$. L'espace euclidien $(\mathfrak{su}(2), -\frac{1}{8}\kappa)$ est alors, par le choix de la base $\{I, J, K\}$, isomorphe à \mathbb{R}^3 muni de son produit scalaire canonique.

Montrons que la forme κ est invariante par l'action adjointe de $\mathbf{SU}(2)$ sur $\mathfrak{su}(2)$. Il nous faut montrer que $(\forall X, Y \in \mathfrak{su}(2)), (\forall g \in \mathbf{SU}(2))$,

$$\kappa(\text{Ad}(g) \cdot X, \text{Ad}(g) \cdot Y) = \kappa(X, Y),$$

c'est-à-dire

$$(III.1.1) \quad \text{Tr}(\text{ad}(gXg^{-1})\text{ad}(gYg^{-1})) = \text{Tr}(\text{ad}(X)\text{ad}(Y)).$$

Or, pour tout $Z \in \mathfrak{su}(2)$,

$$\begin{aligned} \text{ad}(gXg^{-1})\text{ad}(gYg^{-1}) \cdot Z &= [gXg^{-1}, [gYg^{-1}, Z]] \\ &= (gXg^{-1})(gYg^{-1})Z - (gXg^{-1})Z(gYg^{-1}) \\ &\quad - (gYg^{-1})Z(gXg^{-1}) + Z(gYg^{-1})(gXg^{-1}) \\ &= gXY(g^{-1}Zg)g^{-1} - gX(g^{-1}Zg)Yg^{-1} - gY(g^{-1}Zg)Xg^{-1} + g((g^{-1}Zg)YX)g^{-1} \\ &= \text{Ad}(g) \cdot (XY(g^{-1}Zg) - (g^{-1}Zg)Y - Y(g^{-1}Zg)X + (g^{-1}Zg)YX) \\ &= \text{Ad}(g) \cdot (X[Y, \text{ad}(g^{-1}) \cdot Z] - [Y, \text{ad}(g^{-1}) \cdot Z]X) \\ &= \text{Ad}(g) \circ \text{ad}(X) \circ \text{ad}(Y) \circ \text{Ad}(g)^{-1}(Z) \end{aligned}$$

D'où

$$\text{ad}(gXg^{-1}) \circ \text{ad}(gYg^{-1}) = \text{Ad}(g) \circ \text{ad}(X) \circ \text{ad}(Y) \circ \text{Ad}(g)$$

et

$$\text{Tr}(\text{ad}(gXg^{-1})\text{ad}(gYg^{-1})) = \text{Tr}(\text{ad}(X) \circ \text{ad}(Y)).$$

On a donc bien obtenu (III.1.1).

Ainsi Ad est un morphisme du groupe $\mathbf{SU}(2)$ dans le groupe orthogonal $\mathbf{O}(\mathfrak{su}(2), \kappa) \simeq \mathbf{O}(3)$, et de plus, comme $\mathbf{SU}(2)$ est connexe, le déterminant de $\text{Ad}(g)$, $g \in G$ est toujours 1. Donc en fait, l'on obtient un morphisme :

$$\text{Ad} : \mathbf{SU}(2) \longrightarrow \mathbf{SO}(\mathfrak{su}(2), \kappa) \simeq \mathbf{SO}(3).$$

La différentielle de ce morphisme est le morphisme d'algèbre de Lie

$$\text{ad} : \mathfrak{su}(2) \longrightarrow \mathfrak{so}(\mathfrak{su}(2), \kappa) \simeq \mathfrak{so}(3).$$

Un calcul explicite dans la base $\{I, J, K\}$ montre que

$$\text{ad} \begin{pmatrix} iz & -y + ix \\ y + ix & -iz \end{pmatrix} = \begin{pmatrix} 0 & -2z & 2y \\ 2z & 0 & -2x \\ -2y & 2x & 0 \end{pmatrix}$$

et donc ad réalise un isomorphisme entre $\mathfrak{su}(2)$ et $\mathfrak{so}(3)$. Ainsi, Ad est un isomorphisme local entre $\mathbf{SU}(2)$ et $\mathbf{SO}(3)$. Comme $\mathbf{SO}(3)$ est connexe, il est surjectif. Calculons maintenant son noyau. On cherche les $g \in \mathbf{SU}(2)$ tels que

$$\text{Ad}(g) \cdot I = I, \quad \text{Ad}(g) \cdot J = J, \quad \text{Ad}(g) \cdot K = K.$$

Un calcul explicite montre g est une matrice scalaire. La condition $\det g = 1$ impose alors $g = \pm \text{Id}$. Résumons :

Proposition III.1.2. — *La suite courte suivante est exacte :*

$$1 \longrightarrow \{\pm \text{Id}\} \longrightarrow \mathbf{SU}(2) \longrightarrow \mathbf{SO}(3) \longrightarrow 1.$$

Le groupe $\mathbf{SU}(2)$ est un revêtement de $\mathbf{SO}(3)$ (d'ordre 2).

III.2. Représentations de $\mathfrak{sl}(2, \mathbb{C})$

Comme nous l'avons expliqué en II.6, nous allons étudier les représentations de dimension finie de $\mathbf{SO}(3)$ et $\mathbf{SU}(2)$ en étudiant celle de leur algèbre de Lie $\mathfrak{so}(3) \simeq \mathfrak{su}(2)$. Commençons par identifier la complexification de $\mathfrak{su}(2)$:

$$\mathfrak{su}(2)_{\mathbb{C}} = \mathfrak{su}(2) \otimes_{\mathbb{R}} \mathbb{C}.$$

Comme $\{I, J, K\}$ est une base de $\mathfrak{su}(2)$ sur \mathbb{R} , on obtient

$$\begin{aligned} \mathfrak{su}(2)_{\mathbb{C}} &= \mathbb{C}I \oplus \mathbb{C}J \oplus \mathbb{C}K \\ &= \left\{ \begin{pmatrix} iz & -y + ix \\ y + ix & -iz \end{pmatrix}, x, y, z \in \mathbb{C} \right\} \\ &= \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix}, \alpha, \beta, \gamma \in \mathbb{C} \right\} \\ &= \mathfrak{sl}(2, \mathbb{C}) \end{aligned}$$

Il est donc équivalent d'étudier les représentations \mathbb{R} -linéaires de $\mathfrak{su}(2)$ dans des espaces vectoriels complexes ou les représentations \mathbb{C} -linéaires de $\mathfrak{sl}(2, \mathbb{C})$.

Une base (sur \mathbb{C}) de $\mathfrak{sl}(2, \mathbb{C})$ est

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

et l'on a les relations de commutation suivantes :

$$[h, e] = 2e, \quad [h, f] = -2f, \quad [e, f] = h.$$

Soit $\phi : \mathfrak{sl}(2, \mathbb{C}) \rightarrow \mathfrak{gl}(V)$ une représentation \mathbb{C} -linéaire de dimension finie. Pour tout $\lambda \in \mathbb{C}$, soit V_λ le sous-espace propre (éventuellement nul) de $\phi(h)$ pour la valeur propre λ . Comme \mathbb{C} est algébriquement clos, si V n'est pas nul, il existe un V_λ non nul.

Lemme III.2.1. —

$$\phi(e)(V_\lambda) \subset V_{\lambda+2}, \quad \phi(f)(V_\lambda) \subset V_{\lambda-2}.$$

Démonstration. Soit $v \in V_\lambda$. On a :

$$\begin{aligned} \phi(h)(\phi(e)(v)) &= \phi(e)(\phi(h)(v)) + [\phi(h), \phi(e)](v) \\ &= \lambda\phi(e)(v) + \phi([h, e])(v) \\ &= \lambda\phi(e)(v) + 2\phi(e)(v) = (\lambda + 2)\phi(e)(v) \end{aligned}$$

ce qui montre la première inclusion. Le deuxième s'obtient de la même manière. \square

Comme V est de dimension finie, il n'y a qu'un nombre fini de V_λ non nuls. Si V est non nul, il existe donc $\lambda_0 \in \mathbb{C}$ tel que $V_{\lambda_0} \neq \{0\}$ et $V_{\lambda_0+2} = \{0\}$. Soit $0 \neq v_0 \in V_{\lambda_0}$. Posons, pour tout $k \in \mathbb{N}$,

$$v_k = \phi(f)^k(v_0).$$

Alors v_k est dans V_{λ_0-2k} .

On a par exemple $v_1 = \phi(f)(v_0)$ et

$$\begin{aligned} \phi(e)(v_1) &= \phi(e)(\phi(f)(v_0)) \\ &= \phi(f)(\phi(e)(v_0)) + [\phi(e), \phi(f)](v_0) \\ &= 0 + \phi([e, f])(v_0) = \phi(h)(v_0) \\ &= \lambda_0(v_0) \end{aligned}$$

De même :

$$\begin{aligned} \phi(e)(v_2) &= \phi(e)(\phi(f)(v_1)) \\ &= \phi(f)(\phi(e)(v_1)) + [\phi(e), \phi(f)](v_1) \\ &= \lambda_0\phi(f)(v_0) + \phi(h)(v_1) \\ &= \lambda_0v_1 + (\lambda_0 - 2)v_1 = (2\lambda_0 - 2)v_1 \end{aligned}$$

Tentons d'avancer une hypothèse de récurrence :

$$(H) \quad \phi(e)(v_k) = k(\lambda_0 - k + 1)v_{k-1}.$$

Nous avons vu qu'elle est satisfaite pour $k = 1, 2$ et

En utilisant l'hypothèse de récurrence, on obtient, pour $k \geq 2$:

$$\begin{aligned}\phi(e)(v_{k+1}) &= \phi(e)(\phi(f)(v_k)) \\ &= \phi(f)(\phi(e)(v_k)) + [\phi(e), \phi(f)](v_k) \\ &= k(\lambda_0 - k + 1)\phi(f)(v_{k-1}) + \phi(h)(v_k) \\ &= k(\lambda_0 - k + 1)v_k + (\lambda_0 - 2k)v_k \\ &= (k + 1)(\lambda_0 - k)v_k\end{aligned}$$

Ceci prouve l'hypothèse (H).

Tant que les vecteurs v_0, v_1, \dots, v_j sont tous non nuls, ils sont linéairement indépendants, et donc, comme la dimension de V est finie, il existe un plus petit entier n non nul tel que v_{n+1} est nul. En particulier $v_n \neq 0$. On a alors

$$0 = \phi(e)(v_{n+1}) = (n + 1)(\lambda_0 - n)v_n,$$

d'où $\lambda_0 = n$.

D'autre part, v_0, v_1, \dots, v_n engendrent un sous-espace de V stable sous l'action de $\mathfrak{sl}(2, \mathbb{C})$. Si l'on suppose (ϕ, V) irréductible, on voit alors que (v_0, v_1, \dots, v_n) est une base de V . En particulier $\dim_{\mathbb{C}} V = n + 1$.

Réciproquement, si $n \in \mathbb{N}$, et si l'on fixe une base (v_0, v_1, \dots, v_n) de \mathbb{C}^{n+1} , alors on définit une représentation ϕ de $\mathfrak{sl}(2, \mathbb{C})$ par

$$\phi(h)(v_k) = (n - 2k)v_k, \quad \phi(e)(v_k) = k(n - k + 1)v_{k-1}, \quad \phi(f)(v_k) = v_{k+1},$$

pour tout $k = 0, \dots, n$, avec les conventions $v_{-1} = v_{n+1} = 0$. On voit facilement qu'elle est irréductible : tout sous-espace stable W contient un vecteur propre non nul pour $\phi(h)$, et est donc, à un scalaire non nul près, l'un des v_i .

Ceci donne la classification des représentations irréductibles \mathbb{C} -linéaires de $\mathfrak{sl}(2, \mathbb{C})$: notons $(\phi_n, \mathbb{C}^{n+1})$ la représentation irréductible de dimension $n + 1$ de $\mathfrak{sl}(2, \mathbb{C})$ définie ci-dessus. Toute représentation irréductible \mathbb{C} -linéaire de $\mathfrak{sl}(2, \mathbb{C})$ de dimension $n + 1$ est isomorphe à $(\phi_n, \mathbb{C}^{n+1})$.

Exercice III.2.2. — Exprimer I, J, K dans la base (h, e, f) . En déduire l'action de I, J, K dans $(\phi_n, \mathbb{C}^{n+1})$.

Nous allons maintenant montrer que les représentations $(\phi_n, \mathbb{C}^{n+1})$ se remontent en des représentations du groupe $\mathbf{SU}(2)$. Pour cela, nous allons commencer par trouver une autre réalisation de ces représentations.

Soit $\mathbb{C}[z_1, z_2]$ l'espace vectoriel des polynômes à coefficients complexes à deux variables. Posons $\partial_i = \frac{\partial}{\partial z_i}$, $i = 1, 2$, et

$$(III.2.1) \quad D_e = -z_2\partial_1, \quad D_f = -z_1\partial_2, \quad D_h = z_2\partial_2 - z_1\partial_1.$$

Alors D_e, D_f et D_h sont des opérateurs différentiels (linéaires à coefficients constants) qui agissent sur $\mathbb{C}[z_1, z_2]$. On vérifie par le calcul qu'ils satisfont les relations de commutation :

$$[D_h, D_e] = 2D_e, \quad [D_h, D_f] = -2D_f, \quad [D_e, D_f] = D_h.$$

Ainsi

$$e \mapsto D_e, \quad f \mapsto D_f, \quad h \mapsto D_h,$$

définit une représentation D de $\mathfrak{sl}(2, \mathbb{C})$ dans $\mathbb{C}[z_1, z_2]$.

Une base de $\mathbb{C}[z_1, z_2]$ est donnée par les monômes $z_1^k z_2^{n-k}$, avec k, n entiers, $k \leq n$. On calcule

$$D_h(z_1^k z_2^{n-k}) = (n - 2k)z_1^k z_2^{n-k} \quad D_f(z_1^k z_2^{n-k}) = -(n - k)z_1^{k+1} z_2^{n-k-1},$$

$$D_e(z_1^k z_2^{n-k}) = -kz_1^{k-1} z_2^{n-k+1},$$

avec la convention $z_i^{-1} = 0$.

Ceci montre que l'espace $\mathbb{C}[z_1, z_2]_n$ des polynômes homogènes de degré n fixé (qui est de dimension $n+1$), est stable sous l'action de $\mathfrak{sl}(2, \mathbb{C})$. La représentation $(D, \mathbb{C}[z_1, z_2]_n)$ est isomorphe à $(\phi_n, \mathbb{C}^{n+1})$ comme on le voit en normalisant convenablement la base des $z_1^k z_2^{n-k}$, (n fixé), en posant

$$v_k = (-1)^k \frac{z_1^k z_2^{n-k}}{(n-k)!}.$$

Exercice III.2.3. — En utilisant l'exercice précédent, écrire les formules donnant D_I, D_J, D_K et l'action de I, J, K sur les monômes $z_1^k z_2^{n-k}$.

Nous allons maintenant montrer que l'action de $\mathfrak{sl}(2, \mathbb{C})$ sur $\mathbb{C}[z_1, z_2]_n$ est la différentielle d'une représentation de $\mathbf{SL}(2, \mathbb{C})$ dans ce même espace. En effet, $\mathbf{SL}(2, \mathbb{C})$ agit naturellement sur \mathbb{C}^2 , et donc sur $\mathbb{C}[z_1, z_2]$ par

$$(\rho(g) \cdot P)(z_1, z_2) = P(g^{-1} \cdot (z_1, z_2)).$$

De manière explicite, si $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $ad - bc = 1$ est un élément de $\mathbf{SL}(2, \mathbb{C})$, son inverse est $g^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, et donc

$$g^{-1} \cdot (z_1, z_2) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = (dz_1 - bz_2, -cz_1 + az_2).$$

Ainsi

$$(\rho(g) \cdot P)(z_1, z_2) = P(dz_1 - bz_2, -cz_1 + az_2).$$

Nous espérons que le lecteur n'est pas gêné par le passage de l'écriture en colonnes des vecteurs de \mathbb{C}^2 à l'écriture en ligne des variables. Il est clair que cette action préserve les polynômes homogènes de degré n . On en déduit donc une représentation $(\rho, \mathbb{C}[z_1, z_2]_n)$. Nous allons maintenant calculer sa différentielle.

Soit $X = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ dans $\mathfrak{sl}(2, \mathbb{C})$ (donc $\delta = -\alpha$) et posons

$$c(t) = \exp tX = \begin{pmatrix} a(t) & b(t) \\ c(t) & d(t) \end{pmatrix} \in \mathbf{SL}(2, \mathbb{C}), \quad (t \in \mathbb{R})$$

On a $c(0) = \text{Id}$ et $c'(0) = X$, soit $a'(0) = \alpha$, $b'(0) = \beta$... On peut alors calculer la différentielle de ρ :

$$\begin{aligned} (d\rho(X) \cdot P)(z_1, z_2) &= \frac{d}{dt}|_{t=0}(\rho(c(t)) \cdot P)(z_1, z_2) \\ &= \frac{d}{dt}|_{t=0}P((d(t)z_1 - b(t)z_2, -c(t)z_1 + a(t)z_2) \\ &= (\delta z_1 - \beta z_2)(\partial_1 P)(z_1, z_2) + (-\gamma z_1 + \alpha z_2)(\partial_2 P)(z_1, z_2) \end{aligned}$$

En particulier, on trouve :

$$d\rho(h) = -z_1\partial_1 + z_2\partial_2, \quad d\rho(e) = -z_2\partial_1, \quad d\rho(f) = -z_1\partial_2.$$

Nous retrouvons (III.2.1).

On peut maintenant restreindre les représentations $(\rho, \mathbb{C}[z_1, z_2])_n$ à $\mathbf{SU}(2) \subset \mathbf{SL}(2, \mathbb{C})$. Notons (π_n, V_n) ces représentations. Nous avons montré que toute représentation irréductible de dimension $n + 1$ de $\mathbf{SU}(2)$ est isomorphe à (π_n, V_n) .

Les représentations irréductibles de dimension de $\mathbf{SO}(3)$ sont celles obtenues à partir de celles de $\mathbf{SU}(2)$ triviales en $-\text{Id} \in \mathbf{SU}(2)$. Or $-\text{Id}$ agit sur un polynôme P par

$$(\rho(-\text{Id}) \cdot P)(z_1, z_2) = P(-z_1, -z_2).$$

Ainsi les représentations de $\mathbf{SU}(2)$ triviales en $-\text{Id}$ sont les (π_n, V_n) , n pair.

Exercice III.2.4. — (Formule de Clebsch-Gordan). Soient (π_n, V_n) et (π_m, V_m) les représentations irréductibles de $\mathbf{SU}(2)$ de dimension $n + 1$ et $m + 1$ respectivement, avec $n \geq m$. Montrer que

$$\pi_n \otimes \pi_m = \bigoplus_{l=0, m}^{\oplus} \pi_{n+m-2l}.$$

III.3. Les harmoniques sphériques

Le groupe $\mathbf{SO}(3)$ agit sur \mathbb{R}^3 en préservant la sphère unité

$$S^2 = \{\underline{x} = (x_1, x_2, x_3) \in \mathbb{R}^3; \|\underline{x}\|^2 = x_1^2 + x_2^2 + x_3^2 = 1.\}$$

On en déduit des représentations linéaires de $\mathbf{SO}(3)$ dans l'espace des fonctions sur \mathbb{R}^3 et dans l'espace des fonctions sur S^2 (à valeurs complexes) donnée par la formule usuelle

$$g \cdot f(\underline{x}) = f(g^{-1} \cdot \underline{x}).$$

III.3.1. Fonctions harmoniques. — Le Laplacien Δ est l'opérateur différentiel à coefficients constants sur \mathbb{R}^3 donné par

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}.$$

Définition III.3.1. — Une fonction harmonique sur \mathbb{R}^3 (à valeurs complexes) est une fonction f de classe \mathcal{C}^2 telle que $\Delta(f) = 0$.

Lemme III.3.2. — Soit $f \in \mathcal{C}^2(\mathbb{R}^3, \mathbb{C})$ et soit $g \in \mathbf{SO}(3)$. Alors

$$g \cdot \Delta(g^{-1} \cdot f) = \Delta(f).$$

Démonstration. Posons $g = (g_{ij})_{1 \leq i, j \leq 3}$ et $\underline{y} = (y_1, y_2, y_3) = g(\underline{x})$. On a alors, par la formule de changement de variables :

$$\frac{\partial f}{\partial x_i}(f \circ g)(\underline{x}) = \sum_{j=1}^3 g_{ji} \frac{\partial f}{\partial y_j}(f)(\underline{y}),$$

D'ou

$$\Delta(f \circ g)(\underline{x}) = \sum_{i, j, k=1}^3 g_{ji} g_{ki} \frac{\partial^2 f}{\partial y_j \partial y_k}(f)(\underline{y}).$$

Or $g \in \mathbf{SO}(3)$, donc ${}^t g g = \text{Id}_{\mathbb{R}^3}$, et ainsi

$$\sum_{i=1}^3 g_{ji} g_{ki} = \begin{cases} 1 & \text{si } j = k \\ 0 & \text{sinon} \end{cases}.$$

Ceci donne

$$\Delta(f \circ g)(\underline{x}) = \sum_{k=1}^3 \frac{\partial^2 f}{\partial y_k^2}(f)(\underline{y}) = \Delta(f)(\underline{y}) = \Delta(f)(g \cdot \underline{x}).$$

□

Pour dire les choses de manière équivalentes, les espaces $\mathcal{C}^k(\mathbb{R}^3, \mathbb{C})$, $k \in \mathbb{N}$, sont des espaces de représentations du groupe $\mathbf{SO}(3)$ et

$$\Delta : \mathcal{C}^{k+2}(\mathbb{R}^3, \mathbb{C}) \longrightarrow \mathcal{C}^k(\mathbb{R}^3, \mathbb{C})$$

est un opérateur d'entrelacement.

III.3.2. Action de $\mathfrak{so}(3)$ dans $\mathcal{C}^\infty(\mathbb{R}^3, \mathbb{C})$. — Rappelons que

$$\begin{aligned} \mathfrak{so}(3) &= \{X \in \mathfrak{gl}(3, \mathbb{R}) \mid {}^t X + X = 0\} \\ &= \{M(x, y, z) = \begin{pmatrix} 0 & -2z & 2y \\ 2z & 0 & -2x \\ -2y & 2x & 0 \end{pmatrix}, x, y, z \in \mathbb{R}\}, \end{aligned}$$

dont une base est donnée par les matrices

$$\mathcal{I} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -2 \\ 0 & 2 & 0 \end{pmatrix}, \quad \mathcal{J} = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ -2 & 0 & 0 \end{pmatrix}, \quad \mathcal{K} = \begin{pmatrix} 0 & -2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

vérifiant les relations de commutation

$$[\mathcal{I}, \mathcal{J}] = \mathcal{K}, \quad [\mathcal{J}, \mathcal{K}] = \mathcal{I}, \quad [\mathcal{K}, \mathcal{I}] = \mathcal{J}.$$

Soit $M = M(x, y, z)$ dans $\mathfrak{so}(3)$ et soit $c(t)$ une courbe \mathcal{C}^1 dans $\mathbf{SO}(3)$ telle que $c(0) = \text{Id}_{\mathbb{R}^3}$ et $c'(0) = M$. On pose $c(t) = (c_{ij}(t))_{1 \leq i, j \leq 3}$. On a alors pour toute fonction $f \in \mathcal{C}^\infty(\mathbb{R}^3, \mathbb{C})$ et tout $\underline{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$,

$$\begin{aligned} (c(t) \cdot f)(\underline{x}) &= f(c(t)^{-1} \cdot \underline{x}) = f({}^t c(t) \cdot \underline{x}) \\ &= f((c_{11}x_1 + c_{21}x_2 + c_{31}x_3, c_{12}x_1 + c_{22}x_2 + c_{32}x_3, c_{13}x_1 + c_{23}x_2 + c_{33}x_3)). \end{aligned}$$

On dérive par rapport à t , et on évalue en $t = 0$ pour calculer $(M \cdot f)(\underline{x})$, ce qui nous donne

$$(M \cdot f)(\underline{x}) = \left((2zx_2 - 2yx_3) \frac{\partial}{\partial x_1} + (2xx_3 - 2zx_1) \frac{\partial}{\partial x_2} + (2yx_1 - 2xx_2) \frac{\partial}{\partial x_3} \right) \cdot f(\underline{x}).$$

L'algèbre de Lie $\mathfrak{so}(3)$ agit donc dans $\mathcal{C}^\infty(\mathbb{R}^3, \mathbb{C})$ par des opérateurs différentiels de degré 1 à coefficients constants. Notons ϕ cette représentation de $\mathfrak{so}(3)$ dans $\mathcal{C}^1(\mathbb{R}^3, \mathbb{C})$. On a en particulier :

$$\phi(\mathcal{I}) = 2x_3 \frac{\partial}{\partial x_2} - 2x_2 \frac{\partial}{\partial x_3}, \quad \phi(\mathcal{J}) = 2x_1 \frac{\partial}{\partial x_3} - 2x_3 \frac{\partial}{\partial x_1}, \quad \phi(\mathcal{K}) = 2x_2 \frac{\partial}{\partial x_1} - 2x_1 \frac{\partial}{\partial x_2}.$$

III.3.3. Opérateur de Casimir. — Pour toute représentation (ρ, V) de l'algèbre de Lie $\mathfrak{so}(3)$, notons $\Omega_\rho = \frac{1}{4}(\rho(\mathcal{I})^2 + \rho(\mathcal{J})^2 + \rho(\mathcal{K})^2)$. C'est un élément de $\text{End}(V)$, appelé opérateur de Casimir. On a dans $\text{End}(V)$,

$$\begin{aligned} [\rho(\mathcal{I}), \Omega_\rho] &= \frac{1}{4}([\rho(\mathcal{I}), \rho(\mathcal{I})^2] + [\rho(\mathcal{I}), \rho(\mathcal{J})^2] + [\rho(\mathcal{I}), \rho(\mathcal{K})^2]) \\ &= \frac{1}{4}(\rho(\mathcal{I})\rho(\mathcal{J})^2 - \rho(\mathcal{J})^2\rho(\mathcal{I}) + \rho(\mathcal{I})\rho(\mathcal{K})^2 - \rho(\mathcal{K})^2\rho(\mathcal{I})) \\ &= \frac{1}{4}((\rho(\mathcal{I})\rho(\mathcal{J}) - \rho(\mathcal{J})\rho(\mathcal{I}))\rho(\mathcal{J}) - \rho(\mathcal{J})(\rho(\mathcal{J})\rho(\mathcal{I}) - \rho(\mathcal{I})\rho(\mathcal{J}))) \\ &\quad + (\rho(\mathcal{I})\rho(\mathcal{K}) - \rho(\mathcal{K})\rho(\mathcal{I}))\rho(\mathcal{K}) - \rho(\mathcal{K})(\rho(\mathcal{J})\rho(\mathcal{K}) - \rho(\mathcal{I})\rho(\mathcal{K}))) \\ &= \frac{1}{4}(\rho([\mathcal{I}, \mathcal{J}])\rho(\mathcal{J}) - \rho(\mathcal{J})\rho([\mathcal{J}, \mathcal{I}]) + \rho([\mathcal{I}, \mathcal{K}])\rho(\mathcal{K}) - \rho(\mathcal{K})\rho([\mathcal{K}, \mathcal{I}])) \\ &= \frac{1}{4}(\rho(\mathcal{K})\rho(\mathcal{J}) + \rho(\mathcal{J})\rho(\mathcal{K}) - \rho(\mathcal{J})\rho(\mathcal{K}) - \rho(\mathcal{K})\rho(\mathcal{J})) \\ &= 0. \end{aligned}$$

De même, on trouve $[\rho(\mathcal{J}), \Omega_\rho] = [\rho(\mathcal{K}), \Omega_\rho] = 0$. Ceci montre que Ω_ρ est un opérateur d'entrelacement de la représentation (ρ, V) .

Corollaire III.3.3. — Si (ρ, V) est une représentation irréductible de dimension finie de l'algèbre de Lie $\mathfrak{so}(3)$, l'opérateur Ω_ρ est un multiple scalaire de l'identité.

Démonstration. C'est le lemme de Schur. □

Particularisons maintenant ceci à la représentation ϕ de $\mathfrak{so}(3)$ dans $\mathcal{C}^\infty(\mathbb{R}^3, \mathbb{C})$. On a $\Omega_\phi = \frac{1}{4}(\phi(\mathcal{I})^2 + \phi(\mathcal{J})^2 + \phi(\mathcal{K})^2)$. C'est un opérateur différentiel de degré 2 à coefficients constants sur

\mathbb{R}^3 . Plus explicitement, en notant $r^2 = x_1^2 + x_2^2 + x_3^2$,

$$\begin{aligned} \Omega_\phi &= \left(x_3 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_3} \right)^2 + \left(x_1 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_1} \right)^2 + \left(x_2 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_2} \right)^2 \\ &= x_3^2 \frac{\partial^2}{\partial x_2^2} + x_2^2 \frac{\partial^2}{\partial x_3^2} - x_3 \frac{\partial}{\partial x_3} - x_2 \frac{\partial}{\partial x_2} - 2x_2 x_3 \frac{\partial^2}{\partial x_2 \partial x_3} \\ &\quad + x_1^2 \frac{\partial^2}{\partial x_3^2} + x_3^2 \frac{\partial^2}{\partial x_1^2} - x_1 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_1} - 2x_1 x_3 \frac{\partial^2}{\partial x_1 \partial x_3} \\ &\quad + x_2^2 \frac{\partial^2}{\partial x_1^2} + x_1^2 \frac{\partial^2}{\partial x_2^2} - x_2 \frac{\partial}{\partial x_2} - x_1 \frac{\partial}{\partial x_1} - 2x_1 x_2 \frac{\partial^2}{\partial x_1 \partial x_2} \\ &= (r^2 - x_2^2) \frac{\partial^2}{\partial x_2^2} + (r^2 - x_3^2) \frac{\partial^2}{\partial x_3^2} + (r^2 - x_1^2) \frac{\partial^2}{\partial x_1^2} \\ &\quad - 2E - 2x_2 x_3 \frac{\partial^2}{\partial x_2 \partial x_3} - 2x_1 x_3 \frac{\partial^2}{\partial x_1 \partial x_3} - 2x_1 x_2 \frac{\partial^2}{\partial x_1 \partial x_2}, \end{aligned}$$

où E est l'opérateur d'Euler

$$E = \left(x_1 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2} + x_3 \frac{\partial}{\partial x_3} \right) = r \frac{\partial}{\partial r}.$$

Remarquons que

$$r^2 \frac{\partial^2}{\partial r^2} = \left(r \frac{\partial}{\partial r} \right)^2 - r \frac{\partial}{\partial r} = E^2 - E$$

et continuons le calcul

$$\Omega_\phi = r^2 \Delta - 2E - (E^2 - E).$$

Cette formule n'est rien d'autre que celle du Laplacien en coordonnées sphériques ($x_1 = r \sin \theta \cos \phi$, $x_2 = r \sin \theta \sin \phi$, $x_3 = r \cos \theta$), $r \geq 0$, $\phi \in [0, 2\pi[$, $\theta \in [0, \pi]$:

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \Omega_\rho$$

et Ω_ρ est le « Laplacien sphérique » Δ_{S^2} ,

$$\Delta_{S^2} = \frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2}.$$

III.3.4. Polynômes homogènes. — Pour tout $l \in \mathbb{N}$, notons \mathcal{P}_l l'espace des polynômes homogènes de degré l sur \mathbb{R}^3 . Un tel polynôme P s'écrit

$$P(X, Y, Z) = \sum_{k=0}^l P_k(X, Y) Z^{l-k}$$

où P_k est un polynôme homogène de degré k sur \mathbb{R}^2 . Or la dimension de l'espace des polynômes homogènes de degré k sur \mathbb{R}^2 est $k+1$, on a donc

$$\dim \mathcal{P}_l = \sum_{k=0}^l (k+1) = \frac{(l+1)(l+2)}{2}.$$

Notons \mathcal{H}_l le sous-espace de \mathcal{P}_l des polynômes homogènes de degré l et harmoniques sur \mathbb{R}^3 .

Lemme III.3.4. — Pour tout $l \in \mathbb{N}$, l'application linéaire

$$\Delta : \mathcal{P}_{l+2} \longrightarrow \mathcal{P}_l$$

est surjective.

Démonstration. Il est clair que Δ abaisse le degré d'un polynôme homogène de deux. Montrons que pour tout $(a, b, c) \in \mathbb{N}^3$ tel que $a + b + c = l$, $X^a Y^b Z^c$ est dans l'image de Δ . Raisonnons par récurrence sur $q = a + b$. Si $q = 0$, la propriété est vraie car

$$\Delta(Z^{l+2}) = (l+2)(l+1)Z^l.$$

De même si $q = 1$ puisque

$$\Delta(XZ^{l+1}) = (l+1)l XZ^{l-1}, \quad \Delta(YZ^{l+1}) = (l+1)l YZ^{l-1}.$$

En général, la formule

$$\Delta(X^a Y^b Z^c) = a(a-1)X^{a-2}Y^b Z^c + b(b-1)X^a Y^{b-2} Z^c + c(c-1)X^a Y^b Z^{c-2}$$

montre que la propriété au rang $q - 2$ implique la propriété au rang q . \square

Corollaire III.3.5. — $\dim \mathcal{H}_l = l + 1$.

Démonstration. Ce qui précède montre que l'on a une suite exacte

$$(III.3.1) \quad 0 \rightarrow \mathcal{H}_l \rightarrow \mathcal{P}_l \xrightarrow{\Delta} \mathcal{P}_{l-2} \rightarrow 0.$$

On en déduit que $\dim \mathcal{P}_l = \dim \mathcal{H}_l + \dim \mathcal{P}_{l-2}$. On obtient facilement la formule voulue par récurrence. \square

Il est clair que l'action de $\mathbf{SO}(3)$ sur les fonctions sur \mathbb{R}^3 préserve les espaces \mathcal{P}_l . Elle préserve aussi \mathcal{H}_l d'après le lemme III.3.2.

Théorème III.3.6. — L'espace \mathcal{H}_l est une représentation irréductible de $\mathbf{SO}(3)$.

Démonstration. D'après la classification des représentations irréductibles de $\mathbf{SO}(3)$, il n'existe à isomorphisme près qu'une seule représentation irréductible de dimension $2l + 1$. Considérons la représentation ϕ de $\mathfrak{so}(3)$ dans \mathcal{H}_l obtenue par différentiation en l'identité de la représentation de $\mathbf{SO}(3)$. On sait qu'elle est totalement réductible. Elle ne peut avoir comme composantes irréductibles que des représentations de dimension au plus $2l + 1$, et parmi celles-ci, seule celle de dimension $2l + 1$ admet un vecteur propre non nul pour l'opérateur $\phi(i\mathcal{K})$ de valeur propre $2l$. Or on a vu que

$$\phi(i\mathcal{K}) = i \left(2x_2 \frac{\partial}{\partial x_1} - 2x_1 \frac{\partial}{\partial x_2} \right).$$

Appliquons ceci au polynôme $P(x_1, x_2, x_3) = (x_1 + ix_2)^l$ qui est de manière évidente dans \mathcal{H}_l .

$$(\phi(i\mathcal{K}) \cdot P)(x_1, x_2, x_3) = 2il(x_2 - ix_1)(x_1 + ix_2)^{l-1} = 2lP(x_1, x_2, x_3).$$

On en déduit le théorème. \square

Proposition III.3.7. — On a, pour tout $l \geq 2$, en tant que représentation de $\mathbf{SO}(3)$,

$$\mathcal{P}_l = \mathcal{H}_l \oplus r^2 \mathcal{P}_{l-2}.$$

Démonstration. Il est clair que \mathcal{H}_l et $r^2\mathcal{P}_{l-2}$ sont des sous-espaces de \mathcal{P}_l stables sous l'action de G . Par comparaison des dimensions, on voit qu'il suffit de montrer que $\mathcal{H}_l \cap r^2\mathcal{P}_{l-2} = 0$. De manière équivalente, il s'agit de montrer que la suite exacte (III.3.1) admet une section

$$s : \mathcal{P}_{l-2} \longrightarrow \mathcal{P}_l$$

dont l'image est $r^2\mathcal{P}_{l-2}$. Une telle section s est une application linéaire vérifiant $\Delta(s(P)) = P$ pour tout $P \in \mathcal{P}_{l-2}$. La première idée est de prendre $s(P) = r^2P$. Or, pour tout $P \in \mathcal{P}_{l-2}$

$$\Delta(r^2P) = 6P + 4E(P) + r^2\Delta(P).$$

L'opérateur d'Euler E agit sur les polynômes homogènes de degré $l-2$ par le scalaire $l-2$ et donc

$$\Delta(r^2P) = 2(2l+1)P + r^2\Delta(P).$$

Plus généralement,

$$\Delta(r^{2k}P) = 2k(2l+2k-1)r^{2k-2}P + r^{2k}\Delta(P).$$

On se rapproche la section voulue en posant, pour une constante c à choisir :

$$s(P) = \frac{r^2}{2(2l+1)}P - cr^4\Delta(P),$$

ce qui nous donne

$$\Delta(s(P)) = P + \frac{r^2}{2(2l+1)}\Delta(P) - c4(2l+3)r^2\Delta(P) - cr^4\Delta^2(P).$$

On prend donc $c = \frac{1}{2(2l+1)4(2l+3)}$ pour annuler les termes médians. On réitère ce procédé, et l'on voit que l'on doit prendre, pour avoir $\Delta(s(P)) = P$,

$$\begin{aligned} s(P) &= \frac{r^2}{2(2l+1)}P - \frac{r^4}{2(2l+1)4(2l+3)}\Delta(P) \\ &+ \frac{r^6}{2(2l+1)4(2l+3)6(2l+5)}r^6\Delta^3(P) + \dots \\ &+ (-1)^{k+1} \frac{r^{2k}}{\prod_{i=1}^k 2i(2l+2i-1)}\Delta^k(P) + \dots \end{aligned}$$

Cette somme est finie, puisque pour k assez grand $\Delta^k(P)$ est nul. D'autre part il est clair que l'image de s est dans $r^2\mathcal{P}_{l-2}$. De plus, s est injective, puisque c'est une section, et en comparant les dimensions, on voit que l'image de s est exactement $r^2\mathcal{P}_{l-2}$. □

Corollaire III.3.8. — *On a pour tout $l \geq 2$, en tant que représentation de $\mathbf{SO}(3)$,*

$$\mathcal{P}_l = \mathcal{H}_l \oplus r^2\mathcal{H}_{l-4} \oplus \dots \oplus \begin{cases} r^l\mathcal{H}_0 & \text{si } l \text{ est pair} \\ r^{l-1}\mathcal{H}_1 & \text{si } l \text{ est impair.} \end{cases}$$

Remarquons que \mathcal{H}_0 est l'espace des polynômes constants et \mathcal{H}_1 celui des polynômes homogènes de degré 1.

D'après le corollaire III.3.3, l'opérateur $\Omega = \Omega_\phi$ agit par un scalaire sur chaque \mathcal{H}_l . Ce scalaire est facile à calculer, puisque nous avons établi plus haut que

$$\Omega = r^2 \Delta - E^2 - E.$$

En effet, Δ est nul sur \mathcal{H}_l et l'opérateur d'Euler agit par multiplication par le degré d'homogénéité l . On voit donc que Ω agit sur \mathcal{H}_l par le scalaire $-l^2 + l = -l(l+1)$.

III.3.5. Harmoniques sphériques. — La sphère S^2 est munie d'une action naturelle de $\mathbf{SO}(3)$. Cette action est transitive, et le stabilisateur du point $(0, 0, 1)$ de S^2 dans $\mathbf{SO}(3)$ est le sous-groupe des rotations d'axe Ox_3 , que l'on identifie à $\mathbf{SO}(2)$. Ainsi, on obtient un difféomorphisme

$$\mathbf{SO}(3)/\mathbf{SO}(2) \simeq S^2, \quad \bar{g} \mapsto g \cdot (0, 0, 1).$$

Toute fonction f sur S^2 peut-être vue comme une fonction \tilde{f} sur $\mathbf{SO}(3)$ vérifiant

$$\tilde{f}(gk) = \tilde{f}(g), \quad (\forall g \in \mathbf{SO}(3), \forall k \in \mathbf{SO}(2)).$$

Soit μ la mesure de Haar normalisée sur $\mathbf{SO}(3)$. Pour toute fonction continue f sur S^2 , posons

$$\langle \mu_{S^2}, f \rangle = \langle \mu, \tilde{f} \rangle.$$

On a alors, pour tout $g \in \mathbf{SO}(3)$,

$$\langle \mu_{S^2}, g \cdot f \rangle = \langle \mu, \widetilde{g \cdot f} \rangle = \langle \mu, g \cdot \tilde{f} \rangle = \langle \mu, \tilde{f} \rangle = \langle \mu_{S^2}, f \rangle.$$

On a ainsi défini une mesure μ_{S^2} sur S^2 , invariante par l'action de $\mathbf{SO}(3)$. On peut vérifier qu'en coordonnées sphériques, cette mesure est donnée par

$$\frac{1}{4\pi} \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} f(\theta, \phi) \sin \theta \, d\theta \, d\phi.$$

L'espace $H = L^2(S^2, \mu_{S^2})$ est un espace de Hilbert pour le produit scalaire

$$(f_1, f_2) = \int_{S^2} \overline{f_1(\underline{x})} f_2(\underline{x}) \, d\mu_{S^2}(\underline{x}).$$

L'invariance de la mesure μ_{S^2} entraîne que la représentation de $\mathbf{SO}(3)$ dans H est unitaire.

Un polynôme homogène de degré l sur \mathbb{R}^3 est déterminé par sa restriction à S^2 . Appelons respectivement $\tilde{\mathcal{P}}_l$ et $\tilde{\mathcal{H}}_l$ les espaces des restrictions à S^2 des polynômes dans \mathcal{P}_l et \mathcal{H}_l respectivement. Ce sont des espaces de représentations pour $\mathbf{SO}(3)$, et le corollaire III.3.8 nous donne

$$(III.3.2) \quad \tilde{\mathcal{P}}_l = \tilde{\mathcal{H}}_l \oplus \tilde{\mathcal{H}}_{l-4} \oplus \dots \oplus \begin{cases} \tilde{\mathcal{H}}_0 & \text{si } l \text{ est pair} \\ \tilde{\mathcal{H}}_1 & \text{si } l \text{ est impair.} \end{cases}$$

Ceci est une décomposition de $\tilde{\mathcal{P}}_l$ en représentations irréductibles de $\mathbf{SO}(3)$. D'après ce qui précède, ce sont des représentations unitaires. Les éléments de l'algèbre de Lie $\mathfrak{so}(3)$ agissent dans ces espaces par des opérateurs anti-hermitiens. On en déduit que l'opérateur de Casimir Ω est lui un opérateur hermitien. Ses sous-espaces propres sont donc orthogonaux, et la décomposition de $\tilde{\mathcal{P}}_l$ établie ci-dessus est une décomposition orthogonale.

Voici une autre manière de présenter ces résultats. L'opérateur de Casimir Ω est le « Laplacien sphérique » Δ_{S^2} . C'est un opérateur différentiel de degré 2 sur la sphère S^2 , qui agit sur $\tilde{\mathcal{P}}_l$, cette

action commutant à celle de $\mathbf{SO}(3)$. Les sous-espaces propres de Ω dans $\tilde{\mathcal{P}}_l$ sont donc des sous-représentations de $\mathbf{SO}(3)$. Ce qui a été fait ci-dessus montre qu'elles sont irréductibles.

Théorème III.3.9. — *L'espace $H = L^2(S^2, \mu_{S^2})$ est somme directe hilbertienne des $\tilde{\mathcal{H}}_l$ pour $l \in \mathbb{N}$.*

Démonstration. On a vu que les $\tilde{\mathcal{H}}_l$ sont orthogonaux puisque ce sont des sous-espaces propres pour l'opérateur hermitien Ω . (On remarque que Ω étant un opérateur différentiel, il n'est pas bien défini sur $H = L^2(S^2, \mu_{S^2})$, mais il l'est certainement sur le sous-espace des fonctions \mathcal{C}^∞ par exemple, qui contient tout les $\tilde{\mathcal{H}}_l$). Il reste à démontrer la densité de la somme directe algébrique des $\tilde{\mathcal{H}}_l$ dans H . Ceci est une conséquence du théorème d'approximation de Stone-Weierstrass. On peut approcher toute fonction continue sur S^2 par des restrictions de fonctions polynomiales sur \mathbb{R}^3 . On peut ensuite décomposer chaque polynôme en somme de polynômes homogènes et utiliser la décomposition (III.3.2) pour conclure. \square

Remarque III.3.10. — Dans la décomposition Hilbertienne

$$L^2(S^2, \mu_{S^2}) = \widehat{\bigoplus_{l \in \mathbb{N}} \tilde{\mathcal{H}}_l}$$

chaque représentation irréductible de $\mathbf{SO}(3)$ apparaît avec multiplicité 1.

CHAPITRE IV

REPRÉSENTATIONS DU GROUPE DE HEISENBERG

Ce chapitre est constitué de l'énoncé d'un problème d'examen et de son corrigé.

Première Partie : Représentations de dimension finie du groupe de Heisenberg.

Soit $\mathfrak{h} = \left\{ \begin{pmatrix} 0 & x & z \\ 0 & 0 & y \\ 0 & 0 & 0 \end{pmatrix} ; x, y, z \in \mathbb{R} \right\}$.

Posons $X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ et $Z = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

— 1. Montrer que $[X, Y] = Z$, $[X, Z] = [Y, Z] = 0$. En déduire que \mathfrak{h} est une sous-algèbre de Lie de $\mathfrak{gl}(3, \mathbb{R})$.

— 2. Calculer $\exp tX$, $\exp tY$ et $\exp tZ$ pour tout $t \in \mathbb{R}$.

— 3. Calculer $g_{a,b,c} = \exp cZ \exp bY \exp aX$. Exprimer $\exp aX \exp bY$ comme un élément $g_{a,b,c}$. Montrer que $g_{a,b,c}$ est dans l'image de l'exponentielle.

— 4. Quel est le sous-groupe connexe de $\mathbf{GL}(3, \mathbb{R})$ d'algèbre de Lie \mathfrak{h} ? On le note H . Montrer que $\exp : \mathfrak{h} \rightarrow H$ est bijective.

— 5. Quelles sont les sous-algèbres de Lie de \mathfrak{h} ?

— 6. Soit (ϕ, V) une représentation de dimension finie de \mathfrak{h} . Soit λ une valeur propre de l'opérateur $\phi(Z)$. Montrer que le sous-espace propre V_λ associé à cette valeur propre est stable par ϕ .

— 7. On suppose maintenant que (ϕ, V) est irréductible. Que vaut $\phi(Z)$? En déduire que la valeur propre λ est nulle.

— 8. Montrer qu'alors $\dim V = 1$. On pose $\phi(X) = \alpha$, $\phi(Y) = \beta$.

— 9. Montrer que les classes d'équivalence de représentations irréductibles de dimension finie de \mathfrak{h} forment une famille à deux paramètres $\phi_{\alpha,\beta}$, $\alpha, \beta \in \mathbb{C}$. Ces représentations sont-elles

les différentielles en Id de représentations de H . Les expliciter le cas échéant. Quelles sont les représentations unitaires de dimension finie de H ?

Deuxième partie : Représentations unitaires irréductibles de H

Soit $\mathcal{S}(\mathbb{R})$ l'espace de Schwartz sur \mathbb{R} . Rappelons que $\mathcal{S}(\mathbb{R})$ est l'espace des fonctions \mathcal{C}^∞ sur \mathbb{R} , à valeurs complexes, qui sont, ainsi que toutes leurs dérivées, à "décroissance rapide" lorsque $t \rightarrow \pm\infty$. Ceci assure que les intégrales qui apparaissent ci-dessous convergent, et nous ne nous préoccupons plus de ce problème. On munit $\mathcal{S}(\mathbb{R})$ du produit hermitien (défini positif) :

$$(f_1, f_2) = \int_{\mathbb{R}} f_1(t) \overline{f_2(t)} dt$$

La norme associée sera notée $\|\cdot\|_2$.

Considérons sur $\mathcal{S}(\mathbb{R})$ l'opérateur

$$a = \frac{1}{\sqrt{2}} \left(t + \frac{d}{dt} \right),$$

C'est-à-dire que $a(f)(t) = \frac{1}{\sqrt{2}}(tf(t) + f'(t))$. En général, si $g \in \mathcal{C}^\infty(\mathbb{R})$, on confond dans les notations la fonction g et l'opérateur $f \mapsto gf$ sur $\mathcal{S}(\mathbb{R})$.

— 1. Montrer que son adjoint pour le produit hermitien est

$$a^* = \frac{1}{\sqrt{2}} \left(t - \frac{d}{dt} \right).$$

— 2. Vérifier que $[a, a^*] = \text{Id}_{\mathcal{S}(\mathbb{R})}$. En déduire que

$$\phi : X \mapsto a, \quad Y \mapsto a^*, \quad Z \mapsto \text{Id}_{\mathcal{S}(\mathbb{R})}$$

définit une représentation de \mathfrak{h} dans $\mathcal{S}(\mathbb{R})$.

— 3. On pose $E = aa^* + a^*a$ et $N = a^*a$. Montrer que $[E, a] = -2a$, $[E, a^*] = 2a^*$ et que $[N, a] = -a$, $[N, a^*] = a^*$.

— 4. Résoudre $Nf = 0$ dans $\mathcal{S}(\mathbb{R})$ (on pourra chercher une solution de la forme $t \mapsto e^{\alpha t^2}$). Montrer qu'il existe une unique solution v_0 de $Nf = 0$ dans $\mathcal{S}(\mathbb{R})$ vérifiant $(v_0, v_0) = 1$.

— 5. Montrer que si v est vecteur propre de N dans $\mathcal{S}(\mathbb{R})$ pour la valeur propre λ , alors $a \cdot v$ est dans le sous-espace propre de N pour la valeur propre $\lambda - 1$ et $a^* \cdot v$ est dans le sous-espace propre pour la valeur propre $\lambda + 1$.

— 6. On définit par récurrence $v_{k+1} = \frac{a^* \cdot v_k}{\sqrt{k+1}}$. Montrer que les $(v_k)_{k \in \mathbb{N}}$ forment une famille orthonormale de $\mathcal{S}(\mathbb{R})$. Montrer que les v_k sont des fonctions de la forme

$$v_k(t) = \frac{H_k(t)}{\sqrt{k+1}} e^{-\frac{t^2}{2}},$$

où les H_k sont des polynômes (polynômes d'Hermite)

— 7. On pose $p = \frac{a-a^*}{\sqrt{2i}} = -i \frac{d}{dt}$ et $q = \frac{a+a^*}{\sqrt{2}} = t$. Justifiez que si l'on pose $u(s) = e^{isq}$ et $v(s) = e^{isp}$ pour tout $s \in \mathbb{R}$, alors pour tout $f \in \mathcal{S}(\mathbb{R})$

$$(u(s) \cdot f)(t) = e^{ist} f(t), \quad (v(s) \cdot f)(t) = f(t+s).$$

(On demande simplement une justification et non une démonstration, car la démonstration complète est hors de portée du fait que $\mathcal{S}(\mathbb{R})$ est de dimension infinie : on pourra admettre que les propriétés de l'exponentielle d'un opérateur sont les mêmes qu'en dimension finie, par exemple le développement en série où l'équation différentielle vérifiée par l'exponentielle).

— 8. Pour tout $\lambda \in \mathbb{R}$, on pose

$$\pi_\lambda(g_{a,b,c}) = e^{2i\pi\lambda c} u(\lambda b) v(a).$$

Montrer que π_λ définit une représentation de H et qu'elle est continue et unitaire.

L'espace $\mathcal{S}(\mathbb{R})$ n'est pas complet pour la norme $\|\cdot\|_2$. Son complété est l'espace $L^2(\mathbb{R})$. On prolonge π_λ à $L^2(\mathbb{R})$ par continuité. On obtient ainsi une représentation unitaire de H dans l'espace de Hilbert $L^2(\mathbb{R})$.

— 9. Soit A un opérateur sur $L^2(\mathbb{R})$ qui commute avec les translations et la multiplication par t . On admet qu'alors A est un multiple scalaire de l'identité. Dédire des questions précédentes que $(\pi_\lambda, L^2(\mathbb{R}))$ est irréductible.

Corrigé

Première partie.

— 1. On fait le calcul. Tout élément de \mathfrak{h} est une combinaison linéaire de X , Y et Z . Le commutateur de deux éléments de \mathfrak{h} est encore une telle combinaison linéaire, donc \mathfrak{h} est une sous-algèbre de Lie de $\mathfrak{gl}(3, \mathbb{R})$.

— 2. On constate que $X^2 = Y^2 = Z^2 = 0$, donc

$$\exp tX = \text{Id} + tX = \begin{pmatrix} 1 & t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

De même,

$$\exp tY = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \text{ et } \exp tZ = \begin{pmatrix} 1 & 0 & t \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

$$\text{— 3. } g_{a,b,c} = \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}.$$

$$\exp aX \exp bY = \begin{pmatrix} 1 & a & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a & ab \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix},$$

d'où $\exp aX \exp bY = g_{a,b,ab}$.

On a $\exp(aX + bY + cZ) = g_{a,b,c+\frac{ab}{2}}$, d'où

$$g_{a,b,c} = \exp(aX + bY + (c - \frac{ab}{2})Z).$$

— 4. D'après ce qui précède, le sous-groupe de $\mathbf{GL}(3, \mathbb{R})$ engendré par $\exp \mathfrak{h}$ est

$$H = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}, \quad a, b, c \in \mathbb{R}. \right\}$$

Il est clair que l'application exponentielle est injective et surjective.

— 5. Toutes les droites vectorielles dans \mathfrak{h} sont bien sûr des sous-algèbres de Lie. Cherchons maintenant les sous-algèbres de dimension 2. Soient A et B deux éléments engendrant une telle sous-algèbre \mathfrak{a} . Le crochet de deux éléments de \mathfrak{h} est toujours dans la droite engendrée par Z . Supposons $[A, B] \neq 0$. Alors A , B et Z sont liés. On peut alors remplacer soit A , soit B par Z pour obtenir une base de \mathfrak{a} . Mais on voit alors que le crochet est toujours nul : contradiction. Donc $[A, B] = 0$. Si Z n'est pas dans \mathfrak{a} , on voit qu'alors \mathfrak{h} est engendré par A , B et Z : contradiction car \mathfrak{h} serait abélienne. Donc \mathfrak{a} contient Z . Réciproquement, tout espace de dimension 2 qui contient Z est une sous-algèbre abélienne de \mathfrak{h} .

— 6. Soit v un vecteur propre non nul pour la valeur propre λ de $\phi(Z)$. On a, pour tout $A \in \mathfrak{h}$, $[Z, A] = 0$, donc

$$0 = \phi([Z, A]) \cdot v = \phi(Z)\phi(A) \cdot v - \phi(X)\phi(A) \cdot v = \phi(Z)\phi(A) \cdot v - \lambda\phi(A) \cdot v.$$

Ceci montre que $\phi(A) \cdot v \in V_\lambda$.

— 7. Comme ϕ est irréductible, $V = V_\lambda$ et donc $\phi(Z)$ est l'opérateur λId_V . Or $\text{Tr } \lambda \text{Id}_V = \lambda \dim V$ et

$$\begin{aligned} \text{Tr } \phi(Z) &= \text{Tr } (\phi([X, Y])) = \text{Tr } ([\phi(X), \phi(Y)]) \\ &= \text{Tr } (\phi(X)\phi(Y)) - \text{Tr } (\phi(Y)\phi(X)) = 0. \end{aligned}$$

Ceci montre que $\lambda = 0$.

— 8. On a donc $\phi(Z) = 0$ et

$$\phi(Z) = \phi([X, Y]) = [\phi(X), \phi(Y)] = \phi(X)\phi(Y) - \phi(Y)\phi(X) = 0$$

donc les opérateurs $\phi(X)$ et $\phi(Y)$ commutent. Soit v un vecteur propre non nul simultanément de $\phi(X)$ et $\phi(Y)$. Il est clair que le sous-espace engendré par ce vecteur est stable par ϕ , et comme ϕ est irréductible, c'est l'espace V tout entier. Donc V est de dimension 1.

— 9. Réciproquement, si α, β sont dans \mathbb{C} , on pose $\phi(X) = \alpha$, $\phi(Y) = \beta$, $\phi(Z) = 0$. Il est clair que ceci définit une représentation de dimension 1 de \mathfrak{h} . On la note $\phi_{\alpha, \beta}$. Des représentations de dimension 1 sont équivalentes si et seulement si elles sont égales. On cherche maintenant des représentations $\pi_{\alpha, \beta}$ de H telles que $(d\pi_{\alpha, \beta})|_{\text{Id}} = \phi_{\alpha, \beta}$. On doit avoir

$$\pi_{\alpha, \beta}(\exp tZ) = \exp(t\phi_{\alpha, \beta}(Z)) = \exp 0 = \text{Id}$$

$$\pi_{\alpha, \beta}(\exp tX) = \exp(t\phi_{\alpha, \beta}(X)) = \exp t\alpha$$

$$\pi_{\alpha, \beta}(\exp tY) = \exp(t\phi_{\alpha, \beta}(Y)) = \exp t\beta$$

D'où $\pi_{\alpha, \beta}(g_{a, b, c}) = \exp(a\alpha) \exp(b\beta)$.

Comme $g_{a, b, c} g_{a', b', c'} = g_{a+a', b+b', c+c'+ab'}$ on voit que, réciproquement,

$$\pi_{\alpha, \beta}(g_{a, b, c}) = \exp(a\alpha) \exp(b\beta)$$

définit une représentation de H , dont la différentielle est $\phi_{\alpha, \beta}$.

Toute représentation π de H définit par différentiation une représentation ϕ de \mathfrak{h} . Si π est irréductible, alors ϕ aussi (cours). Les seules représentations irréductibles de H sont donc les $\pi_{\alpha, \beta}$. Elles sont unitaires si et seulement si $\alpha \in i\mathbb{R}$, $\beta \in i\mathbb{R}$.

Deuxième partie.

— 1. On calcule à l'aide d'une intégration par partie :

$$\begin{aligned}
(af_1, f_2) &= \int_{\mathbb{R}} \frac{1}{\sqrt{2}} (tf_1(t) + f_1'(t)) \overline{f_2(t)} dt \\
&= \int_{\mathbb{R}} \frac{1}{\sqrt{2}} f_1(t) \overline{tf_2(t)} dt + \int_{\mathbb{R}} \frac{1}{\sqrt{2}} f_1'(t) \overline{f_2(t)} dt \\
&= \int_{\mathbb{R}} \frac{1}{\sqrt{2}} f_1(t) \overline{tf_2(t)} dt - \int_{\mathbb{R}} \frac{1}{\sqrt{2}} f_1(t) \overline{f_2'(t)} dt \\
&= \int_{\mathbb{R}} \frac{1}{\sqrt{2}} f_1(t) (\overline{tf_2(t) - f_2'(t)}) dt = (f_1, a^* f_2)
\end{aligned}$$

— 2.

$$\begin{aligned}
((aa^* - a^*a)(f))(t) &= \frac{1}{2} \left((t + \frac{d}{dt})(t - \frac{d}{dt})f(t) - (t - \frac{d}{dt})(t + \frac{d}{dt})f(t) \right) \\
&= \frac{1}{2} (t^2 f(t) - t \frac{d}{dt} f(t) + \frac{d}{dt} t f(t) - \frac{d^2}{dt^2} f(t)) \\
&\quad - \frac{1}{2} (t^2 f(t) + t \frac{d}{dt} f(t) - \frac{d}{dt} t f(t) - \frac{d^2}{dt^2} f(t)) \\
&= \frac{d}{dt} t f(t) - t \frac{d}{dt} f(t) = f(t) + t f'(t) - t f'(t) = f(t)
\end{aligned}$$

On a $\phi([X, Y]) = \phi(Z) = \text{Id}_{\mathcal{S}(\mathbb{R})}$ et $[\phi(X), \phi(Y)] = [a, a^*] = \text{Id}_{\mathcal{S}(\mathbb{R})}$. De même, pour tout $A \in \mathfrak{h}$, $\phi([A, Z]) = \phi(0) = 0$ et $[\phi(A), \phi(Z)] = [\phi(A), \text{Id}_{\mathcal{S}(\mathbb{R})}] = 0$ car $\text{Id}_{\mathcal{S}(\mathbb{R})}$ commute avec tous les opérateurs sur $\mathcal{S}(\mathbb{R})$. Par linéarité, on en déduit que

$$\phi([A, B]) = [\phi(A), \phi(B)]$$

pour tous $A, B \in \mathfrak{h}$, ce qui signifie que ϕ est une représentation de \mathfrak{h} dans $\mathcal{S}(\mathbb{R})$.

— 3. Comme $aa^* = a^*a + 1$

$$\begin{aligned}
[E, a] &= (aa^* + a^*a)a - a(aa^* + a^*a) = (2a^*a + 1)a - a(2a^*a + 1) \\
&= 2a^*a^2 - 2aa^*a = 2a^*a^2 - 2(a^*a + 1)a = -2a
\end{aligned}$$

De même $[E, a^*] = 2a^*$. Comme $N = \frac{1}{2}(E - [a, a^*]) = \frac{1}{2}(E - \text{Id}_{\mathcal{S}(\mathbb{R})})$, on en déduit

$$[N, a] = -a, \quad [N, a^*] = a^*.$$

— 4. L'équation $N \cdot f = 0$ est l'équation différentielle linéaire du second ordre, homogène

$$\frac{1}{2}((t^2 - 1)f(t) - f''(t)) = 0$$

dont l'espace des solutions est de dimension 2.

Cherchons les solutions sous la forme $t \mapsto e^{\alpha t^2}$. On trouve $\alpha = -\frac{1}{2}$. On cherche une autre solution par la méthode de variation de la constante : $K(t)e^{-\frac{1}{2}t^2}$. On trouve $K''(t) = 2tK'(t)$, soit $K'(t) = Ce^{t^2}$ et l'on obtient une solution linéairement indépendante de la première qui n'est pas à décroissance rapide, donc pas dans $\mathcal{S}(\mathbb{R})$.

Les seules solutions dans $\mathcal{S}(\mathbb{R})$ sont de la forme

$$t \mapsto K_1 e^{-\frac{t^2}{2}}.$$

Si l'on impose de plus la condition $(v_0, v_0) = 1$, ceci fixe la constante K_1 .

— 5. On a, pour $v \in \mathcal{S}(\mathbb{R})$ telle que $N \cdot v = \lambda v$

$$-a \cdot v = [N, a] \cdot v = Na \cdot v - aN \cdot v = Na \cdot v - \lambda a \cdot v$$

d'où $Na \cdot v = (\lambda - 1)a \cdot v$. Ceci montre que $a \cdot v$ dans le sous-espace propre de N pour la valeur propre $\lambda - 1$. Le calcul pour $Na^* \cdot v$ est similaire.

— 6. D'après ce qui précède, v_k est vecteur propre de N pour la valeur propre k . Comme $N = a^*a$ est autoadjoint, ses sous-espaces propres sont donc orthogonaux. Pour montrer que $(v_k, v_k) = 1$, on raisonne par récurrence. On a

$$(v_{k+1}, v_{k+1}) = \frac{1}{k+1} (a^*v_k, a^*v_k) = \frac{1}{k+1} (aa^*v_k, v_k)$$

Or $aa^* = N + 1$ d'où

$$(v_{k+1}, v_{k+1}) = \frac{1}{k+1} ((N+1)v_k, v_k) = (v_k, v_k).$$

On montre v_k est de la forme voulue par récurrence sur k , sans difficulté. Les H_k sont les polynômes d'Hermitte.

— 8. On a :

$$\begin{aligned} (e^{isq} \cdot f)(t) &= (Id_{\mathcal{S}(\mathbb{R})} \cdot f + isq \cdot f + \frac{(isq)^2}{2!} \cdot f + \frac{(isq)^3}{3!} \cdot f + \dots)(t) \\ &= f(t) + istf(t) + \frac{(ist)^2}{2!} f(t) + \frac{(ist)^3}{3!} f(t) + \dots = e^{ist} f(t) \end{aligned}$$

Pour calculer $(e^{isp} \cdot f)(t)$, on peut envisager deux méthodes. La première est que $s \mapsto e^{isp}$ est l'unique solution de $a'(s) = ipa(s)$, telle que $a(0) = Id_{\mathcal{S}(\mathbb{R})}$. Or si l'on pose

$$(a(s) \cdot f)(t) = f(t+s)$$

alors

$$(a'(s) \cdot f)(t) = \frac{d}{ds}(a(s) \cdot f)(t) = \frac{d}{ds}f(t+s) = f'(t+s) = (ipa(s) \cdot f)(t)$$

et bien sûr $a(0) = Id_{\mathcal{S}(\mathbb{R})}$.

Une autre méthode est de calculer formellement

$$\begin{aligned} (e^{isp} \cdot f)(t) &= (Id_{\mathcal{S}(\mathbb{R})} \cdot f + isp \cdot f + \frac{(isp)^2}{2!} \cdot f + \frac{(isp)^3}{3!} \cdot f + \dots)(t) \\ &= (f + sf' + \frac{s^2 f''}{2!} + \frac{s^3 f^{(3)}}{3!} + \dots)(t) \end{aligned}$$

On reconnaît le développement de Taylor de $f(t+s)$.

— 8. On veut montrer que

$$\pi_\lambda(g_{a,b,c}) = e^{2i\pi\lambda c} u(\lambda b)v(a)$$

définit bien un morphisme de groupes de H dans $GL(\mathcal{S}(\mathbb{R}))$. Pour cela, il s'agit de vérifier que

$$\pi_\lambda(g_{a,b,c}g_{a',b',c'}) = \pi_\lambda(g_{a,b,c})\pi_\lambda(g_{a',b',c'}).$$

Or $g_{a,b,c}g_{a',b',c'} = g_{a+a',b+b',c+c'+ab'}$ donc

$$\pi_\lambda(g_{a,b,c}g_{a',b',c'}) = \pi_\lambda(g_{a+a',b+b',c+c'+ab'}) = e^{2i\pi\lambda(c+c'+ab')}u(\lambda(b+b'))v(a+a').$$

D'autre part

$$\begin{aligned} \pi_\lambda(g_{a,b,c})\pi_\lambda(g_{a',b',c'}) &= e^{2i\pi\lambda c}u(\lambda b)v(a)e^{2i\pi\lambda c'}u(\lambda b')v(a') \\ &= e^{2i\pi\lambda(c+c')}u(\lambda b)v(a)u(\lambda b')v(a'). \end{aligned}$$

Il s'agit de faire commuter $v(a)$ et $u(\lambda b')$. Or

$$(v(a)u(\lambda b') \cdot f)(t) = e^{i\lambda b'(t+a)}f(t+a)$$

et

$$(u(\lambda b')v(a) \cdot f)(t) = e^{i\lambda b't}f(t+a)$$

D'où

$$\begin{aligned} \pi_\lambda(g_{a,b,c})\pi_\lambda(g_{a',b',c'}) &= e^{2i\pi\lambda(c+c'+ab')}u(\lambda b)u(\lambda b')v(a)v(a') \\ &= e^{2i\pi\lambda(c+c'+ab')}u(\lambda(b+b'))v(a+a'). \end{aligned}$$

Montrons maintenant que π_λ est unitaire. Il s'agit de montrer que pour tout $g_{a,b,c} \in H$ et pour toutes fonctions $f_1, f_2 \in \mathcal{S}(\mathbb{R})$, on a

$$(\pi_\lambda(g_{a,b,c}) \cdot f_1, \pi_\lambda(g_{a,b,c}) \cdot f_2) = (f_1, f_2).$$

Or

$$\begin{aligned} &(\pi_\lambda(g_{a,b,c}) \cdot f_1, \pi_\lambda(g_{a,b,c}) \cdot f_2) \\ &= \int_{\mathbb{R}} e^{i\lambda bt}f_1(t+a) \overline{e^{i\lambda bt}f_2(t+a)}dt \\ &= \int_{\mathbb{R}} f_1(t+a)\overline{f_2(t+a)}dt = \int_{\mathbb{R}} f_1(t)\overline{f_2(t)}dt = (f_1, f_2) \end{aligned}$$

On a utilisé l'invariance de la mesure de Lebesgue par translation.

Pour la continuité : il s'agit de montrer que

$$H \times \mathcal{S}(\mathbb{R}) \rightarrow \mathcal{S}(\mathbb{R}), \quad (g_{a,b,c}, f) \mapsto \pi_\lambda(g_{a,b,c}) \cdot f$$

est continue. Soient $g_{a,b,c}, g_{a',b',c'} \in H$ et $f_1, f_2 \in \mathcal{S}(\mathbb{R})$. On a

$$\begin{aligned}
 & \|\pi_\lambda(g_{a,b,c}) \cdot f_1 - \pi_\lambda(g_{a',b',c'}) \cdot f_2\|_2 \\
 &= \left(\int_{\mathbb{R}} |e^{2i\pi\lambda c} e^{i\lambda b t} f_1(t+a) - e^{2i\pi\lambda c'} e^{i\lambda b' t} f_2(t+a')|^2 dt \right)^{1/2} \\
 &= \left(\int_{\mathbb{R}} |e^{2i\pi\lambda(c-c')} e^{i\lambda(b-b')(t-a')} f_1(t+a-a') - f_2(t)|^2 dt \right)^{1/2} \\
 &= \|e^{2i\pi\lambda(c-c')} e^{i\lambda(b-b')(t-a')} f_1(t+a-a') - f_2(t)\|_2 \\
 &\leq \|e^{2i\pi\lambda(c-c')} e^{i\lambda(b-b')(t-a')} f_1(t+(a-a')) - f_1(t+(a-a'))\|_2 \\
 &\quad + \|f_1(t+(a-a')) - f_1(t)\|_2 + \|f_1(t) - f_2(t)\|_2
 \end{aligned}$$

Or

$$\begin{aligned}
 & \|e^{2i\pi\lambda(c-c')} e^{i\lambda(b-b')(t-a')} f_1(t+(a-a')) - f_1(t+(a-a'))\|_2 \\
 &= \|(e^{2i\pi\lambda(c-c')} e^{i\lambda(b-b')(t-a)} - 1) f_1(t)\|_2
 \end{aligned}$$

Fixons $\epsilon > 0$. Comme f_1 est de carré intégrable, il existe un réel $R > 0$ tel que

$$\int_{\mathbb{R} \setminus [-R, R]} |f_1(t)|^2 dt \leq \frac{\epsilon^2}{4}$$

et comme

$$|(e^{2i\pi\lambda(c-c')} e^{i\lambda(b-b')(t-a)} - 1)| \leq 2$$

On obtient

$$\int_{\mathbb{R} \setminus [-R, R]} |(e^{2i\pi\lambda(c-c')} e^{i\lambda(b-b')(t-a)} - 1) f_1(t)|^2 dt \leq \epsilon^2$$

D'autre part, on si $|c - c'|$ et $|b - b'|$ sont suffisamment petits, on a par continuité uniforme sur le compact $[-R, R]$:

$$\int_{[-R, R]} |(e^{2i\pi\lambda(c-c')} e^{i\lambda(b-b')(t-a)} - 1) f_1(t)|^2 dt \leq \epsilon^2 \|f_1\|_2^2,$$

d'où :

$$\|(e^{2i\pi\lambda(c-c')} e^{i\lambda(b-b')(t-a)} - 1) f_1(t)\|_2 \leq \epsilon(1 + \|f_1\|_2^2)^{1/2}.$$

La majoration de $\|f_1(t+(a-a')) - f_1(t)\|_2$ est similaire. On coupe l'intégrale en deux : sur $\mathbb{R} \setminus [-R, R]$ et sur $[R, R]$, où l'on utilise la continuité uniforme de f_1 et $|a - a'|$ petit. On obtient ainsi

$$\|f_1(t+(a-a')) - f_1(t)\|_2 \leq \epsilon$$

Au final, on voit que l'on peut rendre

$$\|\pi_\lambda(g_{a,b,c}) \cdot f_1 - \pi_\lambda(g_{a',b',c'}) \cdot f_2\|_2$$

aussi petit que l'on veut en prenant $|c - c'|$, $|a - a'|$, $|b - b'|$ et $\|f_1 - f_2\|_2$ suffisamment petits. Ceci montre la continuité.

— 9. Soit A un opérateur d'entrelacement de π_λ avec elle-même. Alors A commute avec les translations et les multiplications par e^{ist} . On a donc pour toute fonction $f \in \mathcal{S}(\mathbb{R})$, pour tout réel s :

$$e^{ist}(A \cdot f)(t) = A(e^{ist}f)$$

et en dérivant par rapport à s , puis en évaluant en $s = 0$

$$it(A \cdot f)(t) = A(itf).$$

Donc A commute avec la multiplication par t : on en déduit que A est un opérateur scalaire.

Supposons que π_λ ne soit pas irréductible. Alors il existe un sous-espace W fermé de $L^2(\mathbb{R})$ stable par H , qui n'est ni trivial ni $L^2(\mathbb{R})$ tout entier. Comme π_λ est unitaire, on a $L^2(\mathbb{R}) = W \oplus W^\perp$. Mais alors, les projections sur W et W^\perp sont des opérateurs d'entrelacement de π_λ qui ne sont pas des opérateurs scalaires : contradiction.

Remarque : Il est connu que les $(v_k)_{k \in \mathbb{N}}$ forment une base hilbertienne de $L^2(\mathbb{R})$. On peut utiliser ceci pour montrer qu'un opérateur qui commute avec les translations et les multiplications par t est scalaire.

CHAPITRE V

SYMÉTRIES EN MÉCANIQUE QUANTIQUE

Le but est ici de faire le lien entre les symétries en mécanique quantique et la théorie des groupes. On commencera par des exemples simples liés au groupe des rotations, puis on expliquera comment les symétries sont implémentées dans la description quantique d'un système physique. On analysera les conséquences de l'existence de symétries.

V.1. Actions sur les états quantiques

On rappelle que la mécanique quantique est fondée sur les principes suivants :

— L'état d'un système est décrit par un "rayon" d'un espace de Hilbert \mathcal{H} . Un "rayon" est une classe d'équivalence de vecteurs définis à une phase près. On notera les vecteurs de \mathcal{H} par ψ ou, en suivant P.M. Dirac, par le "ket" $|\psi\rangle$. Avec cette notation, les vecteurs $|\psi\rangle$ et $e^{i\alpha}|\psi\rangle$ représentent le même état physique. Le produit hilbertien entre deux vecteurs sera noté par le "bracket" $\langle\psi_1|\psi_2\rangle$ où le "bra" $\langle\psi_1|$ est le dual de $|\psi_1\rangle$.

— Les grandeurs expérimentalement observables sont représentées par des opérateurs A auto-adjoints sur \mathcal{H} : $A = A^\dagger$. Le résultat d'une mesure de la grandeur représentée par A ne peut être qu'un élément du spectre de A . Soit $I = [a_1, a_2] \subset \text{Spec}(A)$ un intervalle inclus dans le spectre de A et P_I le projecteur spectral associé. Si le système est préparé dans l'état $|\psi\rangle$, la probabilité d'observer un élément de I lors d'une mesure de A est $\langle\psi|P_I|\psi\rangle/\langle\psi|\psi\rangle$.

— L'évolution de l'état du système est gouvernée par l'équation de Schroedinger ⁽¹⁾ :

$$i\partial_t|\psi(t)\rangle = H|\psi(t)\rangle$$

où H , opérateur auto-adjoint sur \mathcal{H} , est l'hamiltonien du système. Son spectre donne l'ensemble des énergies accessibles au système.

V.1.1. Translations et rotations. — Commençons par des exemples.

Considérons le système simple constitué d'une particule (massive, non relativiste) se déplaçant dans l'espace tri-dimensionnel, dont la position est notée $\mathbf{x} = (x, y, z)$. En mécanique quantique, l'état du système est décrit par une fonction d'onde $\psi(\mathbf{x})$, vue comme un élément de l'espace $L^2(\mathbb{R}^3)$ dont la structure hilbertienne est définie par $\|\psi\|^2 = \int d\mathbf{x}|\psi(\mathbf{x})|^2$. Le module carré de la

1. Nous posons $\hbar = 1$.

fonction d'onde, $|\psi(\mathbf{x})|^2$, fournit la densité de probabilité de trouver la particule en la position \mathbf{x} .

Pour tout vecteur \mathbf{a} , la translation $T_{\mathbf{a}}$ décrit le déplacement linéaire $T_{\mathbf{a}} : \mathbf{x} \rightarrow \mathbf{x} + \mathbf{a}$. L'ensemble des translations forment un groupe abélien, $T_{\mathbf{a}} T_{\mathbf{b}} = T_{\mathbf{a}+\mathbf{b}}$. L'action de ce groupe sur les fonctions d'onde consiste à déplacer le système physique, donc la fonction d'onde, linéairement le long du vecteur \mathbf{a} . Elle est donnée par $(\hat{T}_{\mathbf{a}} \cdot \psi)(\mathbf{x}) = \psi(\mathbf{x} - \mathbf{a})$. Les opérateurs $\hat{T}_{\mathbf{a}}$ satisfont $\hat{T}_{\mathbf{a}} \hat{T}_{\mathbf{b}} = \hat{T}_{\mathbf{a}+\mathbf{b}}$. Ils forment ainsi une représentation du groupe des translations sur l'espace de Hilbert $L^2(\mathbb{R}^3)$. Cette représentation est unitaire. Les générateurs infinitésimaux des translations sont les impulsions $\mathbf{P} = -i\partial$ qui agissent par dérivation sur les fonctions d'onde, puisque $(\hat{T}_{\mathbf{a}} \cdot \psi)(\mathbf{x}) = \psi(\mathbf{x}) - \mathbf{a} \cdot \partial\psi(\mathbf{x}) + \dots$ de sorte que $\hat{T}_{\mathbf{a}} = \mathbf{1} - i\mathbf{a} \cdot \mathbf{P} + \dots$ pour $|\mathbf{a}| \ll 1$. Ils commutent entre eux⁽²⁾ :

$$[P_j, P_k] = 0.$$

Les translations finies sont données par exponentiation $\hat{T}_{\mathbf{a}} = \exp(-i\mathbf{a} \cdot \mathbf{P})$, comme on peut le vérifier par un développement de Taylor de $\psi(\mathbf{x} - \mathbf{a})$.

De façon similaire, le groupe des rotations agit sur les fonctions d'onde. Une transformation orthogonale est une application linéaire $g : \mathbf{x} \rightarrow g \cdot \mathbf{x}$ préservant la norme du vecteur, $\|g \cdot \mathbf{x}\|^2 = \|\mathbf{x}\|^2$ soit $g^T g = \mathbf{1}$. L'ensemble de ces transformations forment un groupe noté $O(d)$ avec $d = 3$ car nous considérons un espace tri-dimensionnel. Comme $(\det g)^2 = 1$ pour tout $g \in O(3)$, le groupe $O(3)$ contient deux composantes connexes. On note $SO(3)$ la composante connexe à l'identité formée des matrices orthogonales de déterminant 1, c'est l'ensemble des rotations. Une rotation dans l'espace tri-dimensionnel est spécifiée par la donnée de l'axe de rotation, repéré par le vecteur unitaire \mathbf{n} , $\mathbf{n}^2 = 1$, et de l'angle de rotation, noté α . Les rotations de paramètres $(\mathbf{n}; \alpha)$ et $(-\mathbf{n}; -\alpha)$ sont identifiées. La rotation $g_{\mathbf{n};\alpha}$, d'axe \mathbf{n} et d'angle α , agit dans le plan orthogonal à \mathbf{n} selon la matrice $\begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$. Soit $g_{\mathbf{n};\alpha} \cdot \mathbf{x} = (\mathbf{x} \cdot \mathbf{n})\mathbf{n} + (\mathbf{x} - (\mathbf{x} \cdot \mathbf{n})\mathbf{n}) \cos \alpha + \mathbf{n} \wedge \mathbf{x} \sin \alpha$. L'action du groupe des rotations sur les fonctions d'onde est donnée par

$$(R_g \cdot \psi)(\mathbf{x}) = \psi(g^{-1} \cdot \mathbf{x}).$$

On vérifie simplement que cette action définit une représentation de $SO(3)$ sur $L^2(\mathbb{R}^3)$ car $R_{g_1} R_{g_2} = R_{g_1 g_2}$. (Il est important d'agir avec g^{-1} dans la définition de $R_g \cdot \psi$ pour avoir cette propriété). La représentation est unitaire sur $L^2(\mathbb{R}^3) : R_g^\dagger R_g = \mathbf{1}$. Une rotation d'angle d'axe \mathbf{n} et d'angle α infinitésimal agit comme $g_{\mathbf{n};\alpha}^{-1} \cdot \mathbf{x} = \mathbf{x} - \alpha \mathbf{n} \wedge \mathbf{x} + \dots$ pour $\alpha \ll 1$, donc $R_{g_{\mathbf{n};\alpha}} = \mathbf{1} - i\alpha \mathbf{n} \cdot \mathbf{J} + \dots$ où \mathbf{J} sont les opérateurs de moments cinétiques de composantes $J^k = -i\epsilon^{klm} x_l \partial_m$. Ils satisfont les relations :

$$[J^k, J^l] = i\epsilon^{klm} J^m.$$

Les rotations d'angle fini s'obtiennent par exponentiation $R_{g_{\mathbf{n};\alpha}} = \exp(-i\alpha \mathbf{n} \cdot \mathbf{J})$, comme on peut le vérifier par un développement de Taylor de $\psi(g_{\mathbf{n};\alpha}^{-1} \cdot \mathbf{x})$.

L'espace de Hilbert $L^2(\mathbb{R}^3)$ sur lequel agit $SO(3)$ se décompose en une somme directe de représentations irréductibles de ce groupe. Comme les rotations préservent la norme des vecteurs, il est utile d'introduire les coordonnées sphériques (r, θ, φ) et d'extraire la dépendance des fonctions

2. Le commutateur de deux opérateurs A et B est défini par $[A, B] = AB - BA$.

d'onde dans la coordonnée radiale laissée invariante par les rotations. Ceci conduit à l'isomorphisme $L^2(\mathbb{R}^3) \cong L^2(\mathbb{R}_+) \otimes L^2(\mathbb{S}^2)$ où $L^2(\mathbb{S}^2)$ est l'ensemble des fonctions de carré sommable sur la sphère. Toute fonction de $L^2(\mathbb{S}^2)$ se décompose sur les harmoniques sphériques $Y_{l;m}(\theta, \varphi)$, $l \in \mathbb{N}$ et $l \geq m \geq -l$, qui forment ainsi une base de $L^2(\mathbb{S}^2)$. Comme nous le rappelons ci-après, le groupe des rotations agit sur les harmoniques sphériques et, pour l fixé, l'espace vectoriel engendré par les $Y_{l;m}$ supporte une représentation irréductible $R_l^{so(3)}$ de $SO(3)$ de dimension $2l + 1$. Ainsi $L^2(\mathbb{S}^2) \cong \bigotimes_{l=0}^{\infty} R_l^{so(3)}$, soit

$$L^2(\mathbb{R}^3) = \bigotimes_{l=0}^{\infty} L^2(\mathbb{R}_+) \otimes R_l^{so(3)}.$$

Cette égalité signifie que toute fonction $f(\mathbf{x})$ définie sur \mathbb{R}^3 se décompose sur les harmoniques sphériques en : $f(\mathbf{x}) = \sum_{l;m} c_{l;m}(r) Y_{l;m}(\theta, \varphi)$.

L'ensemble des translations et des rotations engendrent le groupe Euclidien, noté $E(d)$ avec ici $d = 3$. Les rotations et les translations ne commutent pas mais satisfont $g \cdot T_{\mathbf{a}} \cdot g^{-1} = T_{g\mathbf{a}}$, qui se traduit par les relations de commutation entre les générateurs d'impulsions et de moments cinétiques :

$$[J^k, P^l] = i\epsilon^{klm} P^m$$

Le groupe Euclidien s'identifie avec le groupe des matrices $M_{g;\mathbf{a}} = \begin{pmatrix} g & \mathbf{a} \\ \mathbf{0} & 1 \end{pmatrix}$ qui correspondent aux produits $T_{\mathbf{a}}g$. La loi de groupe est

$$M_{g;\mathbf{a}} \cdot M_{h;\mathbf{b}} = M_{gh;g\mathbf{b}+\mathbf{a}}.$$

- *Harmoniques sphériques et fonctions homogènes.*

Arrêtons nous un instant pour revoir la construction des harmoniques sphériques $Y_{l;m}(\theta, \varphi)$, $l \in \mathbb{N}$ et $l \geq m \geq -l$. Pour l fixé, elles forment une base de la représentation irréductible $R_l^{so(3)}$ de $SO(3)$ de dimension $2l + 1$. Elles sont fonctions propres du moment cinétique $J^z = i(y\partial_x - x\partial_y)$ et de l'opérateur Casimir $\mathbf{J}^2 = (J^x)^2 + (J^y)^2 + (J^z)^2$:

$$\mathbf{J}^2 \cdot Y_{l;m} = l(l+1) Y_{l;m}, \quad J^z \cdot Y_{l;m} = m Y_{l;m}$$

Ces relations caractérisent les fonctions $Y_{l;m}$ et servent (souvent) de point de départ à leur construction. Mais il est instructif d'identifier les harmoniques sphériques via l'étude des fonctions homogènes dans les coordonnées x, y, z . Considérons donc l'espace vectoriel engendré par les monômes de degré l en les coordonnées x, y, z . Pour $l = 0$, le seul monome est la fonction constant 1, sur laquelle le groupe des rotations agit trivialement. On a $Y_{0;0}(\theta, \varphi) = \sqrt{1/4\pi}$ et $R_{l=0}^{so(3)} = \mathbb{C}$. Il y a trois monomes indépendants de degré un, à savoir les trois coordonnées x, y, z . L'espace vectoriel engendré par ces trois fonctions est clairement irréductible sous les rotations (car les trois coordonnées sont échangées par les rotations), et il est isomorphe à la représentation vectorielle $R_{l=1}^{so(3)}$ de $SO(3)$. On a :

$$r Y_{1;0}(\theta, \varphi) = z\sqrt{3/4\pi} \quad \text{et} \quad r Y_{1;\pm 1}(\theta, \varphi) = (x \pm iy)\sqrt{3/8\pi}.$$

Il y a six monomes indépendants de degré deux : x^2, y^2, z^2 et xy, xz, yz . L'espace vectoriel engendré par ces six fonctions est stable sous l'action des rotations (car les rotations agissent

linéairement sur les coordonnées) mais il n'est pas irréductible car la fonction $r^2 = x^2 + y^2 + z^2$ est invariante sous les rotations. On peut donc décomposer cet espace en la somme de deux espaces qui forment des représentations irréductibles de $SO(3)$: l'espace engendré par la fonction r^2 , qui est isomorphe à $R_{l=0}^{so(3)}$, et l'espace de dimension 5 engendré par les fonctions xy , xz , yz , $2x^2 - y^2 - z^2$ et $2y^2 - z^2 - x^2$, qui est isomorphe à $R_{l=2}^{so(3)}$. Les fonctions $r^2 Y_{2,m}(\theta, \varphi)$ sont des combinaisons linéaires de ces cinq dernières fonctions.

Plus généralement, l'espace vectoriel \mathcal{F}_l des fonctions homogènes de degré l est stable sous les rotations. Un élément de cet espace est une fonction de la forme $G_{j_1, j_2, \dots, j_l} x^{j_1} x^{j_2} \dots x^{j_l}$, où G_{j_1, j_2, \dots, j_l} est un tenseur symétrique d'ordre l , et \mathcal{F}_l est donc de dimension $\frac{(l+1)(l+2)}{2}$. Quitte à les diviser par r^{-l} , nous pouvons voir ces fonctions homogènes comme des fonctions définies sur la sphère de rayon un. L'espace \mathcal{F}_l n'est pas irréductible sous le groupe des rotations car \mathcal{F}_{l-2} se plonge dans \mathcal{F}_l : les fonctions de la forme $r^2 \tilde{G}_{j_1, j_2, \dots, j_{l-2}} x^{j_1} x^{j_2} \dots x^{j_{l-2}}$ constituent une sous-représentation de \mathcal{F}_l . Le complémentaire (orthogonal par rapport au produit scalaire L^2 sur la sphère) de cette sous-représentation est un espace de dimension $\frac{(l+1)(l+2)}{2} - \frac{(l-1)l}{2} = 2l + 1$ formé des fonctions $G_{j_1, j_2, \dots, j_l} x^{j_1} x^{j_2} \dots x^{j_l}$ où le tenseur G est de trace nulle : $\delta^{kn} G_{k, n, j_3, \dots, j_l} = 0$. On peut vérifier qu'il est irréductible sous les rotations et donc isomorphe à $R_l^{so(3)}$. Ainsi les fonctions $r^l Y_{l,m}(\theta, \varphi)$ sont des fonctions homogènes

$$G_{j_1, j_2, \dots, j_l} x^{j_1} x^{j_2} \dots x^{j_l}$$

avec G_{j_1, j_2, \dots, j_l} symétrique et de trace nulle.

V.1.2. Représentations, représentations projectives. — Comment décrire les transformations des états d'un système quantique? E. Wigner a donné une réponse précise à cette question importante. Les transformations des états d'un système quantique sont décrites par *le théorème de Wigner* :

Si $S : |\psi\rangle \rightarrow |S \cdot \psi\rangle$ est une transformation bijective entre les rayons d'un espace de Hilbert \mathcal{H} préservant les modules des produits scalaires, i.e. telle que $|\langle S \cdot \phi | S \cdot \psi \rangle|^2 = |\langle \phi | \psi \rangle|^2$, alors S est une transformation unitaire linéaire, ou anti-linéaire, sur \mathcal{H} :

$$|S \cdot \psi\rangle = U_S |\psi\rangle, \quad U_S^\dagger U_S = \mathbf{1}$$

à une phase près (que l'on peut choisir égale à zéro).

On rappelle qu'une transformation anti-linéaire \hat{U} est telle que $\hat{U}(x|\psi\rangle + y|\phi\rangle) = x^* \hat{U}|\psi\rangle + y^* \hat{U}|\phi\rangle$ où x^* , y^* sont les complexes conjugués de x , y . Son hermitique conjugué est défini par $\langle \hat{U}^\dagger \phi | \psi \rangle = \langle \phi | \hat{U} \psi \rangle^*$. Les transformations anti-linéaires interviennent lorsque le renversement du temps est mis en jeu. Nous ne traiterons pas ces cas dans ce qui suit.

Ainsi, un groupe de symétrie G agit linéairement sur les états d'un système quantique, appartenant à un espace de Hilbert \mathcal{H} , via des opérateurs unitaires :

$$|\psi\rangle \rightarrow |g \cdot \psi\rangle \equiv U(g)|\psi\rangle, \quad g \in G, \quad |\psi\rangle \in \mathcal{H},$$

avec $U(g) \in \text{End} \mathcal{H}$ unitaire, $U(g)^\dagger U(g) = \mathbf{1}$. La condition d'unitarité assure la préservation de la norme $\langle g \cdot \psi | g \cdot \psi \rangle = \langle \psi | \psi \rangle$.

Un groupe est muni d'une loi produit de sorte que le produit de deux éléments du groupe est un élément du groupe. Si un groupe agit sur un système physique, la loi produit doit être implémentée de sorte que les actions successives par deux éléments du groupe sur le système soient équivalentes à l'action de l'élément produit. Le fait que l'état d'un système quantique soit spécifié par un rayon, (i.e. par un vecteur défini à une phase près), laisse une certaine liberté dans le résultat possible des actions successives de deux éléments du groupe sur l'espace de Hilbert.

(i) Soit le produit de deux actions sur un vecteur de \mathcal{H} reproduit identiquement l'action produit sur cet état : $|g_1 \cdot (g_2 \cdot \psi)\rangle = |(g_1 g_2) \cdot \psi\rangle$. Il s'ensuit alors que l'application $g \in G \rightarrow U(g) \in \text{End}\mathcal{H}$ définit une représentation de G sur \mathcal{H} :

$$U(g_1)U(g_2) = U(g_1 g_2).$$

Cette représentation est unitaire puisque nous avons imposé que $U(g)^\dagger U(g) = \mathbf{1}$. Les exemples précédents concernant les translations et les rotations tombent dans cette classe.

Pour un groupe de Lie G , on peut considérer les éléments g au voisinage de l'identité qui correspondent à des transformations infinitésimales. De tels éléments s'écrivent sous la forme $g \simeq 1 - i \sum_a x_a t^a + \dots$ où l'ensemble des t^a forme une base de l'algèbre de Lie \mathfrak{g} de G . Alors

$$U(g) = \mathbf{1} - i \sum_a x_a T^a + \dots$$

où les $T^a = U(t^a)$ définissent une représentation unitaire de l'algèbre de Lie \mathfrak{g} sur \mathcal{H} . La condition d'unitarité se traduit par : $T^a{}^\dagger = T^a$. La représentation des éléments du groupe appartenant à la composante connexe de l'identité s'obtient alors par exponentiation ⁽³⁾ : $U(g) = \exp(-i \sum_a x_a T^a)$ pour $g = e^{-it}$ avec $t = \sum_a x_a t^a \in \mathfrak{g}$.

(ii) Soit le produit de deux actions sur un vecteur de \mathcal{H} reproduit l'action produit sur cet état à une phase près : $|g_1 \cdot (g_2 \cdot \psi)\rangle = e^{i\omega(g_1, g_2)} |(g_1 g_2) \cdot \psi\rangle$ pour tout vecteur $|\psi\rangle$ de sorte que :

$$U(g_1)U(g_2) = \Omega(g_1, g_2) U(g_1 g_2), \quad |\Omega(g_1, g_2)|^2 = 1$$

où l'on a posé $\Omega(g_1, g_2) = e^{i\omega(g_1, g_2)}$. On dit alors que l'application $g \in G \rightarrow U(g) \in \text{End}\mathcal{H}$ est une représentation projective de G sur \mathcal{H} . Les phases $\omega(g_1, g_2)$ ne sont pas arbitraires car demander que le produit des opérateurs $U(g)$ soit associatif impose une condition sur $\Omega(g_1, g_2)$, appelée relation de cocycle :

$$\Omega(g_1, g_2) \Omega(g_1 g_2, g_3) = \Omega(g_1, g_2 g_3) \Omega(g_2, g_3)$$

Une solution simple à cette condition est $\Omega(g_1, g_2) = e^{i(\eta(g_1) + \eta(g_2) - \eta(g_1 g_2))}$. Une telle solution est appelée un cocycle trivial. Dans ce cas, on peut redéfinir les opérateurs $U(g)$ en $\hat{U}(g) \equiv e^{-i\eta(g)} U(g)$ de sorte que les nouveaux opérateurs $\hat{U}(g)$ forment une représentation de G . On identifie ainsi deux représentations projectives dont les phases diffèrent par un cocycle trivial.

Une représentation projective d'un groupe G est une représentation d'un groupe \hat{G} , plus grand que G , appelé une extension (centrale) de G . Les phases $\Omega(g_1, g_2)$ prennent valeurs dans l'ensemble $U(1)$ des nombres complexes de module un, mais il se peut qu'elles n'atteignent qu'un nombre restreint de valeurs possibles (par exemple ± 1). Nous supposons donc que ces phases prennent valeurs dans un sous groupe (abelien) de $U(1)$ que nous notons \mathfrak{Z} (par exemple \mathbb{Z}_2).

3. On fera attention à la différence de convention utilisée par les mathématiciens et les physiciens : les physiciens rajoutent un facteur "i" de sorte que les générateurs infinitésimaux sont hermitiens.

L'extension de \hat{G} de G est alors $G \rtimes \mathfrak{Z}$, les éléments de \hat{G} sont les paires (g, λ) avec $g \in G$ et $\lambda \in \mathfrak{Z}$. La loi produit sur \hat{G} est :

$$(g, \lambda) \cdot (h, \mu) = (gh, \Omega(g, h) \lambda \mu)$$

Cette loi est associative grâce à la relation de cocycle satisfaite par $\Omega(g, h)$. Clairement G est un sous-groupe de \hat{G} . L'application $(g, \lambda) \in \hat{G} \rightarrow \lambda U(g) \in \text{End}\mathcal{H}$ est (par construction) une représentation (unitaire) de l'extension \hat{G} .

• *Exemple : Représentations projectives de $SO(2)$.*

Le groupe $SO(2)$, le groupe des rotations en dimension deux, est isomorphe au groupe $U(1)$, le groupe des nombres complexes de module un. Une rotation dont la matrice est $\begin{pmatrix} \cos 2\pi x & \sin 2\pi x \\ -\sin 2\pi x & \cos 2\pi x \end{pmatrix}$ correspond au nombre $e^{i2\pi x}$ avec x réel défini modulo 1. Ces groupes étant abéliens, toutes leurs représentations sont de dimension 1. Posons $D_j(e^{i2\pi x}) \equiv \exp(i2\pi jx)$. Cette relation définit une application D_j sur $U(1)$ si (et seulement si) le 'spin' j est entier. En effet, traduire le paramètre x d'un entier ne modifie pas l'élément $e^{i2\pi x}$ de $U(1)$ mais modifie son image si j n'est pas un entier (et ne la modifie pas si j est entier). On a évidemment $D_j(e^{i2\pi x}) \cdot D_j(e^{i2\pi y}) = D_j(e^{i2\pi(x+y)})$, de sorte que D_j définit par multiplication une représentation de $U(1)$ sur \mathbb{C} si j est entier. Que se passe-t-il si j n'est pas entier ? D_j n'est alors pas une application sur $U(1)$. Choisissons un représentant de $e^{i2\pi x}$ et posons $\hat{D}_j(e^{i2\pi x}) \equiv \exp(i2\pi j\bar{x})$ où \bar{x} est la partie entière de x , soit $\bar{x} = x - E(x)$. Alors \hat{D}_j est une application de $U(1)$ sur $\text{End}\mathbb{C} \simeq \mathbb{C}$, bien que discontinue en 1. On a :

$$\hat{D}_j(e^{i2\pi x}) \cdot \hat{D}_j(e^{i2\pi y}) = e^{i\omega_j(x,y)} \hat{D}_j(e^{i2\pi(x+y)})$$

avec $\omega_j(x, y) = 2j\pi[E(x+y) - E(x) - E(y)]$. La phase $\omega_j(x, y)$ est bien définie sur $U(1) \times U(1)$, i.e. elle est indépendante du choix du représentant de x ou de y , et elle satisfait la relation de cocycle. Ce n'est pas un cocycle trivial (sauf pour j entier) car $E(x)$ n'est pas définie comme application sur $U(1)$. Donc pour un 'spin' j non-entier, \hat{D}_j définit une représentation projective sur \mathbb{C} .

Pour $j = p/q \in \mathbb{Q}$ avec $p, q \in \mathbb{N}$ premiers entre eux, le cocycle $\Omega_j(x, y) = e^{\omega_j(x,y)}$ est un élément de \mathbb{Z}_q car c'est une racine $q^{\text{ième}}$ de l'unité. L'extension centrale est alors $U(1) \rtimes \mathbb{Z}_q$. L'extension universelle contenant toutes ces extensions (avec q quelconque) est $U(1) \rtimes \mathbb{Z} \cong \mathbb{R}$. Le groupe $U(1)$ est topologiquement un cercle (où la coordonnée $2\pi x$ mesure l'angle le long du cercle). Le groupe $\mathbb{R} \cong U(1) \rtimes \mathbb{Z}$ est le revêtement universel de $U(1)$ (où la coordonnée entière dans $U(1) \rtimes \mathbb{Z}$ mesure le nombre de tours effectués avant d'atteindre un point du cercle). On vérifie bien sûr que les applications D_j définissent des représentations de \mathbb{R} pour j réel quelconque.

• *Exemple : Représentation projective de $SO(3)$ de spin 1/2.*

Comme nous l'avons décrit ci-dessus, en dimension trois les rotations sont paramétrisées par un vecteur unitaire \mathbf{n} , l'axe de rotation, et un nombre α modulo 2π , l'angle de rotation. Les rotations de paramètres (\mathbf{n}, α) et $(-\mathbf{n}, -\alpha)$ sont identiques. Les éléments de $SO(3)$ sont donc déterminés par la donnée de (\mathbf{n}, α) avec $\mathbf{n}^2 = 1$ et $\alpha \in [0, \pi]$, les points (\mathbf{n}, π) et $(-\mathbf{n}, \pi)$ étant identiques. Ainsi, le groupe $SO(3)$ est topologiquement similaire à une boule de rayon π dont les points diamétralement opposés sont identifiés. Ce groupe est non-simplement connexe, car on peut y construire un chemin fermé non-contractible.

Le groupe $SU(2)$ est formé de l'ensemble des matrices 2×2 unitaires et de déterminant un : $UU^\dagger = 1$, $\det U = 1$. Ce groupe est simplement connexe. Nous allons montrer qu'il forme un revêtement de $SO(3)$. A tout vecteur $\mathbf{x} = (x, y, z)$, nous associons une matrice hermitienne 2×2 de trace nulle via :

$$\mathbf{X} \equiv \mathbf{x} \cdot \vec{\sigma} = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix},$$

où $\sigma_x, \sigma_y, \sigma_z$ sont les matrices de Pauli. Notez que $\det \mathbf{X} = -\mathbf{x}^2$. Le groupe $SU(2)$ agit sur l'ensemble de ces matrices par conjugaison :

$$\mathbf{X} \rightarrow U \mathbf{X} U^\dagger, \quad U \in SU(2).$$

Cette action preserve la condition d'hermiticité, la trace et le déterminant : $\det U \mathbf{X} U^\dagger = \det \mathbf{X}$. Elle induit donc une rotation sur le vecteur \mathbf{x} :

$$U \mathbf{X} U^\dagger = (R_U \cdot \mathbf{x}) \cdot \vec{\sigma}.$$

L'application $U \in SU(2) \rightarrow R_U \in SO(3)$ est un morphisme. Son noyau est $\{-1, +1\}$ car R_U est invariant par changement de signe de U . Donc

$$SU(2)/SO(3) \cong \mathbb{Z}_2.$$

La représentation de définition de $SU(2)$, formée des matrices 2×2 unitaires de déterminant un, est sa représentation dite de spin 1/2. Nous laissons au lecteur le plaisir de vérifier que cette représentation définit une représentation projective de $SO(3)$.

V.2. Symétries, dégénérescences et règles de sélection

V.2.1. Symétries et leurs conséquences. — Un groupe G est un groupe de symétrie de la dynamique s'il commute avec celle-ci. En mécanique quantique, la dynamique est régie par l'opérateur hamiltonien H , (qui est un opérateur hermitien $H^\dagger = H$ sur \mathcal{H}), de sorte que l'évolution de l'état d'un système est décrite par l'équation de Schroedinger $i\partial_t|\psi(t)\rangle = H|\psi(t)\rangle$ (nous avons posé $\hbar = 1$). On peut formellement intégrer cette évolution en $|\psi(t)\rangle = e^{-itH}|\psi(0)\rangle$ où $U_t \equiv e^{-itH}$ est l'opérateur d'évolution (par souci de simplicité, nous supposons ici que l'hamiltonien est indépendant du temps). Si le groupe agit sur l'espace des états via les opérateurs $U(g)$, ce groupe sera une symétrie de la dynamique si ces opérateurs commutent avec l'hamiltonien :

$$[H, U(g)] = 0, \quad \forall g \in G$$

Cette relation implique que l'action du groupe G sur \mathcal{H} commute avec l'évolution temporelle. En effet, si $U(g)$ commute avec H , il commute aussi avec e^{-itH} et donc, $|g \cdot \psi(t)\rangle = |(g \cdot \psi)(t)\rangle$. On représente cette relation sous la forme d'un diagramme commutatif :

$$\begin{array}{ccc} |\psi_0\rangle & \xrightarrow{e^{-itH}} & |\psi(t)\rangle \\ \downarrow U(g) & \circlearrowleft & \downarrow \\ |(g \cdot \psi_0)\rangle & \rightarrow & |(g \cdot \psi)(t)\rangle = |g \cdot \psi(t)\rangle \end{array}$$

Ainsi, que G soit un groupe de symétrie signifie qu'il est équivalent d'implémenter l'action de ce groupe sur le système puis de le faire évoluer ou de laisser d'abord le système évoluer puis de lui appliquer l'action du groupe G .

Une conséquence directe d'une symétrie (e.g. via le théorème d'Ehrenfest) est l'existence de lois de conservation. En effet, les valeurs moyennes $\langle \psi(t) | U(g) | \psi(t) \rangle$ des opérateurs représentant le groupe G sont indépendantes du temps (car e^{-itH} commutent avec $U(g)$). Comme en mécanique classique, nous avons donc la correspondance,

$$\text{”symétrie”} \longleftrightarrow \text{”loi de conservation”},$$

qui peut se lire dans les deux sens : l'existence de symétries implique l'existence de lois de conservation (ou de dégénérescences, comme nous allons le voir bientôt) et réciproquement, l'existence de lois de conservation, ou la présence de dégénérescences, est le signe de l'existence de symétries.

Si le groupe de symétrie est un groupe de Lie, la relation de commutation avec l'hamiltonien H est vraie pour tout élément du groupe, en particulier pour des éléments proches de l'identité. Elle est donc aussi vraie pour les générateurs infinitésimaux, éléments de l'algèbre de Lie \mathfrak{g} du groupe G . Ainsi

$$[H, T] = 0, \quad \text{avec } T = U(t), \quad \forall t \in \mathfrak{g}$$

Les générateurs infinitésimaux T , qui sont hermitiens car la représentation est supposée être unitaire, sont donc des bons nombres quantiques car ils commutent avec l'hamiltonien. On peut en particulier diagonaliser simultanément un sous-ensemble commutant de générateurs et l'hamiltonien H . Un état propre de ce sous-ensemble de générateurs restera état propre lors de son évolution temporelle et les valeurs propres correspondantes, appelées nombres quantiques, seront conservées. La symétrie par rotation et la conservation du moment cinétique qui y est associée constitue un des exemples les plus simples de cette correspondance.

V.2.2. Secteurs et règles de sélection. — L'existence d'un groupe de symétrie a des conséquences plus fines que les simples lois de conservation décrites ci-dessus et conduit à une structuration particulière de l'espace de Hilbert. Un outil important de leur analyse est *le lemme de Schur* :

Soient D_1 et D_2 deux représentations irréductibles d'un groupe G , fini ou compact, sur les espaces de Hilbert \mathcal{H}_1 et \mathcal{H}_2 respectivement. Soit S un entrelaceur de \mathcal{H}_1 vers \mathcal{H}_2 , i.e. $SD_1(g) = D_2(g)S$, $\forall g \in G$, alors :

- (i) soit $S = 0$,
- (ii) soit S est une bijection (et les représentations sont équivalentes).

Une conséquence important de ce lemme est :

Si D est une représentation irréductible et unitaire sur un espace de Hilbert \mathcal{H} , et si un endomorphisme S de \mathcal{H} commute avec $D(g)$ pour tout $g \in G$, alors S est un multiple de l'identité : $S = \text{const.} \mathbf{1}$.

Supposons qu'un groupe G , fini ou compact, agit sur l'espace de Hilbert \mathcal{H} d'un système quantique via la représentation unitaire $U(g)$. Cette représentation se décompose en somme directe de représentations irréductibles. Comme plusieurs copies équivalentes de la même représentation

du groupe peuvent apparaître dans cette somme, on écrit :

$$\mathcal{H} = \bigoplus_{\lambda \text{ rep.}} \left(\bigoplus_{r=1}^{m_\lambda} \mathcal{R}_\lambda^{(r)} \right)$$

Dans cette somme, l'indice λ fait référence aux représentations inéquivalentes de G et m_λ est la multiplicité à laquelle la représentation d'indice λ apparaît dans la décomposition de \mathcal{H} . Pour λ fixé, les espaces $\mathcal{R}_\lambda^{(r)}$ sont isomorphes entre eux et à une représentation irréductible de G , notée \mathcal{R}_λ . Dans la base associée à cette décomposition, les matrices représentant les opérateurs $U(g)$ sont bloc-diagonales et s'écrivent par exemple comme :

$$U(g) = \begin{pmatrix} \mathcal{R}_{\lambda_1}^{(1)} & \mathcal{R}_{\lambda_1}^{(2)} & \mathcal{R}_{\lambda_2}^{(1)} & \cdots \\ \tilde{U}_{\lambda_1}(g) & 0 & 0 & \cdots \\ 0 & \tilde{U}_{\lambda_1}(g) & 0 & \cdots \\ 0 & 0 & U_{\lambda_2}(g) & \cdots \end{pmatrix}, \quad \forall g \in G$$

Ici la multiplicité de la représentation indexée par λ_1 a été choisie égale à 2, les matrices $U_{\lambda_1}(g)$ et $\tilde{U}_{\lambda_1}(g)$ sont équivalentes et peuvent être choisies égales par un choix de base appropriées sur $\mathcal{R}_{\lambda_1}^{(1)}$ et $\mathcal{R}_{\lambda_1}^{(2)}$.

On peut écrire cette décomposition différemment en introduisant pour chaque λ un espace de Hilbert auxiliaire W_λ de dimension égale à la multiplicité m_λ tel que si les vecteurs e_λ^r , $r = 1, \dots, m_\lambda$ forment une base de W_λ alors $e_\lambda^r \otimes \mathcal{R}_\lambda \cong \mathcal{R}_\lambda^{(r)}$. Ainsi $W_\lambda \otimes \mathcal{R}_\lambda \cong \bigoplus_{r=1}^{m_\lambda} \mathcal{R}_\lambda^{(r)}$ et

$$\mathcal{H} = \bigoplus_{\lambda \text{ rep.}} W_\lambda \otimes \mathcal{R}_\lambda.$$

Chaque terme $W_\lambda \otimes \mathcal{R}_\lambda$ de cette décomposition constitue par définition un secteur de \mathcal{H} . Suivant cette présentation, les opérateurs $U(g)$ agissent trivialement sur les facteurs W_λ et leur action sur le secteur $W_\lambda \otimes \mathcal{R}_\lambda$ est donnée $\mathbf{1} \otimes U_\lambda(g)$ où les $U_\lambda(g)$ sont les endomorphismes définissant la représentation irréductible de G sur \mathcal{R}_λ .

Il est intéressant de comparer cette décomposition avec celle de l'espace $L^2(\mathbb{R}^3)$ en représentations irréductibles de $SO(3)$: $L^2(\mathbb{R}^3) = \bigotimes_{l=0}^{\infty} L^2(\mathbb{R}_+) \otimes R_l^{so(3)}$. Dans les coordonnées sphériques, le groupe des rotations laisse invariant la coordonnée radiale et n'agit que sur les coordonnées angulaires, autrement dit, le groupe des rotations n'agit pas sur la composante radiale des fonctions d'onde qui est décrite par le facteur $L^2(\mathbb{R}_+)$ mais agit sur leurs composantes angulaires qui correspondent aux facteurs $R_l^{so(3)}$ de cette décomposition.

L'hamiltonien H est un entrelaceur pour l'action du groupe G puisque $U(g)H = HU(g)$. D'après le lemme de Schur, il n'a d'éléments de matrice qu'entre vecteurs appartenant à des représentations équivalentes. L'hamiltonien est donc bloc-diagonal sur les secteurs, soit par exemple :

$$H = \begin{pmatrix} \mathcal{R}_{\lambda_1}^{(1)} & \mathcal{R}_{\lambda_1}^{(2)} & \mathcal{R}_{\lambda_2}^{(1)} & \cdots \\ H_{1;1} & H_{1;\bar{1}} & 0 & \cdots \\ H_{\bar{1};1} & H_{\bar{1};\bar{1}} & 0 & \cdots \\ 0 & 0 & H_{2;2} & \cdots \end{pmatrix}$$

Comme nous allons le montrer, la description de l'hamiltonien est plus explicite dans la présentation de l'espace de Hilbert sous la forme $\mathcal{H} = \bigoplus_{\lambda \text{ rep.}} W_\lambda \otimes \mathcal{R}_\lambda$, puisque H ne mélange pas les

différents secteurs $W_\lambda \otimes \mathcal{R}_\lambda$ et qu'il agit trivialement sur les composantes \mathcal{R}_λ . Autrement dit, $W_\lambda \otimes \mathcal{R}_\lambda$ est stable par H et la restriction de H à un secteur est :

$$H|_{W_\lambda \otimes \mathcal{R}_\lambda} = h_\lambda \otimes \mathbf{1},$$

où h_λ est (souvent) appelé l'hamiltonien réduit, ou effectif, du secteur $W_\lambda \otimes \mathcal{R}_\lambda$. En particulier, la diagonalisation de l'hamiltonien est réduite à celle des hamiltoniens effectifs h_λ , un problème généralement (beaucoup) plus simple car ces derniers opérateurs agissent sur des espaces de dimensions (beaucoup) plus petites que celle de \mathcal{H} . Si ε_λ est une valeur propre de l'hamiltonien réduit h_λ , c'est aussi une valeur propre de l'hamiltonien H et elle apparaît dans le spectre de H avec une multiplicité (au moins) égale à la dimension de la représentation \mathcal{R}_λ du groupe de symétrie.

En effet, soit e_λ^r , $r = 1, \dots, m_\lambda$, une base de W_λ et ψ_λ^j , $j = 1, \dots, \dim \mathcal{R}_\lambda$, une base de \mathcal{R}_λ . Comme le secteur $W_\lambda \otimes \mathcal{R}_\lambda$ est stable sous l'action de H et de celle de $U(g)$, on peut écrire

$$\begin{aligned} H \cdot (e_\lambda^r \otimes \psi_\lambda^j) &= \sum_s e_\lambda^s \otimes (H_\lambda^{sr} \cdot \psi_\lambda^j), \\ U(g) \cdot (e_\lambda^r \otimes \psi_\lambda^j) &= e_\lambda^r \otimes (U_\lambda(g) \cdot \psi_\lambda^j), \end{aligned}$$

où H_λ^{sr} sont des opérateurs agissant sur \mathcal{R}_λ . Le fait que H commute avec $U(g)$ se traduit alors en $H_\lambda^{sr} U_\lambda(g) = U_\lambda(g) H_\lambda^{sr}$. Comme \mathcal{R}_λ est une représentation irréductible de G , il s'en suit d'après le lemme de Schur que les opérateurs H_λ^{sr} sont des multiples de l'identité sur \mathcal{R}_λ , soit : $H_\lambda^{sr} = h_\lambda^{sr} \mathbf{1}_{\mathcal{R}_\lambda}$ avec h_λ^{sr} certaines constantes. Donc

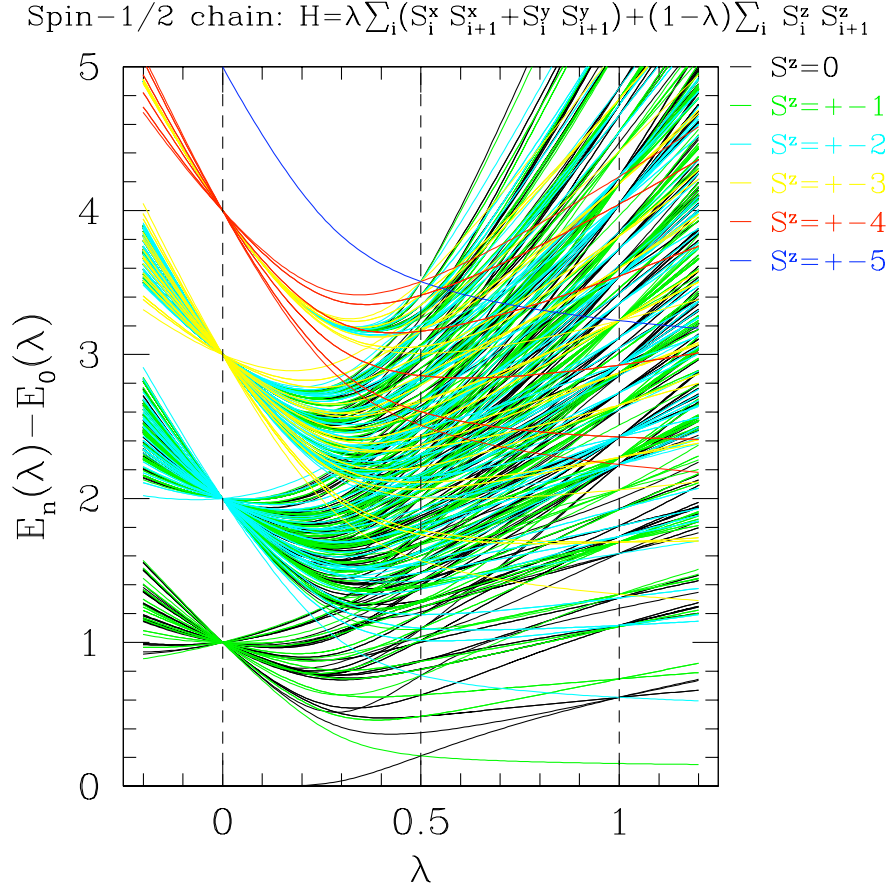
$$H \cdot (e_\lambda^r \otimes \psi_\lambda^j) = \sum_s (e_\lambda^s h_\lambda^{sr}) \otimes \psi_\lambda^j = (h_\lambda \cdot e_\lambda^r) \otimes \psi_\lambda^j$$

avec h_λ^{sr} les éléments de matrice de h_λ dans la base des e_λ^r . On a donc montré que H agit comme $h_\lambda \otimes \mathbf{1}$ sur $W_\lambda \otimes \mathcal{R}_\lambda$.

Le cas d'une particule (non relativiste) de masse m soumise à un potentiel $V(r)$ à symétrie sphérique est un exemple simple d'application de cette remarque. L'hamiltonien est $H = -\frac{1}{2m} \nabla^2 + U(r)$. Il commute avec les opérateurs de moments cinétiques \mathbf{J} qui sont les générateurs du groupe des rotations. Il agit diagonalement sur chacun des secteurs $L^2(\mathbb{R}_+) \otimes R_l^{so(3)}$ de moment cinétique défini. En coordonnées sphériques le laplacien s'écrit $\nabla^2 = \frac{1}{r} \frac{\partial^2}{\partial r^2} r - \frac{1}{r^2} \mathbf{J}^2$, et l'hamiltonien réduit est donc $h_l = -\frac{1}{2m} \frac{1}{r} \frac{\partial^2}{\partial r^2} r + \frac{l(l+1)}{2mr^2} + V(r)$. Par la grâce des symétries, le problème tri-dimensionnel a donc été réduit à un problème unidimensionnel. Le potentiel effectif dans un secteur de moment cinétique l est $V_{\text{eff}}^{(l)}(r) = \frac{l(l+1)}{2mr^2} + V(r)$ où le premier terme en $l(l+1)/r^2$ reflète l'effet de la barrière centrifuge.

Avant de conclure par un exemple issu de la physique de la matière condensée, résumons quelques unes des conséquences de l'existence de symétries :

- (i) L'espace de Hilbert des états du système se décompose en secteurs qui sont indexés par les représentations du groupe de symétrie. Une représentation irréductible de ce groupe peut apparaître dans cette décomposition avec une multiplicité.
- (ii) L'hamiltonien H , et donc l'opérateur d'évolution e^{-itH} , est bloc-diagonal dans la base des secteurs. Il ne peut donc y avoir de transition entre états de secteurs différents. Ce sont les règles de sélection.

FIGURE 1. Spectre en énergie d'une chaîne d'Heisenberg ($N=10$).

(iii) La diagonalisation de l'hamiltonien se réduit à la diagonalisation des hamiltoniens effectifs sur chacun des secteurs. Une valeur propre de l'hamiltonien a une dégénérescence (au moins) égale à la dimension de la représentation du groupe de symétrie du secteur correspondant.

(iv) Réciproquement, une forte dégénérescence dans le spectre des valeurs propres de l'hamiltonien est (souvent) le signe de l'existence d'un groupe de symétrie.

• *Exemple : Analyse du spectre d'une chaîne d'Heisenberg.*

L'hamiltonien d'Heisenberg est celui d'un système modèle adapté à la description des propriétés magnétiques des solides cristallins. Il décrit les interactions entre les moments magnétiques portés par les ions du solide. Ces moments magnétiques sont représentés par des opérateurs S^x , S^y , S^z . En se limitant au cas où seules les interactions entre plus proches voisins sont pertinentes, l'hamiltonien s'écrit

$$H = \lambda \sum_j (S_j^x S_{j+1}^x + S_j^y S_{j+1}^y) + (1 - \lambda) \sum_j S_j^z S_{j+1}^z.$$

dans le cas d'un système uni-dimensionnel. Les opérateurs de moment magnétique forment une représentation de $so(3)$: $[S^x, S^y] = iS^z$ et permutation circulaire. Nous considérons ici le cas le plus simple où cette représentation est la représentation de spin 1/2, donc : $\vec{S} = \vec{\sigma}/2$ où $\vec{\sigma}$

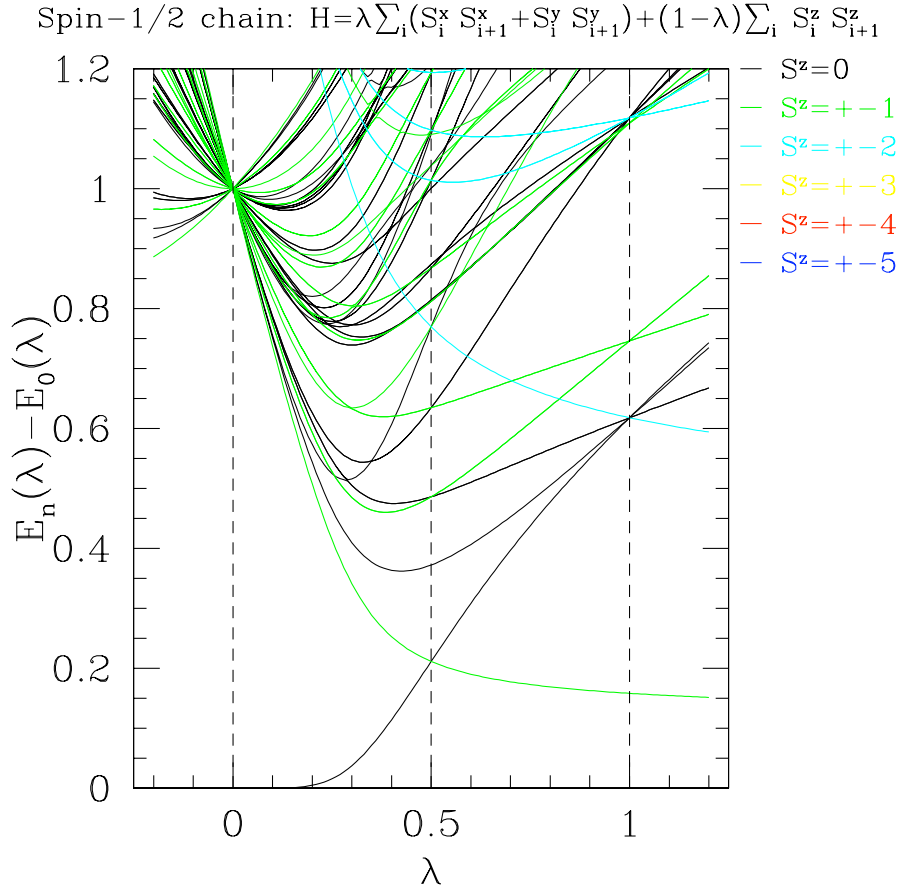


FIGURE 2. Spectre en énergie d'une chaîne d'Heisenberg ($N=10$) : détails.

sont les matrices de Pauli. Les opérateurs \vec{S}_j agissent sur l'espace de Hilbert attaché au moment magnétique fixé à la position j sur le solide. Les opérateurs \vec{S}_j et \vec{S}_k en des positions différentes $j \neq k$ commutent.

L'hamiltonien d'Heisenberg dépend d'un paramètre noté λ . Il est intéressant de comprendre la nature du spectre de cet hamiltonien, sa dépendance en λ , la présence de croisements de niveaux et d'interpréter, en relation avec la théorie des groupes, ses particularités en $\lambda = 0, 1/2$ ou 1 . Le spectre d'énergie pour une chaîne de $N = 10$ sites (avec des conditions aux limites périodiques) est représenté dans les Figures (1,2). Les énergies sont classées suivant la valeur du moment magnétique total $S^z = \sum_j S_j^z$ qui est un bon nombre quantique.

V.3. Observables

Une observable A est un opérateur linéaire hermitique sur \mathcal{H} . Si un groupe de symétrie G agit sur \mathcal{H} par les endomorphismes unitaires $U(g)$, alors il agit sur $\text{End}\mathcal{H}$ par :

$$A \in \text{End}\mathcal{H} \rightarrow g \cdot A \equiv U(g) A U(g)^\dagger, \quad \forall g \in G$$

car $\text{End}\mathcal{H} \simeq \mathcal{H} \otimes \mathcal{H}^*$. Cette relation définit une représentation de G sur $\text{End}\mathcal{H}$ telle que $(g \cdot A)|\psi\rangle = |g \cdot (A\psi)\rangle$.

On dit qu'une famille d'observables A_μ , indexée par un ensemble dénombrable, se transforme de façon covariante sous cette action de groupe si

$$U(g) A_\mu U^\dagger(g) = \sum_{\nu} A_\nu D(g)_{\nu\mu}, \quad \forall g \in G$$

où les matrices $D(g)_{\nu\mu}$ forment une représentation \mathcal{D} de G . En d'autres termes, la donnée d'une famille d'observables covariantes est équivalente à la donnée d'une application linéaire $\hat{A} : \mathbf{v} \in \mathcal{D} \rightarrow A_{\mathbf{v}} \in \text{End}\mathcal{H}$ qui entrelace les actions du groupe G sur \mathcal{D} et sur $\text{End}\mathcal{H}$, i.e.

$$g \cdot A_{\mathbf{v}} \equiv U(g) A_{\mathbf{v}} U(g)^\dagger = A_{g \cdot \mathbf{v}}.$$

En effet, si les vecteurs \mathbf{v}_μ forment une base de la représentation \mathcal{D} de G telle que $g \cdot \mathbf{v}_\mu = \sum_{\nu} \mathbf{v}_\nu D(g)_{\nu\mu}$, alors les observables $A_\mu \equiv A_{\mathbf{v}_\mu}$ se transforment de façon covariante : $g \cdot A_\mu = A_{D(g) \cdot \mathbf{v}_\mu} = \sum_{\nu} A_\nu D(g)_{\nu\mu}$ par linéarité.

Si G est un groupe de Lie, \mathfrak{g} son algèbre de Lie, la loi de transformation covariante pour une famille d'observables se traduit par les relations de commutation

$$[T, A_{\mathbf{v}}] = A_{T \cdot \mathbf{v}}, \quad \text{soit} \quad [T, A_\mu] = \sum_{\nu} A_\nu T_{\nu\mu},$$

pour tout $t \in \mathfrak{g}$ et où $T_{\nu\mu} = D(t)_{\nu\mu}$ sont les éléments de matrice des générateurs représentant l'algèbre de Lie dans la base \mathbf{v}_μ de \mathcal{D} .

Un opérateur appartenant à une famille d'observables covariantes satisfait certaines règles de sélection. Si l'espace de Hilbert se décompose en secteurs $\mathcal{H} = \bigoplus_{\lambda \text{ rep.}} W_\lambda \otimes \mathcal{R}_\lambda$, alors les opérateurs covariants $A_{\mathbf{v}}$ avec $\mathbf{v} \in \mathcal{D}$ n'ont d'éléments de matrices qu'entre les états de secteurs $W_\lambda \otimes \mathcal{R}_\lambda \rightarrow W_{\lambda'} \otimes \mathcal{R}_{\lambda'}$ tel que $\mathcal{R}_{\lambda'}$ apparaisse dans la décomposition du produit tensoriel $\mathcal{D} \otimes \mathcal{R}_\lambda$ en représentations irréductibles de G .

En effet, soit e_λ un vecteur quelconque de W_λ et ψ_λ^j , $j = 1, \dots, \dim \mathcal{R}_\lambda$, une base de \mathcal{R}_λ . Calculons l'action de G sur l'état $A_\mu(e_\lambda \otimes \psi_\lambda^j)$:

$$\begin{aligned} U(g)[A_\mu(e_\lambda \otimes \psi_\lambda^j)] &= \sum_{\nu} A_\nu(e_\lambda \otimes U_\lambda(g)\psi_\lambda^j) D(g)_{\nu\mu} \\ &= \sum_{\nu j} A_\nu(e_\lambda \otimes \psi_\lambda^k) D(g)_{\nu\mu} U(g)^{kj} \end{aligned}$$

Les états $A_\mu(e_\lambda \otimes \psi_\lambda^j)$ se transforment donc comme des éléments du produit tensoriel $\mathcal{D} \otimes \mathcal{R}_{\lambda_1}$, ce qui implique la précédente affirmation.

Par exemple, les opérateurs positions \mathbf{x} , les opérateurs impulsions \mathbf{p} ou les opérateurs moments cinétiques $\mathbf{J} = \mathbf{p} \wedge \mathbf{x}$ sont des opérateurs vectoriels sous le groupe des rotations. Ainsi, une interaction dipolaire de la forme $\sum_i q_i \mathbf{x}_i$ (comme l'interaction de particules chargées avec un champ électrique externe) ne couple que les états d'un système de particules dont les moments cinétiques totaux sont égaux ou ne diffèrent que de ± 1 .

V.4. Permutations et tresses

Le groupe des permutations joue un rôle particulier en physique, en particulier à cause de son implication dans le principe de Pauli.

V.4.1. Le groupe des permutations. — Soit S_n le groupe formé des permutations de n objets. Ce groupe peut être vu comme le groupe des permutations des nombres $1, 2, \dots, n$. C'est donc le groupe des bijections $\sigma : j \rightarrow \sigma(j)$ de l'ensemble $\{1, 2, \dots, n\}$. On représente souvent une permutation graphiquement sous la forme :

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

La loi de groupe est la loi de composition. La dimension de S_n est $n!$.

Les transpositions P_{jk} sont les permutations qui échangent j et k tout en laissant invariant les autres éléments, i.e. $P_{jk}(j) = k$ et $P_{jk}^2 = 1$.

Les cycles sont les permutations circulaires. Un cycle d'ordre k est par exemple formé de la chaîne $1 \rightarrow \sigma(1) \rightarrow \sigma^2(1) \cdots \sigma^{k-1}(1) \rightarrow 1$. Chaque cycle portant sur k nombres engendre un groupe isomorphe à \mathbb{Z}_k .

Toute permutation se décompose en produit de cycles. Pour montrer cette propriété il suffit de présenter la permutation en ordonnant les nombres suivant les cycles auxquels ils appartiennent :

$$\begin{pmatrix} 1 & \sigma(1) & \cdots & \sigma^{k_1-1}(1) & p & \cdots \\ \sigma(1) & \sigma^2(1) & \cdots & 1 & \sigma(p) & \cdots \end{pmatrix}$$

où p est le plus petit entier qui n'appartient pas au cycle du nombre 1, et ainsi de suite... Le produit des cycles intervenant dans cette décomposition est commutatif.

Les classes d'équivalence de S_n sont constituées des permutations dont la décomposition en cycles fait intervenir des cycles de mêmes longueurs k_1, k_2, \dots avec $n = \sum_j k_j$. En effet, si une permutation σ se décompose en cycles de longueurs données, alors clairement toute permutation conjuguée $\tau\sigma\tau^{-1}$ se décompose en cycles de mêmes longueurs. Réciproquement, si deux permutations σ, σ' se décomposent en cycles de longueurs k_1, k_2, \dots , alors il existe une permutation τ telle que $\sigma' = \tau\sigma\tau^{-1}$. Les classes de conjugaison de S_n sont donc indexées par les partitions de n en somme d'entier, $n = \sum_j k_j$. Si q_1 est le nombre de cycle de longueur 1, q_2 celui de cycle de longueur 2, etc., on a $\sum_j j q_j = n$ et on note $\underline{q} = [1^{q_1} 2^{q_2} \cdots]$ la classe de conjugaison correspondante. On montre que sa dimension est $n! / \prod_j q_j! j^{q_j}$.

- *Générateurs et relations.*

Toute permutation s'écrit comme le produit de transpositions (pour le prouver il suffit de le vérifier pour les cycles). Cette décomposition n'est pas unique (comme nous allons le voir sur des exemples) mais la parité du nombre de transpositions est bien définie, ce qui permet de définir la signature d'une permutation ϵ_σ comme étant égale à $+1$ si ce nombre est pair et -1 si il est impair. L'application $\sigma \rightarrow \epsilon_\sigma$ définit une représentation de dimension un de S_n , i.e. $\epsilon_{\sigma_1\sigma_2} = \epsilon_{\sigma_1}\epsilon_{\sigma_2}$.

Plus finement, toute permutation s'écrit comme le produit de transpositions de nombres adjacents, mais cette décomposition n'est pas unique. Vérifions le sur des exemples simples mais génériques. Soit $s_i \equiv P_{i;i+1}$ la transposition de i et $i+1$. Les seules transpositions de S_3 sont P_{12} , P_{23} et P_{13} . Les deux premières sont clairement des permutations de nombres adjacents, quant à la dernière on a :

$$P_{13} = P_{12}P_{23}P_{12} = P_{23}P_{12}P_{23}$$

Par récurrence, cette décomposition s'étend à toute transposition puisque par exemple $P_{1;n} = P_{n;n-1}P_{1;n-1}P_{n;n-1}$.

Ainsi, on déduit de cet exemple que les $s_i \equiv P_{i;i+1}$ satisfont les relations :

$$\begin{aligned} s_i^2 &= 1 \\ s_i s_j &= s_j s_i, \quad |i-j| \geq 2 \\ s_i s_{i+1} s_i &= s_{i+1} s_i s_{i+1} \end{aligned}$$

On montre que ces dernières sont les seules relations satisfaites par les s_i dans le groupe des permutations. Le groupe S_n est donc le groupe engendré par les s_i modulo ces relations. Celles-ci ont une représentation graphique simple que nous donnons ci-après en parallèle avec la présentation du groupe des tresses.

V.4.2. Représentations du groupe des permutations. — Le groupe des permutations agit naturellement sur le produit tensoriel de n copies d'un espace vectoriel V . Notons $V^{\otimes n}$ ce produit tensoriel. Par linéarité, il suffit de définir cette action sur les produits tensoriels de n vecteurs :

$$\sigma \cdot (v_1 \otimes v_2 \otimes \cdots \otimes v_n) \equiv v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \cdots \otimes v_{\sigma^{-1}(n)}$$

pour $\sigma \in S_n$. On vérifie que cette relation définit une représentation de S_n sur $V^{\otimes n}$. (Il est nécessaire de mettre le facteur σ^{-1} pour définir une représentation). Par exemple pour $n=3$, $P_{12} \cdot (a \otimes b \otimes c) = b \otimes a \otimes c$ et $P_{13}P_{12} \cdot (a \otimes b \otimes c) = c \otimes a \otimes b$ qui est une permutation circulaire.

De façon similaire, le groupe des permutations agit aussi naturellement sur l'espace des fonctions de n variables. Soit f une fonction de n variables x_1, \dots, x_n . L'action de S_n sur f est définie par :

$$(\sigma \cdot f)(x_1, \dots, x_n) \equiv f(x_{\sigma^{-1}(1)}, \dots, x_{\sigma^{-1}(n)})$$

Ces représentations ne sont évidemment pas irréductibles.

Deux représentations (de dimension un) jouent un rôle particulier en mécanique quantique : celles associées aux fonctions, ou aux vecteurs de $V^{\otimes n}$, totalement symétriques ou totalement anti-symétriques car, d'après le principe de Pauli, la fonction d'onde d'un système constitué de bosons (respectivement, de fermions) identiques est symétrique (respectivement, anti-symétrique) dans l'échange des particules. Par définition, si f_s (respectivement, f_a) est une fonction, ou un état, symétrique (respectivement, anti-symétrique), alors :

$$\sigma \cdot f_s = f_s \quad , \quad \sigma \cdot f_a = \epsilon_\sigma f_a, \quad \forall \sigma \in S_n$$

Les fonctions d'onde symétriques sont dans l'image du symétriseur $\mathfrak{s} = \frac{1}{n!} \sum_{\sigma \in S_n} \sigma$, les fonctions d'onde anti-symétriques sont dans l'image de l'anti-symétriseur $\mathfrak{a} = \frac{1}{n!} \sum_{\sigma \in S_n} \epsilon_\sigma \sigma$ où ϵ_σ est la

signature de la permutation σ :

$$V_{\mathfrak{s}} \equiv \text{Sym } V^{\otimes n} = \mathfrak{s} \cdot V^{\otimes n}, \quad V_{\mathfrak{a}} \equiv \Lambda^n V = \mathfrak{a} \cdot V^{\otimes n}.$$

Pour $n = 2$, les représentations symétrique et anti-symétrique sont les seules représentations irréductibles de S_2 . Les projecteurs (anti)-symétriseurs

$$\mathfrak{s} = \frac{1}{2}(1 + P_{12}), \quad \mathfrak{a} = \frac{1}{2}(1 - P_{12}),$$

sont complémentaires : $\mathfrak{a}\mathfrak{s} = 0$ et $\mathfrak{s} + \mathfrak{a} = 1$. Ils satisfont $P_{12}\mathfrak{a} = \mathfrak{a}$ et $P_{12}\mathfrak{s} = -\mathfrak{s}$. Leurs espaces images sont donc stables sous le groupe des permutations et on a $V \otimes V = V_{\mathfrak{s}} \oplus V_{\mathfrak{a}}$.

Le groupe S_3 possède deux représentations de dimension un inéquivalentes, les représentations symétrique et anti-symétrique, et une représentation irréductible de dimension 2 (souvent) appelée représentation 'mixte'. La façon la plus simple de la construire consiste à définir l'action de S_3 par permutations sur les triplets de nombres complexes $(z_1, z_2, z_3) \in \mathbb{C}^3$ de somme nulle $z_1 + z_2 + z_3 = 0$. Il s'agit bien d'un espace de dimension 2 et l'on vérifie que cette représentation est irréductible. Comme dans le cas de S_2 , définissons les (anti)-symétriseurs,

$$\begin{aligned} \mathfrak{s} &= \frac{1}{6}(1 + P_{12} + P_{13} + P_{23} + P_{12}P_{13} + P_{13}P_{12}) \\ \mathfrak{a} &= \frac{1}{6}(1 - P_{12} - P_{13} - P_{23} + P_{12}P_{13} + P_{13}P_{12}), \end{aligned}$$

On vérifie que ce sont des projecteurs. Ils satisfont bien sûr $P_{12}\mathfrak{s} = P_{23}\mathfrak{s} = \mathfrak{s}$ et $P_{12}\mathfrak{a} = P_{23}\mathfrak{a} = -\mathfrak{a}$. Ils définissent donc les représentations de dimension 1 par multiplication à gauche dans S_3 . Définissons aussi des projecteurs mixtes,

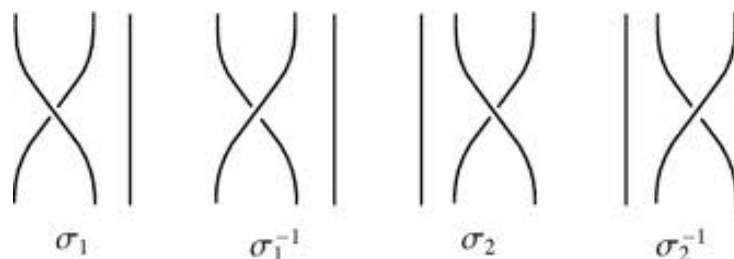
$$\begin{aligned} \mathfrak{m} &= \frac{1}{3}(1 + P_{12} - P_{13} - P_{12}P_{13}) = \frac{1}{3}(1 + P_{12})(1 - P_{13}) \\ \mathfrak{n} &= \frac{1}{3}(1 - P_{12} + P_{13} - P_{13}P_{12}) = \frac{1}{3}(1 + P_{13})(1 - P_{12}). \end{aligned}$$

On vérifie que \mathfrak{m} et \mathfrak{n} sont des projecteurs et, que \mathfrak{m} et $\mathfrak{m}'' \equiv P_{13}\mathfrak{m}$ sont stables par multiplication à gauche dans S_3 , de sorte qu'ils définissent une représentation de dimension 2. En effet, $P_{12}\mathfrak{m} = \mathfrak{m}$, $P_{13}\mathfrak{m} = \mathfrak{m}''$ et $P_{12}\mathfrak{m}'' = -\mathfrak{m} - \mathfrak{m}''$, $P_{13}\mathfrak{m}'' = \mathfrak{m}$. Une construction identique s'applique à \mathfrak{n} et $\mathfrak{n}'' \equiv P_{12}\mathfrak{n}$ puisque $\mathfrak{n} = P_{23}\mathfrak{m}P_{23}$.

Un calcul simple montre que ces projecteurs sont complémentaires, $\mathfrak{a} + \mathfrak{s} + \mathfrak{m} + \mathfrak{n} = 1$, et orthogonaux, $\mathfrak{a}\mathfrak{s} = \mathfrak{a}\mathfrak{m} = \mathfrak{m}\mathfrak{n} = \dots = 0$. Ainsi, en considérant la représentation de S_3 sur $V^{\otimes 3}$ décrite ci-dessus, on décompose ce produit tensoriel en la somme directe des espaces images de chacun de ces projecteurs, soit :

$$V^{\otimes 3} = V_{\mathfrak{s}} \oplus (V_{\mathfrak{m}} \oplus V_{\mathfrak{n}}) \oplus V_{\mathfrak{a}}.$$

Ceci signifie qu'un tenseur d'ordre trois se décompose en la somme de quatre tenseurs, chacun appartenant à l'un des espaces images. Les espaces $V_{\mathfrak{s}}$ et $V_{\mathfrak{a}}$ sont stables sous les permutations, et correspondent respectivement aux tenseurs symétriques et anti-symétriques. Les espaces $V_{\mathfrak{m}}$ et $V_{\mathfrak{n}}$ s'échangent sous l'action des permutations car les relations $P_{12}\mathfrak{m} = \mathfrak{m}$ et $P_{23}\mathfrak{m} = \mathfrak{n}P_{23}$ impliquent que $P_{12}V_{\mathfrak{m}} = V_{\mathfrak{m}}$ et $P_{23}V_{\mathfrak{m}} = V_{\mathfrak{n}}$. Les espaces $V_{\mathfrak{m}}$ et $V_{\mathfrak{n}}$ sont donc isomorphes et s'échangent sous la représentation mixte de S_3 et, $V_{\mathfrak{m}} \oplus V_{\mathfrak{n}} \cong W_{\mathfrak{m}} \otimes V_{\mathfrak{m}}$ où $W_{\mathfrak{m}}$ est un espace vectoriel de dimension

FIGURE 3. Représentation graphique des générateurs σ_j sur trois brins.

2 sur lequel la représentation mixte de S_3 est réalisée. On peut donc écrire la décomposition,

$$V^{\otimes 3} = V_s \oplus (W_m \otimes V_m) \oplus V_a,$$

qui est un cas particulier d'une relation vraie pour $V^{\otimes n}$ pour tout n connue sous le nom dualité de Weyl. Il peut être instructif de comparer cette décomposition à celle de l'espace de Hilbert d'un système quantique en secteurs.

La construction précédente se généralise pour toute valeur de n . Elle permet la construction de toutes les représentations irréductibles de S_n . Elle met en oeuvre une généralisation des projecteurs introduits ci-dessus, appelée symétriseurs d'Young. (Leur étude ainsi que celle des représentations du groupe symétrique peut faire l'objet d'un travail personnel!).

V.4.3. Le groupe des tresses et ses applications. — Le groupe des tresses B_n ('braid group' en anglais) est le groupe engendré par les éléments σ_i , $i = 1, \dots, n-1$, et leur inverse σ_i^{-1} , satisfaisant les relations :

$$\begin{aligned} \sigma_i \sigma_i^{-1} &= \sigma_i^{-1} \sigma_i = 1 \\ \sigma_i \sigma_j &= \sigma_j \sigma_i, \quad |i-j| \geq 2 \\ \sigma_i \sigma_{i+1} \sigma_i &= \sigma_{i+1} \sigma_i \sigma_{i+1} \end{aligned}$$

Celles-ci coïncident avec celles définissant le groupe des permutations à l'exception de la propriété d'involution, i.e. dans le groupe B_n on n'impose pas que les σ_i soient des involutions. Il y a donc un endomorphisme de groupe surjectif de B_n vers S_n : l'image de $\sigma_i \in B_n$ est $s_i \in S_n$.

Comme le groupe des permutations, le groupe des tresses peut se représenter graphiquement. Les générateurs sont représentés comme sur la Figure (3) par leur action sur des 'brins' de cordes. L'opération σ_i consiste à échanger les brins en positions i et $i+1$, en passant un brin au dessus de l'autre. Le produit se fait en 'raboutant' les brins. Alors que la définition précédente donne une caractérisation purement algébrique du groupe des tresses, la description graphique en donne une caractérisation géométrique. Ainsi une tresse peut être vue comme une collection de n chemins, ou brins, joignant n points $\mathbf{x}_1, \dots, \mathbf{x}_n$ fixés du plan à une permutation quelconque de ces points. En choisissant de paramétrer ces chemins par $t \in [0, 1]$, une tresse correspond à la donnée, à déformation continue près, des courbes non-intersectantes $\gamma_i(t)$, $i = 1, \dots, n$, telles que $\gamma_i(0) = \mathbf{x}_i$ et $\gamma_i(1) = \mathbf{x}_{s^{-1}(i)}$ pour $s \in S_n$. Un théorème due à E. Artin montre que les définitions algébrique et géométrique sont équivalentes. Dans le groupe des permutations,

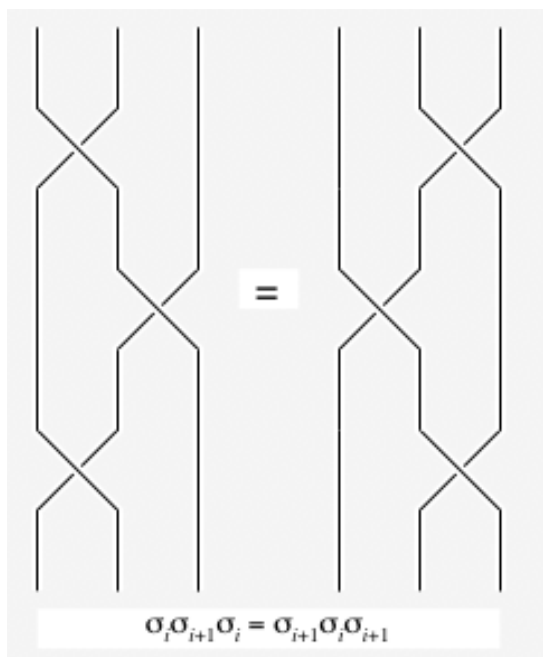


FIGURE 4. Représentation graphique de la relation cubique du groupe des tresses.

contrairement à celui des tresses, on ne distingue pas suivant qu'un brin passe au-dessus ou au-dessous d'un autre. Les relations entre générateurs définissant le groupe des tresses admettent une représentation graphique simple. Voir Figure (4).

- *Le groupe des tresses et les statistiques exotiques.*

Le groupe des tresses apparaît naturellement dans l'analyse des statistiques possibles de particules confinées sur un espace bidimensionnel, c'est-à-dire dans l'analyse du comportement de ces particules lors de l'échange de leurs positions. En effet, les lignes d'univers de ces particules, qui sont des courbes en dimension $2 + 1$, dessinent des tresses. Il existe donc une obstruction topologique à déplacer ces lignes d'univers car elles ne peuvent pas se traverser. Cette contrainte offre la possibilité d'avoir en dimension $2 + 1$ des statistiques plus exotiques que celles des bosons ou des fermions. En dimension supérieure, trois ou plus, cette obstruction n'existe pas et les seules statistiques possibles sont celles des bosons ou fermions.

Comment définir les statistiques exotiques en dimension $2 + 1$? Les fonctions d'onde $\Psi(\mathbf{x}_1, \dots, \mathbf{x}_n)$ dépendent des positions (planaires) des n particules, que l'on suppose distinctes deux à deux. On suppose qu'elles sont à valeurs dans un espace de Hilbert qui supporte une représentation \mathfrak{R} du groupe des tresses (i.e. le groupe des tresses agit sur les vecteurs de cet espace). Par définition le groupe des tresses agit sur les n positions car il suffit de suivre le parcours de ces positions le long des n brins définissant une tresse (cf. la définition géométrique des tresses). Notons $\Psi(\sigma \cdot [\mathbf{x}_1, \dots, \mathbf{x}_n])$ la valeur de la fonction d'onde après avoir déplacé continument les n positions le long des n brins définissant la tresse σ . Notons qu'une tresse σ est un chemin fermé sur $(\mathbb{R}^2)^n / S_n$, l'espace des multiplats $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ définis modulo permutations. (Les x_j sont supposés distincts deux à deux). On dit que les particules satisfont à la statistique

exotique associée à la représentation \mathfrak{R} du groupe des tresses si :

$$\Psi(\sigma \cdot [\mathbf{x}_1, \dots, \mathbf{x}_n]) = (\sigma^{-1} \cdot \Psi)(\mathbf{x}_1, \dots, \mathbf{x}_n)$$

Cette relation impose en particulier que les fonctions d'onde $\Psi(\mathbf{x}_1, \dots, \mathbf{x}_n)$ ne sont pas mono-valuées sur $(\mathbb{R}^2)^n/S_n$. Les statistiques possibles en dimension $2 + 1$ sont donc classées par les représentations du groupe des tresses.

L'existence de systèmes physiques possédant des excitations ayant une statistique exotique est fort probable : les excitations de basse énergie dans les systèmes soumis à l'effet Hall quantique fractionnaire sont des candidats sérieux bien que la nature fractionnaire de leur statistique n'a pas encore été mise en évidence expérimentalement à ce jour (juin 2010). Néanmoins, ces systèmes ont été proposés comme des candidats possibles sur lesquels les notions de calcul quantique pouvaient être testées.

- *Le groupe des tresses et les systèmes quantiques intégrables.*

Il existe un lien profond entre le groupe des permutations et/ou le groupe des tresses et les systèmes quantiques exactement solubles, c'est-à-dire les systèmes quantiques pour lesquels on peut évaluer exactement les énergies propres de l'hamiltonien ainsi que les vecteurs propres.

Commençons par supposer que des particules quantiques, se déplaçant sur une ligne, interagissent entre elles (par une force à courte portée). Ces particules sont caractérisées par leur impulsion p ainsi que par d'autres nombres quantiques tels que la charge, le spin, la couleur, etc. On note $|p, a\rangle$ l'état d'une particule d'impulsion p et de nombre quantique a . Ces états forment une base de l'espace de Hilbert \mathcal{H}_p des états à une particule d'impulsion p . Si elles sont loin les unes des autres, ces particules n'interagissent pas car l'interaction est à courte portée, et elles évoluent librement le long de la ligne. Par contre, lorsque deux particules sont proches, elles interagissent et échangent leurs positions. L'effet de l'interaction se reflète aussi dans une modification des nombres quantiques des particules (autre que les impulsions). La conservation de l'impulsion, $p_1 + p_2 = p'_1 + p'_2$, et celle de l'énergie, $E_1 + E_2 = E'_1 + E'_2$, imposent que les impulsions initiales et finales soient identiques à permutation près. Si $|p_1, a_1\rangle \otimes |p_2, a_2\rangle$ est l'état des deux particules avant la collision, on transcrit le résultat de l'interaction en écrivant l'état des particules après la collision sous la forme

$$\sum_{b_1, b_2} |p_2, b_2\rangle \otimes |p_1, b_1\rangle R(p_1, p_2)_{a_1, a_2}^{b_1, b_2}.$$

Les facteurs $R(p_1, p_2)_{a_1, a_2}^{b_1, b_2}$ sont les amplitudes de transition lors de la collision d'un état à deux particules à un autre état à deux particules. A impulsions fixées, ces amplitudes forment une matrice, appelée matrice de diffusion ("scattering matrix" en anglais),

$$R_{12}(p_1, p_2) : \mathcal{H}_{p_1} \otimes \mathcal{H}_{p_2} \rightarrow \mathcal{H}_{p_2} \otimes \mathcal{H}_{p_1}$$

où \mathcal{H}_{p_1} et \mathcal{H}_{p_2} sont les espaces de Hilbert respectifs des deux particules.

Considérons maintenant la situation où trois particules, numérotées 1, 2, 3, interagissent successivement deux à deux. Cette succession d'interactions peut se faire de deux façons différentes : soit (i) la particule 1 interagit avec 2, puis 1 avec 3 et enfin 2 avec 3, soit (ii) la particule 2 interagit d'abord avec 3, puis 1 avec 3 et enfin 1 avec 2. Que le résultat de ces collisions successives soit indépendant du chemin d'interaction impose aux amplitudes de transitions de satisfaire l'équation

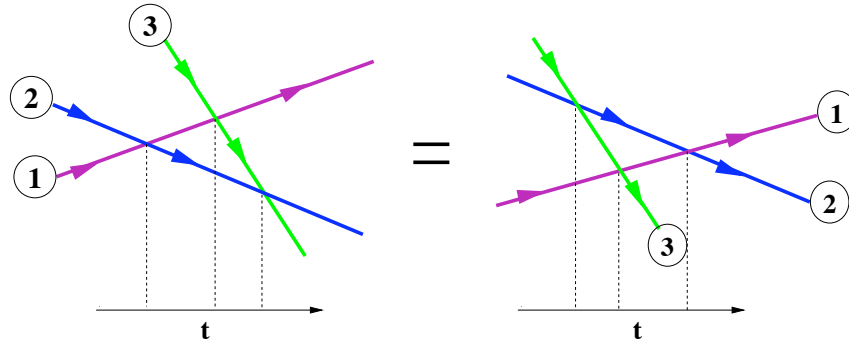


FIGURE 5. Représentation graphique des interactions successives de trois particules.

suivante, dite équation de Yang-Baxter :

$$R_{23}(p_2, p_3) R_{13}(p_1, p_3) R_{12}(p_1, p_2) = R_{12}(p_1, p_2) R_{13}(p_1, p_3) R_{23}(p_2, p_3)$$

Il y a une hypothèse implicite sous-jacente à cette description : on suppose que l'interaction entre les trois particules se décompose en une succession d'interactions à deux particules. Cette hypothèse est liée au caractère exactement soluble du système quantique sous-jacent.

Afin de comprendre la signification géométrique de cette équation, traçons, comme sur la Figure (5), les trajectoires des trois particules lors de cet enchainement de collisions (l'axe temporel est horizontal, l'axe spatial est vertical). Ce diagramme coïncide avec celui définissant les relations dans le groupe des tresses. Il existe donc une relation entre les matrices de diffusion satisfaisant les équations de Yang-Baxter et les représentations du groupe des tresses. En effet, la matrice $S_{12} \equiv P_{12}R_{12}(p_1, p_2)$, où P_{12} est l'opérateur transposition, satisfait l'équation

$$S_{12} S_{23} S_{12} = S_{23} S_{12} S_{23}$$

qui est l'équation de définition du groupe des tresses à trois brins. Comme les relations de définition du groupe des tresses à un nombre de brins quelconque ne fait intervenir que trois brins au plus, la correspondance entre matrices de diffusion satisfaisant l'équation de Yang-Baxter et représentations du groupe de tresses s'étend à un nombre quelconque de particules. Cette correspondance est à l'origine de la théorie des systèmes quantiques exactement solubles et des méthodes de résolution de ces systèmes, ainsi qu'à l'origine de la théorie des groupes quantiques qui sont des déformations des groupes issues de leur quantification.

V.5. Les classes de Wigner

E. Wigner a proposé une classification des hamiltoniens quantiques suivant leurs propriétés par renversement du temps. Nous allons brièvement discuter cette classification, qui trouve de nombreux échos en physique mathématique, notamment dans le cadre de l'étude des grandes matrices aléatoires et de leurs applications.

Un hamiltonien H est un opérateur hermitique sur un espace de Hilbert : $H = H^\dagger$. Si aucune autre spécificité du système modélisé par l'hamiltonien H n'est précisée, cette relation est la seule imposée.

Nous dirons que H est invariant par renversement du temps si il satisfait la propriété suivante : pour toute solution $\psi(t)$ de l'équation de Schroedinger, $i\partial_t\psi = H \cdot \psi$, sa transformée temporelle $\psi^T(t) \equiv C \cdot \psi^*(-t)$ est aussi solution. Ici C est un opérateur linéaire (une matrice) agissant sur l'espace de Hilbert et ψ^* le complexe conjugué de ψ . Cette propriété est assurée si l'hamiltonien satisfait la relation $H = CH^*C^{-1}$, soit $H = CH^TC^{-1}$ car $H = H^\dagger$. Un hamiltonien satisfaisant cette relation sera dit invariant par renversement du temps. L'hamiltonien d'une particule libre fournit évidemment un exemple d'hamiltonien invariant par renversement du temps, par contre la présence d'un champ magnétique externe brise cette symétrie.

La matrice C est unitaire, $CC^\dagger = \mathbf{1}$, afin que la transformation temporelle préserve la norme de la fonction d'onde. Une double itération de la transformation temporelle sur un état doit reproduire celui-ci. Comme un état physique est spécifié par un rayon de l'espace de Hilbert, c'est-à-dire par une fonction d'onde définie à une phase près, cette dernière condition impose que $CC^* = e^{i\theta} \mathbf{1}$ où θ est une phase. La relation d'unitarité $CC^\dagger = \mathbf{1}$ transforme cette condition en $C = e^{i\theta} C^T$. La phase θ est donc égale soit à 0, soit à π , et $C^T = \pm C$.

En résumé, il existe trois classes selon E. Wigner :

- (i) $H = H^\dagger$;
- (ii) $H = H^\dagger = CH^TC^{-1}$, $C^T = C$, $CC^\dagger = \mathbf{1}$;
- (iii) $H = H^\dagger = CH^TC^{-1}$, $C^T = -C$, $CC^\dagger = \mathbf{1}$.

L'étape finale consiste à classer les matrices C à la transformation $C \rightarrow UCU^T$ près, avec U unitaire (transformation qui reflète un changement de base sur l'espace de Hilbert). Commençons par le cas $C^T = C$, $CC^\dagger = \mathbf{1}$. Ces deux conditions imposent que les parties réelle et imaginaire de C sont symétriques par transposition et commutent entre elles. Elles sont donc diagonalisables simultanément par conjugaison par une matrice orthogonale réelle. Ainsi $C = O\Delta O^T$ où $OO^T = \mathbf{1}$ et où Δ est diagonale et unitaire. En posant $U = O\Delta^{1/2}$, on vérifie que $C = UU^T$ et que U est unitaire. Donc, les matrices satisfaisant les deux conditions $CC^\dagger = \mathbf{1}$ et $C = C^T$ sont équivalentes à la matrice identité $C = \mathbf{1}$. Un argument similaire (laissé à la lectrice, au lecteur motivé(e)) conduit à montrer que les matrices qui satisfont aux conditions $CC^\dagger = \mathbf{1}$ et $C^T = -C$ sont équivalentes à $C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \otimes \mathbf{1}$. En résumé, à transformations unitaires près, on a :

$$\text{cas (ii) : } C = \mathbf{1}; \quad \text{cas (iii) : } C = \begin{pmatrix} 0 & \mathbf{1} \\ -\mathbf{1} & 0 \end{pmatrix}.$$

En particulier, le cas (iii) n'existe que pour un espace de Hilbert de dimension paire.

Chacune de ces classes possèdent un groupe de symétrie. Dans le cas (i), la transformation $H \rightarrow UHU^{-1}$ préserve la condition d'hermité $H = H^\dagger$ si U est unitaire, $UU^\dagger = \mathbf{1}$; le groupe de symétrie est donc le groupe unitaire $U(n)$ avec n la dimension de l'espace de Hilbert. Dans le cas (ii), cette transformation préserve la condition $H = H^\dagger = H^T$ si U est une matrice réelle orthogonale; le groupe de symétrie est donc le groupe orthogonal $O(n)$. Dans le cas (iii), cette transformation préserve les relations $H = H^\dagger = CH^TC^{-1}$ avec $C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ si U satisfait $C = UCU^T$, relation qui caractérise le groupe symplectique $Sp(2n)$. En résumé, les groupes de symétries sont :

$$\text{cas (i) : } U(n); \quad \text{cas (ii) : } O(n); \quad \text{cas (iii) : } Sp(2n).$$

Cette classification joue par exemple un rôle important dans l'analyse de l'effet du désordre sur les systèmes électroniques macroscopiques.

V.6. Espaces de Fock

Rappelons d'abord quelques propriétés de l'oscillateur harmonique quantique. Soit q la coordonnée et p l'impulsion, $[q, p] = i$. On a posé $\hbar = 1$. Par définition, l'hamiltonien d'un oscillateur harmonique (de pulsation ω) est $h = \frac{1}{2}(p^2 + \omega^2 q^2)$. On définit deux opérateurs a^\dagger et a , appelés opérateurs création et annihilation, par :

$$q = \frac{1}{\sqrt{2\omega}}(a + a^\dagger), \quad p = -i\sqrt{\frac{\omega}{2}}(a - a^\dagger)$$

Il satisfont les relations de commutation

$$[a, a^\dagger] = 1$$

L'hamiltonien s'écrit $h = \omega(aa^\dagger + 1/2)$. Son spectre est formé des énergies $e_n = (n + 1/2)\omega$ avec n entier positif. L'état fondamental est l'unique état $|0\rangle$ spécifié par $a|0\rangle = 0$. L'état propre d'énergie e_n est proportionnel à $(a^\dagger)^n|0\rangle$. Donc l'opérateur a^\dagger "crée" les états excités par actions successives sur l'état fondamental.

V.6.1. Bosons. — Soit V un espace vectoriel, de dimension D , muni d'une forme sesquilinéaire $g(u, v)$, $u, v \in V$. Les opérateurs créations $a^\dagger(v)$ et annihilations $a(v)$ bosoniques dépendent (anti)-linéairement de $v \in V$ et, par définition, satisfont les relations de commutation :

$$[a(u), a^\dagger(v)] = g(u, v), \quad [a(u), a(v)] = 0 = [a^\dagger(u), a^\dagger(v)], \quad \forall u, v \in V$$

Plus explicitement, si e_i , $i = 1, \dots, D$, est une base de V , on pose $a_i = a(e_i)$ et $a_i^\dagger = a^\dagger(e_i)$. Alors

$$[a_i, a_j^\dagger] = g_{ij}, \quad [a_i, a_j] = 0 = [a_i^\dagger, a_j^\dagger],$$

où $g_{ij} = g(e_i, e_j)$ est égale à δ_{ij} si la base e_i est choisie orthonormée.

L'espace de Fock, noté \mathcal{F}_V^b , sur lesquels les opérateurs agissent, admet plusieurs descriptions. La plus intuitive consiste à définir un état $|0\rangle$, appelé état du vide qui, par définition, est dans le noyau de tous les opérateurs annihilation, $a(u)|0\rangle = 0$ pour tout $u \in V$. Tous les autres états de l'espace de Fock s'obtiennent par actions successives des opérateurs créations, $a^\dagger(u_1) \cdots a^\dagger(u_n)|0\rangle$ avec $u_j \in V$. L'espace de Fock est donc un espace gradué :

$$\mathcal{F}_V^b = \bigoplus_{n=0}^{\infty} \mathcal{F}_V^{b(n)}, \quad \mathcal{F}_V^{b(n)} \equiv \text{Span}\{a^\dagger(u_1) \cdots a^\dagger(u_n)|0\rangle, \quad u_j \in V\}$$

L'espace $\mathcal{F}_V^{b(1)}$ est isomorphe à V et constitue par définition l'espace des états à une particule. L'espace $\mathcal{F}_V^{b(n)}$ est l'espace des états à n bosons. Les opérateurs création-annihilations agissent par multiplication à gauche en utilisant les relations de commutation et la propriété que le vide $|0\rangle$ est annihilé par les opérateurs $a(u)$. En pratique, on déplace les opérateurs annihilations vers la droite en utilisant les relations de commutation jusqu'à ce qu'il agissent sur le vide. La

structure hilbertienne sur \mathcal{F}_V^b est définie en déclarant que $\langle 0|0\rangle = 1$ et que les opérateurs a_i^\dagger sont les conjugués hermitiques des a_i . Par exemple,

$$\|a^\dagger(u)|0\rangle\|^2 = \langle 0|a(u^*)a^\dagger(u)|0\rangle = \|u\|^2.$$

Dans un langage plus algébrique, mais plus abstrait, $\mathcal{F}_V^b \cong \text{Poly}[a_i, a_j^\dagger]/\mathcal{I}[a_i]$ où $\text{Poly}[a_i, a_j^\dagger]$ est l'algèbre des polynômes en les variables $a_i, a_j^\dagger, i, j = 1, \dots, D$ et, où $\mathcal{I}[a_i]$ est l'idéal (à gauche) engendré par les a_i .

Comme les opérateurs créations commutent entre eux, les états $a^\dagger(u_1) \cdots a^\dagger(u_n)|0\rangle$ sont symétriques en u_1, \dots, u_n et s'identifient avec la symétrisation du vecteur $u_1 \otimes \cdots \otimes u_n$. Ainsi, on a l'isomorphisme

$$\mathcal{F}_V^b = \bigoplus_{n=0}^{\infty} \mathcal{F}_V^{b(n)}, \quad \mathcal{F}_V^{b(n)} = \text{Sym } V^{\otimes n}$$

La correspondance entre cette description et la précédente est

$$\sqrt{n!} \mathfrak{s} \cdot u_1 \otimes \cdots \otimes u_n \leftrightarrow a^\dagger(u_1) \cdots a^\dagger(u_n)|0\rangle,$$

où $\mathfrak{s} = \frac{1}{n!} \sum_{\sigma \text{ perm.}} \sigma$ est l'opérateur de symétrisation. De cette correspondance, il est facile de déduire l'action des opérateurs créations et annihilations sur $\text{Sym } V^{\otimes n}$.

Une autre description consiste à réaliser les opérateurs a_i et a_i^\dagger comme des opérateurs différentiels. On introduit les variables $z_i, i = 1, \dots, D$, et on pose $a_i = \partial_{z_i}$ et $a_i^\dagger = z_i$. On vérifie que $[a_i, a_j^\dagger] = \delta_{ij}$. L'espace de Fock est alors isomorphe à l'espace des polynômes dans les z_i :

$$\mathcal{F}_V^b \cong \text{Poly}[z_1, \dots, z_n]$$

qui est bien un espace gradué (suivant le degré du polynôme). Le vide s'identifie avec le polynôme constant 1.

Les opérateurs $N_i = a_i^\dagger a_i$ commutent entre eux, $[N_i, N_j] = 0$. Leurs valeurs propres sont entières et leurs états propres sont les états à n particules car $[N_i, a_j^\dagger] = \delta_{ij} a_j^\dagger$ et $[N_i, a_j] = -\delta_{ij} a_j$. On dit que l'opérateur N_i compte le nombre de bosons dans l'état à une particule e_i et on appelle ces opérateurs les opérateurs 'nombres d'occupation'. La fonction partition $Z_V^b \equiv \text{Tr}_{\mathcal{F}_V^b}(q^{\sum_i u_i N_i})$ est la fonction génératrice des multiplicités de leurs valeurs propres. Comme l'espace de Fock \mathcal{F}_V^b est isomorphe au produit tensoriel des espaces de Fock construits à l'aide des couples d'opérateurs création-annihilations a_i^\dagger, a_i associés à chacun des états à une particule e_i , on a :

$$Z_V^b \equiv \text{Tr}_{\mathcal{F}_V^b}(q^{\sum_i u_i N_i}) = \prod_{i=1}^D \left(\frac{1}{1 - q^{u_i}} \right).$$

En particulier, la dimension de l'espace des états à n particules est

$$\dim \mathcal{F}_V^{b(n)} = \frac{(D + n - 1)!}{n!(D - 1)!}.$$

Pour le vérifier, il suffit de développer $(1 - q^u)^{-D}$ en puissance de q^u .

Plus généralement, considérons l'ensemble des opérateurs biliénaires dans les opérateurs créations ou annihilations :

$$a_i^\dagger a_j^\dagger, \quad a_i^\dagger a_j, \quad a_i a_j.$$

On vérifie que le commutateur de deux de ces opérateurs est encore un opérateur bilinéaire. Ils forment donc une représentation d'une algèbre de Lie, l'algèbre des transformations symplectiques notée $sp(2D)$. Cette algèbre est de dimension $D(2D+1)$. Tous ces opérateurs préservent la parité du nombre d'occupation total $N = \sum_i N_i$. L'espace de Fock \mathcal{F}_V^b se décompose donc en deux représentations (chacune de dimension infinie), respectivement constituée des états à un nombre pair, ou impair, de particules.

Les opérateurs bilinéaires qui commutent avec le nombre d'occupation total, c'est-à-dire les opérateurs de la forme $a_i^\dagger a_j$, $i, j = 1, \dots, D$, engendrent une sous-algèbre de $sp(2D)$ isomorphe à l'algèbre $u(D)$ des transformations unitaires sur V . Chaque sous-espace $\mathcal{F}_V^{b(n)}$ à nombre de particules fixé constitue une représentation (irréductible) de $u(D)$ car ces opérateurs commutent avec $N = \sum_i N_i$.

V.6.2. Fermions. — La construction des espaces de Fock fermioniques est similaire à celle des espaces de Fock bosoniques (nous allons donc nous répéter un peu). Seuls quelques détails changent, notamment les conditions de symétrisation sont transformées en conditions d'anti-symétrisation.

Soit V un espace vectoriel, de dimension D , muni d'une forme sesquilinéaire $g(u, v)$, $u, v \in V$. Les opérateurs créations $b^\dagger(v)$ et annihilations $b(v)$ fermioniques dépendent (anti)-linéairement de $v \in V$ et, par définition, satisfont les relations d'anti-commutation⁽⁴⁾ :

$$\{b(u), b^\dagger(v)\} = g(u, v), \quad \{b(u), b(v)\} = 0 = \{b^\dagger(u), b^\dagger(v)\}, \quad \forall u, v \in V$$

Plus explicitement, si e_i est une base orthonormée de V , $g(e_i, e_j) = \delta_{ij}$, on pose $b_i = b(e_i)$ et $b_i^\dagger = b^\dagger(e_i)$. Alors

$$\{b_i, b_j^\dagger\} = \delta_{ij}, \quad \{b_i, b_j\} = 0 = \{b_i^\dagger, b_j^\dagger\}$$

En particulier, $b_i^2 = 0 = b_i^{\dagger 2}$, une égalité qui reflète le principe d'exclusion de Pauli.

L'espace de Fock fermionique, noté \mathcal{F}_V^f , admet aussi plusieurs descriptions. La plus intuitive consiste à définir un état de vide $|0\rangle$ qui, par définition, est dans le noyau de tous les opérateurs annihilation, $b_i|0\rangle = 0$ pour tout i . Tous les autres états de l'espace de Fock s'obtiennent par actions successives des opérateurs créations, $b^\dagger(u_1) \cdots b^\dagger(u_n)|0\rangle$ avec $u_j \in V$. Pour que cette expression soit non nulle, tous les vecteurs u_j doivent être distincts, car $b(u)^2 = 0$. L'espace de Fock est un espace gradué :

$$\mathcal{F}_V^f = \bigoplus_{n=0}^D \mathcal{F}_V^{f(n)}, \quad \mathcal{F}_V^{f(n)} \equiv \text{Span}\{b^\dagger(u_1) \cdots b^\dagger(u_n)|0\rangle, \quad u_j \in V\}$$

Notez que la graduation s'arrête à $n = D$, en accord avec le principe de Pauli. L'espace $\mathcal{F}_V^{f(1)}$ est isomorphe à V et constitue par définition l'espace des états à une particule. L'espace $\mathcal{F}_V^{f(n)}$ est l'espace des états à n fermions. Les opérateurs création-annihilations agissent par multiplication à gauche. En pratique, on déplace les opérateurs annihilations vers la droite en utilisant les relations d'anti-commutation jusqu'à ce qu'il agissent sur le vide. La structure hilbertienne sur

4. L'anti-commutateur de deux opérateurs A et B est défini par $\{A, B\} = AB + BA$.

\mathcal{F}_V^f est définie en déclarant que $\langle 0|0\rangle = 1$ et que les opérateurs b_i^\dagger sont les conjugués hermitiques des $b(u)$. Par exemple,

$$\|b^\dagger(u)|0\rangle\|^2 = \langle 0|b(u^*)b^\dagger(u)|0\rangle = \|u\|^2.$$

Alternativement, $\mathcal{F}_V^f \cong \text{Poly}[b_i, b_j^\dagger]/\mathcal{I}[b_i]$ où $\text{Poly}[b_i, b_j^\dagger]$ est l'algèbre des polynômes en les b_i, b_j^\dagger , $i, j = 1, \dots, D$ et, où $\mathcal{I}[b_i]$ est l'idéal (à gauche) engendré par les b_i .

Comme les opérateurs créations anti-commutent, les états $b^\dagger(u_1) \cdots b^\dagger(u_n)|0\rangle$ sont anti-symétriques en u_1, \dots, u_n et ils s'identifient avec $u_1 \wedge \cdots \wedge u_n$. Ainsi, on a l'isomorphisme

$$\mathcal{F}_V^f = \bigoplus_{n=0}^D \mathcal{F}_V^{f(n)}, \quad \mathcal{F}_V^{f(n)} = \Lambda^{(n)} V,$$

où $\Lambda^{(n)} V$ est l'espace vectoriel engendré des produits extérieurs de n vecteurs de V . La correspondance entre cette description et la précédente est

$$\sqrt{n!} \mathbf{a} \cdot u_1 \otimes \cdots \otimes u_n \leftrightarrow b^\dagger(u_1) \cdots b^\dagger(u_n)|0\rangle,$$

où $\mathbf{a} = \frac{1}{n!} \sum_{\sigma \text{ perm.}} \epsilon_{\sigma} \sigma$ est l'opérateur d'anti-symétrisation. La dimension de l'espace de Fock fermionique \mathcal{F}_V^f est 2^D , celle de ses sous-espaces à n particules est :

$$\dim \mathcal{F}_V^{f(n)} = \frac{D!}{n!(D-n)!}.$$

Une autre description consiste à introduire des variables z_i , $i = 1, \dots, D$. L'espace de Fock est isomorphe à l'espace des formes de degré quelconque dans les dz_i . Les opérateurs créations sont représentés par les produits extérieurs $b_i^\dagger = dz_i$ et les opérateurs annihilations par les contractions $b_i = \mathbb{B}_{z_i}$. On vérifie que $[b_i, b_j^\dagger] = \delta_{ij}$.

Les opérateurs $N_i = b_i^\dagger b_i$ commutent entre eux, $[N_i, N_j] = 0$. Leurs valeurs propres sont 0 ou 1 car ce sont des projecteurs, $N_i^2 = N_i$. Leurs états propres sont les états à n particules car $[N_i, b_j^\dagger] = \delta_{ij} b_j^\dagger$ et $[N_i, b_j] = -\delta_{ij} b_j$. On appelle ces opérateurs les opérateurs 'nombres d'occupation'. Un état à une particule e_i peut être soit inoccupé ($N_i = 0$), soit occupé par un fermion et un seul ($N_i = 1$) en accord avec le principe de Pauli. La fonction partition $Z_V^f \equiv \text{Tr}_{\mathcal{F}_V^f}(q^{\sum_i u_i N_i})$ est la fonction génératrice des nombres d'occupation. Comme l'espace de Fock \mathcal{F}_V^f est isomorphe au produit tensoriel des espaces de Fock construits à l'aide des couples d'opérateurs création-annihilations b_i^\dagger, b_i associés à chacun des états à une particule e_i , on a :

$$Z_V^f \equiv \text{Tr}_{\mathcal{F}_V^f}(q^{\sum_i u_i N_i}) = \prod_{i=1}^D (1 + q^{u_i}).$$

Plus généralement, considérons l'ensemble des opérateurs bilinéaires dans les opérateurs créations ou annihilations :

$$b_i^\dagger b_j^\dagger, \quad b_i^\dagger b_j, \quad b_i b_j.$$

On vérifie que le commutateur de deux de ces opérateurs est encore un opérateur bilinéaire. Ils forment donc une représentation d'une algèbre de Lie, l'algèbre des transformations orthogonales $so(2D)$. Cette algèbre est de dimension $D(2D - 1)$. Tous ces opérateurs préservent la parité du nombre d'occupation total $N = \sum_i N_i$. L'espace de Fock \mathcal{F}_V^f se décompose donc en deux

représentations (chacune de dimension 2^{D-1}), isomorphes aux représentations spinorielles de $so(2D)$.

Les opérateurs bilinéaires qui commutent avec le nombre d'occupation total, c'est-à-dire les opérateurs de la forme $b_i^\dagger b_j$, $i, j = 1, \dots, D$, engendrent une sous-algèbre de $so(2D)$ isomorphe à l'algèbre $u(D)$ des transformations unitaires sur V . Chaque sous espace $\mathcal{F}_V^{f(n)}$ à nombre de particules fixé constitue une représentation (irréductible) de $u(D)$ car ces opérateurs commutent avec $N = \sum_i N_i$.

Enfin, définissons $2D$ matrices, appelées " matrices gamma " par :

$$\gamma_j = \frac{1}{\sqrt{2}}(b_j + b_j^\dagger), \quad \gamma_{n+j} = \frac{1}{i\sqrt{2}}(b_j - b_j^\dagger)$$

Elles sont hermitiennes et satisfont les relations d'anti-commutations,

$$\{\gamma_i, \gamma_j\} = \delta_{i;j}, \quad \{\gamma_i, \gamma_{n+j}\} = 0, \quad \{\gamma_{n+i}, \gamma_{n+j}\} = \delta_{i;j},$$

relations qui sont constitutives de l'algèbre de Clifford sur \mathbb{R}^{2D} . On laisse l'étude de la théorie des algèbres de Clifford à la lectrice motivée, au lecteur motivé.

CHAPITRE VI

DU MODÈLE DES QUARKS AU MODÈLE STANDARD

Ce chapitre pourrait aussi s'appeler "De la saveur à la couleur". Il a pour but d'introduire le modèle des quarks, fondé sur la symétrie de saveur, ainsi que de décrire l'algèbre de Lie $su(3)$ et ses représentations.

Le modèle des quarks, dû à M. Gell-Mann, trouvent son origine dans les années 60 dans la recherche d'un principe sous-jacent à l'organisation des particules hadroniques. C'est un peu de l'histoire ancienne mais la description de cette recherche offre l'opportunité d'introduire la notion de $su(3)$ -saveur comme symétrie approchée des interactions fortes ainsi que celles de quark et de couleur.

Les hadrons, observés dans les années 60, se classent en deux grandes catégories :

- les mésons, dont les masses sont de l'ordre 200 MeV et,
- les baryons, plus lourds, ayant des masses de l'ordre de 1000 MeV.

Ces particules semblaient se regrouper en multiplets formés de particules de masses avoisinantes et de même spin. On a en particulier observé⁽¹⁾ :

- Un octet de mésons de spin 0 formé de : $\pi^\pm, \pi^0, \eta, K^\pm, K^0, \bar{K}^0$.
- Un octet de baryons de spin 1/2 constitué de : $p, n, \Sigma^\pm, \Sigma^0, \Lambda, \Xi^0, \Xi^-$;

ainsi qu'un autre octet de mésons de spin 1 et un décuplet (incomplet) de baryons de spin 3/2.

Il n'est pas utile à ce stade de faire une liste plus exhaustive des particules hadroniques : le modèle des quarks en fournira une vision synthétique.

VI.1. L'isospin et l'algèbre $su(2)$

Comme souvent en physique, on cherche à interpréter la dégénérescence en masse des particules appartenant à ces multiplets comme la conséquence d'une symétrie (cachée). Les masses n'étant pas exactement identiques cette symétrie ne sera qu'une symétrie approchée.

Le premier pas dans cette direction a été fait par Heisenberg dès les années 30 avec l'introduction de la notion d'isospin afin de rendre compte de la dégénérescence presque exacte entre les

1. Les indices \pm et 0 font référence à la charge électrique des particules

masses du proton p et du neutron n . Le proton et le neutron sont alors vus comme deux états de la même entité distincts par leur charge d'isospin.

La symétrie d'isospin est fondée sur l'algèbre de Lie $su(2)$ que nous avons déjà rencontrée à plusieurs reprises. La symétrie d'isospin est le premier exemple de symétrie interne non-abélienne qui ne soit pas rattachée à une symétrie de l'espace ambiant.

Rappelons que l'algèbre de Lie $su(2)$, isomorphe à $so(3)$, est l'algèbre des matrices hermitiennes 2×2 de trace nulle. Elle est de dimension 3. Une base est fournie par les matrices de Pauli σ_j . Soit $T_j \equiv \sigma_j/2$. Les crochets de Lie (qui, dans ce cas, sont les commutateurs entre les matrices) sont ⁽²⁾ :

$$[T_j, T_k] = i\epsilon_{jkl} T_l,$$

où ϵ_{jkl} est le tenseur d'ordre 3 totalement anti-symétrique, $\epsilon_{123} = 1$. Il est utile de définir $T_{\pm} = T_1 \pm iT_2$ de sorte que

$$[T_3, T_{\pm}] = \pm T_{\pm}, \quad [T_+, T_-] = 2T_3.$$

Les représentations irréductibles unitaires de $su(2)$ sont indexées par un demi-entier, $I \in \mathbb{N}/2$, l'(iso)spin. La représentation d'isospin I est définie sur l'espace \mathbb{C}^{2I+1} sur lequel les générateurs T_j sont représentés par des opérateurs (des matrices) hermitiennes. En particulier, l'opérateur $T_3 = T_3^{\dagger}$ est diagonalisable : ses valeurs propres sont demi-entières et diffèrent de I par un entier. On note $|I; t\rangle$ la base propre diagonalisant la représentation de T_3 sur \mathbb{C}^{2I+1} :

$$T_3 |I; t\rangle = t |I; t\rangle, \quad t = -I, \dots, I.$$

Les relations de commutation $[T_3, T_{\pm}] = \pm T_{\pm}$ impliquent que T_{\pm} transforment un état propre de T_3 en un état qui est, soit nul, soit état propre de T_3 mais avec une valeur propre qui diffère de la valeur propre initiale par ± 1 . La représentation étant de dimension finie, il existe un état de plus grande valeur propre qui est donc aussi annihilé par l'opérateur T_+ . Cet état est $|I; I\rangle$:

$$T_3 |I; I\rangle = I |I; I\rangle, \quad T_+ |I; I\rangle = 0.$$

Les autres états propres de T_3 s'obtiennent par action récursive de T_- sur cet état :

$$T_-^n |I; I\rangle, \quad n = 0, 1, \dots, 2I.$$

On vérifie que $T_-^{2I+1} |I; I\rangle = 0$ dans la représentation de dimension $2I+1$, de sorte que l'ensemble des états $T_-^n |I; I\rangle$ forme une base de cette représentation.

Retournons à la notion d'isospin. L'algèbre de symétrie est $su(2)$. L'opérateur T_3 est par définition l'opérateur de "charge d'isospin". Cette algèbre est supposée, puis vérifiée expérimentalement, être une symétrie (approchée) des interactions fortes. Les particules s'organisent en multiplets, représentations de l'algèbre de symétrie $su(2)$. Ainsi :

— Le proton et le neutron forment un doublet (p, n) d'isospin ($I = 1/2$). Le proton est un état d'isospin $+1/2$ et le neutron est un état d'isospin $-1/2$. Le proton et le neutron sont ainsi vus comme différents états d'une même structure quantique mais de charge d'isospin différente.

— Les pions (π^-, π^0, π^+) forment un triplet d'isospin ($I = 1$). Ce sont les états de charge d'isospin 0 ou ± 1 d'une représentation de dimension 3.

2. Nous prenons ici la convention des physiciens en introduisant le facteur "i" dans les relations de commutation.

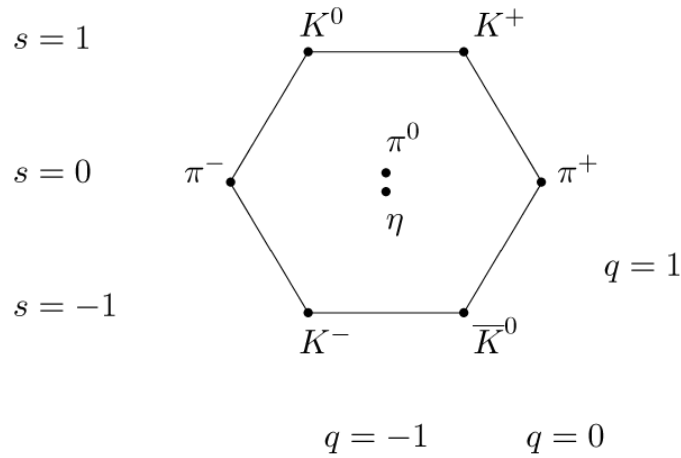


FIGURE 1. Octet de mésons.

Notons que le fait de regrouper les particules en multiplets de dimension supérieure à 1 impose que la symétrie soit non-abelienne (car les représentations d'une algèbre de Lie abélienne sont toutes de dimension 1).

Outre le fait que les masses des particules appartenant à un même multiplet soient (presque) égales, l'existence d'une symétrie, ou la validité de l'hypothèse portant sur l'existence de cette symétrie, peut être testée en étudiant les réactions entre les particules. En effet, l'existence de symétrie implique l'existence de relations entre les amplitudes de probabilité associées à ces réactions. La vérification de ces relations permet de tester l'hypothèse de symétrie.

L'étude de ces réactions permet aussi de déterminer les charges d'isospin des particules et la structure des multiplets qu'elles forment. On trouve ainsi pour les particules de l'octet de baryons :

— un doublet d'isospin (p, n) , un triplet $(\Sigma^-, \Sigma^0, \Sigma^+)$, un singulet Λ , et un second doublet (Ξ^-, Ξ^0) , c.f. Figure (2),

et pour les mésons :

— un doublet d'isospin (K^0, K^+) , un triplet (π^-, π^0, π^+) , un singulet η , et un second doublet (K^-, \bar{K}^0) , c.f. Figure (1).

Notez que les particules et leurs anti-particules ont des charges d'isospin opposées.

VI.2. Symétrie de saveur et l'algèbre $su(3)$

La symétrie d'isospin $su(2)$ ne permet pas d'interpréter les octets de mésons ou de baryons comme des multiplets irréductibles. Il faut pour cela une algèbre de symétrie de dimension plus grande que celle de $su(2)$. Outre la charge d'isospin T_3 , la charge électrique Q_e est conservée par les interactions fortes. Ces deux charges, qui peuvent être mesurées indépendamment, sont représentées sur l'espace de Hilbert des états physiques par des opérateurs hermitiens commutants.

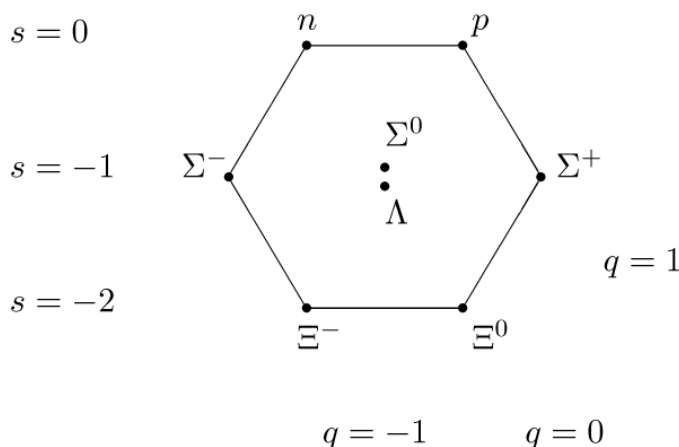


FIGURE 2. Octet de baryons.

Ainsi, si nous recherchons une algèbre de symétrie englobant la charge électrique et celle d'isospin, il nous faut trouver une algèbre de Lie possédant une sous-algèbre abélienne de dimension au moins égale à deux.

VI.2.1. L'algèbre $su(3)$. — L'algèbre $su(3)$ est la plus simple des algèbres de Lie répondant à ces critères. C'est l'algèbre de Lie des matrices hermitiennes 3×3 de trace nulle. Le crochet de Lie est le commutateur, $[X, Y] = XY - YX$, qui satisfait à l'identité de Jacobi :

$$[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0.$$

On définit l'action adjointe d'un élément $X \in su(3)$ sur les vecteurs de $su(3)$ par :

$$(\text{ad}X) \cdot Y = [X, Y], \quad \text{pour tout } Y \in su(3).$$

L'identité de Jacobi assure que cette action définit une représentation de $su(3)$ sur elle-même. En effet, on a : $(\text{ad}X)(\text{ad}Y) - (\text{ad}Y)(\text{ad}X) = (\text{ad}[X, Y])$.

L'algèbre $su(3)$ est de dimension 8. Une base est fournie par les matrices X_{ij} , Y_{ij} et H_{ij} , définies par :

$$X_{ij} = \frac{1}{2}(E_{ij} + E_{ji}), \quad Y_{ij} = \frac{1}{2i}(E_{ij} - E_{ji}), \quad H_{ij} = \frac{1}{2}(E_{ii} - E_{jj}),$$

où $E_{ij} = |i\rangle\langle j|$ sont les matrices canoniques. Les matrices H_{ij} ne sont pas indépendantes car $H_{12} + H_{23} + H_{31} = 0$. Pour former une base de $su(3)$ il faut choisir deux matrices parmi ces trois. L'algèbre $su(2)$ est plongée dans $su(3)$ de différentes manières : les trois sous-algèbres formées des éléments suivants,

$$(X_{12}, Y_{12}, H_{12}), \quad (X_{23}, Y_{23}, H_{23}), \quad (X_{31}, Y_{31}, H_{31}),$$

sont chacune isomorphe à $su(2)$. Afin d'extraire plus simplement la structure de cette algèbre on pose :

$$\begin{aligned} T_3 &\equiv H_{12} \quad , \quad T_{\pm} \equiv X_{12} \pm iY_{12} \\ S_3 &\equiv H_{23} \quad , \quad S_{\pm} \equiv X_{23} \pm iY_{23} \\ R_3 &\equiv H_{31} \quad , \quad R_{\pm} \equiv X_{31} \pm iY_{31} \end{aligned}$$

Ces matrices sont les analogues des matrices de Pauli. On vérifie que $[T_3, T_{\pm}] = \pm T_{\pm}$ et $[T_+, T_-] = 2T_3$. Des relations similaires s'appliquent à S_3, S_{\pm} et à R_3, R_{\pm} . On a bien-sûr $T_3 + S_3 + R_3 = 0$.

Cherchons à décrire les relations de commutation entre ces générateurs. On pourrait évidemment les calculer explicitement mais nous allons procéder différemment. Les générateurs T_3, S_3 et R_3 , forment une sous-algèbre abélienne, on peut donc diagonaliser simultanément les trois matrices représentant leurs actions adjointes ; i.e. les trois applications linéaires adT_3, adS_3 et adR_3 commutent et sont diagonalisables simultanément. Si $X \in su(3)$ est un vecteur propre de ces trois actions adjointes, avec pour valeurs propres (a, b, c) , on pose :

$$[T_3, X] = aX, \quad [S_3, X] = bX, \quad [R_3, X] = cX$$

Evidemment $a + b + c = 0$ puisque $T_3 + S_3 + R_3 = 0$. Les valeurs propres (a, b, c) appartiennent à un plan. Par construction toute combinaison linéaire de T_3, S_3, R_3 est vecteur propre de valeurs propres nulles. De plus on vérifie que (trouver un argument justifiant ces relations sans calcul!) :

$$\begin{aligned} [T_3, T_{\pm}] &= \pm T_{\pm} \quad , \quad [S_3, T_{\pm}] = \mp \frac{1}{2} T_{\pm}, \quad [R_3, T_{\pm}] = \mp \frac{1}{2} T_{\pm} \\ [T_3, S_{\pm}] &= \mp \frac{1}{2} S_{\pm} \quad , \quad [S_3, S_{\pm}] = \pm S_{\pm}, \quad [R_3, S_{\pm}] = \mp \frac{1}{2} S_{\pm} \\ [T_3, R_{\pm}] &= \mp \frac{1}{2} R_{\pm} \quad , \quad [S_3, R_{\pm}] = \mp \frac{1}{2} R_{\pm}, \quad [R_3, R_{\pm}] = \pm R_{\pm} \end{aligned}$$

Ainsi une base de $su(3)$ est formée de deux générateurs parmi T_3, S_3, R_3 et des six autres générateurs vecteurs propres T_{\pm}, S_{\pm} et R_{\pm} . Notez que les valeurs propres non nulles sont non-dégénérées.

Nous allons maintenant coder ces relations de commutation dans un diagramme inscrit dans le plan. Les points de ce diagramme sont en correspondance avec les générateurs de $su(3)$ vecteurs propres des applications adT_3, adS_3 et adR_3 . Afin de préserver la symétrie cyclique, introduisons trois vecteurs planaires $\vec{e}_t, \vec{e}_s, \vec{e}_r$ de norme 1 et d'angle relative $2\pi/3$. Ainsi,

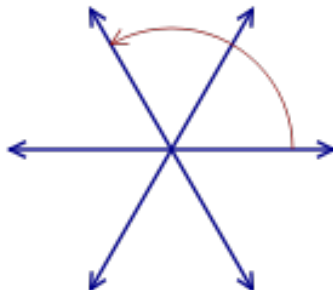
$$\vec{e}_t + \vec{e}_s + \vec{e}_r = 0, \quad \text{et} \quad \vec{e}_t \cdot \vec{e}_t = 1, \quad \vec{e}_t \cdot \vec{e}_s = \vec{e}_t \cdot \vec{e}_r = -1/2.$$

Au générateur X de valeurs propres (a, b, c) on associe le vecteur \vec{x} tel que $a = \vec{x} \cdot \vec{e}_t, b = \vec{x} \cdot \vec{e}_s$ et $c = \vec{x} \cdot \vec{e}_r$. On a donc les associations :

$$T_{\pm} \rightarrow \pm \vec{e}_t, \quad S_{\pm} \rightarrow \pm \vec{e}_s, \quad R_{\pm} \rightarrow \pm \vec{e}_r.$$

Ces vecteurs forment une figure appelée diagramme des racines. Notez la symétrie d'ordre 3 de cette figure. Voir Figure (3).

Le diagramme des racines codent aussi une information importante sur les relations de commutation entre les générateurs vecteurs propres, i.e. entre T_{\pm}, S_{\pm} et R_{\pm} . En effet, l'identité de Jacobi assure que si X et Y sont des vecteurs propres des actions adjointes adT_3, adS_3 et adR_3 , de valeurs propres \vec{x} et \vec{y} , alors le commutateur $[X, Y]$ est, soit nul, soit vecteur propre de ces

FIGURE 3. Vecteurs racines de $su(3)$.

actions adjointes avec pour valeur propre $\vec{x} + \vec{y}$. En particulier le commutateur $[X, Y]$ sera nul si le vecteur $\vec{x} + \vec{y}$ n'est pas sur le diagramme des racines. Par contre, si $\vec{x} + \vec{y}$ est sur le diagramme, alors $[X, Y]$ est proportionnel au générateur associée à ce vecteur (le coefficient de proportionnalité n'est pas codé dans le diagramme). Par exemple :

$$[T_+, R_-] = 0, \quad [T_+, S_+] = R_-, \quad [T_+, S_-] = 0, \quad \text{etc} \dots$$

Bien appréhender ce diagramme et sa signification est un exercice recommandé au lecteur.

Notez que cette description de $su(3)$ a consisté à extraire un ensemble de sous-algèbres $su(2)$ puis à décrire et coder comment ces sous-algèbres sont intriquées les unes parmi les autres. Cette approche se généralise à l'étude de toutes les algèbres de Lie semi-simples.

VI.2.2. Représentations de $su(3)$. — Les représentations de $su(3)$ se construisent de façon assez similaire à celles de $su(2)$ par actions successives de générateurs sur un vecteur particulier, appelé vecteur de plus haut poids.

Dans toute représentation unitaire, les générateurs T_3, S_3, R_3 sont représentés par des opérateurs hermitiens qui commutent entre eux. Ils peuvent donc être diagonalisés simultanément. On note t, s, r , avec $t + s + r = 0$, les valeurs propres respectives. Les relations de commutation entre T_3, S_3, R_3 et T_{\pm}, S_{\pm} et R_{\pm} impliquent que, soit ces derniers annihilent un vecteur propre de T_3, S_3, R_3 , soit ils le transforment en un autre vecteur propre mais en modifiant les valeurs propres.

Comme nous l'avons fait pour la construction du diagramme des racines, à toutes valeurs propres (t, s, r) , avec $t + s + r = 0$, de T_3, S_3, R_3 , nous associons un vecteur $\vec{\mu}$, inscrit sur le plan du diagramme des racines, tel que

$$t = \vec{\mu} \cdot \vec{e}_t, \quad s = \vec{\mu} \cdot \vec{e}_s, \quad r = \vec{\mu} \cdot \vec{e}_r.$$

Les vecteurs $\vec{\mu}$ sont appelés des poids. L'ensemble des poids d'une représentation dépend de celle-ci : on l'appelle le diagramme des poids de la représentation.

Toute représentation de dimension finie de $su(3)$ se décompose en somme de représentations de l'algèbre $su(2)$ engendrée par les générateurs (T_3, T_+, T_-) . Chacune de ces représentations possèdent un vecteur annihilé par T_+ qui est vecteur propre de T_3 de valeur propre maximale. On considère l'espace vectoriel engendré par l'ensemble des vecteurs annihilés par T_+ , ainsi qu'un sous-espace de cet espace formé des vecteurs annihilés à la fois par T_+ et par R_- , soit $\mathcal{O} := \{v, T_+ \cdot v = 0 = R_- \cdot v\}$. L'opérateur T_3 agit sur \mathcal{O} et est diagonalisable sur cet espace. Nous devons nous assurer que cet espace n'est pas trivial. Comme $[T_+, R_-] = 0$, l'opérateur R_- agit sur l'espace des états annihilés par T_+ . Comme $[T_3, R_-] = \frac{1}{2}R_-$, soit il transforme un vecteur propre de T_3 en un autre vecteur propre de T_3 mais de valeur propre strictement plus grande, soit il annihile ce vecteur propre de T_3 . Comme la représentation est de dimension finie non nulle, ceci implique qu'en effet \mathcal{O} est de dimension non nulle. Montrons maintenant que l'espace \mathcal{O} forme une représentation de l'algèbre $su(2)$ engendrée par (S_3, S_+, S_-) . Pour cela il suffit de montrer que chacun des générateurs (S_3, S_+, S_-) agit sur cet espace vectoriel. Puisque $[S_3, T_+] = -\frac{1}{2}T_+$ et $[S_3, R_-] = \frac{1}{2}R_-$, il est clair que S_3 agit sur cet espace. De même, comme $[R_-, S_+] = 0$ et $[T_+, S_+] = R_-$, on en déduit que S_+ agit sur \mathcal{O} . Enfin, les relations $[S_-, T_+] = 0$ et $[S_-, R_-] = T_+$ impliquent que S_- agit aussi sur cet espace qui forme donc une représentation de dimension finie de l'algèbre $su(2)$ engendrée par (S_3, S_+, S_-) . Il existe donc parmi ces vecteurs un vecteur (on montre qu'il est unique) qui est aussi annihilé par S_+ . Cet état est appelé vecteur de plus haut poids et on le note $|t_m; s_m\rangle$. Par définition, il est vecteur propre de T_3 et S_3 et est annihilé par T_+ et S_+ :

$$\begin{aligned} T_3|t_m; s_m\rangle &= t_m |t_m; s_m\rangle, & T_+|t_m; s_m\rangle &= 0; \\ S_3|t_m; s_m\rangle &= s_m |t_m; s_m\rangle, & S_+|t_m; s_m\rangle &= 0. \end{aligned}$$

Si la représentation est une représentation unitaire de dimension finie de $su(3)$, c'est aussi une représentation unitaire de dimension finie de chacune de ces sous-algèbres $su(2)$. Donc, les valeurs propres maximales t_m et s_m sont des demi-entiers :

$$t_m = p/2, \quad s_m = q/2, \quad p, q \in \mathbb{N}$$

D'après la théorie des représentations de $su(2)$ on sait alors que $T_-^{2t_m+1}|t_m; s_m\rangle = 0$ et $S_-^{2s_m+1}|t_m; s_m\rangle = 0$. On obtient tous les états de la représentation par actions successives des générateurs T_-, S_- et R_- sur le vecteur de plus haut poids en tenant compte des relations de commutation et de ces deux dernières relations. On montre que l'on construit ainsi un espace vectoriel de dimension finie et donc une représentation de dimension finie. Nous illustrons cette construction sur des exemples ci-dessous.

Toutes les représentations de dimension finie de $su(3)$ s'obtiennent ainsi. Elles sont donc indexées par deux entiers. On note $D(p, q)$ la représentation de $su(3)$ de plus haut poids $t_m = p/2$ et $s_m = q/2$.

Illustrons cette construction sur quelques exemples :

- $D(1, 0)$: $t_m = 1/2$ et $s_m = 0$.

On construit une représentation de dimension trois, notée [3]. C'est la représentation vectorielle de définition.

- $D(0, 1) : t_m = 0$ et $s_m = 1/2$.

On construit une représentation qui est aussi de dimension trois, notée $[\bar{3}]$. C'est la représentation conjuguée de la représentation vectorielle.

- $D(1, 1) : t_m = 1/2$ et $s_m = 1/2$.

On construit alors une représentation de dimension huit, notée $[8]$. Elle est équivalente à la représentation adjointe.

Nous pourrions continuer ainsi de suite... Par la suite, nous aurons besoin de la représentation $D(3, 0) : elle est de dimension dix et notée $[10]$ (construire son diagramme des poids!).$

VI.3. Le modèle des quarks

Nous allons maintenant utiliser la théorie des représentations de $su(3)$ afin de construire le modèle des quarks dans lequel on interprète les hadrons comme des états liés de quarks q et d'anti-quarks \bar{q} . Les mésons sont des états quark/anti-quark et les baryons des états à trois quarks :

$$\text{mesons : } q\bar{q}; \quad \text{baryons : } qqq$$

Rappelons que nous avons déjà introduit la notion d'isospin dont l'algèbre de symétrie est $su(2)$. Nous identifions les générateurs de la symétrie d'isospin avec les générateurs T_3, T_{\pm} de $su(3)$. Cette algèbre de symétrie $su(3)$ s'appelle la symétrie de saveur. Outre l'isospin t , les interactions fortes conservent la charge électrique q_e . On définit l'hypercharge (forte), notée y , par :

$$q_e = t + y/2.$$

Les charges électrique, d'isospin et d'hypercharge sont représentées sur l'espace des états physiques par des opérateurs hermitiens commutant entre eux (car on peut mesurer indépendamment ces trois charges). Comme l'isospin est identifié avec le générateur T_3 , il nous faut trouver à quelle combinaison linéaire des générateurs T_3, S_3, R_3 l'hypercharge (ou la charge électrique) correspond. L'histoire de la physique des particules nous apprend que la bonne identification est :

$$Y = \frac{2}{3}(T_3 + 2S_3)$$

Comme nous avons associé des vecteurs \vec{e}_t et \vec{e}_s à la mesure des valeurs propres de T_3 et S_3 (par produit scalaire avec le poids), nous pouvons, par combinaison linéaire, définir un vecteur \vec{e}_y associé à la valeur de l'hypercharge : $\vec{e}_y = \frac{2}{3}(\vec{e}_t + 2\vec{e}_s)$. On vérifie que $\vec{e}_y \cdot \vec{e}_t = 0$ et $\vec{e}_y \cdot \vec{e}_y = 2/3$. Le vecteur \vec{e}_y est orthogonal à \vec{e}_t .

- Quarks, $q \in [3]$, et anti-quarks, $\bar{q} \in [\bar{3}]$:

Les quarks sont les éléments de la représentation $[3]$ de $su(3)$. Il y a donc trois quarks de différentes saveurs. On les note $u = \text{"up"}$, $d = \text{"down"}$ et $s = \text{"strange"}$. Les états u, d, s sont les états d'isospin et d'hypercharge bien définis, i.e. ce sont les états propres des opérateurs T_3 et Y , donc

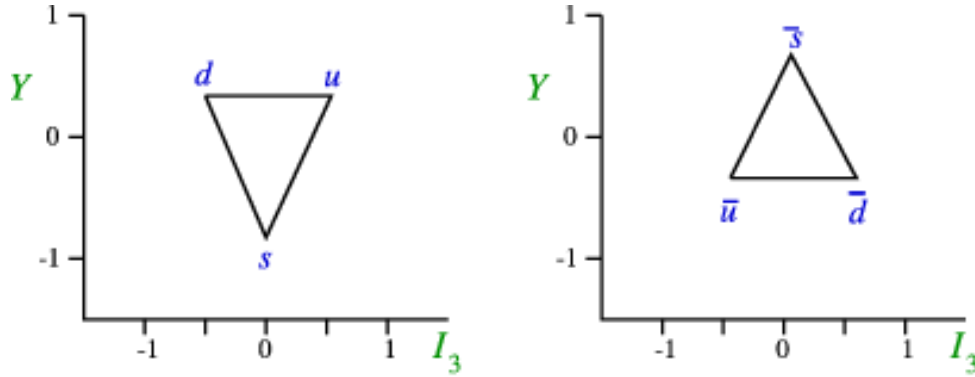


FIGURE 4. Quark and anti-quarks.

des opérateurs T_3 et S_3 dans la représentation $[3]$. Leurs nombres quantiques sont les suivants :

	y	t	q_e
u	$1/3$	$1/2$	$2/3$
d	$1/3$	$-1/2$	$-1/3$
s	$-2/3$	0	$-1/3$

Les anti-quarks sont les éléments de la représentation conjuguée $[\bar{3}]$. On les note \bar{u} , \bar{d} et \bar{s} . Ils ont les nombres quantiques opposés aux quarks u, d, s . Voir Figure (4).

Les quarks ont spin $1/2$ (au sens du groupe de Poincaré). Ils sont (presque) dégénérés en masse : $m_u \simeq 2 \pm 0.5 \text{ MeV}$, $m_d \simeq 5 \pm 1.5 \text{ MeV}$ et $m_s \simeq 95 \pm 25 \text{ MeV}$. Si la symétrie de saveur était une symétrie exacte ils auraient la même masse. Le fait que le quark s a une masse supérieure à celle des quarks d et u exprime le fait que la symétrie de saveur n'est qu'une symétrie approchée. Ce n'est pas une symétrie exacte des interactions fortes.

- Mésons, $q\bar{q} \in [3] \otimes [\bar{3}]$:

Les états liés $q\bar{q}$ sont des éléments du produit tensoriel des représentations $[3]$ et $[\bar{3}]$. Il nous faut donc décomposer ce produit tensoriel en représentations irréductibles. On a (expliquer pourquoi) :

$$[3] \otimes [\bar{3}] = [1] \oplus [8].$$

La représentation $[1]$ est la représentation triviale formée du singulet $u\bar{u} + d\bar{d} + s\bar{s}$ (expliquer pourquoi!). La représentation $[8]$ est la représentation adjointe formée des états ($u\bar{s}, d\bar{s}, u\bar{d}, d\bar{u}, s\bar{d}, s\bar{u}$) de charge non nulle et des états de charge nulle combinaisons linéaires de ($u\bar{u}, d\bar{d}, s\bar{s}$) orthogonales au singulet.

Le spin de ces états découle de celui des quarks et anti-quarks. Les quarks ayant spin $j = 1/2$, les mésons sont des éléments du produit tensoriel $[j = 1/2]_{\text{spin}} \otimes [j = 1/2]_{\text{spin}}$ de représentations de $su(2)$ qui se décompose en la somme d'un triplet de spin 1 et d'un singulet de spin 0 :

$$[j = 1/2]_{\text{spin}} \otimes [j = 1/2]_{\text{spin}} \cong [j = 1]_{\text{spin}} \oplus [j = 0]_{\text{spin}}.$$

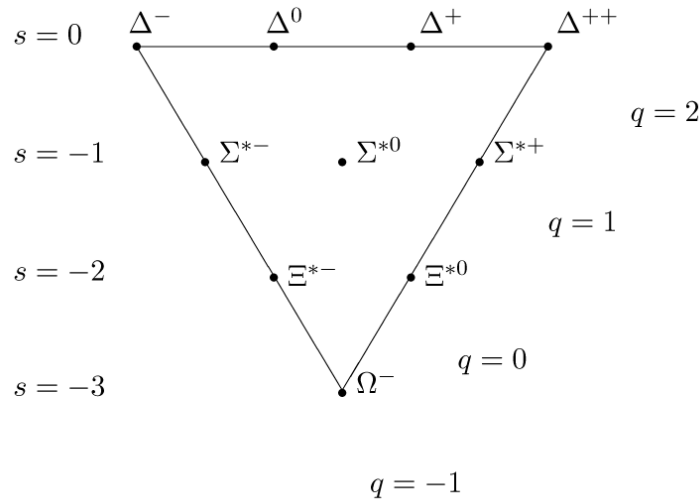


FIGURE 5. Décuplet de baryons.

Ainsi les états liés $q\bar{q}$ forment :

- un octet [8] de mésons de spin 0 ;
- un singulet [1] de spin 0, ce méson était une prédiction du modèle des quarks ;
- un octet [8] de mésons de spin 1 ;
- un singulet [1] mésonique de spin 1.

Ces multiplets de mésons sont en accord avec les particules observées. La différence de masse entre les particules est correctement estimée en tenant compte de la différence de masse entre le quark s et les quarks u et d .

Nous laissons au lecteur l'exercice consistant à décrire explicitement la composition de chacun des mésons des octets de spin 0 ou 1 en termes des états de quarks et anti-quarks.

- Baryons, $qqq \in [3] \otimes [3] \otimes [3]$:

Pour comprendre la structure en saveur des états à trois quarks, nous devons décomposer le produit tensoriel $[3] \otimes [3] \otimes [3]$ en représentations irréductibles de $su(3)$. Le résultat est (expliquer pourquoi!) :

$$[3] \otimes [3] \otimes [3] = [10] \oplus ([8] \oplus [8]) \oplus [1]$$

Dans les notations introduites ci-dessus, la représentation [10] est la représentation $D(3, 0)$. (Donner un argument simple impliquant que celle-ci est présente dans le produit tensoriel $[3] \otimes [3] \otimes [3]$!).

Le modèle des baryons comme états liés de trois quarks prédit donc un décuplet, i.e. un multiplet de dimension 10. Cette prédiction est compatible avec les observations qui ont conduit à mettre évidence ce décuplet dont les particules ont un spin $3/2$ (un fait lui aussi compatible avec le spin $1/2$ des quarks). La différence de masse entre les baryons de ce décuplet est aussi correctement expliquée par la différence de masse du quark s . Nous laissons encore une fois au lecteur l'exercice consistant à décrire explicitement la composition de chacun des baryons en termes de quarks. Voir Figure (5).

Lorsque le modèle des quarks fût proposé, toutes les particules de ce multiplet n'avaient pas été observées. Il en manquait une, le Ω^- portant prédite par ce modèle. Elle fût ultérieurement observée, avec la masse prédite par le modèle des quarks.

Un octet de baryons de spin 1/2, formé des particules $p, n, \Sigma^\pm, \Sigma^0, \Lambda, \Xi^0, \Xi^-$, a aussi été observé. Là encore, l'écart entre les masses est bien expliquée par la différence de masse du quark s . Ces observations semblent donc être en accord avec le modèle des baryons comme états à trois quarks et la décomposition en représentations irréductibles de ces états.

• *La couleur :*

Tout semble bien fonctionner avec le modèle des quarks... mais la situation est en fait plus compliquée :

- On n'observe pas de baryon singulet de saveur (portant prédit par la décomposition des états à trois quarks) ;
- Un seul octet de baryons a été observé (alors que la décomposition des états à trois quarks en prédit deux) ;
- Les états du décuplet sont symétriques en l'échange des quarks en contradiction avec le principe de Pauli et l'hypothèse que les quarks sont des fermions de spin 1/2 (par exemple, la particule Δ^{++} est l'état à trois quarks uuu).

Ces difficultés sont résolues par l'introduction de deux nouveaux concepts :

- (i) la notion de charge de couleur : outre leurs charges électrique et de saveur les quarks portent un autre degré de liberté appelé couleur associé lui aussi à l'algèbre $su(3)$. Donc deux copies de l'algèbre $su(3)$, l'algèbre $su(3)$ -couleur et l'algèbre $su(3)$ -saveur, sont en jeu. Par hypothèse, chacun des quarks u, d, s appartient à la représentation $[3]$ de $su(3)$ -couleur et les anti-quarks $\bar{u}, \bar{d}, \bar{s}$ à la représentation conjuguée $[\bar{3}]$:

$$q^\alpha \in [3]_{\text{couleur}}, \quad \bar{q}^{\bar{\alpha}} \in [\bar{3}]_{\text{couleur}}, \quad \alpha, \bar{\alpha} = 1, 2, 3, \quad q = u, d, s.$$

- (ii) l'hypothèse de confinement qui stipule que seuls les états singulets de couleur sont observables. Cette hypothèse est supposée être une conséquence de la chromodynamique quantique, théorie qui décrit les interactions fortes.

Ainsi, le contenu en couleur des mésons et des baryons est :

$$\text{mesons} = \delta_{\alpha\bar{\alpha}} q^\alpha \bar{q}^{\bar{\alpha}}, \quad \text{baryons} = \epsilon_{\alpha\beta\sigma} q^\alpha q^\beta q^\sigma,$$

où $\epsilon_{\alpha\beta\sigma}$ est le tenseur d'ordre trois totalement anti-symétrique (expliquer pourquoi!).

Le principe de Pauli n'intervient pas dans la construction des mésons car les quarks et anti-quarks ne sont pas des particules identiques. Par contre il intervient dans celle des baryons formés de trois quarks. Ce principe dicte alors que la fonction d'onde des trois quarks doit être totalement anti-symétrique sous l'échange de ces particules. En faisant l'hypothèse que la fonction d'onde est symétrique dans l'échange des coordonnées spatiales (hypothèse raisonnable si on suppose que ces baryons correspondent aux états fondamentaux à trois quarks), la fonction d'onde incluant les degrés de liberté de saveur, de spin et de couleur doit être anti-symétrique. Cette fonction d'onde appartient au produit tensoriel :

$$[3]_{\text{saveur}}^{\otimes 3} \otimes [j = 1/2]_{\text{spin}}^{\otimes 3} \otimes [3]_{\text{couleur}}^{\otimes 3}.$$

		fermions (3 générations de la matière)			bosons (forces)	
		I	II	III		
masse →		2.4 MeV	1.27 GeV	171.2 GeV	0	électromagnétisme interaction forte
charge →		$\frac{2}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	0	
spin →		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
nom →		u up	c charm	t top	γ photon	
	Quarks					
		4.8 MeV	104 MeV	4.2 GeV	0	interaction forte
		$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	0	
		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
		d down	s strange	b bottom	g gluon	
	Leptons					
		<2.2 eV	<0.17 MeV	<15.5 MeV	0	interaction faible
		0	0	0	0	
		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
		ν_e neutrino électronique	ν_μ neutrino muonique	ν_τ neutrino tauique	Z^0 boson Z ⁰	
		0.511 MeV	105.7 MeV	1.777 GeV	80.4 GeV	interaction faible
		-1	-1	-1	± 1	
		$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	1	
		e électron	μ muon	τ tau	W^\pm boson W	

FIGURE 6. Particules élémentaires du modèle standard.

L'hypothèse de confinement force l'anti-symétrie de la partie associée aux degrés de liberté de couleur. La partie associée aux degrés de liberté de saveur et de spin doit donc être symétrique. On vérifie facilement que cela est effectivement le cas pour le décuplet de spin $3/2$. La lectrice motivée, le lecteur motivé, vérifiera qu'une étude plus fine de l'action du groupe des permutations sur les produits tensoriels $[j = 1/2]_{\text{spin}}^{\otimes 3}$ et $[3]_{\text{saveur}}^{\otimes 3}$ montre que les états totalement symétriques du produit $[j = 1/2]_{\text{spin}}^{\otimes 3} \otimes [3]_{\text{saveur}}^{\otimes 3}$ forment un octet de saveur de spin $1/2$, et un seul, en accord avec l'expérience !

VI.4. Eléments du modèle standard

Les quarks sont des éléments fondamentaux du modèle standard et la couleur est l'analogue de la charge électrique pour les interactions fortes. Il n'est pas question ici de décrire entièrement le modèle standard des constituants élémentaires et de leurs interactions. Nous en présentons simplement les éléments essentiels.

La matière apparaît sous deux formes : leptoniques ou baryoniques. Les constituants élémentaires de la matière sont des fermions qui forment trois familles – on dit trois générations – aux propriétés similaires mais distinctes en particulier par les masses de leurs constituants. Les particules de la première famille sont plus légères que celles de la seconde, qui sont elles-mêmes plus légères que celles de la troisième. Décrivons la première famille pour simplifier. Les particules leptoniques sont l'électron et le neutrino, les particules baryoniques sont les quarks u et d . Outre leur masse et leur spin, chacune de ces particules est caractérisée par sa charge électrique (associée à l'interaction électromagnétique), sa couleur (associée aux interactions fortes) et sa charge faible (associée aux interactions faibles).

Il existe quatre types de forces : la gravitation, l'électromagnétisme et les interactions forte et faible. Il n'existe pas à ce jour de description cohérente de la gravitation dans un cadre relativiste et quantique. Toutes les autres forces sont médiées par des bosons de jauge de spin 1. Ces forces

sont décrites par des théories, dites théories de jauge, analogues non-abéliennes de l'électromagnétisme, chacune étant associée à un groupe de Lie. Les médiateurs des interactions fortes sont les gluons et le groupe de jauge est le groupe $SU(3)$ -couleur. Le médiateur de l'électromagnétisme est le photon et le groupe de jauge est le groupe $U(1)$. Les médiateurs des interactions faibles sont les bosons dits intermédiaires Z^0 et W^\pm . Les interactions électromagnétiques et faibles sont unifiées à haute énergie dans le groupe de jauge $SU(2) \times U(1)$. Cette symétrie de jauge est brisée à faible énergie, ne laissant que le groupe $U(1)$ de l'électromagnétisme comme groupe de jauge. Ce mécanisme de brisure symétrie met en jeu une particule scalaire massive appelée le boson de Higgs. Voir Figure (6,7).

• TABLEAU DES PARTICULES ÉLÉMENTAIRES DANS LE CADRE DU MODÈLE STANDARD •

MATIÈRE

ATOME

NOYAU

NUCLÉONS
(Protons & Neutrons)

Quarks

FERMIONS		LEPTONS		QUARKS	
La matière ordinaire est composée de particules de ce groupe		peuvent se déplacer librement		présentes de particules plus grandes, ils ne sont pas observés individuellement.	
ELECTRON	Responsable de l'électricité et des réactions chimiques. Sa charge est -1.	NEUTRINO ELECTRON	Sans charge électrique et interagissant rarement avec le milieu environnant.	BAS	Sa charge électrique est - 1/3e. Le Proton en contient 1, le Neutron 2.
MUON	Un compagnon plus massif de l'électron.	NEUTRINO MUON	Propriétés similaires à celles du Neutrino électron.	ETRANGE	Un compagnon plus lourd du "Bas".
TAU	Un compagnon encore plus lourd que le Muon.	NEUTRINO TAU	Propriétés similaires à celles du Neutrino électron.	HAUT	Sa charge électrique est + 2/3e. Le Neutron en contient 1, le Proton 2.
Troisième Famille		Deuxième Famille		Première Famille	

BOSONS VECTEURS

Particules fondamentales qui assurent la transmission des forces de la nature.

PHOTON	Grain élémentaire de la lumière porteur de la force électromagnétique.	GLUON	Porteur de la force "forte" entre Quarks.	BOSONS INTERMÉDIAIRES : W⁺, W⁻ et Z⁰	Porteurs de la force "faible", responsables de certaines formes de désintégrations radioactives.
---------------	--	--------------	---	--	--

BOSON DE HIGGS ?

Hypothétique

Responsable de la "brisure de symétrie électro-faible"

Hypothétique

GRAVITON ?

Hypothétique

Daniel BONNÉRIE - CEA/DSM/CPN/A-Mixte 1992 (Mars sept. 2005)

CHAPITRE VII

GROUPES DE LORENTZ, DE POINCARÉ ET ÉQUATION DE DIRAC

Le but de ce chapitre est de décrire les caractéristiques essentielles des groupes de Lorentz, de Poincaré ainsi que de leurs algèbres et de leurs représentations.

L'étude des représentations du groupe de Lorentz nous permettra d'introduire les notions de spineurs et d'obtenir ainsi l'équation de Dirac décrivant la dynamique quantique relativiste d'une particule de spin 1/2.

L'étude des représentations du groupe de Poincaré nous conduira à en obtenir une classification fondée sur les notions de masse et de spin. Toute particule élémentaire est alors vue comme un élément d'une représentation irréductible du groupe de Poincaré.

VII.1. Groupes de Lorentz et de Poincaré

VII.1.1. Groupe et algèbre de Lorentz. — Le groupe de Lorentz est le groupe $O(1, 3)$. Soit η la métrique Minkowskienne de signature $(-, +, +, +)$. Pour toute paire de quadri-vecteurs X et Y , on note $X \cdot Y$ le résultat de l'action de η sur X et Y :

$$X \cdot Y \equiv \eta(X, Y) = X^\mu \eta_{\mu\nu} Y^\nu = X^0 Y^0 - \vec{X} \cdot \vec{Y},$$

où l'on utilise la convention d'Einstein de sommation sur les indices répétés et où $\eta_{\mu\nu} = \text{diag}(-, +, +, +)$ sont les composantes de la métrique et $X^\mu = (X^0, \vec{X})$, $Y^\nu = (Y^0, \vec{Y})$ celles des quadri-vecteurs, X^0 est appelée composante temporelle et \vec{X} composantes spatiales.

Une transformation de Lorentz est une transformation linéaire $\Lambda : X \rightarrow \Lambda \cdot X$ telle que $\eta(\Lambda \cdot X, \Lambda \cdot Y) = \eta(X, Y)$. Soit en composantes :

$$(VII.1.1) \quad X^\mu \rightarrow X'^\mu \equiv \Lambda_\nu^\mu X^\nu, \quad \text{tel que} \quad \Lambda_\rho^\mu \eta_{\mu\nu} \Lambda_\sigma^\nu = \eta_{\rho\sigma}.$$

Le groupe de Lorentz est non-simplement connexe : il possède quatre composantes connexes. En effet, les relations satisfaites par la matrice d'une transformation de Lorentz impliquent que $\det \Lambda = \pm 1$ et $(\Lambda_0^0)^2 \geq 1$. On note parfois $O(1, 3)_{\pm}^{\uparrow\downarrow}$ ces composantes connexes selon que $\det \Lambda = \pm 1$ et Λ_0^0 est positif (\uparrow) ou négatif (\downarrow). Une transformation de Lorentz préserve l'orientation si $\det \Lambda = 1$, elle préserve la flèche du temps si $\Lambda_0^0 \geq 1$. La composante connexe à l'identité, qui forme le groupe $SO(1, 3) = O(1, 3)_+^{\uparrow}$, est telle que $\det \Lambda = 1$ et $\Lambda_0^0 \geq 1$: elle est donc constituée de transformations qui préservent l'orientation et la flèche du temps. La parité P , qui laisse

invariante la composante temporelle des quadri-vecteurs mais renverse les composantes spatiales, $P(X^0, \vec{X}) = (X^0, -\vec{X})$, appartient à la composante connexe $O(1, 3)_-^\uparrow$ pour laquelle $\det \Lambda = -1$ et $\Lambda_0^0 \geq 1$. Le renversement du temps $T(X^0, \vec{X}) = (-X^0, \vec{X})$ appartient à la composante $O(1, 3)_-^\downarrow$ pour laquelle $\det \Lambda = -1$ et $\Lambda_0^0 \leq -1$.

Les rotations spatiales sont évidemment des transformations de Lorentz. Rappelons que la matrice d'une rotation d'angle θ autour de l'axe $0z$ est :

$$R_{z;\theta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & \sin \theta & 0 \\ 0 & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Les transformations de Lorentz spéciales ("boost" en anglais) sont aussi évidemment des transformations de Lorentz. La matrice d'une telle transformation d'axe $0z$ et de vitesse $v = \tanh \alpha$ (la vitesse de la lumière est choisie égale à $c = 1$ par convention) est :

$$L_{z;\alpha} = \begin{pmatrix} \cosh \alpha & 0 & 0 & -\sinh \alpha \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh \alpha & 0 & 0 & \cosh \alpha \end{pmatrix}$$

On peut implémenter des rotations autour de chacun des trois axes ou des transformations de Lorentz spéciales le long de chacun des axes. Nous verrons ci-après que l'ensemble de ces six transformations de Lorentz engendre le groupe $SO(1, 3)$.

L'algèbre du groupe de Lorentz, notée $so(1, 3)$, est par définition constituée des transformations de Lorentz infinitésimales, voisines de l'identité. La matrice d'une telle transformation se développe en $\Lambda_\nu^\mu = \delta_\nu^\mu + \epsilon \omega_\nu^\mu + \dots$ avec $\epsilon \ll 1$. Au premier ordre en ϵ , la condition d'invariance (VII.1.1) s'écrit

$$\omega_\mu^\rho \eta_{\rho\nu} + \eta_{\mu\sigma} \omega_\nu^\sigma = 0, \quad \text{soit} \quad \omega_{\nu\mu} + \omega_{\mu\nu} = 0,$$

avec $\omega_{\mu\nu} \equiv \eta_{\mu\sigma} \omega_\nu^\sigma$. Une transformation de Lorentz infinitésimale est donc spécifiée par un tenseur d'ordre deux $\omega_{\mu\nu}$ antisymétrique. La dimension de l'algèbre $so(1, 3)$ est donc égale à la dimension de l'espace des tenseurs d'ordre deux antisymétriques, soit $\dim so(1, 3) = \frac{4 \cdot 3}{2} = 6$.

Un ensemble de 6 générateurs indépendants est formé des 3 générateurs des rotations selon les trois axes, notés J_x, J_y, J_z , et des 3 générateurs des transformations de Lorentz spéciales le long des trois axes, notés K_x, K_y, K_z . Par exemple, les matrices représentant J_z et K_z (dans la représentation de définition de dimension 4), s'obtiennent en prenant θ et α petits dans les matrices des rotations et transformations de Lorentz spéciales données ci-dessus ⁽¹⁾ :

$$J_z = i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad K_z = i \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

1. On adopte la convention des physiciens en introduisant un facteur i dans les générateurs. Ainsi $R_{z;\theta} = \mathbf{1} - i\theta J_z + \dots$ pour $\theta \ll 1$ et $L_{z;\alpha} = \mathbf{1} - i\alpha K_z + \dots$ pour $\alpha \ll 1$.

Beaucoup de zéro, peu de un. Les autres matrices s'obtiennent de manière analogue, par permutation. On peut alors calculer les commutateurs et en déduire les crochets de Lie sur $so(1,3)$:

$$(VII.1.2) \quad \begin{aligned} [J_x, J_y] &= i J_z, & \text{et permutation} \\ [J_x, K_y] &= i K_z, & \text{et permutation} \\ [K_x, K_y] &= i J_z, & \text{et permutation} \end{aligned}$$

La première ligne ne mettant en jeu que les générateurs J_x, J_y, J_z indique qu'ils forment, comme il se doit, l'algèbre des rotations $so(3)$. La seconde ligne montre que les trois générateurs K_x, K_y, K_z se comportent comme les composantes d'un vecteur sous l'action des rotations. La dernière ligne implique que la composition de deux transformations de Lorentz selon deux axes différents n'est pas une transformation de Lorentz spéciale, une propriété impliquée dans la précession de Thomas.

On peut donner une formulation covariante de l'algèbre de Lie $so(1,3)$. Les transformations de Lorentz infinitésimales sont $\Lambda_\nu^\mu = \delta_\nu^\mu + \epsilon \omega_\nu^\mu + \dots$ avec $\omega_{\mu\nu}$ antisymétrique. On introduit donc une base de tenseurs anti-symétriques (avec la convention des physiciens consistant à introduire un facteur i), notée $J^{[\sigma\rho]}$:

$$(J^{[\sigma\rho]})_{\mu\nu} \equiv i(\delta_\mu^\sigma \delta_\nu^\rho - \delta_\mu^\rho \delta_\nu^\sigma).$$

Noter que nous avons bien six tenseurs $J^{[\sigma\rho]}$ indépendants. Tout tenseur antisymétrique $\omega_{\mu\nu}$ peut se décomposer sur la base des $J^{[\sigma\rho]}$: $\omega_{\mu\nu} = -\frac{i}{2} \omega_{[\sigma\rho]} (J^{[\sigma\rho]})_{\mu\nu}$. Une transformation de Lorentz infinitésimale s'écrit donc :

$$\Lambda = \mathbf{1} - i \frac{\epsilon}{2} \omega_{[\sigma\rho]} J^{[\sigma\rho]} + \dots,$$

soit $\Lambda_\nu^\mu = \delta_\nu^\mu - i \frac{\epsilon}{2} \omega_{[\sigma\rho]} (J^{[\sigma\rho]})_\nu^\mu$ avec $(J^{[\sigma\rho]})_\nu^\mu \equiv i(\eta^{\sigma\mu} \delta_\nu^\rho - \eta^{\rho\mu} \delta_\nu^\sigma)$. On calcule aisément les relations de commutations entre les $J^{[\sigma\rho]}$:

$$(VII.1.3) \quad [J^{[\sigma\rho]}, J^{[\alpha\beta]}] = -i \left(\eta^{\sigma\alpha} J^{[\rho\beta]} - \eta^{\rho\alpha} J^{[\sigma\beta]} - \eta^{\sigma\beta} J^{[\rho\alpha]} + \eta^{\rho\beta} J^{[\sigma\alpha]} \right).$$

La relation entre les deux présentations est $J^{[jk]} = \epsilon^{jkn} J_n$, i.e. $J^{[12]} = J_z$ et permutation, et $J^{[0j]} = K_j$, i.e. $J^{[01]} = K_x$. Il est clair que cette dernière présentation s'étend naturellement à toute dimension.

- *Revêtement universel de $SO(1,3)$: $SL(2, \mathbb{C})$*

Le revêtement universel de $SO(1,3)$ est $SL(2, \mathbb{C})$. A tout quadri-vecteur X on associe univoquement une matrice 2×2 hermitienne. Soit $\sigma_0 \equiv \mathbf{1}$ la matrice identité et σ_j les matrices de Pauli. Ces matrices forment une base des matrices hermitiennes 2×2 . Soit $X = (X^0, X^\mu)$ un quadri-vecteur. On pose :

$$\mathbf{X} \equiv \sigma_\mu X^\mu = \begin{pmatrix} X^0 + X^z & X^x - iX^y \\ X^x + iX^y & X^0 - X^z \end{pmatrix}.$$

Cette matrice est hermitienne et de déterminant égal à la norme Minkowskienne du quadri-vecteur :

$$\mathbf{X}^\dagger = \mathbf{X}, \quad \det \mathbf{X} = (X^0)^2 - \vec{X}^2 = X \cdot X.$$

Le groupe $SL(2, \mathbb{C})$, constitué des matrices 2×2 complexes de déterminant 1, agit sur l'espace des matrices hermitiennes selon :

$$\mathbf{X} \rightarrow U \mathbf{X} U^\dagger, \quad U \in SL(2, \mathbb{C}).$$

Cette action préserve la condition d'hermiticité et c'est un homomorphisme. Il préserve aussi le déterminant, $\det U \mathbf{X} U^\dagger = \det \mathbf{X}$, puisque U est de déterminant 1. Il induit donc une transformation de Lorentz sur le quadri-vecteur X définie par :

$$U \mathbf{X} U^\dagger = (\Lambda_U \cdot \mathbf{X}) = \sigma_\mu (\Lambda_U)^\mu_\nu X^\nu.$$

L'application $U \in SL(2, \mathbb{C}) \rightarrow \Lambda_U \in SO(1, 3)$ est un homomorphisme. Son noyau est $\{+1, -1\}$ car Λ_U est invariant par changement de signe de U . Donc

$$SL(2, \mathbb{C})/SO(1, 3) \cong \mathbb{Z}_2.$$

$SL(2, \mathbb{C})$ étant connexe et simplement connexe, c'est donc le revêtement universel de $SO(1, 3)$.

VII.1.2. Groupe et algèbre de Poincaré. — Le groupe de Poincaré s'obtient en complétant le groupe de Lorentz par les translations $T_a : X^\mu \rightarrow X^\mu + a^\mu$. Ces dernières commutent entre elles, $T_a T_b = T_b T_a$, mais elles ne commutent pas avec les transformations de Lorentz : $\Lambda T_a \Lambda^{-1} = T_{\Lambda \cdot a}$, tout $a \in \mathbb{R}^4$ et $\Lambda \in O(1, 3)$. Ainsi, conjuguer une translation par une transformation de Lorentz modifie le vecteur de translation. Le groupe de Poincaré est donc le produit semi-direct $O(1, 3) \rtimes \mathbb{R}^4$. Il s'identifie avec le groupe des matrices

$$M(\Lambda; a) = \begin{pmatrix} \Lambda & a \\ 0 & 1 \end{pmatrix}, \quad a \in \mathbb{R}^4, \quad \Lambda \in O(1, 3).$$

La loi produit est :

$$M(\Lambda; a) M(\Lambda'; a') = M(\Lambda \Lambda'; a + \Lambda \cdot a').$$

On note P_μ les générateurs des translations infinitésimales, i.e. $T_a = 1 - ia^\mu P_\mu + \dots$ pour a petit. Les générateurs de l'algèbre du groupe de Poincaré sont donc les P_μ et les générateurs de transformations de Lorentz $J^{[\sigma\rho]}$. Les relations de commutations sont :

$$(VII.1.4) \quad \begin{aligned} [P_\mu, P_\nu] &= 0 \\ [J^{[\sigma\rho]}, P_\nu] &= -i(\delta_\nu^\sigma P^\rho - \delta_\nu^\rho P^\sigma), \end{aligned}$$

complétées par les relations de commutation entre les générateurs de $so(1, 3)$. Les premières relations reflètent le caractère abélien du groupe des translations, les secondes indiquent que, comme attendu, les générateurs P_μ se transforment comme des quadri-vecteurs sous l'effet des transformations de Lorentz.

• *Action sur les fonctions.*

Il existe bien-sûr une action naturelle du groupe Poincaré sur les fonctions $f(X)$ définies sur l'espace-temps, i.e. pour $X = (X^0, \vec{X}) \in \mathbb{R}^4$:

$$(\Lambda \cdot f)(X) = f(\Lambda^{-1} \cdot x), \quad (T_a \cdot f)(X) = f(X - a).$$

On vérifie facilement que ces actions définissent une représentation du groupe de Poincaré.

On obtient la représentation de l'algèbre du groupe de Poincaré en cherchant les expressions des générateurs des transformations infinitésimales. Pour une translation d'un vecteur a , $|a| \ll 1$, un développement de Taylor au premier ordre fournit $(T_a \cdot f)(X) = f(X) - a^\mu (\partial_\mu f)(X) + \dots$, soit :

$$P_\mu = -i\partial_\mu.$$

Une transformation de Lorentz infinitésimale s'écrit $\Lambda = 1 - i\frac{\epsilon}{2}\omega_{[\sigma\rho]}J^{[\sigma\rho]} + \dots$ avec $(J^{[\sigma\rho]} \cdot X)^\mu = i(\eta^{\sigma\mu}X^\rho - \eta^{\rho\mu}X^\sigma)$. Le développement de Taylor de $(\Lambda \cdot f)(x)$ s'écrit comme $f(X) - \frac{\epsilon}{2}\omega_{[\sigma\rho]}(X^\sigma\partial^\rho - X^\rho\partial^\sigma)f(X) + \dots$. Ainsi :

$$J^{[\sigma\rho]} = i(X^\sigma\partial^\rho - X^\rho\partial^\sigma).$$

On reconnaît ici l'analogie de la formule du moment cinétique en mécanique quantique, $\mathbf{L} = \mathbf{r} \wedge \mathbf{p}$ avec $p = -i\nabla$.

VII.2. Représentations du groupe de Lorentz

Nous construisons ici les représentations de dimensions finies du groupe de Lorentz, ou plus exactement, de son complexifié (une notion que nous allons expliquer ci-après). Ces représentations ne sont pas unitaires.

VII.2.1. Complexification : $so(1, 3)_{\mathbb{C}} \simeq sl(2)_{\mathbb{C}} \oplus sl(2)_{\mathbb{C}}$. — L'algèbre de Lie $so(1, 3)$ est une algèbre définie sur le corps des réels \mathbb{R} , dont les crochets de Lie sont spécifiés par les relations (VII.1.2) que l'on peut aussi écrire sous la forme

$$\begin{aligned} [J_j, J_k] &= i\epsilon_{jkl}J_l, \\ [J_j, K_k] &= i\epsilon_{jkl}K_l, \\ [K_j, K_k] &= i\epsilon_{jkl}J_l. \end{aligned}$$

La complexification de $so(1, 3)$ est l'algèbre $so(1, 3)_{\mathbb{C}} \equiv so(1, 3) \otimes \mathbb{C}$, i.e. dans l'algèbre $so(1, 3)_{\mathbb{C}}$ on s'autorise à utiliser des combinaisons linéaires de générateurs avec des coefficients complexes. La dimension de $so(1, 3)_{\mathbb{C}}$ vue comme un espace vectoriel sur \mathbb{C} est donc égale à la dimension de $so(1, 3)$ vue comme un espace vectoriel sur \mathbb{R} . Dans $so(1, 3)_{\mathbb{C}}$ on peut définir les éléments suivants ($j = 1, 2$ ou 3) :

$$M_j \equiv \frac{1}{2}(J_j - iK_j), \quad N_j \equiv \frac{1}{2}(J_j + iK_j).$$

Ces générateurs forment une base $so(1, 3)_{\mathbb{C}}$. Il est facile de vérifier les relations de commutation suivantes :

$$(VII.2.1) \quad [M_j, M_k] = i\epsilon_{jkl}M_l,$$

$$(VII.2.2) \quad [N_j, N_k] = i\epsilon_{jkl}N_l,$$

$$(VII.2.3) \quad [M_j, N_k] = 0$$

Ainsi, chacune des familles M_j et N_j engendre indépendamment une sous-algèbre $sl(2)_{\mathbb{C}}$, et ces deux sous-algèbres commutent entre elles. Nous avons donc l'isomorphisme :

$$so(1, 3)_{\mathbb{C}} \cong sl(2)_{\mathbb{C}} \oplus sl(2)_{\mathbb{C}}.$$

Les représentations de $so(1, 3)_{\mathbb{C}}$ s'obtiennent en tensorisant deux représentations de $sl(2)_{\mathbb{C}}$. Plus précisément, si V et W sont deux espaces de représentation de $sl(2)_{\mathbb{C}}$, et $\rho_V : sl(2)_{\mathbb{C}} \rightarrow \text{End}V$, $\rho_W : sl(2)_{\mathbb{C}} \rightarrow \text{End}W$ les deux morphismes de représentation, alors une représentation de $so(1, 3)_{\mathbb{C}}$ est définie sur l'espace $V \otimes W$ avec pour morphisme de représentation $\rho_{V \otimes W}(M_j) = \rho_V(M_j) \otimes 1$ et $\rho_{V \otimes W}(N_j) = 1 \otimes \rho_W(N_j)$. Nous savons que les représentations de dimension finie de $sl(2)_{\mathbb{C}}$ sont indexées par un entier $n \in \mathbb{N}$, égale au double du "spin". Une représentation

d'indice n est définie sur \mathbb{C}^{n+1} . Les représentations de $so(1,3)_{\mathbb{C}}$ sont donc indexées par deux entiers. Nous les notons $(\frac{m}{2}, \frac{n}{2})$:

$$\text{Représentation : } (\frac{m}{2}, \frac{n}{2}), \quad \text{sur } \mathbb{C}^{m+1} \otimes \mathbb{C}^{n+1}.$$

Dans la représentation $(\frac{m}{2}, \frac{n}{2})$, les générateurs M_j et N_j agissent non-trivialement sur les vecteurs des espaces \mathbb{C}^{m+1} et \mathbb{C}^{n+1} , respectivement. Ainsi, les générateurs de rotations J_j et de transformations de Lorentz spéciales K_j sont représentés par :

$$\begin{aligned} J_j &= M_j \otimes 1 + 1 \otimes N_j, \\ K_j &= i(M_j \otimes 1 - 1 \otimes N_j). \end{aligned}$$

Notez que si les matrices M_j et N_j sont hermitiques, alors les générateurs J_j le sont aussi mais les générateurs K_j le sont pas.

VII.2.2. Exemples de représentations. — Donnons quelques exemples :

- $(0, 0)$: C'est la représentation triviale, de dimension 1. Elle correspond au scalaire.
- $(1/2, 0)$: Il s'agit d'une représentation de dimension 2, dont les vecteurs, appelés spineurs droits par convention et notés χ_R , possèdent deux composantes. Les générateurs sont représentés par :

$$J_j = \frac{1}{2} \sigma_j, \quad K_j = \frac{i}{2} \sigma_j,$$

où σ_j sont les matrices 2×2 de Pauli : $[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\sigma_l$. Notez la présence du facteur i dans la représentation de K_j , qui est donc non-unitaire.

- $(0, 1/2)$: Il s'agit aussi d'une représentation de dimension 2, qui n'est pas équivalente à la précédente. C'est sa complexe conjuguée. Elle agit sur des vecteurs à deux composantes, appelés spineurs gauches et notés χ_L . Les générateurs sont représentés par :

$$J_j = \frac{1}{2} \sigma_j, \quad K_j = -\frac{i}{2} \sigma_j,$$

où σ_j sont les matrices 2×2 de Pauli. Notez la différence de signe entre la représentation de K_j dans les représentations spinorielles $(1/2, 0)$ et $(0, 1/2)$.

- $(1/2, 1/2)$: Il s'agit d'une représentation de dimension 4. L'espace de représentation est $\mathbb{C}^2 \otimes \mathbb{C}^2 \simeq \mathbb{C}^4$. Les générateurs sont représentés par :

$$J_j = \frac{1}{2} (\sigma_j \otimes 1 + 1 \otimes \sigma_j), \quad K_j = \frac{i}{2} (\sigma_j \otimes 1 - 1 \otimes \sigma_j),$$

Elle s'identifie avec la représentation vectorielle de définition de $so(1,3)$ agissant sur les quadri-vecteurs A_μ (expliciter cet isomorphisme!).

Et ainsi de suite, on laissera au lecteur le plaisir de comprendre la signification géométrique des autres représentations...

- *Effet de la parité :*

L'opération de parité P , $X^0 \rightarrow X^0$ et $\vec{X} \rightarrow -\vec{X}$, appartient à la composante $O(1,3)_-^\uparrow$ qui n'est pas connexe à l'identité : ces représentations ne sont donc pas décrites par les représentations de

l'algèbre de Lie $so(1, 3)$. Les relations de commutations entre P et une transformation de Lorentz quelconque impliquent les relations entre P et les générateurs de $so(1, 3)$:

$$P J_j P = J_j, \quad P K_j P = -K_j.$$

Alternativement $P N_j P = M_j$ et $P M_j P = N_j$.

Les représentations $(0, 0)$ et $(1/2, 1/2)$ sont stables sous l'action de la parité. La parité peut donc prendre deux valeurs, soit $P = +$, soit $P = -$, car $P^2 = 1$.

Les représentations $(1/2, 0)$ et $(0, 1/2)$ ne sont pas stables sous l'action de la parité mais s'échangent l'une l'autre. Une représentation du groupe $O(1, 3)^\uparrow$, contenant la parité P et les transformations de Lorentz connexes à l'identité, sera donc définie sur la somme directe de ces deux représentations, i.e. sur $(1/2, 0) \oplus (0, 1/2)$. Elle agit sur un espace de dimension 4 constitué de vecteurs à quatre composantes $\chi \equiv \begin{pmatrix} \chi_R \\ \chi_L \end{pmatrix}$, somme directe de spineurs gauche et droit. Dans cette base, la parité est représentée par :

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Les vecteurs à quatre composantes $\chi \equiv \begin{pmatrix} \chi_R \\ \chi_L \end{pmatrix}$ sont appelés *spineurs de Dirac*. On peut évidemment choisir de les écrire dans différentes bases. La base que nous avons utilisée ici est appelée base de Weyl.

VII.3. Equation de Dirac

Avant d'appliquer les résultats de théorie des groupes que nous venons d'obtenir, mentionnons brièvement l'approche historique à l'équation de Dirac. Peu après l'avènement de la mécanique quantique, P.M. Dirac chercha une alternative à l'équation de Klein-Gordon (équation de la forme $(\partial_t^2 - \nabla^2 + m^2)\phi = 0$ dont l'interprétation quantique est problématique hors du cadre de la théorie des champs)⁽²⁾ :

- qui soit une équation du premier ordre similaire à l'équation de Schrödinger ;
- qui code une relation de dispersion relativiste : $E^2 = p^2 + m^2$.

Par invariance de Lorentz, si l'équation est du premier ordre en la dérivée temporelle, elle l'est aussi dans les dérivées spatiales. On écrit donc :

$$i\partial_t\psi = H\psi \quad \text{avec} \quad H = \vec{\alpha} \cdot \vec{p} + m\beta,$$

où $\vec{p} = -i\nabla$ et α^j et β des matrices. Imposer la relation de dispersion $E^2 = p^2 + m^2$ équivaut alors aux relations suivantes :

$$\alpha^j \alpha^k + \alpha^k \alpha^j = 2\delta^{jk}, \quad \alpha^j \beta + \beta \alpha^j = 0, \quad \beta^2 = 1.$$

En quelque sorte, la recherche d'une équation matricielle permet de prendre convenablement la racine carré de l'équation de dispersion relativiste $E^2 = p^2 + m^2$. Une (la) représentation de ces matrices est de dimension 4 avec :

$$\beta = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}, \quad \alpha^j = \begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix},$$

2. Nous posons $c = 1$, $\hbar = 1$, par convention

où σ_j sont les matrices de Pauli. L'équation est écrite plus élégamment (et de façon plus covariante) en introduisant les matrices de Dirac γ_μ définie par :

$$\gamma^0 = \gamma_0 = \beta = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}, \quad \gamma^j = -\gamma_j = \beta\alpha^j = \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix}.$$

Elles satisfont aux relations d'anticommutation :

$$\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2\eta_{\mu\nu},$$

où $\eta_{\mu\nu}$ est la métrique Minkowskienne. L'équation de Dirac s'écrit alors :

$$(VII.3.1) \quad (i\gamma^\mu\partial_\mu - m)\psi = 0$$

Nous allons par la suite retrouver cette équation par des considérations issues de la théorie des représentations du groupe de Lorentz. Nous laissons néanmoins à la lectrice, au lecteur, l'exercice important consistant à comprendre les propriétés de covariance de cette équation.

VII.3.1. Spineurs et équation de Dirac. — Au lieu de suivre la route historique, nous allons maintenant déduire l'équation de Dirac à l'aide de la théorie des représentations de $so(1,3)$ que nous avons décrite ci-dessus. On considère un spineur de Dirac $\chi \equiv \begin{pmatrix} \chi_R \\ \chi_L \end{pmatrix}$ qui appartient à la représentation spinorielle $(1/2, 0) \oplus (0, 1/2)$. On suppose que ce spineur est associée à une particule de masse m . Dans un référentiel \mathcal{R} , la quadri-impulsion de cette particule est $P = (P^0, \vec{P})$ avec $P^2 \equiv (P^0)^2 - (\vec{P})^2 = m^2$. Le spineur associé à cette particule dépend de l'impulsion, i.e. $\chi(P)$.

Dans le référentiel propre \mathcal{R}_* de la particule (dans lequel elle est au repos), la quadri-impulsion de la particule est $P_* = (m, \vec{0})$. Sans perdre en généralité, nous faisons le choix que les composantes droite et gauche du spineur sont alors égales, i.e. $\chi_L(P_*) = \chi_R(P_*)$.

L'hypothèse de covariance que nous allons appliquer impose que les spineurs droit et gauche associés à la particule dans un référentiel \mathcal{R} quelconque s'obtiennent à partir de ceux dans le référentiel propre \mathcal{R}_* via la transformation de Lorentz permettant de passer du référentiel \mathcal{R}_* au référentiel \mathcal{R} . L'équation de Dirac code les contraintes issues de cette hypothèse.

Dans un référentiel \mathcal{R} quelconque, la quadri-impulsion P s'écrit sous la forme $P^0 = m \cosh \alpha$ et $\vec{P} = m \sinh \alpha \vec{n}$ avec \vec{n} un vecteur unitaire et α un paramètre réel appelé rapidité. La transformation de Lorentz permettant de passer de \mathcal{R}_* à \mathcal{R} est une transformation de Lorentz spéciale d'axe \vec{n} et de rapidité α . Elle se représente par $\exp(+\frac{\alpha}{2}\vec{n} \cdot \vec{\sigma})$ et $\exp(-\frac{\alpha}{2}\vec{n} \cdot \vec{\sigma})$ sur les spineurs droit et gauche, respectivement. Donc

$$\begin{aligned} \chi_R(P) &= \exp(+\frac{\alpha}{2}\vec{n} \cdot \vec{\sigma}) \chi_R(P_*) \\ \chi_L(P) &= \exp(-\frac{\alpha}{2}\vec{n} \cdot \vec{\sigma}) \chi_L(P_*) \end{aligned}$$

Afin d'éliminer $\chi_L(P_*)$ et $\chi_R(P_*)$ de ces équations, multiplions les respectivement par $\exp(\pm\alpha\vec{n} \cdot \vec{\sigma})$, de sorte que $e^{-\alpha\vec{n} \cdot \vec{\sigma}} \chi_R(P) = \chi_L(P)$ et $e^{+\alpha\vec{n} \cdot \vec{\sigma}} \chi_L(P) = \chi_R(P)$. En utilisant la formule

$$\exp(\alpha\vec{n} \cdot \vec{\sigma}) = \cosh \alpha + \sinh \alpha \vec{n} \cdot \vec{\sigma},$$

et les expressions $P^0 = m \cosh \alpha$ et $\vec{P} = m \sinh \alpha \vec{n}$, on déduit la relation matricielle entre les composantes droite et gauche du spineur $\chi(P)$:

$$(VII.3.2) \quad \begin{pmatrix} -m & P^0 + \vec{P} \cdot \vec{\sigma} \\ P^0 - \vec{P} \cdot \vec{\sigma} & -m \end{pmatrix} \begin{pmatrix} \chi_R(P) \\ \chi_L(P) \end{pmatrix} = 0.$$

Ceci est l'équation de Dirac (écrite dans la base de Weyl).

On définit les matrices de Dirac (dans la base de Weyl) par :

$$\gamma_0 = \begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix}, \quad \gamma_j = \begin{pmatrix} 0 & \sigma_j \\ -\sigma_j & 0 \end{pmatrix}.$$

Ces matrices satisfont les relations d'anti-commutation :

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\eta_{\mu\nu}.$$

L'équation de Dirac s'écrit alors :

$$(\gamma_\mu P^\mu - m) \chi(P) = 0,$$

L'équation de Dirac est ici écrite en considérant le spineur comme fonction de la quadri-impulsion. Par transformée de Fourier, on peut la réécrire comme une équation pour un spineur de Dirac $\psi(X)$ défini sur l'espace-temps avec $X = (X^0, \vec{X})$. En effet, en posant $\psi(X) = \int dP e^{-iP \cdot X} \chi(P)$, l'équation précédente devient :

$$(VII.3.3) \quad (i\gamma_\mu \partial^\mu - m) \psi(X) = 0$$

Ceci est l'équation de Dirac pour le spineur $\psi(X)$. Elle coïncide avec l'équation (VII.3.1). Nous avons écrit ici l'équation de Dirac dans la base dite de Weyl mais tout changement de base est évidemment autorisé.

- *Interprétation :*

Une description détaillée des conséquences de l'équation de Dirac est au-delà du propos de ces notes (les étudiants ayant souvent un travail personnel à effectuer sur ce thème). Mentionnons uniquement que la découverte de l'équation de Dirac mena P.M. Dirac (après plusieurs années) à proposer l'existence du positron, anti-particule de l'électron, de même spin que l'électron mais de charge électrique opposée. P.M. Dirac fût conduit à cette conclusion en interprétant la nature du spectre de l'hamiltonien de Dirac constitué d'énergie positive et négative.

Dans la limite non-relativiste, l'équation de Dirac se réduit à l'équation de Pauli (encore un thème de travail personnel!). Elle prédit alors la valeur du moment magnétique intrinsèque de l'électron $\vec{\mu} = \frac{q_e}{m} \frac{\vec{\sigma}}{2}$, soit un facteur de Landé $g = 2$. Une valeur en très bon accord avec l'expérience. Les déviations par rapport à cette valeur sont dues aux interactions quantiques entre l'électron et le champ électromagnétique. Elles sont remarquablement prédites et évaluées dans le cadre de l'électrodynamique quantique.

VII.4. Représentations du groupe de Poincaré

Nous discutons maintenant des représentations unitaires irréductibles du groupe de Poincaré. Ces représentations sont de dimension infinie, car le groupe est non-compact. La discussion que

nous proposons sera brève, nous ne donnerons pas les détails de la construction explicite de ces représentations.

Afin d'appréhender les représentations du groupe de Poincaré on étudie d'abord celles de son algèbre de Lie. On rappelle que les générateurs de l'algèbre de Poincaré sont les générateurs des translations P_μ et les générateurs des transformations de Lorentz $J^{[\mu\nu]} = -J^{[\nu\mu]}$. Leurs relations de commutation sont décrites par les équations (VII.1.3, VII.1.4), que nous rappelons ici :

$$\begin{aligned} [J^{[\sigma\rho]}, J^{[\alpha\beta]}] &= -i(\eta^{\sigma\alpha} J^{[\rho\beta]} - \eta^{\rho\alpha} J^{[\sigma\beta]} - \eta^{\sigma\beta} J^{[\rho\alpha]} + \eta^{\rho\beta} J^{[\sigma\alpha]}) \\ [P_\mu, P_\nu] &= 0 \\ [J^{[\sigma\rho]}, P_\nu] &= -i(\delta_\nu^\sigma P^\rho - \delta_\nu^\rho P^\sigma). \end{aligned}$$

Pour des représentations unitaires, les générateurs seront représentés par des opérateurs hermitiques.

Afin de classer les représentations du groupe de Poincaré on cherche les opérateurs de "Casimir" de l'algèbre enveloppante de l'algèbre de Poincaré. Ces opérateurs sont des polynômes dans les générateurs de l'algèbre qui commutent avec tous les générateurs P_ρ et $J^{[\mu\nu]}$ de cette algèbre. D'après le lemme de Schur, (qui s'étend aux cas des groupes non-compacts mais pour des représentations unitaires irréductibles), ces opérateurs de Casimir seront représentés par des constantes dans toute représentation unitaire irréductible. Les représentations unitaires irréductibles seront donc classées par les valeurs possibles prises par ces opérateurs de Casimir.

L'algèbre de Poincaré possèdent les deux opérateurs de Casimir suivant :

(i) L'opérateur de masse

$$C_1 \equiv P^2 = P_\mu P^\mu.$$

Il est facile de vérifier que P^2 commutent avec tous les générateurs P_ρ et $J^{[\mu\nu]}$. On notera m^2 la valeur prise par cet opérateur dans une représentation unitaire irréductible. On considèrera les cas des représentations massives $m > 0$ et des représentations de masse nulle, $m = 0$. On pourrait aussi étudier les représentations pour lesquelles $m^2 < 0$, mais celles-ci ne semblent pas jouer de rôle en physique.

(ii) Le second opérateur de Casimir est

$$C_2 \equiv W^2 = W_\mu W^\mu,$$

où W^μ est l'opérateur de Pauli-Lubanski défini par $W^\mu = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma} P_\nu J_{[\rho\sigma]}$. Explicitement, $W_0 = -\vec{P} \cdot \vec{J}$ et $\vec{W} = -P_0 \vec{J} + \vec{P} \wedge \vec{K}$. Les valeurs propres prises par cet opérateur dans une représentation unitaire irréductible spécifient la notion de spin ou d'hélicité d'une particule.

L'opérateur W^2 est un opérateur de Casimir. On vérifie aussi facilement que $[W_\mu, P_\nu] = 0$ et que $[J^{[\sigma\rho]}, W_\nu] = -i(\delta_\nu^\sigma W^\rho - \delta_\nu^\rho W^\sigma)$ de sorte que W^ν se transforme comme un quadri-vecteur.

• *Représentation massive : $m > 0$.*

Les opérateurs P_μ forment une famille d'opérateurs hermitiques commutant entre eux. On peut donc les diagonaliser simultanément. Soient p_μ leurs valeurs propres. Puisque l'opérateur de Casimir P^2 agit comme la multiplication par m^2 sur tout état de la représentation, ces valeurs propres satisfont $p^2 = m^2$: on dit que ces valeurs propres sont sur la couche de masse, elles s'identifient avec la quadri-impulsion d'une particule de masse m . Dans le cas massif, $m > 0$,

tout point sur la couche de masse est dans l'orbite du groupe de Lorentz de la quadri-impulsion $p_* = (m, \vec{0})$ associée à la particule de masse m au repos. Autrement dit, toute quadri-impulsion sur la couche de masse est la transformée de Lorentz de la quadri-impulsion de la particule au repos.

Comme $[W_\mu, P_\nu] = 0$, on peut étudier l'action de W_μ sur les sous-espaces propres des P_μ . D'après la discussion précédente, il suffit, modulo une transformation de Lorentz, de se placer sur le sous-espace propre où la quadri-impulsion est $p_* = (m, \vec{0})$. Ce choix revient à se placer dans le référentiel du centre de masse de la particule. Dans ce sous-espace, on a alors

$$W_0 = 0, \quad \vec{W} = -m \vec{J}.$$

Ainsi, sur ce sous-espace propre, les opérateurs \vec{W} sont proportionnels aux générateurs des rotations dans le référentiel de centre de masse. Ces derniers engendrent le groupe $SO(3)$. Les représentations de ces générateurs sont donc indexés par un demi-entier, $s \in \frac{1}{2}\mathbb{N}$, le spin de la particule.

Les opérateurs W^μ sont les versions covariantes des générateurs des rotations dans le centre de masse. Leurs représentations sur un sous-espace propre de P_μ associée à un point quelconque de la couche de masse s'obtiennent par transformation de Lorentz à partir de celles des générateurs des rotations dans le centre de masse. Puisque l'opérateur de Casimir $C_2 = W^2$ est invariant sous les transformations de Lorentz, sa valeur sur tout sous-espace propre de P_μ est, comme il se doit, égale à celle prise dans le sous-espace propre du centre de masse, soit $C_2 = m^2 s(s+1)$ pour une représentation de spin s .

Ainsi, les représentations unitaires massives du groupe de Poincaré sont indexées par deux nombres : m , réel positif et s , demi-entier, qui s'interprète comme

$$m = \text{masse}, \quad s = \text{spin}.$$

Les vecteurs de l'espace de Hilbert de ces représentations s'identifient avec les états quantiques possibles pour une particule de masse m et de spin s . Une base de cet espace de Hilbert est fournie par les états $|p; s, \sigma\rangle$ avec p , un point sur la couche de masse et $\sigma = -s, -s+1, \dots, +s$. L'impulsion p , $p^2 = m^2$, s'identifie avec celle de la particule de masse m , s avec son spin et σ avec la valeur du moment cinétique intrinsèque J^z dans le référentiel du centre de masse. Nous ne donnerons pas ici la construction explicite de ces représentations. Celle-ci peut se faire par la méthode des représentations induites dont la description sort du cadre de ces notes (cette construction peut faire l'objet d'un travail personnel!).

- *Représentation de masse nulle : $m = 0$.*

Dans ce cas, on étudie encore l'action des opérateurs W_ν sur les sous-espaces propres de P_μ . Si la masse est nulle, ces dernières satisfont $p^2 = 0$. Il n'existe alors pas de référentiel de centre de masse (car une particule de masse nulle se propage à la vitesse de la lumière dans tout référentiel), mais il existe un référentiel tel que $p^0 = k$ et $\vec{p} = k\vec{e}_z$, i.e. ($p^0 = k, p^x = 0, p^y = 0, p^z = k$). Tout point sur la couche de masse $p^2 = 0$ est l'image par une transformation de Lorentz d'une quadri-impulsion de la forme précédente. Dans ce référentiel, on a alors $W^0 = W^z = -kJ^z$, $W^x = -k(J^x + K^x)$ et $W^y = -k(J^y - K^y)$. Les relations de commutations satisfaites par ces générateurs sont celles du groupe Euclidien en dimension deux. Le Casimir C_2 est alors égal à $W^2 = W_x^2 + W_y^2$.

Il semble que seules les représentations de masse nulle pour lesquelles le Casimir C_2 s'annule jouent un rôle en physique. On se restreint donc à ce cas, i.e. $C_2 = 0$. Alors, sur le sous-espace propre où $p^0 = k$, $\vec{p} = k\vec{e}_z$, les opérateurs W^x et W^y sont identiquement nuls et $W^z = W^0 = -k J^z$. L'opérateur J^z est donc le seul opérateur restant pour lequel il faut spécifier la représentation. J^z est le générateur de $SO(2)$, le groupe des rotations autour de l'axe Oz , qui est un groupe Abélien. Ses représentations sont de dimension 1 et indexées par la valeur prise par J^z que l'on note λ . Ainsi, dans ce sous-espace et pour ces représentations, les deux opérateurs $P^\mu = (k, 0, 0, k)$ et $W^\mu = -\lambda(k, 0, 0, k)$ sont proportionnels. Comme P^μ et W^μ se transforment identiquement sous les transformations de Lorentz, cette relation de proportionnalité est vraie dans tout référentiel :

$$W^\mu = \lambda P^\mu.$$

On appelle "hélicité" le paramètre λ .

Les représentations physiques de masse nulle du groupe de Poincaré sont donc indexées par la masse $m = 0$ et l'hélicité λ . Une base des états est fournie par des vecteurs $|p; \lambda\rangle$ avec p un point de la couche de masse, $p^2 = 0$. Nous ne construirons pas explicitement ses représentations, mais celles-ci peuvent s'obtenir par la méthode des représentations induites mentionnée ci-dessus.

- *Parité.*

Pour discuter de l'effet de la parité, il faut distinguer les deux cas, de masse nulle ou non, séparément.

Pour $m > 0$, la parité agit sur les coordonnées spatiales du référentiel du centre de masse par $P \cdot \vec{X} = -\vec{X}$, ce qui a pour effet de remplacer le groupe $SO(3)$ par le groupe $O(3)$. La parité commute avec les rotations spatiales du référentiel du centre de masse. On a $O(3)/SO(3) = \{1, P\}$ et $P^2 = 1$. Pour spécifier la représentation de $O(3)$ du référentiel du centre de masse, il faut donc se donner le spin s et le signe de parité \pm .

Pour $m = 0$, on se place comme ci-dessus dans un référentiel où $p = (k, 0, 0, k)$. L'effet de la parité est donc de remplacer le groupe $SO(2)$ par $O(2)$. La transformation $(x, y) \rightarrow (-x, -y)$ est un élément de $SO(2)$ mais la réflexion $S : (x, y) \rightarrow (x, -y)$ ne l'est pas. Cette réflexion ne commute pas avec les rotations : $S R_\alpha = R_{-\alpha} S$, si R_α est une rotation d'angle α . Ainsi, les représentations de $O(2)$ se construisent sur les sommes directes de deux représentations de $SO(2)$ d'hélicité opposée $\pm\lambda$. L'opération de parité échange les états d'hélicité opposée. Ainsi, une théorie modèle de la physique des particules élémentaires de masse nulle qui préserve la parité contient nécessairement des particules d'hélicité opposée.

- *Petits groupes et représentations induites*

Les représentations unitaires du groupe de Poincaré se construisent par la méthode dite du petit groupe et des représentations induites. Ce thème est hors du propos de ces notes (car il est souvent le sujet d'un travail personnel des étudiants!).

CHAPITRE VIII

INVARIANCE CONFORME EN PHYSIQUE

Dans ce chapitre, nous nous intéressons aux transformations conformes et à leurs applications à la physique des systèmes bidimensionnels. On traitera peu la technique associée à l'étude de ces systèmes, mais on tâchera de dresser un panorama des applications physiques de l'invariance conforme. Celles-ci couvrent des domaines aussi variés que les transitions de phases du second ordre, les systèmes électroniques quasi unidimensionnels ou les théories des cordes, candidats théoriques à une description de la gravitation quantique.

VIII.1. Transformations conformes en physique classique

VIII.1.1. Transformations conformes en 2d. — Les transformations conformes bidimensionnelles sont des transformations géométriques d'un domaine planaire qui préservent les angles. Soient $(x, y) \in \mathbb{R}^2$ les coordonnées d'un point du plan et $z = x + iy$ sa coordonnée complexe. Les transformations conformes sont des transformations holomorphes :

$$z \rightarrow w = f(z)$$

Soient $z_1 = z_0 + \delta z_1$ et $z_2 = z_0 + \delta z_2$ deux points voisins distants du point z_0 . Les deux petits vecteurs δz_1 et δz_2 joignant z_0 à z_1 et z_2 sont transformés en deux vecteurs δw_1 et δw_2 joignant $w_0 = f(z_0)$ à $w_1 = f(z_1)$ et $w_2 = f(z_2)$. Au premier ordre, on a :

$$\delta w_1 = f'(z_0) \delta z_1, \quad \delta w_2 = f'(z_0) \delta z_2$$

L'angle entre les vecteurs δw_1 et δw_2 est donc égal à celui entre les vecteurs δz_1 et δz_2 . Chacun de ces vecteurs subit une rotation d'angle égal à l'argument de $f'(z_0)$ et une dilatation proportionnelle au module de $f'(z_0)$. Ainsi, les transformations conformes agissent localement comme des rotations et des dilatations, mais le coefficient de dilatation ou l'angle de rotation varient de point en point.

Les transformations homographiques,

$$z \rightarrow f(z) = \frac{az + b}{cz + d}, \quad a, b, c, d \in \mathbb{C},$$

sont les seules bijections holomorphes du plan complexe complété du point à l'infini. Par composition, elles forment un groupe isomorphe au groupe des transformations linéaires de \mathbb{C}^2 de déterminant un, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbb{C})/\{\pm 1\}$.

La transformation $z \rightarrow w = \frac{z-i}{z+i}$ est holomorphe, sans singularité, sur le demi-plan supérieur $\mathbb{H} = \{z \in \mathbb{C}, \text{Im}z > 0\}$, et l'image de ce demi-plan est le disque de rayon unité centré en 0, i.e. $|w| < 1$ pour $\text{Im}z > 0$. Plus généralement, d'après le théorème de Riemann, tout domaine planaire ayant la topologie du disque unité, i.e. tout ouvert de \mathbb{C} connexe et simplement connexe distinct de \mathbb{C} , est en bijection conforme avec le disque unité, et donc aussi en bijection conforme avec le demi-plan supérieur.

Les transformations conformes trouvent de nombreuses applications dans tous les problèmes de physique classique régis par l'équation de Poisson-Laplace en deux dimensions,

$$\Delta_z \phi(z, \bar{z}) = 0 \quad \text{avec} \quad \Delta_z = \partial_x^2 + \partial_y^2 = 4 \partial_z \partial_{\bar{z}}.$$

Toute solution de cette équation est appelée une fonction harmonique. Comme le laplacien se factorise en $\Delta_z = 4 \partial_z \partial_{\bar{z}}$, les fonctions harmoniques se décomposent en la somme d'une fonction holomorphe et d'une fonction anti-holomorphe,

$$\phi(z, \bar{z}) = \varphi(z) + \bar{\varphi}(\bar{z}).$$

De cette décomposition découle l'invariance conforme de l'équation de Poisson-Laplace. En effet si $\phi(z, \bar{z})$ est solution, alors $\phi(f(z), \bar{f}(\bar{z}))$ est aussi solution. Alternativement, si $w = f(z)$, le laplacien dans les coordonnées w s'écrit simplement en terme du laplacien dans les coordonnées z : $\Delta_w = 4 \partial_w \partial_{\bar{w}} = |f'(z)|^2 \Delta_z$.

Cette équation s'applique évidemment aux problèmes d'électrostatique où ϕ est le potentiel électrostatique dans le vide. Ces derniers sont généralement formulés sous la forme de problèmes de Dirichlet pour lesquels on cherche la valeur du potentiel électrostatique dans un domaine planaire \mathbb{D} , qui est l'extérieur d'un conducteur parfait, sachant la valeur $\phi|_{\mathbb{D}}$ de ce potentiel au bord du domaine, c'est à dire sur le conducteur.

Elle s'applique également en mécanique des fluides à la description des mouvements bidimensionnels de fluides incompressibles et irrotationnels. En effet, si la vorticité $\omega = \nabla \wedge \mathbf{u}$ est nulle, le champ des vitesses \mathbf{u} dérive d'un gradient $\mathbf{u} = -\nabla \phi$ et la condition d'incompressibilité $\nabla \cdot \mathbf{u} = 0$ impose alors $\Delta \phi = 0$. Les conditions aux limites, appelées conditions de Neumann, dépendent de la forme du domaine dans lequel le fluide se déplace. Elles imposent que le gradient de ϕ soit tangent à la surface du domaine de sorte que l'écoulement soit tangent à cette surface.

VIII.1.2. En quoi la dimension deux est-elle spéciale ?— Une transformation conforme sur une variété métrique – une variété munie d'une métrique permettant de mesurer localement les distances et les angles – est une transformation modifiant localement la métrique par un facteur de dilatation dépendant a priori du point sur la variété. Pour faire simple, considérons l'espace euclidien de dimension d muni de la métrique euclidienne plate $ds^2 = dx_1^2 + \dots + dx_d^2$. Le tenseur métrique $g_{\mu\nu}$, défini par $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ où la sommation sur les indices répétés est implicite, est $g_{\mu\nu} = \delta_{\mu\nu}$, où $\delta_{\mu\nu}$ est le symbole de Kronecker. Lors d'un difféomorphisme $x \rightarrow y = f(x)$, le tenseur métrique se transforme comme $g_{\mu\nu} \rightarrow \hat{g}_{\mu\nu} := \frac{\partial x^\sigma}{\partial y^\mu} \frac{\partial x^\rho}{\partial y^\nu} g_{\sigma\rho}$. Une transformation conforme est donc une transformation $x \rightarrow y = f(x)$, localement définie sur \mathbb{R}^d , telle que

$$\hat{g}_{\mu\nu} = e^{\phi(x)} g_{\mu\nu},$$

où ϕ est le facteur de dilatation local, appelé facteur conforme. Evidemment une telle transformation préserve localement les angles.

Le but est maintenant d'identifier les transformations conformes en dimension d . Elles forment évidemment un groupe continu qui, comme nous allons le vérifier, est de dimension $\frac{(d+2)(d+1)}{2}$ pour $d > 2$, et de dimension infinie pour $d = 2$.

Considérons donc des transformations conformes infinitésimales $x^\mu \rightarrow x^\mu + \epsilon \xi^\mu(x) + \dots$. Au premier ordre en ϵ , la condition d'invariance conforme écrite ci-dessus se traduit en

$$(VIII.1.1) \quad \partial_\mu \xi_\nu + \partial_\nu \xi_\mu = (\delta\phi) \delta_{\mu\nu},$$

avec $\xi_\mu := \xi^\nu \delta_{\nu\mu}$ et $(\delta\phi)$ la variation infinitésimale du facteur de dilatation. En sommant sur les indices on trouve $d(\delta\phi) = 2(\partial_\mu \xi^\mu) = 2(\partial \cdot \xi)$. Dans un premier temps, montrons que, si $d > 2$, le facteur conforme $(\delta\phi)$ est une fonction affine. Dériver l'équation précédente implique $d\Delta\xi_\nu = (2-d)\partial_\nu(\partial \cdot \xi)$. En dérivant de nouveau cette équation, soit par ∂_ν soit par ∂_μ , on obtient deux nouvelles équations : $(d-1)\Delta(\partial \cdot \xi) = 0$, soit $\Delta(\partial \cdot \xi) = 0$ pour $d > 1$, et

$$(2-d)\partial_\mu \partial_\nu (\partial \cdot \xi) = \frac{d}{2}\Delta(\partial_\mu \xi_\nu + \partial_\nu \xi_\mu).$$

On reconnaît ci-dessus le laplacien du membre de gauche de l'équation (VIII.1.1) qui est nul comme le laplacien du facteur conforme, de sorte que

$$(2-d)\partial_\mu \partial_\nu (\partial \cdot \xi) = 0.$$

Ainsi, en dimension $d > 2$, le facteur conforme est une fonction linéaire : $d(\partial \cdot \xi) = k + b_\nu x^\nu$ où k et b_ν sont des constantes d'intégration.

Pour déterminer les transformations conformes infinitésimales, il nous faut donc maintenant résoudre

$$\partial_\mu \xi_\nu + \partial_\nu \xi_\mu = 2(k + b_\sigma x^\sigma) \delta_{\mu\nu}.$$

Il nous faut déterminer la différence $\partial_\mu \xi_\nu - \partial_\nu \xi_\mu$. Pour cela, dérivons l'équation précédente par ∂_σ et permutons les indices. On obtient alors deux équations dont la différence est $\partial_\nu(\partial_\sigma \xi_\mu - \partial_\mu \xi_\sigma) = 2(b_\sigma \delta_{\mu\nu} - b_\mu \delta_{\nu\sigma})$. Par intégration,

$$\partial_\sigma \xi_\mu - \partial_\mu \xi_\sigma = 2(b_\sigma x_\mu - b_\mu x_\sigma) + 2\theta_{\mu\sigma},$$

où $\theta_{\sigma\mu} = -\theta_{\mu\sigma}$ sont des constantes d'intégration. Sommer les deux dernières équations donnent accès à $\partial_\mu \xi_\nu$ et par une dernière intégration,

$$(VIII.1.2) \quad \xi_\nu(x) = a_\nu + kx_\nu + \theta_{\nu\sigma} x^\sigma + [(b \cdot x)x_\nu - \frac{1}{2}(x \cdot x)b_\nu],$$

où a_ν sont des dernières constantes d'intégration. Rappelons que cette expression a été obtenue en supposant que $d > 2$.

Chacun des termes de l'équation précédente possède une interprétation simple : le premier correspond aux translations, le deuxième aux dilatations, le troisième aux rotations et le dernier aux transformations conformes dites spéciales. Rappelons que l'eq.(VIII.1.2) décrit les transformations infinitésimales. Les paramètres de ces transformations sont les constantes d'intégration a_ν , k , b_ν et $\theta_{\nu\sigma} = -\theta_{\sigma\nu}$. La dimension du groupe des transformations conformes est donc $d+1+d+\frac{d(d-1)}{2} = \frac{1}{2}(d+1)(d+2)$. On peut monter – comment ? – que ce groupe est isomorphe à $so(d+1, 1)$.

Enfin, en dimension $d = 2$, le facteur de dilatation $(\partial \cdot \xi)$ est une fonction harmonique : $\Delta(\partial \cdot \xi) = 0$. Ainsi localement $(\partial \cdot \xi) = \partial_z v(z) + \partial_{\bar{z}} \bar{v}(\bar{z})$, ce qui correspond à une transformation infinitésimale $z \rightarrow z + \epsilon v(z)$ avec $v(z)$ localement holomorphe.

VIII.1.3. Diff \mathbb{S}^1 et son extension centrale. — En dimension deux, les transformations conformes sont les transformations locales holomorphes, comme nous venons de le voir. Elles agissent sur les fonctions par composition : $\phi(z) \rightarrow (f \cdot \phi)(z) = \phi(f^{-1}(z))$. Si f est une transformation proche de l'identité, $f(z) = z + \epsilon v(z) + \dots$ avec $\epsilon \ll 1$, la fonction transformée s'écrit $(f \cdot \phi)(z) = \phi(z) + \epsilon \delta(v)\phi(z) + \dots$ avec

$$\delta(v)\phi(z) = -v(z)\partial_z\phi(z).$$

En particulier pour $v(z) = z^{n+1}$ avec n entier, la variation $\delta(v)\phi$ est $\delta(v)\phi(z) = \ell_n\phi(z)$ avec $\ell_n \equiv -z^{n+1}\partial_z$. Ces champs vectoriels peuvent être vus comme les générateurs des transformations conformes infinitésimales. Ils satisfont les relations de commutation suivantes, qui définissent l'algèbre de Witt :

$$(VIII.1.3) \quad [\ell_n, \ell_m] = (n - m)\ell_{n+m}$$

Ces relations de commutation sont aussi celles de l'algèbre de Lie du groupe des difféomorphismes du cercle $\text{Diff } \mathbb{S}^1$. Les difféomorphismes du cercle sont les applications différentiables de \mathbb{S}^1 sur lui-même. Ils forment un groupe par composition. L'algèbre de Lie de $\text{Diff } \mathbb{S}^1$ est l'ensemble des difféomorphismes infinitésimaux $x \rightarrow x + \epsilon v(x)$, $\epsilon \ll 1$, avec $v(x)$ périodique, où $x \in [0, \pi]$ est une coordonnée sur le cercle \mathbb{S}^1 . Ils agissent sur les fonctions φ définies sur \mathbb{S}^1 par l'intermédiaire des champs vectoriels $-v(x)\partial_x$, i.e. $\varphi \rightarrow \varphi + \epsilon \delta(v)\varphi$ avec $\delta(v)\varphi(x) = -v(x)\partial_x\varphi(x)$. Le commutateur de deux de ces transformations est

$$[\delta(v_1), \delta(v_2)] = \delta([v_1, v_2]),$$

où $[v_1, v_2]$ est le crochet de Lie des champs vectoriels défini par

$$[v_1, v_2](x) \equiv (v_1(x)v_2'(x) - v_2(x)v_1'(x))\partial_x.$$

Par décomposition en série de Fourier, une base de champs vectoriels est donnée par les $\ell_n = -e^{-inx}\partial_x$ qui satisfont les relations de commutation de l'algèbre de Witt (VIII.1.3).

C'est rarement l'algèbre de Witt qui joue un rôle en mécanique quantique mais une extension centrale de cette algèbre, nommée algèbre de Virasoro. Comme nous l'avons vu dans les chapitres précédents, la mécanique quantique autorise à considérer les représentations projectives des groupes de symétrie qui peuvent se réinterpréter comme des représentations d'extensions centrales de ces groupes. Dans le cas des difféomorphismes du cercle, l'extension centrale de l'algèbre de Witt est l'algèbre de Lie construite sur l'espace vectoriel $\text{Lie}[\text{Diff } \mathbb{S}^1] \oplus c\mathbb{R}$ où c denote l'élément central – un élément central est un élément qui commute avec tout autre élément. Ici $\text{Lie}[\text{Diff } \mathbb{S}^1]$ est l'algèbre de Lie des champs vectoriels sur \mathbb{S}^1 dont les relations de commutation sont celles de l'algèbre de Witt. Les relations de commutations dans l'extension centrale sont :

$$[\delta(v_1), \delta(v_2)] = \delta([v_1, v_2]) + c\eta(v_1, v_2), \quad [c, \delta(v_1)] = 0,$$

pour tous champs vectoriels v_1 et v_2 sur \mathbb{S}^1 . Ici, $\eta(v_1, v_2)$ est une forme bilinéaire antisymétrique sur $\text{Lie}[\text{Diff } \mathbb{S}^1]$ à valeur réelle. Par l'identité de Jacobi, elle est contrainte à vérifier,

$$\eta(v_1, [v_2, v_3]) + \eta(v_2, [v_1, v_3]) + \eta(v_3, [v_2, v_1]) = 0,$$

et est appelée cocycle. Un cocycle proportionnel au commutateur, $\eta(v_1, v_2) = f([v_1, v_2])$ avec f linéaire, satisfait la relation de cocycle précédente mais est un cocycle dit trivial car il suffit de redéfinir $\delta(v) \rightarrow \delta(v) + cf(v)$ pour absorber l'extension centrale.

Dans le cas de l'algèbre de Witt, l'unique extension centrale, à un cocycle trivial près et à un coefficient multiplication près, est

$$\eta(v_1, v_2) = \frac{1}{24} \int_0^{2\pi} \frac{dx}{2\pi} (v_1''' v_2 - v_1 v_2''').$$

Il est facile de vérifier que cette définition satisfait la relation de cocycle, il est plus difficile de montrer que c'est l'unique solution. Choisissons comme ci-dessus la base des champs vectoriels, $\ell_n = -e^{-inx} \partial_x$. Leurs relations de commutation dans l'extension centrale sont

$$(VIII.1.4) \quad [\ell_n, \ell_m] = (n - m)\ell_{n+m} + \frac{c}{12}(n^3 - n)\delta_{n+m;0}.$$

Cette algèbre est appelée l'algèbre de Virasoro. Le terme de l'extension central proportionnel à n est un cocycle trivial car il peut être absorbé en redéfinissant $\ell_n \rightarrow \ell_n - \frac{c}{24}\delta_{n+m;0}$. Il est choisit de sorte à ce que l'extension centrale s'annule sur ℓ_0 et $\ell_{\pm 1}$ qui sont les générateurs des transformations conformes globales. Comme nous le verrons par la suite, cette algèbre joue un rôle important dans l'analyse des systèmes physiques invariants conformes.

• *Représentations de l'algèbre de Virasoro :*

La théorie des représentations de l'algèbre de Virasoro est riche, avec de multiples ramifications en physique et en mathématique. Le lecteur motivé pourra vérifier qu'il n'existe pas de représentation unitaire (telle qu'il existe une forme hermitienne pour laquelle $\ell_n^\dagger = \ell_{-n}$) et de plus haut poids (telle qu'il existe un vecteur ω satisfaisant $\ell_n \omega = 0$ pour tout $n > 0$) pour $c \leq 0$. En particulier, cela explique indirectement pourquoi c'est l'algèbre de Virasoro et non l'algèbre de Witt qui joue un rôle en physique. Un autre point que le lecteur pourrait être amené à développer est le lien entre la théorie des représentations de l'algèbre de Virasoro et celle des extensions centrales des algèbres de boucles sur une algèbre de Lie, i.e. des algèbres des applications de \mathbb{S}^1 à valeur dans une algèbre de Lie.

VIII.2. Le mouvement brownien bidimensionnel

Le but de cette section est de définir le mouvement brownien (en dimension deux) et de montrer l'invariance conforme de ses trajectoires. Nous discuterons aussi l'extension de ces concepts aux systèmes critiques.

VIII.2.1. Marches aléatoires et le mouvement brownien. — Nous allons construire le mouvement brownien comme la limite continue des marches aléatoires. Soit $\Lambda_a \equiv a^2\mathbb{Z}^2$ le réseau carré dont chacune des mailles est un carré d'aire a^2 . On considère un marcheur se déplaçant aléatoirement sur ce réseau. On repère sa position après N pas par les coordonnées cartésiennes

$\mathbf{X}_N = (x_N, y_N)$. A chaque étape, ce marcheur a une probabilité identique, égale à $1/4$, de faire un pas vers la droite/gauche ou vers le haut/bas. Donc

$$x_N - x_0 = a \sum_{i=1}^{N_1} \epsilon_i, \quad y_N - y_0 = a \sum_{i=1}^{N_2} \hat{\epsilon}_i, \quad N = N_1 + N_2,$$

où ϵ_i et $\hat{\epsilon}_i$ sont des variables aléatoires indépendantes et identiquement distribuées (i.i.d.) valant chacune ± 1 avec probabilité $1/2$, et N_1 (resp. N_2) sont les nombres de pas vers la droite/gauche (resp. vers le haut/bas). Par construction, la position moyenne est constante, $\mathbb{E}[\mathbf{X}_N] = \mathbf{X}_0$, mais l'écart type croît linéairement avec N :

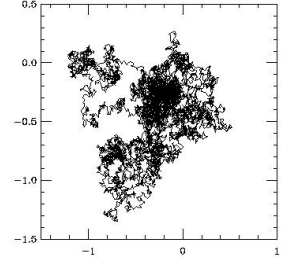
$$\mathbb{E}[(\mathbf{X}_N - \mathbf{X}_0)^2] = a^2 \mathbb{E}\left[\left(\sum_{i=1}^{N_1} \epsilon_i\right)^2 + \left(\sum_{i=1}^{N_2} \hat{\epsilon}_i\right)^2\right] = a^2 N.$$

La limite continue, qui correspond à observer cette marche de loin, s'obtient en prenant la limite où la taille de la maille du réseau s'annule tout en laissant le marcheur effectuer un nombre infini de pas, c'est-à-dire $a \rightarrow 0$, $N \rightarrow \infty$ avec Na^2 fixé. On pose $t \equiv Na^2/2$ que l'on identifie comme le temps pendant lequel le marcheur se déplace.

Dans cette limite on obtient alors un processus continu,

$$t \rightarrow \mathbf{X}_t = (x_t, y_t),$$

dont les trajectoires forment des courbes aléatoires inscrites sur le plan issues du point \mathbf{X}_0 . D'après la loi des grands nombres, les variables aléatoires x_t et y_t sont gaussiennes de variance t , $\mathbb{E}[(\mathbf{X}_t - \mathbf{X}_0)^2] = 2t$.



Plus précisément, calculons la probabilité que le marcheur partant de \mathbf{X}_0 se trouve en un point d'un domaine Ω du plan à l'instant t , i.e. la probabilité que $\mathbf{X}_t \in \Omega$. Notons $\mathbb{P}_t[\mathbf{X}_0; \Omega]$ cette probabilité. Commençons par le modèle discret et notons $\mathbb{P}_N[\mathbf{X}_0; \Omega]$ la probabilité que le marcheur partant de \mathbf{X}_0 soit dans le domaine Ω après N pas. Le marcheur aura fait le premier pas soit vers la droite/gauche ou soit vers le haut/bas avec probabilité $1/4$, donc

$$\mathbb{P}_{N+1}[\mathbf{X}_0; \Omega] = \frac{1}{4} \sum_{k=1}^4 \mathbb{P}_N[\mathbf{X}_0 + a\mathbf{e}_k; \Omega],$$

où \mathbf{e}_k sont les quatre vecteurs unités pointant vers la droite/gauche ou vers le haut/bas. Dans la limite continue, $a \rightarrow 0$, $N \rightarrow \infty$ avec $t = Na^2/2$, on a :

$$2\left(\mathbb{P}_{N+1}[\mathbf{X}_0; \Omega] - \mathbb{P}_N[\mathbf{X}_0; \Omega]\right) \simeq a^2 \partial_t \mathbb{P}_t[\mathbf{X}_0; \Omega]$$

$$\sum_{k=1}^4 \left(\mathbb{P}_N[\mathbf{X}_0 + a\mathbf{e}_k; \Omega] - \mathbb{P}_N[\mathbf{X}_0; \Omega]\right) \simeq a^2 \Delta_{X_0} \mathbb{P}_t[\mathbf{X}_0; \Omega]$$

avec $\Delta_{X_0} = \partial_{x_0}^2 + \partial_{y_0}^2$ le laplacien bidimensionnel agissant sur la variable \mathbf{X}_0 . La probabilité $\mathbb{P}_t[\mathbf{X}_0; \Omega]$ est donc solution de l'équation de la chaleur :

$$\partial_t \mathbb{P}_t[\mathbf{X}_0; \Omega] = \frac{1}{2} \Delta_{X_0} \mathbb{P}_t[\mathbf{X}_0; \Omega].$$

Si Ω est petit et centré autour d'un point \mathbf{Y} on peut écrire $\mathbb{P}_t[\mathbf{X}_0; \Omega]$ comme une densité de probabilité sous la forme $p_t[\mathbf{X}_0; \mathbf{Y}] d^2\mathbf{Y}$ et,

$$\mathbb{P}_t[\mathbf{X}_0; \Omega] = \int_{\Omega} d^2\mathbf{Y} p_t[\mathbf{X}_0; \mathbf{Y}]$$

pour tout domaine Ω . Par linéarité, la densité de probabilité est aussi une solution de l'équation de la chaleur et la condition initiale à $t = 0$ est $p_{t=0}[\mathbf{X}_0; \mathbf{Y}] = \delta(\mathbf{X}_0 - \mathbf{Y})$. Donc

$$p_t[\mathbf{X}_0; \mathbf{Y}] d^2\mathbf{Y} = \frac{d^2\mathbf{Y}}{2\pi t} \exp\left(-\frac{(\mathbf{X}_0 - \mathbf{Y})^2}{2t}\right).$$

Le processus $t \rightarrow \mathbf{X}_t$ est donc bien gaussien.

Cette formule révèle quelques propriétés d'invariance ou de symétrie du mouvement brownien. Il est invariant :

- par translation, $\mathbf{X}_t \rightarrow \mathbf{X}_t + \mathbf{a}$,
- par rotation, $\mathbf{X}_t \rightarrow R \cdot \mathbf{X}_t$,
- par dilatation, si une dilatation spatiale est accompagnée d'une dilatation temporelle, $\mathbf{X}_t \rightarrow \lambda\mathbf{X}_t$, $t \rightarrow \lambda^2 t$.

Le mouvement brownien possède une autre propriété, une sorte de localité temporelle appelée propriété de Markov, qui indique que le futur de la trajectoire au-delà d'un temps s est indépendant du passé de cette trajectoire et ne dépend que de la position \mathbf{X}_s à cet instant s . De plus, l'écart entre la position à un temps $t + s$ et la position à l'instant s est distribué comme l'écart entre la position à l'instant t et la position initiale :

$$\mathbf{X}_{t+s} - \mathbf{X}_s \equiv_{\text{en loi}} \mathbf{X}_t - \mathbf{X}_0$$

On dit que les incréments sont distribués identiquement, ils sont aussi indépendants. Ces propriétés découlent de la définition des marches aléatoires : leurs comportements après M pas sont indépendants du passé, ne dépendent que de la position \mathbf{X}_M après ces M pas et sont identiques à ceux des marches aléatoires partant du point \mathbf{X}_M .

VIII.2.2. Invariance conforme et systèmes critiques. — Montrons maintenant l'invariance conforme des trajectoires browniennes suivant un argument dû à P. Lévy. Cette invariance ne concerne que l'ensemble des traces laissées par les trajectoires, i.e. leurs graphes, elle ne concerne pas les trajectoires paramétrisées. Ces traces ne sont pas invariantes conformes réalisation par réalisation mais leur ensemble statistique l'est. L'invariance conforme se manifeste par exemple de la manière suivante. Imaginons réaliser des échantillons du mouvement brownien dans le disque unité partants de l'origine et arrêtés dès qu'ils touchent le bord du disque. Par transformation conforme, on peut transporter ces courbes dans un autre domaine planaire ayant une topologie identique à celle du disque, par exemple un rectangle, on obtient alors un ensemble statistique de courbes inscrites dans ce nouveau domaine planaire. L'invariance conforme affirme que cet ensemble de courbes possède une distribution statistique identique à celle du mouvement brownien partant d'un point du nouveau domaine et arrêté à l'instant où il touche le bord de ce domaine. On ne peut donc distinguer l'image des courbes browniennes par transformations conformes des courbes browniennes.

L'argument de Lévy consiste à promouvoir les invariances globales par rotation et dilatation en une invariance locale (donc en une invariance conforme) en utilisant les propriétés de localité du mouvement brownien. Ce principe a un domaine de validité qui s'étend au-delà de l'étude des courbes browniennes. Considérons une trajectoire brownienne arrêtée à un instant T . Divisons l'intervalle de temps $[0, T]$ en un grand nombre \mathcal{N} d'intervalles, $0 = t_0 < t_1 < \dots < t_{\mathcal{N}} = T$, et décomposons la trajectoire en la somme des incréments pendant les intervalles de temps $[t_j, t_{j+1}]$,

$$\mathbf{X}_t - \mathbf{X}_0 = (\mathbf{X}_{t_1} - \mathbf{X}_{t_0}) + (\mathbf{X}_{t_2} - \mathbf{X}_{t_1}) + \dots + (\mathbf{X}_t - \mathbf{X}_{t_i}),$$

pour $t \in [t_i, t_{i+1}]$. Les incréments $\mathbf{X}_{t_{j+1}} - \mathbf{X}_{t_j}$ sont indépendants les uns des autres et statistiquement distribués comme $\mathbf{X}_{t_{j+1}-t_j}$. Cette décomposition signifie que l'on peut reconstruire les courbes browniennes en concaténant les incréments.

Transformons maintenant chacun de ces incréments par une rotation R_j et une dilatation de facteur λ_j , qui peuvent varier d'un incrément à l'autre, pour obtenir,

$$(\mathbf{X}_{t_{j+1}} - \mathbf{X}_{t_j}) \rightarrow \lambda_j R_j \cdot (\mathbf{X}_{t_{j+1}} - \mathbf{X}_{t_j}) \equiv_{\text{en loi}} \mathbf{X}_{\lambda_j^2(t_{j+1}-t_j)}$$

où dans la dernière équivalence nous avons utilisé les invariances globales de la statistique du mouvement brownien. Ainsi, après concaténation des incréments ayant subis une rotation et une dilatation, nous obtenons des courbes ayant la même statistique que les courbes browniennes mais avec une paramétrisation temporelle différente (car les incréments temporels $t_{j+1} - t_j$ ont été transformés en $\lambda_j^2(t_{j+1} - t_j)$). On peut appliquer une transformation conforme quelconque aux courbes browniennes en prenant naïvement une limite où la discrétisation est de plus en plus fine. Cet argument, qui pourrait être rendu rigoureux, indique que l'image d'une courbe brownienne par transformation conforme est une courbe brownienne à reparamétrisation temporelle près. L'invariance conforme des courbes browniennes est à l'origine de nombreuses de leurs propriétés particulières.

Pour conclure, notons que cette propriété repose sur le principe suivant : "invariance globale sous les dilatations et les rotations plus localité implique invariance sous les dilatations et les rotations locales, i.e. sous les transformations conformes."

Un principe qui s'applique à d'autres (nombreux) systèmes physiques.

- *Systèmes critiques bidimensionnels :*

Un grand nombre de systèmes physiques sont invariants conformes, en particulier tous les modèles de physique statistique bidimensionnels ayant une transition de phase du second ordre. Le prototype de ces systèmes est le modèle d'Ising, un modèle de magnétisme, qui décrit le comportement thermodynamique d'une large collection de petits moments magnétiques, alias des spins, en interaction. Plus précisément, soit $\Lambda_a = a^2 \mathbb{Z}^2$ un réseau plan carré de maille a^2 . Sur chacun des sites de ce réseau on dispose une variable de spin s_i , $i \in \Lambda_a$, prenant deux valeurs $s_i = \pm 1$. Une configuration est la donnée des valeurs des spins s_i en tout point i du réseau. L'énergie d'une configuration décrit l'interaction entre les spins voisins, elle est donnée par la formule

$$E[\mathbf{s}] = -J \sum_{\langle ij \rangle} s_i s_j - h \sum_i s_i$$

où la somme porte sur toutes les paires de spins situés en des sites i, j voisins sur le réseau. La constante $J > 0$ mesure la force de l'interaction et h est un champ magnétique externe. La

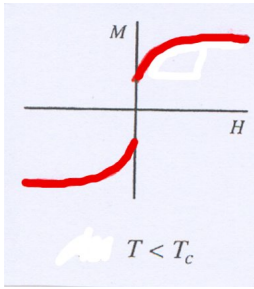
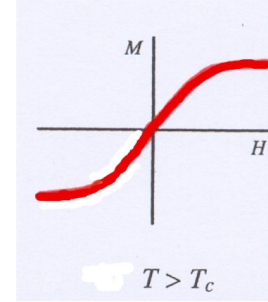
statistique est spécifiée par les règles de Boltzmann : la probabilité d'observer une configuration s est proportionnelle à $e^{-E[s]/T}$ où T est la température. Comme J est positive, les configurations où les spins sont alignés, i.e. où ils prennent les mêmes valeurs, sont les plus probables à faible température.

Le modèle d'Ising présente une transition de phase (en champ nul $h = 0$) à une certaine température T_c , appelée température critique.

A forte température, $T > T_c$, le système est dans une phase désordonnée dans laquelle les valeurs des spins varient rapidement et aléatoirement de site en site. Par exemple, à température infinie, i.e. $T \gg J$, les spins en des sites différents sont des variables binaires aléatoires indépendantes. A $T > T_c$, la réponse du système à un faible champ magnétique est linéaire : la magnétisation moyenne $\bar{M} = \mathbb{E}[s_i]$ est linéaire dans le champ et,

$$\chi(T) \equiv \partial \bar{M}(h) / \partial h|_{h=0},$$

la susceptibilité magnétique, est finie.



A faible température, $T < T_c$, le système s'ordonne. Les spins, qui ont une forte probabilité d'être alignés, sont corrélés sur de grandes distances. Par exemple, à température quasi nulle, les configurations les plus probables sont celles où tous les spins sont identiques. Le système possède une magnétisation spontanée : le système est dans une phase ferromagnétique. Plus précisément,

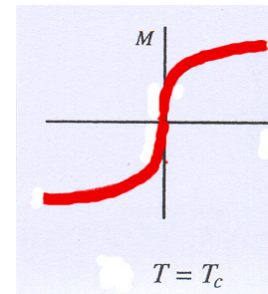
$$\lim_{h \rightarrow 0^+} \bar{M}(h) \neq 0.$$

On dit qu'il y a brisure de symétrie, un concept important dans la théorie des transitions de phase mais aussi en physique des particules.

Au point critique, $T = T_c$, le système possède des propriétés particulières et anormales. La magnétisation à faible champ est singulière,

$$\bar{M}(h)|_{T=T_c} \propto h^{1/\delta},$$

la susceptibilité magnétique diverge, $\chi(T) \propto |T - T_c|^{-\gamma}$, etc. Les exposants δ , γ , etc, sont non triviaux. Ces comportements proviennent de l'existence de fortes fluctuations des spins à toutes échelles de longueur de sorte que la physique du système est gouvernée par des phénomènes collectifs.



Les systèmes critiques (i.e. à température critique) sont caractérisés par l'existence de fortes fluctuations sans échelle de longueur caractéristique. La longueur de corrélation entre degrés de liberté locaux est infinie. En l'absence de longueur caractéristique, le système est donc invariant

par dilatation globale. En appliquant le "dogme", que nous avons rencontré dans le cas du mouvement brownien, "invariance globale plus localité implique invariance conforme", on peut conjecturer que ces systèmes sont invariants conformes. Cette conjecture est vraie (et même prouvée rigoureusement dans certains cas).

L'invariance conforme des systèmes critiques bidimensionnels a d'importantes conséquences. Elle permet de contrôler la limite continue de ces modèles, (i.e. la limite de grande échelle, ou alternativement, celle où la maille du réseau tend vers zéro), de déterminer les corrélations entre les degrés de liberté locaux dans la limite continue, de déterminer les exposants anormaux, etc. Une classification (partielle) des points critiques bidimensionnels a aussi été obtenue.

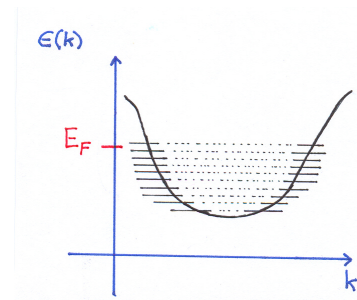
VIII.3. Systèmes électroniques mésoscopiques 1D

A basse énergie les systèmes électroniques unidimensionnels (1D) manifestent des propriétés d'invariance conforme. Nous expliquons d'abord le mécanisme sous-jacent à ces propriétés puis nous l'illustrons dans le cadre de l'effet Hall quantique.

VIII.3.1. Excitations et mer de Fermi-Dirac. — Supposons que nous puissions négliger les interactions entre électrons dans un système électronique unidimensionnel. Nous devons alors décrire le comportement d'un gaz d'électrons indépendants (que l'on supposera sans spin pour souci de simplicité) se propageant sur une ligne. Afin de quantifier ce système, imposons des conditions aux limites périodiques de sorte que le gaz électronique se déplace sur un cercle de rayon L . Les impulsions des électrons prennent alors des valeurs discrètes, $k = \frac{2\pi}{L}n$ avec n entier, afin que leurs fonctions d'ondes soient périodiques de période L . Un électron d'impulsion k aura une énergie $\epsilon(k)$ dont l'expression en fonction de k dépend de l'hamiltonien gouvernant la dynamique des électrons. La forme explicite de la fonction $k \rightarrow \epsilon(k)$, appelée relation de dispersion, intervient peu dans la suite de la discussion; nous supposons seulement que le système est invariant par parité de sorte que $\epsilon(k)$ est paire en k et est minimale en $k = 0$.

Si les électrons sont sans interaction, l'état fondamental (à température nulle $T = 0$) du système quantique constitué de tous les électrons du gaz est obtenu en remplissant les niveaux d'énergie les plus faibles en accord avec le principe de Pauli. On rappelle que ce principe interdit à deux électrons identiques d'occuper un même état quantique. Nous devons placer un électron dans chaque état accessible (si on tenait compte du spin des électrons on pourrait placer deux électrons par état).

L'état de plus basse énergie à une impulsion nulle $k = 0$, le suivant à une impulsion $k = \pm \frac{2\pi}{L}$, et ainsi de suite. Ainsi l'état fondamental du gaz d'électrons est obtenu en remplissant les niveaux d'énergie jusqu'à une énergie E_F , appelée énergie de Fermi, qui correspond à une impulsion $\pm k_F$, appelée impulsion de Fermi, $E_F = \epsilon(k_F)$. L'ensemble des états occupés par les électrons dans l'état fondamental est appelée la mer de Fermi-Dirac.



Quelles sont les excitations de basse énergie de ce système ? Elles s'obtiennent en rajoutant, ou en enlevant, un électron à un niveau d'énergie proche du niveau de Fermi. On appelle ces excitations des excitations particules/trous : les particules correspondent à rajouter un électron, les trous à en enlever un. Puisque ces excitations sont faites au voisinage de l'énergie de Fermi E_F , on peut estimer leur énergie en linéarisant le spectre au voisinage de l'impulsion de Fermi k_F ⁽¹⁾ :

$$\epsilon(k) - E_F \simeq v_F (k - k_F)$$

où $v_F = \left. \frac{\partial \epsilon(k)}{\partial k} \right|_{k=k_F}$ est appelée vitesse de Fermi. Autrement dit, l'énergie relative $\hat{\epsilon}(k) \equiv \epsilon(k) - E_F$ des excitations de basse énergie est linéaire dans l'impulsion relative $p \equiv k - k_F$, soit $\hat{\epsilon}(p) \simeq v_F p$. Il s'agit d'une relation de dispersion "relativiste" — l'énergie est proportionnelle à l'impulsion — analogue à celle d'un photon avec v_F jouant un rôle analogue à la vitesse de la lumière. La relation $\hat{\epsilon}(p) \simeq v_F p$ est indépendante des détails du spectre d'énergie $\epsilon(k)$ car elle est obtenue par linéarisation. La seule dépendance est via la vitesse de Fermi v_F . Les excitations de basse énergie auront ainsi un comportement "universel" indépendant des détails de l'hamiltonien microscopique.

La relation de dispersion relativiste ne fait intervenir aucune échelle de masse, ni aucune échelle de longueur. La physique de basse énergie du gaz d'électrons sera donc invariante par dilatation car il n'existe aucune échelle de longueur caractéristique.

Suivant le "dogme" que nous avons rencontré dans l'étude des courbes browniennes ou des systèmes critiques, la présence de l'invariance par dilatation globale et les propriétés de localité de la physique électronique suggèrent une invariance conforme de la physique de basse énergie. Cette suggestion est en fait vraie : ces systèmes quantiques sont invariants conformes. Les transformations conformes agissent sur les points de l'espace-temps bidimensionnel dont les coordonnées sont constituées de la coordonnée spatiale et du temps. Suivant les règles de la mécanique quantique, les générateurs des transformations conformes infinitésimales ℓ_n sont représentés par des opérateurs L_n qui agissent sur l'espace de Hilbert des excitations de basse énergie (formé des états excités particules/trous). Comme le système est défini sur un cercle, les L_n peuvent aussi s'interpréter comme représentant l'action de $\text{Diff } \mathbb{S}^1$. Ces opérateurs, qui peuvent s'écrire explicitement à l'aide d'opérateurs créant ou annihilant des particules/trous, satisfont les relations de commutations suivantes,

$$[L_n, L_m] = (n - m)L_{n+m} + \frac{c}{12}n(n^2 - 1)\delta_{n+m,0}$$

où c est un nombre, i.e. il commute avec les opérateurs L_n . Ces relations définissent une algèbre de Lie, appelée algèbre de Virasoro. En comparant avec l'algèbre de Witt (VIII.1.3), dont les relations de commutations,

$$[\ell_n, \ell_m] = (n - m)\ell_{n+m},$$

sont satisfaites par les générateurs des transformations conformes, ou de $\text{Diff } \mathbb{S}^1$, on voit que l'algèbre de Virasoro est une extension (centrale) de l'algèbre des transformations conformes. Ainsi les transformations conformes, ou les difféomorphismes de \mathbb{S}^1 , sont représentées projectivement sur l'espace de Hilbert. (Un thème qui peut faire l'objet d'un travail personnel!).

1. Une discussion analogue s'applique pour des impulsions voisines de $-k_F$.

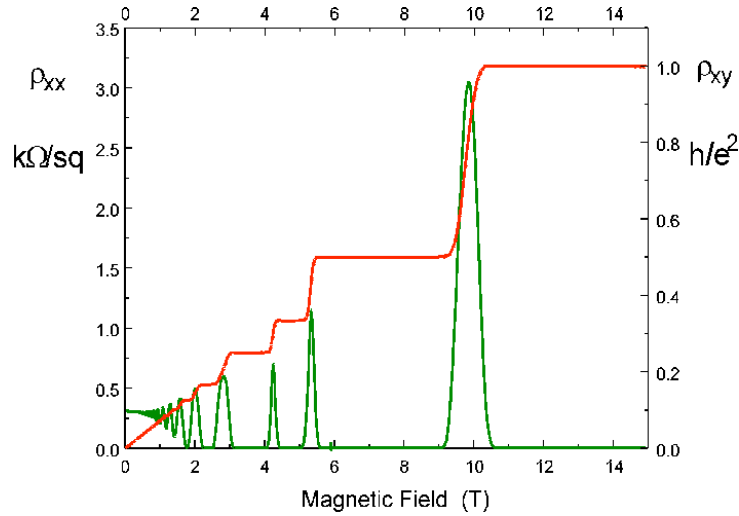


FIGURE 1. Résistivité dans l'effet Hall quantique.

L'algèbre de Virasoro, via la théorie de ses représentations, possède un caractère incontournable dans la physique des systèmes bidimensionnels invariants conformes. Elle gouverne la physique des systèmes quantiques électroniques unidimensionnels, celle des systèmes statistiques classiques critiques, mais aussi celle des champs quantiques bidimensionnels comme nous le verrons dans une prochaine section.

VIII.3.2. Effet Hall quantique. — L'effet Hall est un phénomène classique bien connu. On soumet une bande conductrice planaire, disons un rectangulaire de taille $L_x \times L_y$ avec $L_y \gg L_x$, à un champ magnétique transverse. Lors du passage d'un courant I selon la direction L_y apparaît une différence de potentiel $V_H = R_H I$ dans la direction orthogonale L_x . La résistance R_H , appelée résistance Hall, est classiquement proportionnelle au champ magnétique B : $R_H = B/en$ où n est la densité d'électrons par unité de surface. Les électrons qui transportent le courant électrique sont soumis à la force de Lorentz $\mathbf{f} = e(\mathbf{E} + \mathbf{u} \wedge \mathbf{B})$. Un champ électrique E_x doit être engendré afin d'annuler la force selon L_x de sorte que les électrons puissent se déplacer selon la direction L_y .

Dans les années 1980, des mesures de R_H ont été faites sur des échantillons, suffisamment propres, mais pas trop, refroidis à des températures extrêmement faibles et soumis à des champs magnétiques importants. Les résultats expérimentaux, suprenants, montrent une quantification de la résistance Hall en contradiction avec l'interprétation classique du phénomène :

$$R_H = \frac{1}{n} \frac{h}{e^2}, \quad n \text{ entier}, \quad n = 1, 2, 3, \dots$$

où h est la constante de Planck et e la charge de l'électron. La quantité h/e^2 est le quanta universel de résistance. Cette mesure est la mesure la plus précise (à ce jour) de $h/e^2 \simeq 25812,8$ Ohms. Voir Figure (1).

Cet effet est (bien entendu) une conséquence du comportement quantique des électrons soumis à un fort champ magnétique. Rappelons que l'hamiltonien gouvernant la dynamique d'un électron

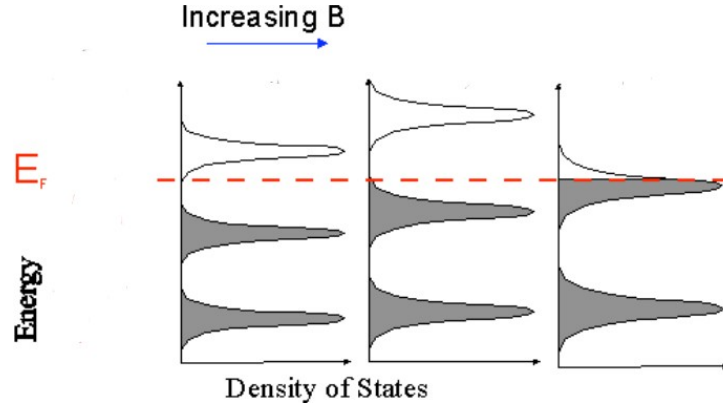


FIGURE 2. Niveaux de Landau en présence d'impuretés.

soumis à un champ magnétique est

$$H = \frac{(\mathbf{p} - e\mathbf{A})^2}{2m_e} + V$$

où \mathbf{A} est le potentiel vecteur, $\mathbf{B} = \nabla \wedge \mathbf{A}$ et V un potentiel externe auquel pourrait être soumis l'électron. Dans le cas d'un système infini uniforme (en l'absence d'un potentiel), le spectre de cet hamiltonien a été déterminé par L. Landau. Les énergies propres sont égales à $(m + 1/2)\hbar\omega_B$ avec m entier, équidistantes de $\hbar\omega_B$ où $\omega_B = eB/m_e$ est la fréquence cyclotron. Ces niveaux d'énergie sont grandement dégénérés : dans chaque niveau le nombre d'états accessibles par unité de surface est eB/h , proportionnel à B , un état "occupant" une aire de l'ordre de ℓ_B^2 avec $\ell_B^2 = h/eB$. En présence d'impuretés (un échantillon n'est jamais totalement propre) ces niveaux d'énergie s'élargissent un peu pour former des bandes d'énergie centrées autour des niveaux de Landau. Voir Figure (2).

Pour un gaz d'électrons à température nulle, on remplit les niveaux d'énergie les plus faibles en accord avec le principe de Pauli (un seul électron par niveau, si on néglige le spin) jusqu'à une énergie appelée énergie de Fermi. La quantification de la résistance Hall se produit lorsque l'énergie de Fermi est entre deux bandes de Landau. Si n bandes de Landau sont remplies, la densité électronique par unité de surface est $\mathbf{n} = neB/h$ et, en appliquant la formule classique pour R_H , on trouve en effet une résistance Hall $R_H = h/ne^2$ quantifiée.

La réalité physique est un peu plus compliquée. La quantification de la résistance Hall est observable lorsque l'on varie le champ magnétique. La présence de plateaux sur lesquels R_H est quantifiée indique que l'énergie de Fermi est bloquée entre deux bandes de Landau lorsque le champ magnétique varie. Ce blocage est assuré par l'existence d'états électroniques aux bords des bandes de Landau dont les fonctions d'onde sont localisées et qui ne participent donc pas à la conduction électronique. Ces états localisés servent de réservoir d'états permettant à l'énergie de Fermi de rester bloquée entre les niveaux de Landau.

L'invariance conforme se manifeste lorsque l'on observe les excitations de basse énergie. Quelles sont ces excitations? Pour répondre à cette question il faut considérer un système de taille finie. Supposons que l'échantillon soit rectangulaire de taille $L_x \times L_y$ avec $L_y \gg L_x$, le courant se propageant selon la grande (infinie) direction L_y . En choisissant $A_x = 0$ et $A_y = Bx$

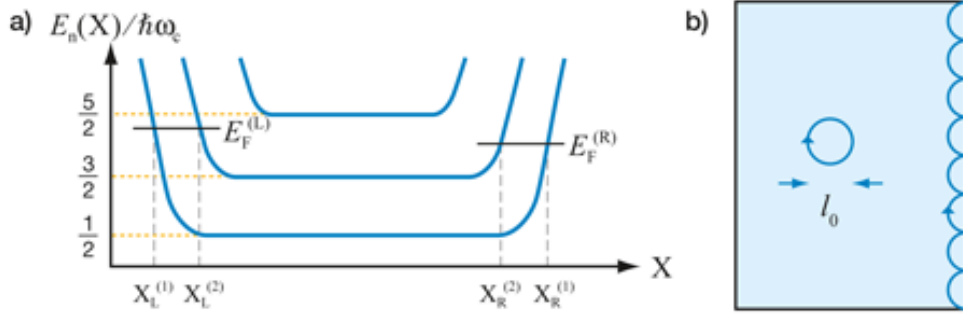


FIGURE 3. Excitations de bord dans l'effet Hall quantique.

comme potentiel, l'hamiltonien est alors

$$H = \frac{1}{2m_e}(p_x^2 + (p_y - eBx)^2) + V(x)$$

où $V(x)$ est le potentiel représentant le confinement des électrons à l'intérieur d'une boîte de largeur L_x , i.e. $V(x)$ est nulle à l'intérieur de la boîte et tend vers l'infini au bord de la boîte. Cet hamiltonien commute avec p_y , qui est donc un bon nombre quantique. A p_y fixé, le premier terme est l'hamiltonien d'un oscillateur harmonique centré en $x = p_y/eB$, dont le spectre est $(m + 1/2)\hbar\omega_B$ et dont les fonctions d'onde ont une extension spatiale de largeur ℓ_B (qui est d'autant plus petite que le champ magnétique est grand). Le spectre d'énergie est approximativement donné par

$$E_{m;p_y} \simeq (m + 1/2)\hbar\omega_B + V(p_y/eB).$$

Un électron d'énergie $E_{m;p_y}$ a une fonction d'onde localisée autour de $x = p_y/eB$. Comme le potentiel est constant et nulle dans le coeur de la boîte, exciter un électron à l'intérieur de la boîte coûte une énergie $\hbar\omega_B$, d'autant plus grande que B est grand. Si l'énergie de Fermi est entre deux niveaux de Landau, les excitations de basse énergie sont donc celles qui impliquent les électrons au bord de la boîte. Ainsi à basse énergie (i.e. à des énergie plus petites que $\hbar\omega_B$), le système est effectivement unidimensionnel, avec des excitations localisés sur les bords de l'échantillon. Ces dernières ont des impulsions longitudinales de l'ordre de $p_F \simeq \hbar L_x/2\ell_B^2$ fixé par la relation $E_F = E_{m;p_F}$. Elles sont obtenues en variant l'impulsion transverse au voisinage de p_F . Leurs énergies sont déterminées par linéarisation du spectre $E_{m;p_y}$ autour de E_F . Elles sont donc données par une relation de dispersion "relativiste". On est ainsi dans une situation identique à celle des systèmes unidimensionnels décrits ci-dessus à laquelle l'invariance conforme s'applique. Voir Figure (3).

VIII.3.3. Bosonisation, fermionisation. — Les systèmes électroniques unidimensionnels ont par nature un caractère fermionique car ils sont constitués d'électrons. Ils ont aussi un caractère bosonique car leurs excitations de basse énergie sont formées de paires particule/trous, or une paire électron/trou obéit à une statistique bosonique. Cette dualité de description entre bosons et fermions peut être appréhender de manière combinatoire.

Rappelons que, dans l'approximation linéaire au voisinage de l'énergie de Fermi, le spectre d'excitations de basse d'énergie est de la forme $\hat{\epsilon}(p) = v_F p$ où $p = k - k_F$ est l'impulsion relative

par rapport à l'impulsion de Fermi k_F . Les impulsions k sont quantifiées $k = \frac{2\pi}{L}n$ avec n entier. Par convention choisissons de situer l'impulsion de Fermi entre deux de ces valeurs de sorte que $k_F = \frac{\pi}{L}(2n_F + 1)$ avec n_F entier. Les impulsions relatives prennent donc les valeurs $p = \frac{\pi}{L}(2j + 1)$ avec j entier positif ou négatif. Les énergies (relatives à l'énergie de Fermi) des états accessibles sont donc

$$\hat{\epsilon}_j = \varepsilon(2j + 1), \quad j \in \mathbb{Z}$$

avec $\varepsilon = \frac{\pi v_F}{L}$ une échelle d'énergie caractéristique. Par convention, l'énergie de Fermi est nulle et l'état fondamental du gaz électronique s'obtient en remplissant tous les états d'énergie négative. Cette collection d'états forme la mer de Fermi-Dirac.

Nous allons maintenant énumérer les états excités du gaz électronique au dessus de la mer de Fermi-Dirac. Les états excités seront repérés par deux nombres quantiques : (i) leur énergie, calculée en associant une énergie $\hat{\epsilon}_j$ pour chaque électrons dans le $j^{\text{ième}}$ niveau et, (ii) leur charge, calculée en associant une charge $+1$ à chaque particule et charge -1 à chaque trou. On énumère tous les états formés d'un nombre quelconque de particules ou de trous. Pour cela nous allons calculer une fonction de partition (grande canonique) qui est la fonction génératrice de cette énumération.

Commençons par la description fermionique de l'espace des excitations. Dans la mer de Fermi-Dirac tous les états d'énergie $\hat{\epsilon}_j$ avec $j < 0$ sont occupés et tous les états avec $j > 0$ sont vides. Donc, si on rajoute une particule dans le $j^{\text{ième}}$ niveau, $j > 0$, on augmente l'énergie de $\hat{\epsilon}_j$ et la charge d'une unité ; si on rajoute un trou dans le $j^{\text{ième}}$ niveau, $j < 0$, on augmente l'énergie de $-\hat{\epsilon}_j$ et on diminue la charge d'une unité. Introduisons les nombres d'occupation n_j , $j > 0$, des états d'énergie positive par des particules, et les nombres d'occupation \tilde{n}_j , $j < 0$, des états d'énergie négative par des trous. Le principe de Pauli impose que n_j et \tilde{n}_j soient égaux à 0 ou 1. Un état excité est univoquement spécifié par la donnée de ces nombres d'occupation. Son énergie E et sa charge Q sont :

$$\begin{aligned} E[\mathbf{n}, \tilde{\mathbf{n}}] &= \sum_{j>0} n_j \hat{\epsilon}_j - \sum_{j<0} \tilde{n}_j \hat{\epsilon}_j \\ Q[\mathbf{n}, \tilde{\mathbf{n}}] &= \sum_{j>0} n_j - \sum_{j<0} \tilde{n}_j \end{aligned}$$

Par définition, la fonction de partition (grande canonique) est la somme sur toutes les configurations $Z = \sum_{[\mathbf{n}, \tilde{\mathbf{n}}]} \exp(-\beta E[\mathbf{n}, \tilde{\mathbf{n}}] - \mu Q[\mathbf{n}, \tilde{\mathbf{n}}])$. Elle se factorise en produits :

$$Z = \prod_{j>0} (1 + e^{-\beta \hat{\epsilon}_j - \mu}) \prod_{j<0} (1 + e^{\beta \hat{\epsilon}_j + \mu}) = \prod_{j>0} (1 + x^{2j+1} y) \prod_{j>0} (1 + x^{2j+1} y^{-1})$$

où nous avons posé $x \equiv e^{-\beta \varepsilon}$ et $y \equiv e^{-\mu}$.

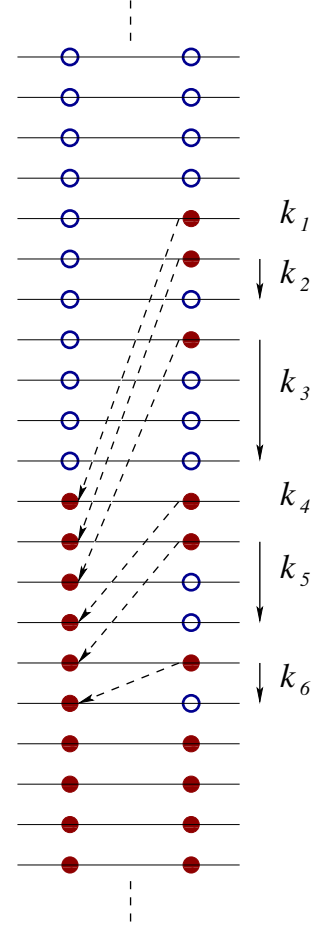
Passons maintenant à la description bosonique des états excités qui consiste à énumérer le nombre d'excitations particule/trous. Une excitation particule/trou ne change pas la charge, car une particule a une charge $+1$ et un trou une charge -1 . Il nous faut donc d'abord identifier les états de plus basse énergie de charge donnée. Si la charge est $Q = q$, q entier positif, cet état est obtenu en remplissant de particules les q premiers niveaux au dessus de la mer de Fermi-Dirac et son énergie est $\hat{\epsilon}_1 + \dots + \hat{\epsilon}_q = 2\varepsilon q^2$. Si la charge est négative, $Q = -q$, on remplit les q premiers niveaux dans la mer de Fermi-Dirac par des trous et son énergie est $-\hat{\epsilon}_{-1} - \dots - \hat{\epsilon}_{-q} = 2\varepsilon q^2$. Ces

états, avec q particules ou q trous adjacents à la mer de Fermi-Dirac, forment aussi des mers de Fermi-Dirac mais décalées par rapport à la mer de Fermi-Dirac originelle ; ils ont même énergie mais une charge opposée. Les autres états sont obtenus en excitant des paires particule/trous neutres, c'est-à-dire en transformant des particules de ces mers de Fermi-Dirac chargées en trous et en déplaçant les particules ainsi obtenues sur des niveaux d'énergie plus élevée. Il est clair que l'on obtient de cette façon tous les états excités.

Réciproquement, à toute configuration fermionique, on associe une charge q et des entiers positifs $m_1 \leq m_2 \leq \dots \leq m_l$ qui codent de combien d'échelons (sur l'échelle des niveaux d'énergie) il faut déplacer les particules (en commençant par les particules les moins énergétiques) afin de transformer cette configuration en la mer de Fermi-Dirac de charge q . En pratique, on déplace les particules présentes dans les états d'énergie élevée de manière à remplir tous les trous qui sont en dessous. Chaque groupe de k trous adjacents impose de déplacer de k niveaux vers le bas toute la configuration de particules et de trous au dessus de ce groupe. On pose donc

$$m_1 = k_l, \quad m_2 = k_l + k_{l-1}, \dots, \\ m_l = k_l + \dots + k_2 + k_1$$

Dans la figure ci-contre, $k_1 = 0, k_2 = 1, k_3 = 3, k_4 = 0, k_5 = 2$ et $k_6 = 1$.



Comme l'écart d'énergie entre deux niveaux est 2ε , l'énergie d'une configuration de charge q est

$$E[q, \mathbf{k}] = 2\varepsilon q^2 + \sum_j 2\varepsilon m_j = 2\varepsilon q^2 + \sum_j 2\varepsilon j k_j.$$

Cette énergie coïncide avec celle d'une collection de bosons, chacun d'énergie $2\varepsilon j$, avec j entier. Il n'y a pas de restriction sur les nombres d'occupation bosoniques k_j . La fonction de partition (grande canonique), $Z = \sum_{[q, \mathbf{k}]} \exp(-\beta E[q, \mathbf{k}] - \mu q)$, s'écrit alors

$$Z = \sum_{q \in \mathbb{Z}} e^{-2\beta\varepsilon q^2 - \mu q} \sum_{\mathbf{k}} e^{-2\beta\varepsilon \sum_{j \geq 1} j k_j} = \left(\sum_{q \in \mathbb{Z}} x^{q^2} y^q \right) \prod_{j \geq 1} \left(\frac{1}{1 - x^{2j}} \right)$$

où nous avons posé $x \equiv e^{-\beta\varepsilon}$ et $y \equiv e^{-\mu}$ comme précédemment.

Cette dualité de description conduit à l'identité combinatoire remarquable

$$\prod_{j>0} (1 + x^{2j+1}y) \prod_{j<0} (1 + x^{2j+1}y^{-1}) = \left(\sum_{q \in \mathbb{Z}} x^{q^2} y^q \right) \prod_{j \geq 1} \left(\frac{1}{1 - x^{2j}} \right)$$

qui est due à Jacobi.

Finalement, notez l'analogie entre la discussion sur la mer de Fermi-Dirac et ses excitations que nous avons présentée ci-dessus et celle liée à la caractérisation du vide et des excitations en théorie des champs quantiques.

VIII.4. Champs quantiques et algèbre de Virasoro

Après avoir donné des exemples d'invariance conforme en physique statistique (classique) et en physique électronique, on donne ici un exemple en physique (quantique) des hautes énergies via l'étude de la quantification d'un champ libre de masse nulle en dimension deux. Ce sera aussi l'occasion d'avoir un très (très) léger aperçu sur les théories des cordes.

VIII.4.1. Equation d'onde et quantification. — L'équation d'onde

$$(\partial_t^2 - \partial_x^2) \phi(x, t) = 0$$

dont les solutions sont de la forme $\phi(x, t) = \varphi(x - t) + \bar{\varphi}(x + t)$, décrit des ondes qui se propagent vers la droite ou vers la gauche. On suppose que le champ ϕ est réel. Nous allons chercher à quantifier ces modes de propagation (afin de décrire les quantas associés à ces ondes). Par quantification, le champ $\phi(x, t)$, qui satisfait toujours à l'équation d'onde ci-dessus, devient un opérateur agissant sur un espace de Hilbert que nous spécifions. Il faut imposer des conditions aux limites. Par choix, on suppose que $x \in [0, \pi]$ et on impose $\phi(0, t) = 0 = \phi(\pi, t)$. Ces conditions sont (parfois) appelées les conditions de "cordes ouvertes" : le champ $\phi(x, t)$ est vu comme la position d'une corde fixée aux extrémités, la coordonnée x est l'abscisse le long de la corde et t est le temps.

Le résultat de la quantification va se décrire à l'aide d'une série d'oscillateurs harmoniques associés aux différents modes de vibration. Avec les conditions aux limites ci-dessus, $\phi(x, t)$ se décompose en série de Fourier

$$\phi(x, t) = \sqrt{2} \sum_{n>0} Q_n(t) \sin(nx)$$

L'équation d'onde se traduit en les équations de mouvement des amplitudes Q_n ,

$$\partial_t^2 Q_n + n^2 Q_n = 0.$$

Donc, chaque mode Q_n est la coordonnée d'un oscillateur harmonique de pulsation n , et les différents oscillateurs sont indépendants les uns des autres. Quantifier le champ consiste à quantifier chacun de ces oscillateurs harmoniques. On rappelle que si q est la coordonnée d'un oscillateur harmonique, les équations de mouvement sont $\partial_t^2 q + \omega^2 q = 0$, $\partial_t q = p$ et l'hamiltonien est

$h = p^2/2 + \omega^2 q^2$. Par quantification, $q = \frac{1}{\sqrt{2\omega}}(a + a^\dagger)$, $p = -i\sqrt{\frac{\omega}{2}}(a - a^\dagger)$, et $h = \omega(aa^\dagger + 1/2)$, avec $[a, a^\dagger] = 1$. Donc, par quantification :

$$Q_n = \frac{1}{\sqrt{2n}}(a_n + a_n^\dagger), \quad \text{avec} \quad [a_n, a_m^\dagger] = \delta_{n;m} \text{ et } [a_n, a_m] = 0.$$

L'évolution temporelle est donnée par conjugaison par e^{-iht} où h est l'hamiltonien. Puisque les pulsations sont les entiers n , on a :

$$Q_n(t) = \frac{1}{\sqrt{2n}}(a_n e^{+int} + a_n^\dagger e^{-int}).$$

Le champ ϕ s'écrit donc comme

$$\phi(x, t) = \sum_{n>0} \frac{1}{\sqrt{n}}(a_n e^{+int} + a_n^\dagger e^{-int}) \sin(nx)$$

L'espace de Hilbert sur lequel agit ce champ quantique est l'espace de Fock engendré par tous les opérateurs de création-annihilation a_n^\dagger , a_n associés aux modes d'oscillation. Cet espace est construit par actions successives des opérateurs création a_n^\dagger sur un état de vide $|0\rangle$ tel que $a_n|0\rangle = 0$, pour tout $n > 0$. L'état de vide décrit une situation où aucun mode n'est excité. Les états de la forme $a_n^\dagger|0\rangle$ sont des états contenant un seul quanta dans un seul mode de vibration, ceux de la forme $a_n^\dagger a_m^\dagger|0\rangle$ sont ceux contenant, soit deux quantas dans le même mode de vibration si $n = m$, soit deux quantas répartis dans deux modes de vibration différents si $n \neq m$, etc.

Il est usuel de décomposer le champ ϕ est la somme de deux champs, l'un se propageant vers la droite et l'autre vers la gauche : $\phi(x, t) = \varphi(x-t) + \bar{\varphi}(x+t)$. Comme les conditions aux limites $\phi(0, t) = 0 = \phi(\pi, t)$ fixent la corde aux deux extrémités, ces champs ne sont pas indépendants, $\bar{\varphi}(y) = -\varphi(-y)$, car les modes de vibrations sont réfléchis aux extrémités. Ainsi :

$$\varphi(x-t) = \frac{i}{2} \sum_{n \neq 0} \frac{1}{n} \alpha_n e^{-in(x-t)},$$

où $\alpha_n = \sqrt{n}a_n$ et $\alpha_{-n} = \sqrt{n}a_n^\dagger$ pour $n > 0$. Les relations de commutation des a_n et a_m^\dagger se traduisent en

$$[\alpha_n, \alpha_m] = n\delta_{n+m;0}.$$

L'état de vide satisfait $\alpha_n|0\rangle = 0$ pour $n > 0$. Le champ φ est appelé composante chirale du champ ϕ .

VIII.4.2. Invariance conforme et algèbre de Virasoro. — L'équation d'onde $(\partial_t^2 - \partial_x^2)\phi = 0$ est l'équation de Klein-Gordon pour une particule de masse nulle. Elle n'implique pas d'échelle de masse ou de longueur. D'après le dogme déjà invoqué dans l'étude des systèmes critiques ou électroniques, on s'attend à ce que la théorie soit invariante sous les dilatations, et par localité, invariante sous les transformations conformes.

Fixons le temps $t = 0$. Le champ $\phi(x, 0)$ est défini sur l'intervalle $[0, \pi]$ mais la composante chirale $\varphi(x)$ est définie sur $[0, 2\pi]$ avec des conditions aux limites périodiques, donc sur le cercle \mathbb{S}^1 . Si les difféomorphismes de \mathbb{S}^1 agissent sur la théorie, ils agissent sur le champ φ par l'intermédiaire des champs vectoriels $\ell_n \equiv -e^{inx} \partial_x$ via

$$\varphi(x) \rightarrow \varphi(x) + \epsilon(\ell_n \cdot \varphi)(x), \quad \epsilon \ll 1$$

On rappelle que les ℓ_n sont les générateurs infinitésimaux de $\text{Diff}(\mathbb{S}^1)$ et qu'ils satisfont aux relations de commutation $[\ell_n, \ell_m] = (n - m)\ell_{n+m}$ de l'algèbre de Witt.

On peut promouvoir les transformations infinitésimales de $\text{Diff}(\mathbb{S}^1)$ en des transformations conformes infinitésimales en posant $z = e^{i(x-t)}$ de sorte que le champ φ est fonction de z :

$$\varphi(z) = \frac{1}{2i} \sum_{n \neq 0} \frac{1}{n} \alpha_n z^{-n}.$$

Dans la coordonnée z , les champs vectoriels ℓ_n deviennent $\ell_n = -z^{n+1}\partial_z$ et les transformations conformes agissent alors comme $\varphi \rightarrow \varphi + \epsilon\delta\varphi$ avec $\delta\varphi(z) = (\ell_n \cdot \varphi)(z)$.

De manière à implémenter la symétrie conforme dans la théorie quantique, on cherche à définir des opérateurs L_n , agissant sur l'espace de Hilbert (ici l'espace de Fock) tel que :

$$[L_n, \varphi(z)] = (\ell_n \cdot \varphi)(z)$$

Cette relation impose que $[L_n, \alpha_m] = -m\alpha_{n+m}$. Une solution est :

$$(VIII.4.1) \quad L_0 = \sum_{n>0} \alpha_{-n}\alpha_n$$

$$(VIII.4.2) \quad L_n = \frac{1}{2} \sum_{k \neq 0} \alpha_{-k}\alpha_{n+k}$$

Cette construction s'appelle la construction de Sugawara. Dans L_0 , on reconnaît l'hamiltonien $\sum_{n>0} n a_n^\dagger a_n$ de la collection des oscillateurs harmoniques associés à chacun des modes, car $\alpha_n = \sqrt{n}a_n$ et $\alpha_{-n} = \sqrt{n}a_n^\dagger$. L'hamiltonien est l'intégrale de la densité d'énergie du champ. Les opérateurs L_n sont les modes de Fourier de cette densité d'énergie. (Vérifier les relations $[L_n, \alpha_m] = -m\alpha_{n+m}$!)

Nous laissons au lecteur le plaisir de vérifier que ces opérateurs ont les bonnes relations de commutation avec le champ φ et qu'ils satisfont les relations de commutation suivantes (un calcul plus difficile!) :

$$[L_n, L_m] = (n - m)L_{n+m} + \frac{c}{12}n(n^2 - 1)\delta_{n+m;0}.$$

Cette algèbre s'appelle l'algèbre de Virasoro, et elle est donc représentée sur l'espace de Fock du champ libre (bosonique, de masse nulle). Ici $c = 1$. Ces éléments sont les générateurs de transformations de $\text{Diff}\mathbb{S}^1$ et des transformations conformes. Celles-ci sont représentées projectivement car l'algèbre de Virasoro est une extension centrale de l'algèbre de Witt.

• *B.A.-BA. de théories des cordes :*

En théorie des cordes, le champ ϕ est interprété comme la position de la corde. Si celle-ci se déplace dans un espace de dimension D , on repère la position de la corde par une collection de D champs $\Phi = (\phi^1, \dots, \phi^D)$. Il existe deux types de cordes : des cordes dites ouvertes dont les extrémités sont fixées (comme nous avons discuté ci-dessus) ou des cordes fermées pour lesquels les champs ϕ^j sont périodiques. L'application $x \in \mathbb{S}^1 \rightarrow \Phi(x) \in \mathbb{R}^D$ décrit le plongement de la corde dans l'espace \mathbb{R}^D , et $t \rightarrow \Phi(\cdot, t)$ décrit son évolution.

Les applications physiques de la théorie des cordes sont fondées sur le principe que les constituants élémentaires sont des cordes et que les particules s'identifient avec les modes de vibrations de ces cordes.

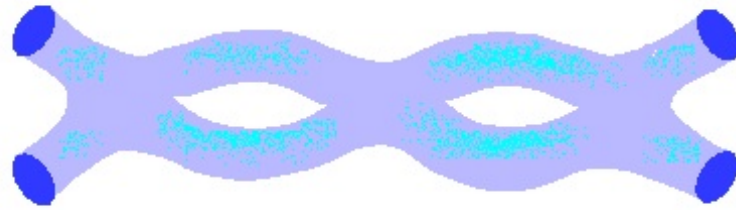


FIGURE 4. Exemple de surface de Riemann induite par l'évolution de deux cordes fermées en interaction.

Lors de son évolution, une corde décrit une surface. Les cordes interagissent entre elles en se joignant ou en se détachant. En conséquence, il peut y avoir des "trous" ou des "poignées" dans ces surfaces — mathématiquement, on dit que la surface est une surface de Riemann de genre élevé. Voir Figure (4). Décrire la dynamique quantique des cordes consiste à définir les amplitudes de probabilité associées à chacune des surfaces parcourues par les cordes lors de leurs évolutions et de leurs interactions. Ces surfaces peuvent être localement paramétrisées par des coordonnées complexes (z, \bar{z}) . Afin que la dynamique des cordes soit intrinsèque, on impose que les amplitudes de probabilité soient indépendantes, autant que faire se peut, du choix de la métrique sur la surface, i.e. qu'elles soient invariantes sous les transformations conformes sur la surface. Ainsi, la dynamique des cordes est décrite par les fluctuations quantiques des surfaces qu'elles parcourent et ces fluctuations sont invariantes conformes. La suite de la théorie des cordes est un peu plus compliquée...

CHAPITRE IX

OSCILLATEURS HARMONIQUES ET REPRÉSENTATIONS DU GROUPE DE HEISENBERG

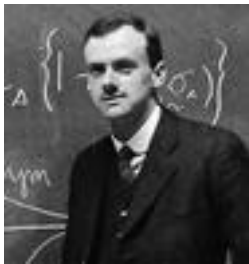


FIGURE 1. Paul Dirac

The steady progress of physics requires for its theoretical formulation a mathematics which get continually more advanced. This is only natural and to be expected. What however was not expected by the scientific workers of the last century was the particular form that the line of advancement of mathematics would take, namely it was expected that mathematics would get more and more complicated, but would rest on a permanent basis of axioms and definitions, while actually the modern physical developments have required a mathematics that continually shifts its foundation and gets more abstract. Non-euclidean geometry and noncommutative algebra, which were at one time were considered to be purely fictions of the mind and pastimes of logical thinkers, have now been found to be very necessary for the description of general facts of the physical world. It seems likely that this process of increasing abstraction will continue in the future and the advance in physics is to be associated with continual modification and generalisation of the axioms at the base of mathematics rather than with a logical development of any one mathematical scheme on a fixed foundation (Paul Dirac).



The universe is an enormous direct product of representations of symmetry groups. (attribué à Hermann Weyl).

FIGURE 2. Hermann Weyl

IX.1. Introduction

IX.1.1. Introduction. — Le but de cette séance est de montrer les liens entre oscillateurs harmoniques et représentations du groupe de Heisenberg. On montre les difficultés mathématiques qui se posent dans la mesure où la représentation de Schrödinger fait apparaître des opérateurs non bornés, *i.e.* non partout définis. On explique pourquoi les points de vue de Heisenberg et de Schrödinger sont équivalents et enfin on introduit l'algèbre de Heisenberg de dimension infinie, algèbre des symétries d'une infinité d'oscillateurs et l'action correspondante de l'algèbre de Virasoro, fenêtre vers la théorie des cordes qui sera ébauchée au cours suivant. On voit apparaître un phénomène très courant en théorie des champs, l'apparition de grandeurs infinies dans les calculs, qui nécessite une « renormalisation ».

IX.1.2. L'oscillateur harmonique à un degré de liberté. — Physiquement, on a une masse m de quantité de mouvement p dans un potentiel quadratique $\frac{1}{2}m\omega^2x^2$ où x est la position. On a donc deux variables quantiques normalisées, l'impulsion p (quantifiant $\frac{\hat{p}}{\sqrt{m\hbar\omega}}$) et la position q (quantifiant $\sqrt{\frac{m\omega}{\hbar}}\hat{x}$). L'énergie (l'hamiltonien) est (à un facteur $\frac{1}{2}$ près) égal à

$$H = p^2 + q^2.$$

On choisit l'unité de longueur de sorte que de plus $\hbar = 1$ (rappelons que \hbar est homogène à m^2kgs^{-3}). En mécanique quantique (cf. cours de tronc commun), p, q sont des observables, à savoir des opérateurs auto-adjoints agissant sur un espace de Hilbert (séparable) complexe \mathcal{H} .



FIGURE 3. David Hilbert

Les équations fondamentales sont

$$(IX.1.1) \quad [p, q] = -i \text{ et } \dot{\pi} = -i[H, \pi]$$

pour tout observable π et où H est donc l'observable

$$H = p^2 + q^2 = p \circ p + q \circ q$$

(on note 1 l'identité).

Remarque IX.1.1. — Par analogie au cas de la dimension finie, l'algèbre de Lie du groupe unitaire $\mathbf{U}(\mathcal{H})$ devrait être l'algèbre de Lie $\mathfrak{u}(\mathcal{H})$ des opérateurs anti-hermitiens de sorte qu'il n'y

a pas lieu de s'étonner que le crochet $[p, q] = -[ip, iq]$ soit l'opérateur anti-hermitien i et non un opérateur hermitien.



FIGURE 4. Sophus Lie

D'un point de vue formel, une récurrence sans mystère donne

$$p^n q - q p^n = -i n p^{n-1}$$

pour tout $n \geq 0$ d'où en sommant après multiplication par $\frac{(ix)^n}{n!}$, $x \in \mathbf{R}$

$$\exp(ixp)q - q \exp(ixp) = x \exp(ixp)$$

à savoir

$$\exp(ixp)q \exp(-ixp) = q + x.$$

De même, on obtient

$$\exp(ixp) \exp(iyq) \exp(-ixp) = \exp(iy(q + x))$$

et donc

$$(IX.1.2) \quad (\exp(ixp), \exp(iyq)) = \exp(ixy) \text{ pour tout } x, y \in \mathbf{R}$$

où $(g_1, g_2) = g_1 g_2 g_1^{-1} g_2^{-1}$ désigne le commutateur de deux éléments g_1, g_2 d'un groupe (ici $\mathbf{U}(\mathcal{H})$).

IX.2. Groupe et algèbre d'Heisenberg

IX.2.1. Rappels sur le groupe de Heisenberg. — Rappelons que le groupe de Heisenberg H_n à n degrés de libertés peut-être vu comme le sous-groupe de $\mathbf{GL}_{2n+1}(\mathbf{R})$ des matrices $\exp(M(x, y, z)) = \text{Id} + M(x, y, z)$ avec

$$M(x, y, z) = \begin{pmatrix} 0 & {}^t x & {}^t y & 2z \\ 0 & 0 & 0 & x \\ 0 & 0 & 0 & -y \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ avec } x, y \in \mathbf{R}^n, z \in \mathbf{R}$$

avec

$$\exp(M(x, y, z)) \exp(M(x', y', z')) = \exp(M(x + x', y + y', z + z' + \frac{1}{2}\omega(x, y, x', y')))$$

où $\omega(x, y, x', y') = {}^t x y' - {}^t y x'$. L'algèbre de Lie \mathfrak{h}_n de H_n est

$$\mathbf{R}^{2n+1} = \mathbf{R}.E \oplus \mathbf{R}P_i \oplus \mathbf{R}Q_i$$

avec

$$P_i = M(e_i, 0, 0), Q_i = M(0, e_i, 0), E = M(0, 0, 1).$$



FIGURE 5. Werner Heisenberg

Le crochet de l'algèbre de Lie \mathfrak{h}_n de H_n est donc donné par

$$[M(x, y, z), M(x', y', z')] = M(0, 0, \omega(x, y, x', y')),$$

ou encore est *déterminé* par les conditions E central et

$$(IX.2.1) \quad [P_i, P_j] = [Q_i, Q_j] = 0, [P_i, Q_j] = \delta_{i,j}E.$$

Du point de vue groupe, H_n est engendré par les groupes à 1 paramètre

$$\exp(xP_i), \exp(xQ_i), \exp(xE)$$

soumis aux seules relations $\exp(xE)$ central et

$$(IX.2.2) \quad (\exp(xP_i), \exp(yP_j)) = (\exp(xQ_i), \exp(yQ_j)) = \text{Id},$$

$$(\exp(xP_i), \exp(yQ_j)) = \exp(\delta_{i,j}xyE)$$

pour tout $x, y \in \mathbf{R}$. En effet, les relations de commutation (IX.2.2) entraînent immédiatement la formule

$$\prod \exp(xP_i) \prod \exp(yQ_j) \cdot \exp(zE) \prod \exp(x'P_i) \prod \exp(y'Q_j) \cdot \exp(z'E)$$

$$= \prod \exp((x + x')P_i) \prod \exp((y + y')Q_j) \exp((z + z' + \frac{1}{2}\omega(x, y, x', y'))E)$$

qui est la loi de groupe.

IX.2.2. Oscillateur harmonique et représentations de \mathfrak{h}_1, H_1 . — La formule (IX.2.1) montre, au moins formellement que se donner p, q satisfaisant (IX.1.1) n'est rien d'autre que de se donner une représentation

$$\delta : \mathfrak{h}_1 \rightarrow \mathfrak{u}(\mathcal{H})$$

grâce à la règle

$$(IX.2.3) \quad \delta(P_1) = ip, \delta(Q_1) = iq, \delta(E) = i.$$

Une manière équivalente est de définir δ par l'opérateur d'annihilation

$$a = \frac{1}{\sqrt{2}}(q + ip)$$

et l'opérateur de création (son adjoint)

$$a^* = \frac{1}{\sqrt{2}}(q - ip)$$

en respectant la relation

$$[a, a^*] = 1.$$

De même, se donner p, q satisfaisant (IX.2.2), c'est formellement se donner ⁽¹⁾

$$\rho : H_1 \rightarrow \mathbf{U}(\mathcal{H})$$

intégrant δ , c'est-à-dire faisant commuter le diagramme

$$\begin{array}{ccc} \mathfrak{h}_n & \xrightarrow{\delta} & \mathfrak{u}(\mathcal{H}) \\ \exp_{H_1} \downarrow & & \downarrow \exp_{\mathbf{U}(\mathcal{H})} \\ H_1 & \xrightarrow{\rho} & \mathbf{U}(\mathcal{H}) \end{array}$$

ou, plus concrètement, tel que

$$(IX.2.4) \quad \rho(\exp(xP_1)) = \exp(ixp), \rho(\exp(yQ_1)) = \exp(iyq), \rho(\exp(zE)) = \exp(iz).$$

IX.2.3. Représentations de Schrödinger, de Fock. — On pose

$$\mathcal{H} = L^2(\mathbf{R}, \mathbf{C}) \text{ avec } (f, g) = \int_{\mathbf{R}} \overline{f(t)}g(t)dt \text{ et } p = -i\frac{d}{dt}, q = t$$

et on tombe sur la représentation δ_S de \mathfrak{h}_1 dite de Schrödinger.

Dans cette représentation, il y a (à une phase près) un unique état de plus basse énergie (ce qui est conforme à l'expérience), associé à la valeur propre 1 de l'Hamiltonien $H = \text{Id}_{\mathcal{H}} + 2aa^*$: il s'agit de la gaussienne

$$|0\rangle = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$$

1. Évidemment, on ne regarde que les représentations continues, ce qui par locale compacité de H_n équivaut au fait que pour tout $v \in \mathcal{H}$, l'application $g \mapsto \rho(g).v$ de H_n dans \mathcal{H} est continue.



FIGURE 6. Erwin Schrödinger

C'est l'unique (à une phase près) *vecteur de vide* unitaire, à savoir annulé par l'opérateur d'annihilation $a = \frac{1}{\sqrt{2}}(t + \frac{d}{dt})$.

Mais on peut aussi considérer \mathcal{F} l'espace des fonctions holomorphes entières (c'est-à-dire définies et développables en séries entières sur tout le plan complexe de variable $z = x + iy$, $x, y \in \mathbf{R}$), muni du produit hermitien

$$(f, g) = \frac{1}{\pi} \int_{\mathbf{C}} \overline{f(z)} g(z) \exp(-|z|^2) dx dy$$

(exercice : vérifier que l'espace de Fock \mathcal{F} est bien un Hilbert). On pose alors

$$a = \frac{d}{dz}$$

dont l'adjoint a^* est l'opérateur z de multiplication par z définissant une représentation δ_F dite de Fock d'après le paragraphe précédent.



FIGURE 7. Vladimir Fock

Dans cette représentation, « le vecteur de vide » $|0\rangle = 1/\sqrt{\pi}$ est l'unique (à une phase près) état de plus basse énergie $H = 1 + 2a^*a$.

Ces deux représentations semblent très différentes. Il n'en est rien, et, comme on le verra plus bas (Thm. IX.3.5), ce n'est pas un hasard. Pour l'instant, contentons-nous de l'énoncé suivant.

Exercice IX.2.1. — Montrer que $a : \mathcal{F} \rightarrow \mathcal{H}$ défini par

$$a(f)(t) = \frac{1}{\pi} \int_{\mathbf{C}} f(z) \exp\left(-\frac{t^2}{2} + t\bar{z} - \frac{\bar{z}^2}{2}\right) dx dy$$

est une isométrie \mathfrak{h}_1 -équivariante d'inverse

$$a^{-1}(\phi)(z) = \int_{\mathbf{R}} \phi(t) \exp\left(-\frac{t^2}{2} + tz\sqrt{2} - \frac{z^2}{2}\right) dt.$$

Mais comme dit un auteur célèbre, « on ne nous dit pas tout ».



FIGURE 8. Un auteur célèbre

IX.2.4. « On ne nous dit pas tout » : domaine de définition de p, q . — Regardons par exemple la représentation de Schrödinger. L'opérateur p n'est hélas pas défini partout : il vaut mieux définir les dérivées de fonctions qui ont tendance à être dérivables ! Introduisons l'espace de Sobolev

$$H^1 = \{f \in L^2 \text{ tels que } f' \in L^2\}$$

où f' est par exemple la dérivée au sens des distributions.

Une autre manière de voir est d'utiliser la transformée de Fourier \hat{f} de sorte que

$$H^1 = \{f \in L^2 \text{ tels que } x\hat{f} \in L^2\}.$$

On vérifie que H^1 est dense dans \mathcal{H} et que le graphe $\{(f, f', f \in H^1\}$ de p est un fermé du Hilbert $\mathcal{H} \times \mathcal{H}$. Un tel opérateur (défini sur un espace dense à graphe fermé) est dit opérateur fermé. Notons (exercice) que p n'est pas continu, *i.e.* non borné sur la boule unité : on parle d'opérateur non borné. Il est facile de montrer que p n'a pas de prolongement fermé sur un espace plus gros. On est donc sorti du cadre des opérateurs au sens usuel du terme.

La question qui se pose alors est la suivante : peut-on trouver une représentation

$$\mathfrak{h}_1 \rightarrow \mathfrak{u}(\mathcal{H})$$

à valeurs dans les *endomorphismes bornés anti-hermitiens*. La réponse est négative !

Lemme IX.2.2. — *Il n'existe pas de \mathbf{C} -algèbre normée A contenant deux éléments a, b tels que $[a, b] = -i$.*



FIGURE 9. Joseph Fourier



FIGURE 10. Serguei Sobolev

Démonstration. — Une récurrence immédiate donne pour tout $n \in \mathbf{N}$ la relation

$$(IX.2.5) \quad a^n b - b a^n = -i n a^{n-1}.$$

Prenons la norme. Tenant compte de la sous-multiplicativité de la norme, on a

$$n \|a^{n-1}\| \leq 2 \|a\| \|b\| \|a^{n-1}\|.$$

Comme a, b sont nécessairement non nuls, on déduit l'existence de $n > 0$ tel que $a^{n-1} = 0$. Mais la relation (IX.2.5) entraînerait par récurrence descendante la nullité de a , une contradiction avec $[a, b] = -i \neq 0$. \square

La considération d'opérateurs non bornés, *i.e.* non définis partout et/ou non continus est donc nécessaire. On peut alors se demander, dans la mesure où p, q ne sont pas définis partout, ce que signifie $[p, q]$, au cas où q ne serait pas partout défini sur l'image de p ...

IX.2.5. « On ne nous dit pas tout » : définition de $\exp(ixp), \exp(ixq)$. — Usuellement, l'exponentielle d'un endomorphisme continu u est défini par

$$\exp(ixu) = \sum_{n \geq 0} \frac{(ixu)^n}{n!}.$$

Mais dans l'exemple de Schrödinger, u n'est pas défini... On va plutôt partir de l'équation différentielle et caractériser l'exponentielle $\exp(ixu)$ comme étant l'unique solution de

$$f'(x) = i u f(x), \quad f(0) = \text{Id}$$

où $f(x)$ est à valeurs dans $\text{End}(\mathcal{H})$.

Exercice IX.2.3. — Regardons les opérateurs $p = -i \frac{d}{dx}, q = t$ de Schrödinger. Pour calculer $\exp(ixp)(f_0), f_0 \in L^2$, on doit résoudre

$$\frac{\partial f(x, t)}{\partial x} = \frac{\partial f(x, t)}{\partial t}, \quad f(0, t) = f_0$$

autrement dit

$$\exp(ixp)(f_0(t)) = f_0(t + x)$$

qui est un élément bien défini de $\mathbf{U}(\mathcal{H})$ au contraire de p lui-même! De même, pour calculer $\exp(ixq)(f_0), f_0 \in L^2$, on doit résoudre

$$\frac{\partial f(x, t)}{\partial x} = i t f(x, t), \quad f(0, t) = f_0$$

autrement dit

$$\exp(ixq)(f_0) = \exp(ixt)f_0$$

et là encore est un élément bien défini de $\exp(iq) \in \mathbf{U}(\mathcal{H})$ au contraire de q . Un calcul direct donne

$$(\exp(ixp), \exp(iyq)) = \exp(ixy)$$

de sorte qu'on a bien défini une représentation unitaire

$$\rho_S : H_1 \rightarrow \mathbf{U}(\mathcal{H})$$

grâce à (IX.2.2).

Exercice IX.2.4. — Montrer que la représentation de Fock s'intègre en une représentation unitaire $H_1 \rightarrow \mathbf{U}(\mathcal{F})$ qu'on explicitera.

L'exemple précédent suggère que le passage à l'exponentielle améliore la situation. Cette constatation n'a absolument rien d'exceptionnel.

IX.3. Le théorème de Stone - von Neumann



FIGURE 11. Marshall Stone



FIGURE 12. Un lien de famille ?

IX.3.1. Le théorème de Stone. — Commençons par un « rappel ». Soit u un opérateur défini sur un sous-espace dense $\text{dom}(u)$ de \mathcal{H} . Pour tout $x \in \text{dom}(u)$, posons $\phi_y(x) = \langle u(x), y \rangle$: on définit ainsi une forme linéaire, en général pas continue, sur $\text{dom}(u)$. Posons alors

$$\text{dom}(u^*) = \{y \in \mathcal{H} \text{ tels que } \phi_y \text{ continue sur } \text{dom}(u)\}.$$

Si $y \in \text{dom}(u^*)$, la forme ϕ_y se prolonge par continuité en une unique forme continue définie sur $\mathcal{H} = \overline{\text{dom}(u)}$. D'après le théorème de Riesz, elle s'écrit de façon unique $\phi_y(x) = \langle x, u^*(y) \rangle$ définissant sans ambiguïté une application linéaire

$$u^* : \begin{cases} \text{dom}(u^*) & \rightarrow & \mathcal{H} \\ y & \mapsto & u^*(y) \end{cases}$$

On a donc la formule

$$\text{pour tout } x \in \text{dom}(u), y \in \text{dom}(u^*), \langle u(x), y \rangle = \langle x, u^*(y) \rangle.$$

On dit que u est *symétrique* si on a

$$\text{pour tout } x, y \in \text{dom}(u) \quad \langle u(x), y \rangle = \langle x, u(y) \rangle,$$

ce qui entraîne évidemment $\text{dom}(u) \subset \text{dom}(u^*)$. On dit que u est *autoadjoint* s'il est symétrique et si $\text{dom}(u) = \text{dom}(u^*)$.

Exercice IX.3.1. — Montrer que le graphe de u^* est

$$\{(x, y) \in \mathcal{H}^2 \text{ tels que pour tout } z \in \text{dom}(u) \quad \langle y, z \rangle = \langle x, u(z) \rangle\}.$$

En déduire que le graphe de u^* ainsi que de tout opérateur auto-adjoint est fermé.

On a alors le théorème, en un sens assez étonnant, suivant.

Théorème IX.3.2 (Stone). — Soit $u(t) : \mathbf{R} \rightarrow \mathbf{U}(\mathcal{H})$ un groupe à un paramètre continu⁽²⁾. Alors, il existe u'_0 auto-adjoint défini sur un sous-espace dense⁽³⁾ W de \mathcal{H} tel que $u|_W$ soit l'unique solution sur W de

$$(*) \quad f'(x) = iu'_0 f(x), \quad f(0) = \text{Id}.$$

Inversement, si u est autoadjoint, il existe une unique groupe à un paramètre continu $\mathbf{R} \rightarrow \mathbf{U}(\mathcal{H})$ noté $x \mapsto \exp(ixu)$ unique solution sur $\text{dom}(u)$ de (*).

Démonstration. — C'est la réciproque qui nous intéresse. Expliquons heuristiquement pourquoi l'exponentielle améliore la situation. Imaginons, comme dans le cas de la dimension finie, que \mathcal{H} admette une base hilbertienne $e_n, n \geq 0$ de vecteurs propres pour u (c'est le cas pour p, q par exemple, mais c'est faux en général). On note λ_i les valeurs propres correspondantes. Si $x = \sum x_n e_n$ avec $\sum |x_n|^2 < \infty$, $u(x)$ n'est défini que si $\sum |\lambda_n x_n|^2 < \infty$. Mais, comme u autoadjoint, λ_n est réel. Formellement, $\exp(ixu)$ est diagonalisé sur les e_n de valeur propre $\exp(i\lambda_n)$, nombre complexe de module 1. Du coup, si on a $\sum |x_n|^2 < \infty$ alors

$$\sum |\exp(ix\lambda_n)x_n|^2 = \sum |x_n|^2 < \infty$$

ce qui montre que $\exp(iu)$ est bien défini et continu sur tout \mathcal{H} (et solution de l'équation différentielle sur le domaine de u). La preuve dans le cas général est tout à fait abordable et jolie, mais est un peu longue : on renvoie en appendice (IX.6) pour plus de détails. \square

Exercice IX.3.3. — Soit $H = L^2([0, 1], \mathbf{C})$ et u l'opérateur non borné défini par

$$\text{dom}(u) = \{f \in H \text{ continue} \mid f' \in H\} \text{ et } u(f) = if'.$$

Montrer que u est symétrique mais pas autoadjoint. En revanche, montrer que l'opérateur non borné défini par

$$\text{dom}(v) = \{f \in H \text{ continue} \mid f' \in H, f(0) = f(1)\} \text{ et } v(f) = if'$$

est autoadjoint.

3. Au sens où pour tout $v \in \mathcal{H}$, l'application $t \mapsto u(t).v$ est continue.

3. En fait, on montre que $\text{dom}(u)$ est l'ensemble des $v \in \mathcal{H}$ tels que $t \mapsto u(t)v$ est dérivable.

IX.3.2. Irréductibilité de la représentation de Schrödinger. — On sait (lemme de Schur) qu'une représentation (continue) de dimension finie d'un groupe compact $G \rightarrow \mathbf{GL}(V)$ est irréductible si et seulement si les seuls endomorphismes de V commutant à l'action de G sont les homothéties. Ce résultat est encore vrai pour les représentations unitaires à valeurs dans les Hilbert (une représentation sera dite irréductible si elle n'a pas de sous-espace fermé stable non trivial), pour autant qu'on se restreigne aux endomorphismes continus commutant à l'action du groupe.

Proposition IX.3.4. — *La représentation de Schrödinger est irréductible.*

Démonstration. — Soit donc α un endomorphisme de $L^2(\mathbf{R}, \mathbf{C})$ commutant à $\exp(ixp)$, $\exp(ixq)$. En dérivant (soigneusement), on se ramène à prouver que si α commute avec q (sur son domaine de définition) et $\exp(ixp)$, alors α est scalaire. La formule $\alpha(q^n f) = q^n \alpha(f)$, $n \geq 0$ appliquée à $f = \exp(-t^2)$ prouve que pour tout polynôme

$$\alpha(P(t) \exp(-t^2)) = P(t) \alpha(\exp(-t^2)).$$

Posons $\phi(t) = \exp(t^2) \alpha(\exp(-t^2))$ qui est bien défini presque partout et soit $f \in \mathcal{H}$. Comme les fonctions $P(t) \exp(-t^2)$ sont denses dans L^2 (voir cours de tronc commun de mathématiques ou de physique), il existe une suite f_n de telles fonctions convergeant vers f dans L^2 . On a donc, par continuité de α ,

$$\alpha(f) = \lim f_n \phi,$$

la limite étant à comprendre dans L^2 .

Quitte à extraire une sous-suite, on peut supposer que f_n converge simplement presque partout vers f et $f_n \phi$ vers $\alpha(f)$. En regardant ces sous suites, on déduit qu'on a $\alpha(f) = f \phi$ presque partout. Donc, la multiplication par ϕ respecte L^2 et est continue, ce qui entraîne que $\phi \in L^\infty$ et plus précisément que $\|\phi\|_\infty \leq \|\alpha\|$ car si

$$|\phi| \geq \|\alpha\| + \epsilon$$

sur un ensemble de mesure positive E , on aurait

$$\|\alpha(1_E)\|^2 = \int_E |\phi|^2 dt \geq (\|\alpha\| + \epsilon)^2 \|1_E\|^2$$

ce qui est absurde car $\|1_E\| > 0$ par hypothèse.

En utilisant que α commute avec $\exp(ixp)$, on obtient

$$(\phi(t+x) - \phi(t))f(t) = 0$$

pour tout $f \in L^2$ et $x \in \mathbf{R}$ de sorte que $\phi(t+x) = \phi(t)$ p.p. et donc ϕ constante, ce qu'on voulait. \square

IX.3.3. Unicité de la description quantique : le théorème de Stone-von Neumann.

— La question qui se pose est de savoir si l'axiomatique de Heisenberg caractérise la description de l'oscillateur harmonique, autrement dit, la description de Schrödinger est-elle unique à un changement de base unitaire près ?

La réponse est négative (cf. section suivante) si on se contente des relations (IX.2.1), c'est-à-dire qu'on regarde les représentations de \mathfrak{h}_1 , et positive si on pose

convenablement le problème, à savoir si on regarde les relations (IX.2.2), c'est-à-dire qu'on regarde les représentations unitaires de H_1 .

Commençons par quelques remarques formelles.

D'abord, si ρ est une représentation unitaire irréductible de H_1 , le centre commute à l'action de H_1 , et donc agit par homotéthies grâce au lemme de Schur. Autrement dit, il existe un caractère unitaire $\chi : \mathbf{R} \rightarrow \mathbf{U}_1$ tel que $\rho(\exp(xE)) = \chi(x)$ où \mathbf{U}_1 est le cercle unité dans \mathbf{C} (en tant que groupe multiplicatif). Or, c'est un exercice de taupe que de vérifier qu'il existe un unique $\lambda \in \mathbf{R}$ tel que $\chi(x) = \chi_\lambda(x) = \exp(i\lambda x)$. On dit que λ est le *niveau* de la représentation. Physiquement, changer $\lambda \neq 0$, c'est simplement changer la constante de Planck ou si on veut les unités.

La représentation de Schrödinger est de niveau 1. En général, on peut la modifier pour obtenir une représentation ρ_λ de niveau λ par les formules

$$\rho_\lambda(\exp(ixP))(f(t)) = \exp(i\lambda xt)f(t), \rho_\lambda(\exp(ixQ))(f(t)) = f(t+x).$$

La preuve de IX.3.4 s'adapte immédiatement pour montrer l'irréductibilité de ρ_λ pour λ non nul.



FIGURE 13. Max Planck

Bien entendu, lorsqu'on a n degré de libertés, on définit l'espace de Schrödinger par $L^2(\mathbf{R}^n, dt)$ et la représentation ρ_λ est définie « coordonnée par coordonnée » par

$$\rho_\lambda(\exp(ixP_i))(f(\underline{t})) = \exp(i\lambda xt_i)f(\underline{t}), \rho_\lambda(\exp(ixQ_i))(f(\underline{t})) = f(\underline{t} + xe_i)$$

où e_k est la base canonique de \mathbf{R}^n .

La description de Fock est analogue : on regarde l'espace des fonctions holomorphes sur \mathbf{C}^n muni de la mesure $\prod_{k=1}^n \frac{1}{\pi} dx_k dy_k$ et on pose $a_k = \lambda \frac{\partial}{\partial z_k}$ de sorte que $a_k^* f(z) = z_k f(z)$.

Le cas $\lambda = 0$ est le cas de la mécanique classique qui correspond à une représentation de \mathbf{R}^{2n} vu comme le quotient de H_n par son centre qui est de dimension 1 et pas très intéressante de ce point de vue.

Théorème IX.3.5 (von Neumann, 1931). — *Pour tout $\lambda \neq 0$, la représentation de Schrödinger ρ_λ du groupe H_n dans l'espace $L^2(\mathbf{R}^n, dt)$ est, à isomorphisme unitaire près, l'unique représentation unitaire irréductible de niveau λ .*

Il n'est pas question de donner la preuve (cf. [10], ou, plus lisible, [2]) ici - bien soit du niveau de ce cours -.



FIGURE 14. John von Neumann

IX.3.4. Lien entre représentation du groupe et de l'algèbre de Heisenberg. — Bien que H_n soit simplement connexe, on peut se demander si le calcul formel aboutissant à (IX.2.2) est justifié, ce qui entraînerait le théorème de Stone-von Neumann au niveau des algèbres de Lie. **Il n'en est rien.** En effet, on peut même trouver des opérateurs auto-adjoints u, v (non bornés) d'un Hilbert qui commutent sur un espace stable dense D tels que $\exp(ixu)$ et $\exp(iyq)$ ne commutent pas (cf. [7, VIII.5.Ex 1]). Mais on a un résultat positif lorsque que l'hamiltonien est raisonnable (cf. [5] et ses références).

Théorème IX.3.6 (Nelson, 1959). — *Supposons que p, q sont auto-adjoints, laissent stable un sous-espace dense $D \subset \text{dom}(p) \cap \text{dom}(q)$ et vérifient (IX.1.1) sur D . Supposons que l'adhérence du graphe de $H = p^2 + q^2$ sur D soit le graphe d'un opérateur autoadjoint (condition de Nelson). Alors, les relations (IX.2.2) sont vérifiées.*

IX.3.5. La représentation métaplectique. — D'un point de vue sans coordonnée, on a considéré $V = \mathbf{R}^n \oplus \mathbf{R}^n$ muni de sa forme symplectique (i.e. alternée non dégénérée) standard (unique à changement de base près). Le groupe de Heisenberg est alors $H_\omega = \mathbf{R} \oplus V$ muni de la loi

$$(t, e).(t', e) = (t + t' + \frac{1}{2}\omega(e, e'), e + e') \text{ avec } e, e' \in V$$

tandis que son algèbre de Lie $\mathfrak{h}_\omega = \mathbf{R} \oplus V$ muni du crochet

$$(IX.3.1) \quad [(t, e), (t', e)] = (\omega(e, e'), 0)$$

de sorte que $E = (1, 0)$ est une base du centre de \mathfrak{h}_n . L'espace \mathbf{R}^n est simplement un lagrangien L de ω (sous-espace totalement isotrope maximal) et la représentation de Schrödinger $\rho : H_\omega \rightarrow \mathbf{U}(\mathcal{H})$ sur $\mathcal{H} = L^2(L, \mathbf{C})$ est l'unique représentation unitaire de niveau 1. Soit alors $g \in \mathbf{Sp}(V)$, le groupe symplectique de ω . Comme g préserve la forme ω , l'application $(t, e) \rightarrow (t, g(e))$ est un isomorphisme de G préservant le centre de H_ω . On déduit une nouvelle représentation de niveau 1

$$\rho_g : H_\omega \rightarrow \mathbf{U}(\mathcal{H})$$

défini par

$$\rho_g(t, e) = \rho(t, g(e)).$$

Par construction, ρ_g est irréductible de niveau 1. Le théorème de Stone-von Neumann assure qu'il existe un endomorphisme unitaire μ_g bien défini à une phase près (lemme de Schur) identifiant ρ et ρ_g , *i.e.* tel que

$$(IX.3.2) \quad \mu_g(\rho(h).v) = \rho_g(h)\mu_g(v) \text{ avec } h \in H_\omega, v \in \mathcal{H}.$$

Comme $\mu_{g_1} \circ \mu_{g_2}$ et $\mu_{g_1 g_2}$ sont solutions de (IX.3.2), ils coïncident à une phase près d'après le lemme de Schur, donc coïncident dans $\mathbf{U}(\mathcal{H})/\mathbf{U}_1$. On a donc un morphisme de groupes

$$\mu : \mathbf{Sp}(V) \rightarrow \mathbf{U}(\mathcal{H})/\mathbf{U}_1,$$

ce qu'on appelle une représentation *projective* du groupe symplectique. Lorsqu'on a une telle représentation $G \rightarrow \mathbf{U}(\mathcal{H})/\mathbf{U}_1$, on construit formellement un autre groupe de Lie

$$\tilde{G} = \{(g, u) \in G \times \mathbf{U}(\mathcal{H}) \text{ tels que } \rho(g) = u \pmod{\mathbf{U}_1}\}.$$

Il est muni d'un morphisme surjectif sur G (la première projection) dont le noyau s'identifie à \mathbf{U}_1 et est central par construction. On dira que \tilde{G} est une extension centrale de G par \mathbf{U}_1 . La représentation ρ induit trivialement une représentation unitaire

$$\tilde{\rho} : \tilde{G} \rightarrow \mathbf{U}(\mathcal{H}).$$

Dans le cas précédent, $\widetilde{\mathbf{Sp}(V)}$ s'appelle le groupe métaplectique et la représentation correspondante la représentation métaplectique. On peut prouver que $\widetilde{\mathbf{Sp}(V)}$ **n'est pas un groupe linéaire**. Cette construction pour un espace vectoriel réel est due au physicien Shale ([9]) et a ensuite été considérablement développée par Weil (cf. [11]) qui a notamment découvert des liens avec... la théorie des nombres. En fait, il existe un sous-groupe $\mathbf{Sp}_2(V)$ de $\widetilde{\mathbf{Sp}(V)}$ extension centrale de $\mathbf{Sp}(V)$ par ± 1 qui donc par restriction joue le même rôle que le groupe métaplectique (*loc. cit.* page 197, ou, plus lisible, [4]).



FIGURE 15. André Weil

IX.4. Algèbre d'Heisenberg de dimension infinie

IX.4.1. L'oscillateur harmonique à une infinité de degrés de liberté et sa représentation de Fock. — L'algèbre de Heisenberg opère par restriction de la représentation de Fock sur l'espace des polynômes $\mathbf{C}[z_1, \dots, z_n]$ grâce aux opérateurs de création a_k^* de multiplication par z_k et d'annihilation a_k de dérivation par rapport à z_k . Le polynôme constant 1 est le vecteur de vide normalisé $|0\rangle$, autrement dit est tué par tous les opérateurs d'annihilation.

Considérons V_∞ l'espace vectoriel des polynômes en z, z^{-1} sans terme constant. On pose

$$\omega(P, Q) = \text{Res}(Q dP) = \text{Res}(Q(z)P'(z)dz)$$

où $\text{Res}(P(z)dz)$ est le résidu en zéro, à savoir le coefficient du terme en $z^{-1}dz$ de P . Comme $\text{Res}(dP) = 0$, on a

$$\text{Res}(QdP + PdQ) = 0$$

de sorte que ω est alternée. La formule

$$\text{Res}\left(\left(\sum a_k z^k\right)z^{-1-n}dz\right) = a_n$$

prouve que ω est non dégénérée.

Ceci permet de considérer comme plus haut (cf. (IX.3.1)) l'algèbre de Heisenberg

$$\mathfrak{h} = \mathfrak{h}_\infty = \mathbf{R}.E \oplus V_\infty$$

avec E central et

$$[e, e'] = \omega(e, e')E \text{ avec } e, e' \in V_\infty.$$

On a donc

$$[z^n, z^m] = n \text{Res}(z^{n+m-1}dz) = n\delta_{n,m}E$$

qui est bien une algèbre de Lie.

De même que dans le cas $V = \mathbf{R}^n \oplus \mathbf{R}^n$, où l'algèbre \mathfrak{h}_n opérait sur les fonctions holomorphes à n variables, on va construire une représentation (de Fock)

$$\delta : \mathfrak{h}_\infty \rightarrow \text{End}(\mathcal{F}_\infty)$$

sur les polynômes sur « la moitié » de V_∞ . Posons donc

$$\mathcal{F}_\infty = \mathbf{C}[z_1, \dots, z_n, \dots],$$

l'algèbre de polynômes à une infinité de variables $z_n, n > 0$. On a

$$\mathcal{F}_\infty = \mathbf{C}[z_1, \dots, z_n, \dots] \xrightarrow{\sim} \text{Sym}L^+$$

où L^+ est le lagrangien

$$L^+ = \langle z^n, n > 0 \rangle$$

de \mathfrak{h}_∞ , ce qui est la version algébrique de $\mathcal{F} = \{\text{fonctions entières sur } \mathbb{R}^n\}$ dans le cas de la dimension finie (les polynômes sont denses dans \mathcal{F} pour la norme L^2).

En s'inspirant de la dimension finie, on pose

$$\delta(z^n) = \frac{\partial}{\partial z_n}, \delta(z^{-n}) = nz_n, n > 0, \delta(E) = 1.$$

Ces formules définissent un morphisme d'algèbre de Lie

$$\delta_\infty : \mathfrak{h}_\infty \rightarrow \text{End}(\mathcal{F}_\infty),$$

la représentation de Fock de \mathfrak{h}_∞ . Notons que le polynôme constant $|0\rangle = 1 \in \mathcal{F}_\infty$ est un vecteur de vide, c'est-à-dire est tué par tous les opérateurs d'annihilation $\delta(z^n), n > 0$. La représentation est de niveau 1, à savoir $\delta(E) = \text{Id}$.

Remarque IX.4.1. — Le lecteur esthète décrira cette action indépendamment des coordonnées.

Cette situation algébrique⁽⁴⁾ est en tout point analogue à ce qui précède.

Proposition IX.4.2. — *La représentation de Fock de \mathfrak{h}_∞ est irréductible, engendrée par le vecteur de vide $|0\rangle$ comme \mathfrak{h}_∞ -module. Le lemme de Schur est valable dans cette situation. De plus (forme faible de Stone-von Neumann), toute représentation irréductible de \mathfrak{h}_∞ engendrée par un vecteur de vide (annulé par les $z^k, k > 0$) et de niveau 1 est isomorphe à la représentation de Fock.*

Démonstration. — Soit $P = P[z_1, \dots, z_n] \in \mathcal{F}_\infty$ non nul contenu dans un espace stable S . Par applications successives de $\delta(z^n)$, on peut supposer que P n'a pas de terme faisant apparaître z_n . Par récurrence sur n , on peut supposer que P est une constante non nulle. Mais alors par application successive de $\delta(z^{-k}, k > 0)$, on obtient que tous les monômes sont dans S et donc $S = \mathcal{F}_\infty$.

Le même argument prouve que $|0\rangle$ engendre.

Comme

$$\bigcap_{k>0} \ker(\delta(z^k)) = \mathbf{C}|0\rangle,$$

un morphisme f de \mathcal{F}_∞ commutant à δ préserve $\mathbf{C}|0\rangle$: soit $\lambda \in \mathbf{C}$ tel que $f(|0\rangle) = \lambda|0\rangle$. Alors, $\text{Ker}(f - \lambda)$ est stable par \mathfrak{h}_∞ et non nul, c'est donc \mathcal{F}_∞ , ce qu'on voulait.

Si maintenant \mathcal{F}' est irréductible engendrée par $|0'\rangle$, on définit

$$f(P(z_1, \dots, z_n)) = f(P(\delta(z^{-1}), \dots, \delta(z^{-n})|0\rangle)) = P(\delta'(z^{-1}), \dots, \delta'(z^{-n})|0'\rangle).$$

On vérifie que f commute à l'action de \mathfrak{h}_∞ . Comme f est non nulle, son noyau est, qui est invariant par \mathfrak{h}_∞ , est nul (irréductibilité de \mathcal{F}). De même, son image est \mathcal{F}' par irréductibilité. \square

IX.4.2. Action de \mathbf{C}^* . — Si on change z en $\lambda^{-1}z, \lambda \in \mathbf{C}^*$, le coefficient de $z^{-1}dz$ reste inchangé et donc le résidu de $P(z)dz$ l'est aussi. On déduit que l'action induite sur $V_\infty = \mathbf{C}[z^{\pm 1}]$ préserve ω et donc induit une transformation symplectique de V_∞ .

On déduit comme dans la section précédente une action de \mathbf{C}^* sur \mathfrak{h}_∞ et donc pour tout $\lambda \in \mathbf{C}$ une nouvelle représentation définie par

$$\delta_\lambda(P(z)).v = \delta(P(\lambda^{-1}z)).v, P(z) \in \mathbf{C}[z^\pm], v \in \mathcal{F}_\infty.$$

Comme $\mathbf{C}[z]$ est préservé, $|0\rangle$ est encore un vecteur de vide pour δ_λ et la représentation est encore de niveau 1. Grâce à (IX.4.2), on déduit comme plus haut l'existence d'un automorphisme

$$\mu_\lambda : \mathcal{F} \rightarrow \mathcal{F},$$

a priori seulement bien défini à un scalaire près, identifiant δ_λ et δ , c'est-à-dire vérifiant

$$(IX.4.1) \quad \mu_\lambda(P(z)v) = P(\lambda^{-1}z)\mu_\lambda(v), P(z) \in \mathbf{C}[z^\pm], v \in \mathcal{F}.$$

4. Si on veut vraiment une situation d'action sur un Hilbert, il faut un peu compléter. C'est technique et n'apporte pas grand-chose pour comprendre. Le lecteur intéressé pourra aller consulter le très beau livre [6].

Comme μ_λ doit envoyer $|0\rangle$ sur un de ses multiples, on peut normaliser μ_λ en lui imposant

$$\mu_\lambda(|0\rangle) = |0\rangle.$$

Dans ces conditions, on a donc une représentation $\mathbf{C}^* \rightarrow \mathbf{GL}(\mathcal{F})$ définie par $\lambda \rightarrow \mu_\lambda$. La formule (IX.4.1) et le fait que $|0\rangle$ engendre \mathcal{F} comme \mathfrak{h}_∞ module prouve que pour tout v , on peut écrire $\mu_\lambda(v)$ comme un polynôme en les z_i à coefficients dans $\mathbf{C}[\lambda]$. Comme μ_λ est polynomial en λ , on a au niveau des algèbres de Lie une flèche

$$\mathbf{C} \rightarrow \text{End}(\mathcal{F})$$

définie par l'image L_0 de 1 obtenue en écrivant $\lambda = 1 + \epsilon$ et en regardant à l'ordre 1 en ϵ . Écrivons donc

$$\mu_{1+\epsilon} = 1 + \epsilon L_0 + O(\epsilon)$$

et remplaçons dans (IX.4.1). On trouve alors

$$(1 + \epsilon L_0)(P(z).v) = (P(z) - \epsilon z \frac{dP}{dz}).(1 + \epsilon L_0)(v) + O(\epsilon)$$

et donc

$$L_0(P(z).v) = P(z).L_0(v) - z \frac{dP}{dz}.v.$$

Ainsi, L_0 vérifie

$$(IX.4.2) \quad [L_0, \delta(P)] = \delta(-z \frac{dP}{dz}), \quad P(z) \in \mathbf{C}[z^\pm].$$

IX.4.3. Action par transformation holomorphe. — On a utilisé dans la section précédente le fait que la transformation $z \mapsto \lambda z$ préservait $\mathbf{C}[z]$ et le résidu pour définir une action de \mathbf{C}^* puis L_0 .

Généralement, on a envie de faire la même chose en utilisant une transformation $z \mapsto h(z)$ inversible préservant l'origine. L'idéal serait qu'elle soit polynomiale. Hélas, le théorème de d'Alembert interdit cette possibilité dès que h est de degré > 1 (exercice : pourquoi ?). L'idée est d'autoriser h à être holomorphe et de développer h en série entière. Mais du coup, il faut un peut agrandir \mathfrak{h}_∞ pour autoriser des séries de Laurent car $P(h(z))$ ne sera plus un polynôme en z^{-1} , z mais une série. On définit simplement

$$\hat{\mathfrak{h}} = \mathbf{C}.E \oplus \mathbf{C}((z))$$

où $\mathbf{C}((z)) = \mathbf{C}[[z]][z^{-1}]$ est l'anneau (en fait le corps) des séries de Laurent formelles. Si $h(z) = \lambda z + O(z^2)$, on a

$$z^{-n} = \lambda^{-n} z^{-n} (1 + zO(z))^{-n} = \lambda^{-n} z^{-n} (1 - nzO(z) + \dots) \in \mathbf{C}((z))$$

de sorte que

$$P(z) \in \mathbf{C}((z)) \Rightarrow P(h(z)) \in \mathbf{C}((z)).$$

Le résidu est toujours défini sur $\mathbf{C}((z))dz$ ce qui permet de définir le crochet de Lie par E central et

$$[P, Q] = \text{Res}(QdP)E.$$

Comme pour tout $v \in \mathcal{F}_\infty = \mathbf{C}[z_1, \dots, z_n, \dots]$, on a $z^n.v = \frac{\partial}{\partial z_n}v = 0$ pour $n \gg 0$, l'algèbre de Lie $\hat{\mathfrak{h}}$ agit encore sur \mathcal{F}_∞ puisque seuls un nombre fini de termes de la série

$$P(z) = \sum_{n > -N} a_n z^n$$

contribuent donc au calcul de $P(z).v$. Bien entendu, comme $\mathfrak{h}_\infty \subset \hat{\mathfrak{h}}$, les analogues des énoncés (IX.4.2) sont encore vrais pour $\hat{\mathfrak{h}}$.

Pour qu'une transformation holomorphe

$$z \mapsto h(z), h(0) = 0$$

agisse sur $\hat{\mathfrak{h}}$ par reparamétrisation

$$P(z) \mapsto P(h(z)).$$

encore faut-il vérifier qu'on a invariance du résidu afin qu'elle soit symplectique.

Lemme IX.4.3. — Soit $h(z) = \sum_{n \geq 1} a_n z^n, a_1 \neq 0$ une série entière. Pour tout $P \in \mathbf{C}((z))$, on a

$$(*) \quad \text{Res}(P(h(z))dh(z)) = \text{Res}(P(z)dz).$$

Démonstration. — Si $P \in \mathbf{C}[[z]]$, les deux membres de (*) sont nuls. Par linéarité, il suffit de le vérifier pour $P = z^m, m < 0$. Si $m < -1$, on a

$$h(z)^m h'(z) dz = \frac{1}{m+1} dh(z)^{m+1}$$

de résidu nul de sorte que les deux membres de (*) sont nuls. Si $m = -1$, on a

$$h(z)^m h'(z) dz = \frac{\lambda + O(z)}{\lambda z + O(z^2)} = \frac{1 + O(z)}{z + O(z^2)} = \frac{1}{z} + O(1)$$

de sorte que les deux membres ont résidu 1. □

Soit Γ l'ensemble des séries entières de la forme

$$h(z) = \sum_{n \geq 1} a_n z^n, a_1 \neq 0,$$

supposées convergentes⁽⁵⁾ pour simplifier, pensées comme des applications holomorphes définies au voisinage de zéro et \mathbf{C} préservant l'origine. Le théorème d'inversion locale assure l'existence pour tout h dans Γ d'une série convergente

$$h^{-1}(z) \in \Gamma$$

tel que

$$h(h^{-1}(z)) = h^{-1}(h(z)) = z$$

au voisinage de l'origine, et donc avec l'égalité correspondante des séries.

L'action de Γ sur $\mathbf{C}((z))$

$$P \mapsto P(h^{-1}(z))$$

5. c'est-à-dire de rayon de convergence non nul.

s'étend en une action sur $\hat{\mathfrak{h}}$ grâce au lemme précédent. On a donc un groupe Γ qui agit sur $\hat{\mathfrak{h}}$. Comme dans le cas $\lambda \in \mathbf{C}^*$ de la section précédente, on peut « tordre » l'action de $\hat{\mathfrak{h}}$ sur \mathcal{F}_∞ grâce à $\mathfrak{h} \in \Gamma$ et ainsi définir comme plus haut

$$\delta_h : \hat{\mathfrak{h}} \rightarrow \text{End}(\mathcal{F}_\infty)$$

par la règle⁽⁶⁾

$$\delta_h(P(z)).v = \delta(P(h^{-1}(z))).v, \quad P(z) \in \mathbf{C}[[z]], \quad v \in \mathcal{F}_\infty.$$

Comme l'action de Γ préserve par construction $\mathbf{C}[[z]]$, le vecteur de vide $|0\rangle$ est encore annulé par $\delta_h(z\mathbf{C}[[z]])$. Comme plus haut, (IX.4.2) permet de définir $\mu_h : \mathcal{F}_\infty \rightarrow \mathcal{F}_\infty$ normalisé par $\mu_h(|0\rangle) = |0\rangle$ identifiant δ_h et δ , c'est-à-dire vérifiant

$$(IX.4.3) \quad \mu_h(P(z)v) = P(h^{-1}z)\mu_h(v), \quad P(z) \in \mathbf{C}[z^\pm], \quad v \in \mathcal{F}_\infty.$$

On a donc obtenu une représentation

$$\mu : \begin{cases} \Gamma & \rightarrow \mathbf{GL}(\mathcal{F}_\infty) \\ h & \mapsto \mu_h \end{cases}$$

On peut montrer que, dans un sens convenable, l'algèbre de Lie \mathfrak{L}^+ de Γ s'identifie à l'algèbre des champs de vecteurs holomorphes à coefficients polynomiaux de base $d_n = -z^{n+1} \frac{d}{dz}$, $n \geq 0$ avec comme crochet celui des champs des vecteurs, à savoir

$$[d_n, d_m]_0 = (n - m)d_{n+m}$$

de sorte qu'on doit avoir un morphisme d'algèbre de Lie

$$d\mu : \mathfrak{L}^+ \rightarrow \text{End}(\mathcal{F}_\infty).$$

Pour éviter trop d'abstraction inutile, voyons comment faire apparaître les endomorphismes

$$L_n, n \geq 0$$

images de d_n par $d\mu$. On regarde pour $n \geq 0$ l'élément

$$h_\epsilon^{-1}(z) = z - \epsilon z^{n+1}$$

(si $n = 0$ on suppose $\epsilon \neq 1$, ce qui n'aura pas d'influence car on va développer formellement en ϵ , c'est-à-dire si on veut à ϵ petit). En regardant à l'ordre 1 la relation (IX.4.3) après avoir écrit

$$\mu_{h_\epsilon} = 1 + \epsilon L_n,$$

on trouve exactement comme plus haut la relation

$$L_n(P(z)).v = P(z).L_0(v) - z^{n+1} \frac{dP}{dz}.v$$

c'est-à-dire

$$(IX.4.4) \quad [\delta(z^m), L_n] = m\delta(z^{n+m}), \quad m \in \mathbf{Z}.$$

Remarque IX.4.4. — La formule (IX.4.4) caractérise L_n puisqu'elle traduit la relation d'entrelacement (IX.4.3).

6. On est contraint de prendre l'inverse de h pour avoir une action à gauche.

Comme μ est un morphisme de groupes, $d\mu$ doit être un morphisme d'algèbre de Lie, c'est-à-dire qu'on doit avoir

$$(IX.4.5) \quad [L_n, L_m] = (n - m)L_{n+m} \text{ pour tout } n, m \geq 0$$

IX.5. Algèbre de Virasoro

IX.5.1. La formule de Sugawara. — Il est clair que la formule (IX.4.4) permet de calculer $L_n, n \geq 0$ de proche en proche en partant de $L_n(|0\rangle) = |0\rangle$. Il se trouve qu'il existe une formule remarquable permettant de les calculer. Commençons par introduire l'ordre normal sur les

$$a_n \stackrel{\text{déf}}{=} \delta(z^n), \quad n \in \mathbf{Z}^* \text{ et } a_0 = 0$$

(dans ce qui suit, on aurait pu supposer a_0 égal à n'importe quel homotétrie) de sorte qu'on a

$$[a_m, a_n] = m\delta_{m,-n}$$

puisque δ est de niveau 1.

Avec ces notations, (IX.4.4) devient

$$(IX.5.1) \quad [a_m, L_n] = ma_{n+m}.$$

On pose

$$: a_i a_j := \begin{cases} a_i \circ a_j & \text{si } i \leq j \\ a_j \circ a_i & \text{sinon.} \end{cases}$$

On utilisera sans plus de précaution la formule triviale

$$[a, bc] = [a, b]c + b[a, c]$$

valable pour 3 endomorphisme a, b, c d'un espace vectoriel quelconque.

Proposition IX.5.1 (Sugawara). — Pour tout $v \in \mathcal{F}_\infty$, l'ensemble des $j \in \mathbf{Z}$ tels que

$$: a_{-j} a_{j+n} : (v) \neq 0$$

est fini et on a $L_n = \frac{1}{2} \sum_j : a_{-j} a_{j+n} :$

Démonstration. — D'après la remarque IX.4.4, il suffit de vérifier que $\frac{1}{2} : a_{-j} a_{j+n} :$ satisfait les relations (IX.4.4). Prouver la proposition par un calcul brutal n'est pas difficile, mais on peut aller très vite en utilisant une idée chère aux physiciens. Considérons l'expression formelle

$$\Lambda_n = \frac{1}{2} \sum_j a_{-j} a_{j+n},$$

qui elle a un nombre infini de termes non nuls lorsqu'on cherche à l'évaluer sur un élément de \mathcal{F}_∞ . La remarque est que la différence

$$: a_i a_j : - a_i a_j$$

est soit nul soit égal à $[a_i, a_j]$ et donc **est toujours central**. Ainsi, on a formellement

$$L_n = \Lambda_n + N \text{Id}$$

où N est une constante... infinie, ou plutôt une somme infinie

$$N = \sum_I c_i$$

de constantes c_i . Évidemment, ceci ne veut rien dire, sauf si on calcule un crochet $[a, L_n]$ puisqu'alors on a

$$[a, \sum_I c_i \text{Id}] = \sum_I c_i [a, \text{Id}] = \sum_I 0 = 0$$

de sorte qu'on a

$$[a, L_n] = \frac{1}{2} \sum_j [a, a_{-j} a_{j+n}] \ll = \gg [a, \Lambda_n]$$

Et ce raisonnement est en fait... parfaitement correct même si expliqué de façon un peu provocatrice. On a donc

$$\begin{aligned} [a_k, L_n] &= \frac{1}{2} \sum_j [a_k, a_{-j} a_{j+n}] \\ &= \frac{1}{2} \sum_j [a_k, a_{-j}] a_{j+n} + \frac{1}{2} \sum_j a_{-j} [a_k, a_{j+n}] \\ &= \frac{1}{2} k a_{k+n} + \frac{1}{2} k a_{k+n} \\ &= k a_{k+n} \end{aligned}$$

ce qu'on voulait. □

Comme on voit, la preuve n'est pas difficile. C'est trouver la formule qui était aussi difficile qu'inattendu. Outre la possibilité de calculer les L_n apparaît un autre phénomène : les calculs n'utilisent pas la positivité de n .

Définition IX.5.2. — Pour tout $n \in \mathbf{Z}$, on définit l'endomorphisme de L_n de \mathcal{F}_∞ par

$$L_n = \frac{1}{2} \sum_j : a_{-j} a_{j+n} : .$$

Ainsi, pour tout n la formule (IX.4.4) est valable.

IX.5.2. La représentation de l'algèbre de Virasoro. — On va définir une nouvelle algèbre de Lie, dite de Virasoro, en deux étapes.

D'abord, on considère l'algèbre de Lie \mathcal{L} des champs de vecteurs *méromorphes* base

$$d_n = -z^{n+1} \frac{d}{dz}, n \in \mathbf{Z}$$

avec comme plus haut le crochet celui des champs des vecteurs, à savoir

$$[d_n, d_m]_0 = (n - m) d_{n+m}.$$

On définit ensuite une extension centrale $\mathfrak{Vir} = \mathbf{C}.c \oplus \mathcal{L}$ de \mathcal{L} par $\mathbf{C}.c$ par les règles c central et

$$(IX.5.2) \quad [d_n, d_m] = (n - m) d_{n+m} + \delta_{n,-m} \frac{m^3 - m}{12} c.$$



FIGURE 16. Miguel Virasoro

Lemme IX.5.3. — \mathfrak{Vir} est une algèbre de Lie.

Démonstration. — On doit vérifier l'identité de Jacobi, autrement dit

$$[[d_p, d_q], d_r] + [[d_q, d_r], d_p] + [[d_r, d_p], d_q] = 0$$

pour tout $p, q, r \in \mathbf{Z}$. C'est un calcul. \square

Notons que \mathfrak{L}^+ est bien une sous-algèbre de Lie de \mathfrak{Vir} puisque la partie centrale s'annule sur \mathfrak{L}^+ . En fait, et c'est un point clef de la théorie des cordes, la représentation $\mathfrak{L}^+ \rightarrow \text{End}(\mathcal{F}_\infty)$ se prolonge à \mathfrak{Vir} . Précisément, on a

Proposition IX.5.4 (Sugawara). — Pour tout $n \in \mathbf{Z}$, les opérateurs L_n vérifient

$$[L_n, L_m] = (n - m)L_{n+m} + \delta_{n,-m} \frac{m^3 - m}{12} \text{Id}$$

de sorte que l'application linéaire

$$\begin{cases} \mathfrak{Vir} & \rightarrow \text{End}(\mathcal{F}_\infty) \\ d_n & \mapsto L_n \\ c & \mapsto \text{Id} \end{cases}$$

est un morphisme d'algèbre de Lie.

Démonstration. — On utilise la même astuce pour se débarrasser de l'ordre normal (pourtant indispensable pour définir les L_n sans sommer une infinité de termes non nuls!). On écrit d'abord, avec la convention d'écriture que les sommes en jeu sont des sommes finies une fois évaluées sur $v \in \mathcal{F}_\infty$ donné, l'ensemble d'indice étant aussi grand qu'on veut dans \mathbf{Z} .

$$(IX.5.3) \quad [L_m, L_n] = \frac{1}{2} \sum_j [a_{-j} a_{j+m}, L_n]$$

$$(IX.5.4) \quad = \frac{1}{2} \sum_j a_{-j} [a_{j+m}, L_n] + \frac{1}{2} \sum_j [a_{-j}, L_n] a_{j+m}$$

$$(IX.5.5) \quad = \frac{1}{2} \sum_j (j + m) a_{-j} a_{j+m+n} - \frac{1}{2} \sum_j j a_{-j+n} a_{j+m}$$

On va remettre dans l'ordre normal. Dans la première somme, on sépare les indices j tels que

- 1) $-j \leq j + m + n$ pour lesquels l'ordre est l'ordre normal des indices j tels que
 2) $-j > j + m + n$. Pour ceux là, on observe que a_{-j} et a_{j+m+n} commutent sauf si $m + n = 0$ de sorte qu'on a

$$a_{-j}a_{j+m+n} =: a_{-j}a_{j+m+n} : + \delta_{n,-m}[a_{-j}, a_{j+m+n}] =: a_{-j}a_{j+m+n} : - j\delta_{n,-m}.$$

Mais si $n + m$ est nul, $j < 0$ de sorte qu'on a

$$\frac{1}{2} \sum_j (j + m) a_{-j} a_{j+m+n} = \frac{1}{2} \sum_j (j + m) : a_{-j} a_{j+m+n} : - \frac{1}{2} \sum_{j < 0} j(j + m) \delta_{n,-m}.$$

Exactement de la même manière dans la seconde somme, on obtient

$$\frac{1}{2} \sum_j j a_{-j+n} a_{j+m} = \frac{1}{2} \sum_j j : a_{-j+n} a_{j+m} : - \frac{1}{2} \sum_{j \geq -m} j(j + m) \delta_{n,-m}$$

où l'on rappelle que les sommes sont finies, mais sommées sur un intervalle de \mathbf{Z} aussi grand qu'on veut.

En additionnant et en changeant d'indice $j \mapsto j + n$ dans la seconde somme, on obtient

$$[L_m, L_n] = (m - n)L_{m+n} - \frac{1}{2} \sum_{j < -m}^{-1} j(j + m) \delta_{n,-m}.$$

On calcule alors $-\frac{1}{2} \sum_{j < -m}^{-1} j(j + m) = (m^3 - m)/12$ pour conclure. \square

Le lecteur intéressé par ce sujet, notamment pour le lien entre cette théorie et la correspondance bosons-fermions, pourra consulter [3], référence de laquelle les derniers présentés calculs sont essentiellement issus.

IX.5.3. Remarques finales. — On pourrait penser que l'action de la partie négative de \mathfrak{Vir} engendrée par les $d_n, n \leq 0$ devrait s'obtenir aussi simplement que celle de la partie positive, simplement en changeant z en z^{-1} . C'est plus compliqué que cela, car cette opération ne préserve pas $\mathbf{C}[z]$ et donc on n'a plus préservation du vecteur de vide. On pourrait obtenir une représentation de cette manière, mais il faudrait montrer qu'on a encore un vecteur de vide pour l'action tordue, ce qui est lourd. C'est pourquoi l'on a procédé de cette manière.

Par ailleurs, comme on le voit, ce n'est pas \mathfrak{L} qui agit, mais \mathfrak{Vir} , à savoir une extension centrale avec action scalaire du centre. Autrement dit, on a une action projective de \mathfrak{L} : c'est une situation en tout point analogue à la situation métaplectique.

On peut se demander si l'étrange facteur central $\frac{m^3 - m}{12} \delta_{m,-n}$ a une propriété particulière, autrement dit s'il y a d'autres extensions centrales non triviales de \mathfrak{L} par \mathbf{C} . En fait, il n'y en a essentiellement qu'une (voir [3, 1.3]).

IX.6. Appendice

On veut donner ici quelques détails de plus sur la définition de $\exp(itu)$ dans le cas auto-adjoint. La clef est d'approximer u par des opérateurs continus auto-adjoints, qui donc eux ont une exponentielle simplement définie par la série entière habituelle, puis de passer à la limite. Il n'est pas question de donner la preuve dans son intégralité, qui est assez longue quoique élémentaire, mais de donner les idées clefs.

Soit donc u auto-adjoint.

Lemme IX.6.1. — Soit $\lambda \in \mathbf{C} - \mathbf{R}$. Alors $u + \lambda$ est une bijection de D sur \mathcal{H} . Son inverse $(u + \lambda)^{-1}$ est continu de norme $\leq 1/|\Im(\lambda)$.

Démonstration. — Comme $(u(x), x) = (x, u(x)) = \overline{(u(x), x)}$, on a $(u(x), x) \in \mathbf{R}$. On déduit que ses valeurs propres sont réelles et donc $u + \lambda$ est injectif. Montrons la surjectivité.

On a donc

$$(IX.6.1) \quad |(u + \lambda)x, x| \geq |\Im((u + \lambda)x, x)| = |\Im(\lambda)|(x, x)$$

de sorte que (Cauchy-Schwarz)

$$\|(u + \lambda)x\| \geq |\Im(\lambda)| \|x\|.$$

Montrons que l'image I_λ de $(u + \lambda)$ est fermée. Soit donc $y_n = (u + \lambda)(x_n) \in I_\lambda$ convergeant vers $y \in \mathcal{H}$. On a donc

$$\|x_n - x_m\| \leq \frac{\|y_n - y_m\|}{|\Im(\lambda)|}$$

prouvant que x_n est de Cauchy, donc converge vers $x \in \mathcal{H}$. On déduit que la suite $(x_n, u(x_n) = y_n - \lambda x_n)$ du graphe de u converge vers $(x, y - \lambda x)$. Comme le graphe est fermé, on a $y - \lambda x = u(x)$ et donc $y \in I_\lambda$, ce qu'on voulait. Montrons que I_λ est dense. Soit donc y dans l'orthogonal. La formule

$$\langle u(x) + \lambda(x), y \rangle = 0$$

prouve que $x \mapsto \langle u(x), y \rangle$ est continue sur $\text{dom}(u)$ donc $y \in \text{dom}(u^*) = \text{dom}(u)$ puisque u autoadjoint. On a alors

$$\langle x, u(y) + \bar{\lambda}(y) \rangle = 0 \text{ pour tout } x \in \text{dom}(u)$$

et donc $u(y) = -\bar{\lambda}y$ par densité de $\text{dom}(u)$. Comme les valeurs propres de u sont réelles, y est nul puisque $\bar{\lambda} \notin \mathbf{R}$ de sorte que $u + \lambda$ est bijectif. Ainsi, $I_\lambda = \mathcal{H}$ puisque dense et fermé. L'inégalité (IX.6.1) se réécrit en posant $x = (u + \lambda)^{-1}(y)$

$$\|(u + \lambda)^{-1}(y)\| \leq \frac{\|y\|}{|\Im(\lambda)|}$$

ce qui achève la preuve du lemme. □

L'idée est alors de considérer

$$u_\epsilon = \frac{(1 + i\epsilon u)^{-1} - 1}{i\epsilon}.$$

En développant formellement à l'ordre 1 en ϵ , on s'attend à $\lim_{\epsilon \rightarrow 0} u_\epsilon = u$, ce qui est bien le cas au sens où pour tout $v \in \text{dom}(u)$, on a

$$\lim_{\epsilon \rightarrow 0} u_\epsilon(v) = u(v).$$

On montre alors que les exponentielles usuelles $\exp(itu_\epsilon)(v)$ convergent pour tout $v \in \mathcal{H}$ vers un vecteur noté $\exp(itu)(v)$ qui a les bonnes propriétés. Pour les (longues) vérifications, adapter [1, VII.7].

BIBLIOGRAPHIE

- [1] H. Brezis. *Analyse fonctionnelle*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983. Théorie et applications. [Theory and applications].
- [2] F. Bruhat. *Représentations des groupes localement compacts, Notes de cours polycopiées*. ENS, 45 rue d'Ulm, Paris.
- [3] V. G. Kac and A. K. Raina. *Bombay lectures on highest weight representations of infinite-dimensional Lie algebras*, volume 2 of *Advanced Series in Mathematical Physics*. World Scientific Publishing Co. Inc., Teaneck, NJ, 1987.
- [4] G. Lion and M. Vergne. *The Weil representation, Maslov index and theta series*, volume 6 of *Progress in Mathematics*. Birkhäuser Boston, Mass., 1980.
- [5] N. S. Poulsen. On the canonical commutation relations. *Math. Scand.*, 32 :112–122, 1973.
- [6] A. Pressley and G. Segal. *Loop groups*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 1986. Oxford Science Publications.
- [7] M. Reed and B. Simon. *Methods of modern mathematical physics. I*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, second edition, 1980. Functional analysis.
- [8] D. Renard. *Groupes et représentations*. Editions de l'Ecole Polytechnique, Palaiseau, 2010.
- [9] D. Shale. Linear symmetries of free boson fields. *Trans. Amer. Math. Soc.*, 103 :149–167, 1962.
- [10] J. von Neumann. Die Eindeutigkeit der Schrödingerschen Operationen. *Math. Annalen.*, 1931.
- [11] A. Weil. Sur certains groupes d'opérateurs unitaires. *Acta Math.*, 111 :143–211, 1964.