



M2/International Centre for Fundamental Physics
Parcours de Physique Théorique

Invariances en physique et théorie des groupes

Jean-Bernard Zuber



Quelques contributeurs à la théorie des groupes mentionnés dans la première partie de ce cours

Avertissement

Le chapitre 0 couvre essentiellement le cours de “prérentrée”.

Les chapitres 1 à 5 suivent fidèlement le contenu de mon cours proprement dit. Ils contiennent aussi dans des paragraphes en petits caractères et dans des appendices quelques compléments non traités en cours.

Bibliographie générale

- [BC] N.N. Bogolioubov et D.V. Chirkov, *Introduction à la théorie quantique des champs*, Dunod.
- [BDm] J.D. Bjorken and S. Drell : *Relativistic Quantum Mechanics*, McGraw Hill.
- [BDf] J.D. Bjorken and S. Drell : *Relativistic Quantum Fields*, McGraw Hill.
- [Bo] N. Bourbaki, *Groupes et Algèbres de Lie*, Chap. 1-9, Hermann 1960-1983.
- [Bu] D. Bump, *Lie groups*, Series “Graduate Texts in Mathematics”, vol. **225**, Springer 2004.
- [DFMS] P. Di Francesco, P. Mathieu et D. Sénéchal, *Conformal Field Theory*, Springer,
- [DNF] B. Doubrovine, S. Novikov et A. Fomenko, *Géométrie contemporaine*, 3 volumes, Éditions de Moscou 1982, réédité en anglais par Springer.
- [FH] W. Fulton and J. Harris, *Representation Theory*, Springer.
- [Gi] R. Gilmore, *Lie groups, Lie algebras and some of their applications*, Wiley.
- [Ha] M. Hamermesh, *Group theory and its applications to physical problems*, Addison-Wesley
- [IZ] C. Itzykson et J.-B. Zuber, *Quantum Field Theory*, McGraw Hill 1980 ; Dover 2006.
- [Ki] A.A. Kirillov, *Elements of the theory of representations*, Springer.
- [LL] L. Landau et E. Lifschitz, *Théorie du Champ*, Editions Mir, Moscou ou *The Classical Theory of Fields*, Pergamon Pr.
- [M] A. Messiah, *Mécanique Quantique*, 2 tomes, Dunod.
- [OR] L. O’ Raifeartaigh, *Group structure of gauge theories*, Cambridge Univ. Pr. 1986.
- [PS] M. Peskin and D.V. Schroeder, *An Introduction to Quantum Field Theory*, Addison Wesley.
- [Po] L.S. Pontryagin, *Topological Groups*, Gordon and Breach, 1966.
- [St] S. Sternberg, *Group theory and physics*, Cambridge University Press.
- [W] H. Weyl, *Classical groups*, Princeton University Press.
- [Wf] S. Weinberg, *The Quantum Theory of Fields*, vol. 1, 2 and 3, Cambridge University Press.
- [Wg] S. Weinberg, *Gravitation and Cosmology*, John Wiley & Sons.
- [Wi] E. Wigner, *Group Theory and its Applications to Quantum Mechanics*. Academ. Pr. 1959.
- [Z-J] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, Oxford Univ. Pr.

Table des matières

0	Quelques éléments de base sur les groupes $SO(3)$, $SU(2)$ et $SL(2, \mathbb{C})$	1
0.1	Rotations de \mathbb{R}^3 , les groupes $SO(3)$ et $SU(2)$	1
0.1.1	Le groupe $SO(3)$, groupe à trois paramètres	1
0.1.2	Du groupe $SO(3)$ au groupe $SU(2)$	3
0.2	Générateurs infinitésimaux. L'algèbre de Lie $\mathfrak{su}(2)$	4
0.2.1	Générateurs infinitésimaux de $SO(3)$	4
0.2.2	Générateurs infinitésimaux de $SU(2)$	6
0.2.3	Algèbre de Lie $\mathfrak{su}(2)$	7
0.3	Représentations de $SU(2)$	9
0.3.1	Représentations des groupes $SO(3)$ et $SU(2)$	9
0.3.2	Représentations de l'algèbre $\mathfrak{su}(2)$	9
0.3.3	Construction explicite	13
0.4	Produit direct de représentations de $SU(2)$	15
0.4.1	Produit direct de représentations et l'"addition de moments angulaires"	15
0.4.2	Coefficients de Clebsch-Gordan, symboles $3-j$ et $6-j$	16
0.5	Une application physique : l'isospin	19
0.6	Représentations de $SO(1,3)$ et $SL(2, \mathbb{C})$	21
0.6.1	Petit rappel sur le groupe de Lorentz	21
0.6.2	Algèbre de Lie des groupes de Lorentz et Poincaré	22
0.6.3	Groupes de recouvrement de \mathcal{L}_+^\uparrow et \mathcal{P}_+^\uparrow	23
0.6.4	Représentations irréductibles de dimension finie de $SL(2, \mathbb{C})$	24
0.6.5	Représentations irréductibles unitaires du groupe de Poincaré. États à une particule.	26
1	Groupes. Groupes et algèbres de Lie	31
1.1	Généralités sur les groupes	31
1.1.1	Définitions de base et premiers exemples	31
1.1.2	Classes de conjugaison d'un groupe	33
1.1.3	Sous-groupes	33
1.1.4	Homomorphisme d'un groupe G dans un groupe G'	34
1.1.5	Classes par rapport à un sous-groupe	34
1.1.6	Sous-groupe invariant	35

1.1.7	Groupe simple, groupe semi-simple	36
1.2	Groupes continus. Propriétés topologiques. Groupes de Lie	36
1.2.1	Connexité	37
1.2.2	Simple connexité. Groupe d'homotopie. Recouvrement universel	37
1.2.3	Groupes compacts et non compacts	40
1.2.4	Mesure invariante de Haar	40
1.2.5	Groupes de Lie	41
1.3	Étude locale d'un groupe de Lie. Algèbre de Lie	42
1.3.1	Algèbres et algèbres de Lie. Définitions	42
1.3.2	Espace tangent d'un groupe de Lie G	43
1.3.3	Relations entre l'espace tangent \mathfrak{g} et le groupe G	44
1.3.4	L'espace tangent comme algèbre de Lie	45
1.3.5	Un exemple explicite : l'algèbre de Lie $\mathfrak{so}(n)$ de $\mathrm{SO}(n)$	46
1.3.6	Un exemple de dimension infinie : l'algèbre de Virasoro	48
1.4	Relations entre les propriétés de \mathfrak{g} et de G	48
1.4.1	Simplicité, semi-simplicité	48
1.4.2	Compacité. Complexification	49
1.4.3	Connexité, simple-connexité	50
1.4.4	Constantes de structure. Forme de Killing. Critères de Cartan	50
1.4.5	Opérateur(s) de Casimir	52
2	Représentations linéaires des groupes	65
2.1	Définitions et propriétés générales.	65
2.1.1	Définitions de base	65
2.1.2	Représentations équivalentes. Caractères	66
2.1.3	Représentations réductibles et irréductibles	67
2.1.4	Représentations conjuguée et contragrédiente	68
2.1.5	Représentations unitaires	69
2.1.6	Lemme de Schur	70
2.1.7	Produit tensoriel de représentations; décomposition de Clebsch-Gordan	71
2.1.8	Décomposition d'une représentation d'un groupe en représentations irréductibles d'un sous-groupe	75
2.2	Représentations des groupes et représentations des algèbres de Lie	75
2.2.1	Définition. Universalité	75
2.2.2	Représentations d'un groupe de Lie et de son algèbre de Lie	76
2.3	Représentations des groupes de Lie compacts	77
2.3.1	Orthogonalité et complétude	77
2.3.2	Conséquences	81
2.3.3	Cas des groupes finis	81
2.3.4	Récapitulation	83
2.4	Représentations projectives. Théorème de Wigner	83

2.4.1	Définition	83
2.4.2	Théorème de Wigner	84
2.4.3	Invariances d'un système quantique	86
2.4.4	Transformations des observables. Théorème de Wigner–Eckart	87
2.4.5	Forme infinitésimale d'une représentation projective. Extension centrale	89
3	Algèbres de Lie simples, classification et représentations	105
3.1	Sous-algèbre de Cartan. Racines. Forme canonique de l'algèbre.	105
3.1.1	Sous-algèbre de Cartan	105
3.1.2	Base canonique de l'algèbre de Lie	106
3.2	Géométrie des systèmes de racines	109
3.2.1	Produits scalaires de racines. La matrice de Cartan	109
3.2.2	Systèmes de racines des algèbres simples. La classification de Cartan	113
3.2.3	Base de Chevalley de l'algèbre	114
3.2.4	Coracines. Plus haute racine. Nombres, exposants de Coxeter	114
3.3	Représentations des algèbres semi-simples	115
3.3.1	Poids. Réseau des poids	115
3.3.2	Racines et poids de $\mathfrak{su}(n)$	120
3.4	Produit tensoriel des représentations de $\mathfrak{su}(n)$	123
3.4.1	Règles de Littlewood–Richardson et de Racah–Speiser	123
3.4.2	Construction tensorielle explicite des représentations de $SU(2)$ et $SU(3)$	125
3.5	Tableaux d'Young et représentations de $GL(n)$ et $SU(n)$	127
4	Symétries globales en physique des particules	139
4.1	Symétries globales exactes ou brisées. Brisure spontanée	139
4.1.1	Panorama. Symétries exactes ou brisées	139
4.1.2	Brisure de la symétrie chirale	142
4.1.3	Brisures quantiques. Anomalies	143
4.2	La symétrie de saveur $SU(3)$ et le modèle des quarks	144
4.2.1	Pourquoi le groupe $SU(3)$?	144
4.2.2	Conséquences de la symétrie $SU(3)$	146
4.2.3	Brisures électromagnétiques de la symétrie $SU(3)$	148
4.2.4	Écarts de masses “forts”. Formule de masse de Gell-Mann–Okubo	150
4.2.5	Les quarks	151
4.2.6	Courants hadroniques et interactions faibles	152
4.3	De $SU(3)$ à $SU(4)$ et aux six saveurs	153
4.3.1	Nouvelles saveurs	153
4.3.2	Introduction de la couleur	154
5	Théories de jauge. Modèle standard	161
5.1	Invariance de jauge. Couplage minimal. Lagrangien de Yang–Mills	161
5.1.1	Invariance de jauge de l'électrodynamique	161

5.1.2	Extension non abélienne de Yang–Mills	162
5.1.3	Géométrie des champs de jauge	165
5.1.4	Lagrangien de Yang–Mills	166
5.1.5	Quantification. Règles de Feynman.	167
5.2	Champs de jauge massifs	168
5.2.1	Interactions faibles et bosons intermédiaires	168
5.2.2	Théorie de jauge à symétrie brisée spontanément. Mécanisme de Brout– Englert–Higgs	169
5.3	Le modèle standard	170
5.3.1	Le secteur fort	170
5.3.2	Le secteur électro-faible, une esquisse.	171
5.4	Compléments	175
5.4.1	Modèle standard et au delà.	175
5.4.2	Théories grand-unifiées ou GUTs	175
5.4.3	Anomalies	177

Chapitre 0

Quelques éléments de base sur les groupes $\text{SO}(3)$, $\text{SU}(2)$ et $\text{SL}(2, \mathbb{C})$

0.1 Rotations de \mathbb{R}^3 , les groupes $\text{SO}(3)$ et $\text{SU}(2)$

0.1.1 Le groupe $\text{SO}(3)$, groupe à trois paramètres

On considère l'espace euclidien à trois dimensions et le groupe des rotations. Ces rotations laissent invariante la norme carrée du rayon vecteur $\mathbf{OM}^2 = x_1^2 + x_2^2 + x_3^2 = x^2 + y^2 + z^2$ ¹ et préservent l'orientation. Elles sont représentées dans une base orthonormée par des matrices 3×3 orthogonales réelles, de déterminant 1 : elles forment le groupe "spécial orthogonal" $\text{SO}(3)$.

Formule d'Olinde Rodrigues

Toute rotation de $\text{SO}(3)$ est une rotation d'un angle ψ autour d'un axe de vecteur directeur unitaire \mathbf{n} , et les rotations associées à (\mathbf{n}, ψ) et à $(-\mathbf{n}, -\psi)$ sont identiques. On notera $R_{\mathbf{n}}(\psi)$ cette rotation. De façon très explicite, on écrit $\mathbf{x} = \mathbf{x}_{\parallel} + \mathbf{x}_{\perp} = (\mathbf{x} \cdot \mathbf{n})\mathbf{n} + (\mathbf{x} - (\mathbf{x} \cdot \mathbf{n})\mathbf{n})$ et $\mathbf{x}' = \mathbf{x}_{\parallel} + \cos \psi \mathbf{x}_{\perp} + \sin \psi \mathbf{n} \wedge \mathbf{x}_{\perp}$, d'où la formule d'O. Rodrigues

$$\mathbf{x}' = R_{\mathbf{n}}(\psi)\mathbf{x} = \cos \psi \mathbf{x} + (1 - \cos \psi)(\mathbf{x} \cdot \mathbf{n})\mathbf{n} + \sin \psi (\mathbf{n} \times \mathbf{x}) . \quad (0.1)$$

Comme un vecteur unitaire \mathbf{n} dans \mathbb{R}^3 dépend de deux paramètres, par exemple l'angle θ qu'il fait avec l'axe Oz et l'angle ϕ que fait sa projection dans le plan Ox, Oy avec l'axe Ox (voir Fig. 1) un élément de $\text{SO}(3)$ est paramétrisé par 3 variables continues. On prendra ainsi

$$0 \leq \theta \leq \pi, \quad 0 \leq \phi < 2\pi, \quad 0 \leq \psi \leq \pi . \quad (0.2)$$

Mais il demeure une petite redondance d'apparence anodine, $R_{\mathbf{n}}(\pi) = R_{-\mathbf{n}}(\pi)$, à suivre ...

1. Dans tout ce chapitre, nous utilisons alternativement les notations (x, y, z) ou (x_1, x_2, x_3) pour désigner les coordonnées dans un repère orthonormé.

$SO(3)$ est donc une variété de dimension 3. Pour la rotation d'axe \mathbf{n} colinéaire à l'axe Oz , on a la matrice

$$\mathcal{R}_z(\psi) = \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (0.3)$$

tandis qu'autour des axes Ox et Oy

$$\mathcal{R}_x(\psi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & -\sin \psi \\ 0 & \sin \psi & \cos \psi \end{pmatrix} \quad \mathcal{R}_y(\psi) = \begin{pmatrix} \cos \psi & 0 & \sin \psi \\ 0 & 1 & 0 \\ -\sin \psi & 0 & \cos \psi \end{pmatrix}. \quad (0.4)$$

Conjugaison de $R_{\mathbf{n}}(\psi)$ par une autre rotation

Une relation que nous allons abondamment utiliser est que

$$\boxed{RR_{\mathbf{n}}(\psi)R^{-1} = R_{\mathbf{n}'}(\psi)} \quad (0.5)$$

où \mathbf{n}' est le transformé de \mathbf{n} par la rotation R , $\mathbf{n}' = R\mathbf{n}$ (la vérifier !). Inversement toute rotation d'angle ψ autour d'un vecteur \mathbf{n}' peut se mettre sous la forme (0.5) : on dira plus tard que les "classes de conjugaison" du groupe $SO(3)$ sont caractérisées par l'angle ψ .

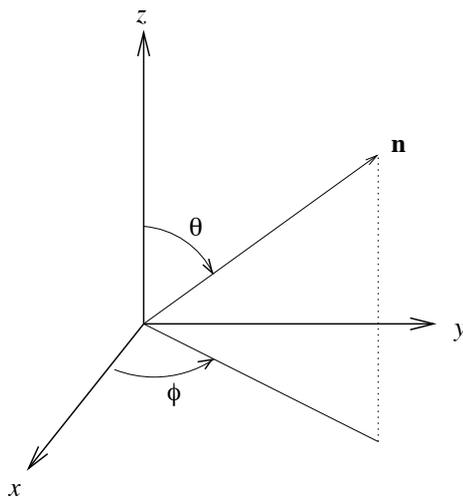


Fig. 1

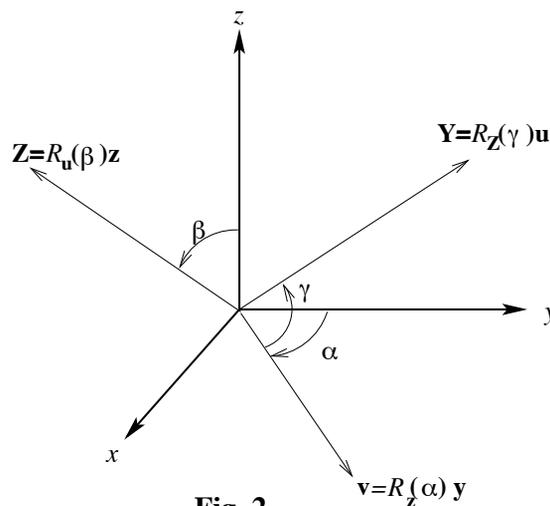


Fig. 2

Angles d'Euler

Une autre description fait appel aux *angles d'Euler* : étant donné un repère orthonormé (Ox, Oy, Oz) , toute rotation autour de O qui envoie ce repère sur (OX, OY, OZ) peut être considérée comme résultant de la composition d'une rotation d'angle α autour de Oz qui amène le repère sur (Ou, Ov, Oz) , suivie d'une rotation d'angle β autour de Ov l'amenant sur (Ou', Ov, OZ) , et enfin d'une rotation d'angle γ autour de OZ qui amène le repère sur (OX, OY, OZ) , (voir Fig. 2). On prend donc $0 \leq \alpha < 2\pi$, $0 \leq \beta \leq \pi$, $0 \leq \gamma < 2\pi$ et on écrit

$$R(\alpha, \beta, \gamma) = R_Z(\gamma)R_v(\beta)R_z(\alpha) \quad (0.6)$$

mais selon (0.5)

$$R_Z(\gamma) = R_v(\beta)R_z(\gamma)R_v^{-1}(\beta) \quad R_v(\beta) = R_z(\alpha)R_y(\beta)R_z^{-1}(\alpha)$$

d'où en reportant dans (0.6)

$$R(\alpha, \beta, \gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma) . \quad (0.7)$$

où on a utilisé le fait que $R_z(\alpha)R_z(\gamma)R_z^{-1}(\alpha) = R_z(\gamma)$ car les rotations autour d'un même axe commutent (elles forment un sous-groupe abélien, isomorphe à $SO(2)$).

Exercice : En utilisant (0.5), écrire l'expression d'une matrice R qui amène le vecteur unitaire \mathbf{z} porté par Oz sur le vecteur unitaire \mathbf{n} , en termes par exemple de $R_z(\phi)$ et de $R_y(\theta)$, puis l'expression de $R_{\mathbf{n}}(\psi)$ en termes de R_y et R_z . Écrire l'expression explicite de cette matrice et de (0.7) et en déduire les relations entre θ, ϕ, ψ et les angles d'Euler. (Voir aussi plus bas, équ. (0.66).)

0.1.2 Du groupe $SO(3)$ au groupe $SU(2)$

Considérons une autre paramétrisation des rotations. À la rotation $R_{\mathbf{n}}(\psi)$, nous associons le vecteur unitaire à quatre dimensions $u : (u_0 = \cos \frac{\psi}{2}, \mathbf{u} = \mathbf{n} \sin \frac{\psi}{2})$; on a $u^2 = u_0^2 + \mathbf{u}^2 = 1$, et u appartient à la sphère unité S^3 dans l'espace \mathbb{R}^4 . Le changement de détermination de ψ par un multiple impair de 2π change u en $-u$. Il y a donc bijection entre $R_{\mathbf{n}}(\psi)$ et la paire $(u, -u)$, c'est-à-dire entre $SO(3)$ et S^3/\mathbb{Z}_2 , la sphère dans laquelle on identifie les paires de points opposés. On dira que la sphère S^3 est un "groupe de recouvrement" de $SO(3)$. En quel sens cette sphère est-elle un groupe? Pour répondre à cette question, introduisons les matrices de Pauli $\sigma_i, i = 1, 2, 3$

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \quad (0.8)$$

Avec la matrice identité \mathbf{I} , elles constituent une base de l'espace des matrices 2×2 hermitiques. Elles satisfont l'identité

$$\sigma_i \sigma_j = \delta_{ij} \mathbf{I} + i \epsilon_{ijk} \sigma_k , \quad (0.9)$$

avec ϵ_{ijk} le tenseur complètement antisymétrique, $\epsilon_{123} = +1$, ϵ_{ijk} = signature de la permutation (ijk) .

Pour u un vecteur unitaire réel à quatre dimensions (c'est-à-dire un point de S^3), formons la matrice

$$U = u_0 \mathbf{I} - i \mathbf{u} \cdot \boldsymbol{\sigma} \quad (0.10)$$

qui est unitaire et de déterminant 1 (le vérifier et montrer aussi la réciproque : toute matrice unitaire unimodulaire (= de déterminant 1) 2×2 est de la forme (0.10), avec $u^2 = 1$). Ces matrices forment le groupe $SU(2)$ qui est donc isomorphe à S^3 . En développant l'exponentielle en puissances et en utilisant que $(\mathbf{n} \cdot \boldsymbol{\sigma})^2 = \mathbf{I}$, conséquence de (0.9), on peut vérifier (Exercice!) que

$$e^{-i \frac{\psi}{2} \mathbf{n} \cdot \boldsymbol{\sigma}} = \cos \frac{\psi}{2} - i \sin \frac{\psi}{2} \mathbf{n} \cdot \boldsymbol{\sigma} . \quad (0.11)$$

Il est suggéré que la multiplication des matrices

$$U_{\mathbf{n}}(\psi) = e^{-i \frac{\psi}{2} \mathbf{n} \cdot \boldsymbol{\sigma}} = \cos \frac{\psi}{2} - i \sin \frac{\psi}{2} \mathbf{n} \cdot \boldsymbol{\sigma}, \quad 0 \leq \psi \leq 2\pi, \quad \mathbf{n} \in S^2 \quad (0.12)$$

fournit la loi de groupe cherchée dans S^3 . Montrons qu'en effet à une matrice de $SU(2)$ on peut associer une rotation de $SO(3)$ et qu'au produit de deux matrices de $SU(2)$ correspond le produit des rotations de $SO(3)$. Au point x de \mathbb{R}^3 de coordonnées x_1, x_2, x_3 , associons la matrice hermitique

$$X = \mathbf{x} \cdot \boldsymbol{\sigma} = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix}, \quad (0.13)$$

avec inversement $x_i = \frac{1}{2} \text{tr}(X\sigma_i)$, et agissons sur cette matrice selon

$$X \mapsto X' = UXU^\dagger, \quad (0.14)$$

ce qui définit une transformation linéaire $x \mapsto x' = \mathcal{T}x$. On calcule aisément que

$$\det X = -(x_1^2 + x_2^2 + x_3^2) \quad (0.15)$$

et comme $\det X = \det X'$, la transformation linéaire $x \mapsto x' = \mathcal{T}x$ est une isométrie, donc $\det \mathcal{T} = 1$ ou -1 . Pour se convaincre qu'il s'agit bien d'une rotation, c'est-à-dire que la transformation a un déterminant 1, il suffit de calculer ce déterminant pour $U = \mathbf{I}$ où $\mathcal{T} =$ l'identité donc $\det \mathcal{T} = 1$, puis d'utiliser la connexité de la variété $SU(2) (\cong S^3)$ pour conclure que la fonction continue $\det \mathcal{T}(U)$ ne peut sauter à la valeur -1 . En fait, en utilisant l'identité (0.9), le calcul explicite de X' conduit après un peu d'algèbre à

$$\begin{aligned} X' &= \left(\cos \frac{\psi}{2} - i\mathbf{n} \cdot \boldsymbol{\sigma} \sin \frac{\psi}{2} \right) X \left(\cos \frac{\psi}{2} + i\mathbf{n} \cdot \boldsymbol{\sigma} \sin \frac{\psi}{2} \right) \\ &= \left(\cos \psi \mathbf{x} + (1 - \cos \psi)(\mathbf{x} \cdot \mathbf{n}) \mathbf{n} + \sin \psi (\mathbf{n} \wedge \mathbf{x}) \right) \cdot \boldsymbol{\sigma} \end{aligned} \quad (0.16)$$

sur lequel on reconnaît la formule (0.1). On en conclut que la transformation $x \rightarrow x'$ effectuée par les matrices de $SU(2)$ dans (0.14) est bien la rotation d'angle ψ autour de \mathbf{n} . Au produit $U_{\mathbf{n}'}(\psi')U_{\mathbf{n}}(\psi)$ dans $SU(2)$ correspond dans $SO(3)$ la composition des deux rotations $R_{\mathbf{n}'}(\psi')R_{\mathbf{n}}(\psi)$ de $SO(3)$. Il y a donc un "homomorphisme" du groupe $SU(2)$ dans $SO(3)$. Cet homomorphisme envoie les deux matrices U et $-U$ sur la même rotation.

Résumons les acquis de ce paragraphe. Nous avons montré que le groupe $SU(2)$ est un groupe de recouvrement (d'ordre 2) du groupe $SO(3)$ (le sens topologique précis en sera donné par la suite), et que l'homomorphisme de $SU(2)$ dans $SO(3)$ est fourni par les équations (0.12)-(0.14).

Exercice : vérifier que toute matrice de $SU(2)$ peut s'écrire $\begin{pmatrix} a & -b \\ b^* & a^* \end{pmatrix}$ avec $|a|^2 + |b|^2 = 1$. Quelle est la relation avec (0.10) ?

0.2 Générateurs infinitésimaux. L'algèbre de Lie $\mathfrak{su}(2)$

0.2.1 Générateurs infinitésimaux de $SO(3)$

Les rotations $R_{\mathbf{n}}(\psi)$ autour d'un axe \mathbf{n} donné forment un *sous-groupe à un paramètre* isomorphe à $SO(2)$. Dans ce chapitre, nous suivons l'usage des physiciens en écrivant les générateurs infinitésimaux des rotations comme des opérateurs hermitiens $J = J^\dagger$. Ainsi on écrit

$$R_{\mathbf{n}}(d\psi) = (I - id\psi J_{\mathbf{n}}) \quad (0.17)$$

où $J_{\mathbf{n}}$, le “générateur” de ces rotations d’axe \mathbf{n} , est une matrice hermitique 3×3 . Montrons d’abord que l’on peut reconstruire les rotations finies à partir de ces générateurs infinitésimaux. Par la propriété de groupe,

$$R_{\mathbf{n}}(\psi + d\psi) = R_{\mathbf{n}}(d\psi)R_{\mathbf{n}}(\psi) = (I - id\psi J_{\mathbf{n}})R_{\mathbf{n}}(\psi) , \quad (0.18)$$

ou encore

$$\frac{\partial R_{\mathbf{n}}(\psi)}{\partial \psi} = -iJ_{\mathbf{n}}R_{\mathbf{n}}(\psi) \quad (0.19)$$

qui, compte tenu de $R_{\mathbf{n}}(0) = I$, s’intègre en

$$\boxed{R_{\mathbf{n}}(\psi) = e^{-i\psi J_{\mathbf{n}}} .} \quad (0.20)$$

Pour être plus explicites, introduisons les trois matrices de base J_1 , J_2 et J_3 décrivant les rotations infinitésimales autour des axes correspondants². De la version infinitésimale de (0.3) on tire

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad J_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \quad J_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (0.21)$$

ce que l’on peut exprimer par une formule unique

$$\boxed{(J_k)_{ij} = -i\epsilon_{ijk}} \quad (0.22)$$

à l’aide du tenseur complètement antisymétrique ϵ_{ijk} .

Montrons que les trois matrices (0.21) forment une base des générateurs infinitésimaux et que $J_{\mathbf{n}}$ s’exprime simplement comme

$$J_{\mathbf{n}} = \sum_k J_k n_k \quad (0.23)$$

ce qui permet d’écrire (0.20) sous la forme

$$\boxed{R_{\mathbf{n}}(\psi) = e^{-i\psi \sum_k n_k J_k} .} \quad (0.24)$$

L’expression (0.23) découle simplement de la version infinitésimale de la formule d’O. Rodrigues, $R_{\mathbf{n}}(d\psi) = (\mathbf{I} + d\psi \mathbf{n} \wedge)$ donc $-iJ_{\mathbf{n}} = \mathbf{n} \wedge$ ou si on préfère $-i(J_{\mathbf{n}})_{ij} = \epsilon_{ikj} n_k = n_k (-iJ_k)_{ij}$, cqfd. (Ici et dans la suite, convention de sommation sur les indices répétés : $\epsilon_{ikj} n_k \equiv \sum_k \epsilon_{ikj} n_k$, etc.)

Un commentaire sur (0.24) : on n’a évidemment pas le droit d’écrire en général $R_{\mathbf{n}}(\psi) = e^{-i\psi \sum_k n_k J_k} \stackrel{?}{=} \prod_{k=1}^3 e^{-i\psi n_k J_k}$. Par ailleurs on voit que par la formule (0.7), on peut écrire toute rotation de $\text{SO}(3)$ sous la forme

$$R(\alpha, \beta, \gamma) = e^{-i\alpha J_3} e^{-i\beta J_2} e^{-i\gamma J_3} . \quad (0.25)$$

Les trois matrices J_i , $i = 1, 2, 3$ satisfont les très importantes relations de commutation

$$\boxed{[J_i, J_j] = i\epsilon_{ijk} J_k} \quad (0.26)$$

2. Ne pas confondre $J_{\mathbf{n}}$ indexé par le vecteur \mathbf{n} , avec J_k , $k^{\text{ième}}$ composante de \mathbf{J} . La relation entre les deux va être donnée plus bas.

qui découlent de l'identité suivante (Jacobi) vérifiée par le tenseur ϵ

$$\epsilon_{iab}\epsilon_{bjc} + \epsilon_{icb}\epsilon_{baj} + \epsilon_{ijb}\epsilon_{bca} = 0 . \quad (0.27)$$

Exercice : bien comprendre la structure de cette identité et vérifier qu'elle implique (0.26).

Au vu de l'importance des relations (0.23–0.26), il est utile de les retrouver par une autre route. Notons d'abord que l'équation (0.5) implique que pour tout R

$$R e^{-i\psi J_{\mathbf{n}}} R^{-1} = e^{-i\psi R J_{\mathbf{n}} R^{-1}} = e^{-i\psi J_{\mathbf{n}'}} \quad (0.28)$$

avec $\mathbf{n}' = R\mathbf{n}$, donc

$$R J_{\mathbf{n}} R^{-1} = J_{\mathbf{n}'}. \quad (0.29)$$

Le tenseur ϵ_{ijk} est invariant par l'action des rotations

$$\epsilon_{lmn} \mathcal{R}_{il} \mathcal{R}_{jm} \mathcal{R}_{kn} = \epsilon_{ijk} \det \mathcal{R} = \epsilon_{ijk} \quad (0.30)$$

puisque la matrice \mathcal{R} est de déterminant 1. Cette matrice étant aussi orthogonale, on peut faire passer un \mathcal{R} au membre de droite

$$\epsilon_{lmn} \mathcal{R}_{jm} \mathcal{R}_{kn} = \epsilon_{ijk} \mathcal{R}_{il} \quad (0.31)$$

ce qui au vu de (0.22) exprime que

$$\mathcal{R}_{jm} (J_l)_{mn} \mathcal{R}_{nk}^{-1} = (J_i)_{jk} \mathcal{R}_{il} \quad (0.32)$$

c'est-à-dire, pour tout R et sa matrice \mathcal{R} ,

$$R J_l R^{-1} = J_i \mathcal{R}_{il}. \quad (0.33)$$

Soit R une rotation qui amène le vecteur unitaire \mathbf{z} porté par Oz sur le vecteur \mathbf{n} , on a donc $n_k = \mathcal{R}_{k3}$ et

$$J_{\mathbf{n}} \stackrel{(0.29)}{=} R J_3 R^{-1} \stackrel{(0.33)}{=} J_k \mathcal{R}_{k3} = J_k n_k, \quad (0.34)$$

qui n'est autre que (0.23). Noter que les équations (0.33) et (0.34) sont bien compatibles avec (0.29)

$$J_{\mathbf{n}'} \stackrel{(0.29)}{=} R J_{\mathbf{n}} R^{-1} \stackrel{(0.34)}{=} R J_k n_k R^{-1} \stackrel{(0.33)}{=} J_l \mathcal{R}_{lk} n_k = J_l n'_l.$$

Comme on le verra de façon plus systématique par la suite, la relation (0.26) de commutation des générateurs infinitésimaux J code une version infinitésimale de la loi de groupe. Considérons par exemple une rotation d'angle infinitésimal $d\psi$ autour de Oy agissant sur J_1

$$R_2(d\psi) J_1 R_2^{-1}(d\psi) \stackrel{(0.33)}{=} J_k [\mathcal{R}_2(d\psi)]_{k1} \quad (0.35)$$

mais au premier ordre, $R_2(d\psi) = \mathbf{I} - id\psi J_2$, donc le membre de gauche de (0.35) est égal à $J_1 - id\psi [J_2, J_1]$ et au membre de droite, $[\mathcal{R}_2(d\psi)]_{k1} = \delta_{k1} - id\psi (J_2)_{k1} = \delta_{k1} - d\psi \delta_{k3}$ d'après (0.22), d'où $i[J_1, J_2] = -J_3$, qui est l'une des relations (0.26).

0.2.2 Générateurs infinitésimaux de $SU(2)$

Examinons maintenant les choses du point de vue de $SU(2)$. Toute matrice unitaire U (ici 2×2) peut se diagonaliser dans une base orthonormée $U = V \exp\{i \text{diag}(\lambda_k)\} V^\dagger$, V unitaire, et donc s'écrire

$$U = \exp iH = \sum_0^\infty \frac{(iH)^n}{n!} \quad (0.36)$$

avec H hermitique, $H = V \text{diag}(\lambda_k) V^\dagger$. La somme converge (pour la norme $\|M\|^2 = \text{tr} MM^\dagger$). La condition d'unimodularité $1 = \det U = \exp i \text{tr} H$ est garantie si $\text{tr} H = 0$. L'ensemble de ces

matrices hermitiques de trace nulle forme un espace vectoriel \mathcal{V} de dimension 3 sur \mathbb{R} . Or les matrices hermitiques 2×2 de trace nulle sont des combinaisons linéaires à coefficients réels des 3 matrices de Pauli

$$H = \sum_{k=1}^3 \eta_k \frac{\sigma_k}{2}, \quad (0.37)$$

ce que l'on peut reporter dans (0.36). On a en fait déjà observé plus haut que toute matrice unitaire 2×2 peut s'écrire sous la forme (0.11). En comparant cette forme avec celle obtenue en (0.24), ou encore en comparant sa version infinitésimale $U_n(d\psi) = (I - i d\psi \mathbf{n} \cdot \frac{\boldsymbol{\sigma}}{2})$ avec (0.17-0.34), on voit que les matrices $\frac{1}{2}\sigma_j$ jouent ici dans $\text{SU}(2)$ le rôle joué par les générateurs infinitésimaux J_j dans $\text{SO}(3)$. Or ces matrices $\frac{1}{2}\sigma_j$ vérifient les relations de commutation

$$\left[\frac{\sigma_i}{2}, \frac{\sigma_j}{2} \right] = i\epsilon_{ijk} \frac{\sigma_k}{2}. \quad (0.38)$$

avec les mêmes *constantes de structure* ϵ_{ijk} que dans (0.26). Autrement dit, nous venons de découvrir que les générateurs infinitésimaux J_i (éq. (0.21) de $\text{SO}(3)$) et $\frac{1}{2}\sigma_i$ de $\text{SU}(2)$ satisfont aux mêmes relations de commutation (on dira plus tard qu'ils forment deux représentations de la même algèbre de Lie $\mathfrak{su}(2) = \mathfrak{so}(3)$). Cela implique que des calculs menés avec les $\frac{1}{2}\vec{\sigma}$ et faisant appel *uniquement aux règles de commutation des générateurs* demeurent valables avec les \vec{J} , et vice versa. Ainsi, des relations (0.33), par exemple $R_2(\beta)J_kR_2^{-1}(\beta) = J_l\mathcal{R}_y(\beta)_{lk}$, il découle sans aucun calcul supplémentaire que pour les matrices de Pauli, on a

$$e^{-i\frac{\beta}{2}\sigma_2}\sigma_k e^{i\frac{\beta}{2}\sigma_2} = D_2(\beta)\sigma_k D_2^{-1}(\beta) = \sigma_l \mathcal{R}_y(\beta)_{lk} \quad (0.39)$$

avec $D_k(\psi) := e^{-i\psi\frac{\sigma_k}{2}}$, où on lit les éléments de matrice \mathcal{R}_y en (0.4). On a en effet l'identité générale $e^A B e^{-A} = B + \sum_{n=1}^{\infty} \frac{1}{n!} \underbrace{[A[A, [\dots, [A, B] \dots]]]}_{n \text{ commutateurs}}$, cf Chap. 1, (1.27), et ce calcul ne fait donc appel qu'à des commutateurs. En revanche, la relation

$$\sigma_i \sigma_j = \delta_{ij} + i\epsilon_{ijk} \sigma_k$$

(qui ne fait pas appel qu'aux commutateurs) est spécifique à la représentation de dimension 2 de l'algèbre $\mathfrak{su}(2)$.

0.2.3 Algèbre de Lie $\mathfrak{su}(2)$

Récapitulons : nous venons d'introduire l'algèbre de commutation des générateurs infinitésimaux (ou algèbre de Lie) du groupe $\text{SU}(2)$ (ou $\text{SO}(3)$), notée $\mathfrak{su}(2)$ ou $\mathfrak{so}(3)$. Elle est définie par les relations (0.26), que nous récrivons ici

$$[J_i, J_j] = i\epsilon_{ijk} J_k. \quad (0.26)$$

On utilise aussi beaucoup les trois combinaisons

$$J_z \equiv J_3, \quad J_+ = J_1 + iJ_2, \quad J_- = J_1 - iJ_2. \quad (0.40)$$

Il est alors immédiat de calculer

$$\begin{aligned} [J_3, J_+] &= J_+ \\ [J_3, J_-] &= -J_- \\ [J_+, J_-] &= 2J_3. \end{aligned} \tag{0.41}$$

On vérifie aussi que l'opérateur de Casimir défini par

$$\mathbf{J}^2 = J_1^2 + J_2^2 + J_3^2 = J_3^2 + J_3 + J_- J_+ \tag{0.42}$$

commute avec tous les J

$$[\mathbf{J}^2, J_i] = 0, \tag{0.43}$$

ce qui signifie qu'il est invariant par rotation.

Anticipant un peu sur la suite, nous nous intéresserons le plus souvent aux "représentations unitaires", où les générateurs J_i , $i = 1, 2, 3$ sont hermitiques, donc

$$J_i^\dagger = J_i, \quad i = 1, 2, 3 \quad J_\pm^\dagger = J_\mp. \tag{0.44}$$

Pour terminer, mentionnons l'interprétation des J_i comme opérateurs différentiels agissant sur les fonctions différentiables des coordonnées de l'espace \mathbb{R}^3 . Dans l'espace \mathbb{R}^3 , l'effet d'une rotation infinitésimale sur le vecteur \mathbf{x} est de le changer en

$$\mathbf{x}' = \mathbf{x} + \delta\psi \mathbf{n} \wedge \mathbf{x}$$

donc une fonction *scalaire* de \mathbf{x} , $f(\mathbf{x})$, est changée en $f'(\mathbf{x}') = f(\mathbf{x})$ soit

$$\begin{aligned} f'(\mathbf{x}) &= f(R^{-1}\mathbf{x}) = f(\mathbf{x} - \delta\psi \mathbf{n} \wedge \mathbf{x}) \\ &= (1 - \delta\psi \mathbf{n} \cdot \mathbf{x} \wedge \nabla) f(\mathbf{x}) \\ &= (1 - i\delta\psi \mathbf{n} \cdot \mathbf{J}) f(\mathbf{x}). \end{aligned} \tag{0.45}$$

On identifie donc

$$\mathbf{J} = -i\mathbf{x} \wedge \nabla, \quad J_i = -i\epsilon_{ijk} x_j \frac{\partial}{\partial x_k} \tag{0.46}$$

ce qui permet de le calculer dans des coordonnées quelconques, par exemple sphériques (Appendice 0 de ce chapitre). (Comparer aussi (0.46) avec l'expression du moment angulaire en Mécanique Quantique $L_i = \frac{\hbar}{i}\epsilon_{ijk} x_j \frac{\partial}{\partial x_k}$). Exercice : vérifier que ces opérateurs différentiels ont bien les relations de commutation (0.26).

Parmi les combinaisons de J que l'on peut construire, l'une doit jouer un rôle particulier, le laplacien sur la sphère S^2 , opérateur différentiel du second ordre invariant par changement de coordonnées (cf Appendice 0). Il doit en particulier être invariant par rotation, être de degré 2 dans les J_i , ce ne peut être que l'opérateur de Casimir \mathbf{J}^2 (à un facteur près). De fait le laplacien dans \mathbb{R}^3 s'écrit en coordonnées sphériques

$$\begin{aligned} \Delta_3 &= \frac{1}{r} \frac{\partial^2}{\partial r^2} r - \frac{\mathbf{J}^2}{r^2} \\ &= \frac{1}{r} \frac{\partial^2}{\partial r^2} r + \frac{\Delta_{\text{sphère } S^2}}{r^2}. \end{aligned} \tag{0.47}$$

Nous nous sommes restreints ici pour plus de simplicité au cas de fonctions scalaires, mais on pourrait aussi s'intéresser plus généralement à la transformation d'une collection de fonctions des coordonnées de \mathbb{R}^3 "formant une représentation" de $SO(3)$, c'est-à-dire se transformant linéairement entre elles sous l'action de ce groupe

$$A'(\mathbf{x}') = D(R)A(\mathbf{x})$$

soit encore

$$A'(\mathbf{x}) = D(R)A(R^{-1}\mathbf{x}) ,$$

par exemple un champ vectoriel se transformant par

$$\mathbf{A}'(\mathbf{x}) = R\mathbf{A}(R^{-1}\mathbf{x}) .$$

Le produit scalaire de deux tels champs vectoriels est une fonction scalaire. Que devient la discussion qui précède sur les générateurs infinitésimaux pour de tels objets? Montrer qu'ils sont sommes de deux contributions, l'une donnée par (0.46) et l'autre venant de la forme infinitésimale de R ; en termes physiques, ces deux contributions correspondent aux moments angulaires orbital et intrinsèque (ou de spin).

0.3 Représentations de SU(2)

0.3.1 Représentations des groupes SO(3) et SU(2)

En géométrie de l'espace \mathbb{R}^3 , les notions de vecteur ou de tenseur sont familières. Il s'agit d'objets se transformant de façon linéaire sous l'effet des rotations

$$V_i \mapsto \mathcal{R}_{ii'}V_{i'} \quad (V \otimes W)_{ij} = V_iW_j \mapsto \mathcal{R}_{ii'}\mathcal{R}_{jj'}(V \otimes W)_{i'j'} = \mathcal{R}_{ii'}\mathcal{R}_{jj'}V_{i'}W_{j'} \quad \text{etc.}$$

D'une façon générale, on appelle *représentation* d'un groupe G dans un espace vectoriel E un homomorphisme de G dans le groupe des transformations linéaires $GL(E)$ (cf. Chap. 2). Ainsi, comme on vient de le voir, le groupe SO(3) admet une représentation dans l'espace \mathbb{R}^3 (les vecteurs V de l'exemple ci-dessus), une représentation dans l'espace des tenseurs de rang deux, etc. Nous allons maintenant nous intéresser à la construction des représentations générales de SO(3) et SU(2). Pour les besoins de la physique, en particulier de la mécanique quantique, on a surtout besoin de représentations unitaires, dans lesquelles les matrices de représentation sont unitaires. En fait, comme on le verra, il suffit d'étudier les représentations de SU(2) pour avoir aussi celles de SO(3), et mieux encore, il suffira d'étudier la façon dont sont représentés les éléments du groupe au voisinage de l'identité, c'est-à-dire d'étudier les représentations des générateurs infinitésimaux de SU(2) (et SO(3)) (qui respectent les relations de commutation (0.26)).

Il suffit donc pour trouver les représentations unitaires du groupe SU(2) de trouver les représentations par des matrices hermitiques de son algèbre de Lie $\mathfrak{su}(2)$.

0.3.2 Représentations de l'algèbre $\mathfrak{su}(2)$

Procédons à la construction classique des représentations de l'algèbre $\mathfrak{su}(2)$. Comme précédemment, J_{\pm} et J_z désignent les représentants des générateurs infinitésimaux dans une certaine représentation. Ils satisfont aux relations de commutation (0.41) et d'hermiticité (0.44). La commutation des opérateurs J_z et \mathbf{J}^2 garantit que l'on peut en chercher des vecteurs propres communs. Les valeurs propres de ces opérateurs hermitiques étant réelles et \mathbf{J}^2 étant semi-défini positif, on peut toujours écrire ses valeurs propres sous la forme $j(j+1)$, j réel positif ou nul

et on considère donc un vecteur propre commun $|j m\rangle$

$$\begin{aligned} \mathbf{J}^2|j m\rangle &= j(j+1)|j m\rangle \\ J_z|j m\rangle &= m|j m\rangle . \end{aligned} \quad (0.48)$$

avec m un réel a priori arbitraire. Par abus de langage, on dira que $|j m\rangle$ est un “vecteur propre de valeurs propres (j, m) ”.

(i) Agissons avec J_+ et $J_- = J_+^\dagger$ sur $|j m\rangle$. Utilisant la relation $J_\pm J_\mp = \mathbf{J}^2 - J_z^2 \pm J_z$ (conséquence de (0.41)), on calcule la norme carrée de $J_\pm|j m\rangle$:

$$\begin{aligned} \langle j m|J_-J_+|j m\rangle &= (j(j+1) - m(m+1)) \langle j m|j m\rangle \\ &= (j-m)(j+m+1) \langle j m|j m\rangle \\ \langle j m|J_+J_-|j m\rangle &= (j(j+1) - m(m-1)) \langle j m|j m\rangle \\ &= (j+m)(j-m+1) \langle j m|j m\rangle . \end{aligned} \quad (0.49)$$

Ces normes carrées ne peuvent être négatives donc

$$\begin{aligned} (j-m)(j+m+1) \geq 0 & : -j-1 \leq m \leq j \\ (j+m)(j-m+1) \geq 0 & : -j \leq m \leq j+1 \end{aligned} \quad (0.50)$$

qui impliquent

$$-j \leq m \leq j . \quad (0.51)$$

De plus $J_+|j m\rangle = 0$ si et seulement si $m = j$ et $J_-|j m\rangle = 0$ si et seulement si $m = -j$

$$J_+|j j\rangle = 0 \quad J_-|j -j\rangle = 0 . \quad (0.52)$$

(ii) Si $m \neq j$, $J_+|j m\rangle$ est non nul, vecteur propre de valeurs propres $(j, m+1)$. En effet

$$\begin{aligned} \mathbf{J}^2 J_+|j m\rangle &= J_+ \mathbf{J}^2|j m\rangle = j(j+1) J_+|j m\rangle \\ J_z J_+|j m\rangle &= J_+(J_z+1)|j m\rangle = (m+1) J_+|j m\rangle . \end{aligned} \quad (0.53)$$

De même si $m \neq -j$, $J_-|j m\rangle$ est un vecteur propre (non nul) de valeurs propres $(j, m-1)$.

(iii) Considérons la suite des vecteurs

$$|j m\rangle, J_-|j m\rangle, J_-^2|j m\rangle, \dots, J_-^p|j m\rangle \dots$$

S'ils sont non nuls ils constituent des vecteurs propres de J_z de valeurs propres $m, m-1, m-2, \dots, m-p \dots$. Les valeurs propres autorisées de J_z étant bornées par (0.51), cette suite doit s'arrêter au bout d'un nombre fini d'étapes. Soit p l'entier tel que $J_-^p|j m\rangle \neq 0$, $J_-^{p+1}|j m\rangle = 0$. En vertu de (0.52), $J_-^p|j m\rangle$ est un vecteur propre de valeurs propres $(j, -j)$ donc $m-p = -j$ c'est-à-dire

$$(j+m) \text{ est un entier non négatif.} \quad (0.54)$$

Opérant de même avec J_+, J_+^2, \dots sur $|j m\rangle$, on est mené à la conclusion que

$$(j-m) \text{ est un entier non négatif.} \quad (0.55)$$

et par conséquent j et m sont simultanément entiers ou demi-entiers. Pour chaque valeur de j

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$$

m peut prendre les $2j + 1$ valeurs³

$$m = -j, -j + 1, \dots, j - 1, j. \quad (0.56)$$

Partant du vecteur $|j m = j\rangle$, (“vecteur de plus haut poids”), choisi de norme 1, on construit la base orthonormée $|j m\rangle$ par application répétée de J_- et on a

$$\begin{aligned} J_+ |j m\rangle &= \sqrt{j(j+1) - m(m+1)} |j m+1\rangle \\ J_- |j m\rangle &= \sqrt{j(j+1) - m(m-1)} |j m-1\rangle \\ J_z |j m\rangle &= m |j m\rangle. \end{aligned} \quad (0.57)$$

Ces $2j + 1$ vecteurs forment la base de la “représentation de spin j ” de l’algèbre $\mathfrak{su}(2)$.

En fait, cette représentation de l’algèbre $\mathfrak{su}(2)$ s’étend en une représentation du groupe SU(2), comme on va le voir maintenant.

La discussion précédente a fait jouer un rôle central à l’unitarité de la représentation et donc à l’hermiticité des générateurs infinitésimaux, donc à la positivité : $\|J_{\pm} |j m\rangle\|^2 \geq 0 \implies -j \leq m \leq j$, etc, et a permis de conclure que la représentation est nécessairement de dimension finie. Inversement on peut insister sur cette dernière condition, et montrer qu’elle suffit à assurer les conditions précédentes. Partant d’un vecteur propre $|\psi\rangle$ de J_z , la suite $J_+^p |\psi\rangle$ produit des vecteurs propres de J_z de valeur propre croissante, donc linéairement indépendants s’ils sont non nuls. Si par hypothèse la représentation est de dimension finie, cette suite est finie, et il existe un vecteur noté $|j\rangle$ tel que $J_+ |j\rangle = 0$, $J_z |j\rangle = j |j\rangle$. Par la relation $\mathbf{J}^2 = J_- J_+ + J_z(J_z + 1)$, c’est aussi un vecteur propre de valeur propre $j(j+1)$ de \mathbf{J}^2 . Il s’identifie donc avec le vecteur de plus haut poids noté précédemment $|j j\rangle$, notation que nous adoptons donc dans la suite de cette discussion. A partir de ce vecteur, les $J_-^p |j j\rangle$ forment une suite qui doit elle aussi être finie

$$\exists q \quad J_-^{q-1} |j j\rangle \neq 0 \quad J_-^q |j j\rangle = 0. \quad (0.58)$$

On démontre aisément par récurrence que

$$J_+ J_-^q |j j\rangle = [J_+, J_-^q] |j j\rangle = q(2j+1-q) J_-^{q-1} |j j\rangle = 0 \quad (0.59)$$

donc $q = 2j + 1$. Le nombre j est donc entier ou demi-entier, les vecteurs de la représentation ainsi construite sont vecteurs propres de \mathbf{J}^2 de valeur propre $j(j+1)$ et de J_z de valeur propre m satisfaisant (0.56). On a bien retrouvé tous les résultats précédents. Sous cette forme, la construction de ces “représentations de plus haut poids” se généralise à d’autres algèbres de Lie, (même de dimension infinie, telle l’algèbre de Virasoro, voir Chap. 1, § 1.3.6).

Les matrices \mathcal{D}^j de la représentation de spin j sont telles que sous l’action de la rotation $U \in SU(2)$

$$|j m\rangle \mapsto D^j(U) |j m\rangle = |j m'\rangle \mathcal{D}_{m'm}^j(U). \quad (0.60)$$

3. En fait, on vient de trouver une condition nécessaire sur les j, m . Le fait que tous ces j donnent effectivement des représentations va être vérifié au paragraphe suivant.

Selon la paramétrisation $((\mathbf{n}, \psi)$, angles d'Euler, ...), on écrira aussi $\mathcal{D}_{m'm}^j(\mathbf{n}, \psi)$, $\mathcal{D}_{m'm}^j(\alpha, \beta, \gamma)$, etc. Par (0.7), on a donc

$$\begin{aligned} \mathcal{D}_{m'm}^j(\alpha, \beta, \gamma) &= \langle j m' | D(\alpha, \beta, \gamma) | j m \rangle \\ &= \langle j m' | e^{-i\alpha J_z} e^{-i\beta J_y} e^{-i\gamma J_z} | j m \rangle \\ &= e^{-i\alpha m'} d_{m'm}^j(\beta) e^{-i\gamma m} \end{aligned} \quad (0.61)$$

où la matrice d^j est définie par

$$d_{m'm}^j(\beta) = \langle j m' | e^{-i\beta J_y} | j m \rangle . \quad (0.62)$$

Une formule explicite pour d^j sera donnée au paragraphe suivant. On a encore

$$\begin{aligned} \mathcal{D}_{m'm}^j(\mathbf{z}, \psi) &= e^{-i\psi m} \delta_{mm'} \\ \mathcal{D}_{m'm}^j(\mathbf{y}, \psi) &= d_{m'm}^j(\psi) . \end{aligned} \quad (0.63)$$

Les matrices \mathcal{D}^j et d^j sont appelées *matrices de Wigner*.

Exercice : Calculer $\mathcal{D}^j(\mathbf{x}, \psi)$. (On pourra utiliser (0.5).)

On note que $\mathcal{D}^j(\mathbf{z}, 2\pi) = (-1)^{2j} I$, puisque $(-1)^{2m} = (-1)^{2j}$, compte tenu de (0.55), et la propriété est vraie pour tout axe \mathbf{n} par conjugaison (0.5)

$$\mathcal{D}^j(\mathbf{n}, 2\pi) = (-1)^{2j} I . \quad (0.64)$$

Cela montre qu'une rotation de 2π dans $SO(3)$ est représentée par $-I$ dans une représentation de spin demi-entier de $SU(2)$. Les représentations de spin demi-entier de $SU(2)$ sont des représentations "projectives", (c'est-à-dire ici à un signe près), de $SO(3)$; on reviendra au chapitre 2 sur la notion de représentation projective.

On vérifie aussi l'unimodularité des matrices \mathcal{D}^j (ou de façon équivalente, le fait que les représentants des générateurs infinitésimaux sont de trace nulle). Si $\mathbf{n} = R\mathbf{z}$, $\mathcal{D}(\mathbf{n}, \psi) = \mathcal{D}(R)\mathcal{D}(\mathbf{z}, \psi)\mathcal{D}^{-1}(R)$, donc

$$\det \mathcal{D}(\mathbf{n}, \psi) = \det \mathcal{D}(\mathbf{z}, \psi) = \det e^{-i\psi J_z} = \prod_{m=-j}^j e^{-im\psi} = 1 . \quad (0.65)$$

Il peut être utile d'écrire explicitement ces matrices dans les cas $j = \frac{1}{2}$ et $j = 1$. Le cas de $j = \frac{1}{2}$ est très simple, puisque

$$\begin{aligned} \mathcal{D}^{\frac{1}{2}}(U) &= U = e^{-i\frac{1}{2}\psi \mathbf{n} \cdot \boldsymbol{\sigma}} = \begin{pmatrix} \cos \frac{\psi}{2} - i \cos \theta \sin \frac{\psi}{2} & -i \sin \frac{\psi}{2} \sin \theta e^{-i\phi} \\ -i \sin \frac{\psi}{2} \sin \theta e^{i\phi} & \cos \frac{\psi}{2} + i \cos \theta \sin \frac{\psi}{2} \end{pmatrix} \\ &= e^{-i\frac{\alpha}{2}\sigma_3} e^{-i\frac{\beta}{2}\sigma_2} e^{-i\frac{\gamma}{2}\sigma_3} = \begin{pmatrix} \cos \frac{\beta}{2} e^{-i\frac{\alpha+\gamma}{2}} & -\sin \frac{\beta}{2} e^{-i\frac{\alpha-\gamma}{2}} \\ \sin \frac{\beta}{2} e^{i\frac{\alpha-\gamma}{2}} & \cos \frac{\beta}{2} e^{i\frac{\alpha+\gamma}{2}} \end{pmatrix} \end{aligned} \quad (0.66)$$

résultat attendu puisque les matrices U du groupe en forment bien évidemment une représentation. (Au passage, on a obtenu des relations entre les deux paramétrisations, $(\mathbf{n}, \psi) = (\theta, \phi, \psi)$)

et les angles d'Euler (α, β, γ) . Pour $j = 1$, dans la base $|1, 1\rangle$, $|1, 0\rangle$ et $|1, -1\rangle$ où J_z est diagonale (qui n'est pas la base (0.21)!) :

$$J_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad J_+ = \sqrt{2} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad J_- = \sqrt{2} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad (0.67)$$

d'où

$$d^1(\beta) = e^{-i\beta J_y} = \begin{pmatrix} \frac{1+\cos\beta}{2} & -\frac{\sin\beta}{\sqrt{2}} & \frac{1-\cos\beta}{2} \\ \frac{\sin\beta}{\sqrt{2}} & \cos\beta & -\frac{\sin\beta}{\sqrt{2}} \\ \frac{1-\cos\beta}{2} & \frac{\sin\beta}{\sqrt{2}} & \frac{1+\cos\beta}{2} \end{pmatrix} \quad (0.68)$$

comme le lecteur le vérifiera.

Dans le paragraphe qui suit on écrit plus explicitement ces matrices de représentation du groupe SU(2), et dans l'Appendice E du Chap. 2, on détaillera les équations différentielles qu'elles satisfont et leurs relations avec les "fonctions spéciales", polynômes orthogonaux et harmoniques sphériques. . .

Irréductibilité

Une notion centrale dans l'étude des représentations est celle d'irréductibilité. Une représentation est irréductible si elle n'admet aucun sous-espace invariant. Montrons que la représentation de spin j de SU(2) que nous venons de construire est irréductible. On montrera plus bas au Chap. 2 que la représentation étant unitaire, elle est soit irréductible soit "complètement réductible" (il existe un sous-espace invariant et son espace supplémentaire l'est aussi) ; dans ce dernier cas, il existerait nécessairement des opérateurs diagonaux par blocs, différents de l'identité et commutant avec les matrices de la représentation, en particulier avec les générateurs J_i . Or dans la base (0.5) toute matrice M commutant avec J_z est diagonale, $M_{mm'} = \mu_m \delta_{mm'}$, (le vérifier!), et la commutation avec J_+ force tous les μ_m à être égaux : la matrice M est multiple de l'identité et la représentation est bien irréductible.

On peut aussi se demander pourquoi l'étude des représentations de dimension finie que nous venons de construire suffit aux besoins du physicien, par exemple en mécanique quantique, où la scène se passe en général dans un espace de Hilbert de dimension infinie. On démontrera plus bas (Chap. 2) que

Toute représentation de SU(2) ou SO(3) dans un espace de Hilbert est équivalente à une représentation unitaire, et est complètement réductible en une somme (finie ou infinie) de représentations irréductibles de dimension finie.

0.3.3 Construction explicite

Soient ξ et η deux variables complexes sur lesquelles les matrices $U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ de SU(2) agissent selon $\xi' = a\xi + c\eta$, $\eta' = b\xi + d\eta$. En d'autres termes, ξ et η sont les vecteurs de base de la représentation de dimension 2 (représentation de spin $\frac{1}{2}$) de SU(2). Une construction explicite

des représentations précédentes est obtenue en considérant les polynômes homogènes de degré $2j$ dans les deux variables ξ et η , dont une base est donnée par les $2j + 1$ polynômes

$$P_{jm} = \frac{\xi^{j+m}\eta^{j-m}}{\sqrt{(j+m)!(j-m)!}} \quad m = -j, \dots, j. \quad (0.69)$$

(En fait, les considérations qui suivent demeurent valables si U est une matrice quelconque du groupe $GL(2, \mathbb{C})$ et en fournissent une représentation.) Sous l'action de U sur ξ et η , les $P_{jm}(\xi, \eta)$ se transforment en $P_{jm}(\xi', \eta')$, eux aussi homogènes de degré $2j$ en ξ et η , qui se développent donc sur les $P_{jm}(\xi, \eta)$. Ces derniers portent donc une représentation de dimension $2j + 1$ de $SU(2)$ (ou de $GL(2, \mathbb{C})$), qui n'est autre que la représentation de spin j précédente. Cela permet d'écrire des formules très explicites pour les \mathcal{D}^j .

$$P_{jm}(\xi', \eta') = \sum_{m'} P_{jm'}(\xi, \eta) \mathcal{D}_{m'm}^j(U). \quad (0.70)$$

On obtient

$$\mathcal{D}_{m'm}^j(U) = ((j+m)!(j-m)!(j+m')!(j-m')!)^{\frac{1}{2}} \sum_{\substack{n_1, n_2, n_3, n_4 \geq 0 \\ n_1+n_2=j+m'; \quad n_3+n_4=j-m' \\ n_1+n_3=j+m; \quad n_2+n_4=j-m}} \frac{a^{n_1} b^{n_2} c^{n_3} d^{n_4}}{n_1! n_2! n_3! n_4!}. \quad (0.71)$$

Pour $U = -\mathbf{I}$, on vérifie à nouveau que $\mathcal{D}^j(-\mathbf{I}) = (-1)^{2j} I$. Dans le cas particulier de $U = e^{-i\psi \frac{\sigma_2}{2}} = \cos \frac{\psi}{2} \mathbf{I} - i \sin \frac{\psi}{2} \sigma_2$, on a donc

$$d_{m'm}^j(\psi) = ((j+m)!(j-m)!(j+m')!(j-m')!)^{\frac{1}{2}} \sum_{k \geq 0} \frac{(-1)^{k+j-m} \cos \frac{\psi}{2}^{2k+m+m'} \sin \frac{\psi}{2}^{2j-2k-m-m'}}{(m+m'+k)!(j-m-k)!(j-m'-k)!k!} \quad (0.72)$$

où la somme court sur les $k \in [\inf(0, -m-m'), \sup(j-m, j-m')]$. L'expression des générateurs infinitésimaux sur les polynômes P_{jm} s'obtient en considérant des U proches de l'identité. On trouve

$$J_+ = \xi \frac{\partial}{\partial \eta} \quad J_- = \eta \frac{\partial}{\partial \xi} \quad J_z = \frac{1}{2} \left(\xi \frac{\partial}{\partial \xi} - \eta \frac{\partial}{\partial \eta} \right) \quad (0.73)$$

dont il est facile de vérifier les relations de commutation ainsi que l'action sur les P_{jm} en accord avec (0.57). Cela achève l'identification de (0.69) avec la représentation de spin j .

Remarques et exercices

1. Répéter la preuve de l'irréductibilité de la représentation de spin j dans cette nouvelle forme.
2. Noter que ce que les polynômes homogènes de degré $2j$ dans les variables ξ et η ont construit n'est autre que la puissance tensorielle $2j$ symétrisée de la représentation de dimension 2. (Voir ci-dessous la définition de ce concept.)
3. Écrire la forme explicite de la matrice \mathcal{D}^1 de spin 1 en utilisant (0.71).

0.4 Produit direct de représentations de SU(2)

0.4.1 Produit direct de représentations et l'“addition de moments angulaires”

Intéressons-nous au produit de deux représentations de spin j_1 et j_2 et à leur décomposition sur des vecteurs de spin total donné (“décomposition en représentations irréductibles”). On part donc de la représentation produit engendrée par les vecteurs

$$|j_1 m_1\rangle \otimes |j_2 m_2\rangle \equiv |j_1 m_1; j_2 m_2\rangle \quad \text{abrégé en} \quad |m_1 m_2\rangle \quad (0.74)$$

sur lesquels agissent les générateurs infinitésimaux sous la forme

$$\mathbf{J} = \mathbf{J}^{(1)} \otimes \mathbf{I}^{(2)} + \mathbf{I}^{(1)} \otimes \mathbf{J}^{(2)}. \quad (0.75)$$

L'indice supérieur indique sur quel espace agissent les opérateurs. Par abus de notation, on écrit souvent au lieu de (0.75)

$$\mathbf{J} = \mathbf{J}^{(1)} + \mathbf{J}^{(2)} \quad (0.75')$$

et (en Mécanique Quantique), on parle de l'“addition des moments angulaires” $J^{(1)}$ et $J^{(2)}$. Il s'agit donc de décomposer les vecteurs (0.74) sur une base de vecteurs propres de \mathbf{J} et J_z . Comme $\mathbf{J}^{(1)2}$ et $\mathbf{J}^{(2)2}$ commutent entre eux et avec \mathbf{J}^2 et J_z , on peut chercher des vecteurs propres communs que l'on notera

$$|(j_1 j_2) J M\rangle \quad \text{ou plus simplement} \quad |J M\rangle \quad (0.76)$$

étant entendu que l'on s'est fixé la valeur de j_1 et j_2 . La question est donc double : quelles valeurs J et M peuvent-ils prendre et quelle est la matrice du changement de base $|m_1 m_2\rangle \rightarrow |J M\rangle$? En d'autres termes quelle est la décomposition (de Clebsch-Gordan) et quels sont les coefficients de Clebsch-Gordan ?

Les valeurs possibles de M , valeur propre de $J_z = J_z^{(1)} + J_z^{(2)}$ sont aisées à trouver

$$\begin{aligned} \langle m_1 m_2 | J_z | J M \rangle &= (m_1 + m_2) \langle m_1 m_2 | J M \rangle \\ &= M \langle m_1 m_2 | J M \rangle \end{aligned} \quad (0.77)$$

et la seule valeur de M telle que $\langle m_1 m_2 | J M \rangle \neq 0$ est donc

$$M = m_1 + m_2. \quad (0.78)$$

A j_1 , j_2 et M fixés, il y a autant de vecteurs indépendants ayant cette valeur de M qu'il y a de couples (m_1, m_2) satisfaisant (0.78), soit

$$n(M) = \begin{cases} 0 & \text{si } |M| > j_1 + j_2 \\ j_1 + j_2 + 1 - |M| & \text{si } |j_1 - j_2| \leq |M| \leq j_1 + j_2 \\ 2 \inf(j_1, j_2) + 1 & \text{si } 0 \leq |M| \leq |j_1 - j_2| \end{cases} \quad (0.79)$$

(voir Fig. 3 pour laquelle $j_1 = 5/2$ et $j_2 = 1$). Soit N_J le nombre de fois où la représentation de spin J apparaît dans la décomposition du produit des représentations de spin j_1 et j_2 . Les $n(M)$ vecteurs de valeur propre M pour J_z peuvent aussi s'interpréter comme provenant des N_J vecteurs $|JM\rangle$ pour les différentes valeurs de J compatibles avec cette valeur de M

$$n(M) = \sum_{J \geq |M|} N_J \tag{0.80}$$

soit en retranchant membre à membre deux telles relations

$$\begin{aligned} N_J &= n(J) - n(J+1) \\ &= 1 \quad \text{si et seulement si} \quad |j_1 - j_2| \leq J \leq j_1 + j_2 \\ &= 0 \quad \text{sinon.} \end{aligned} \tag{0.81}$$

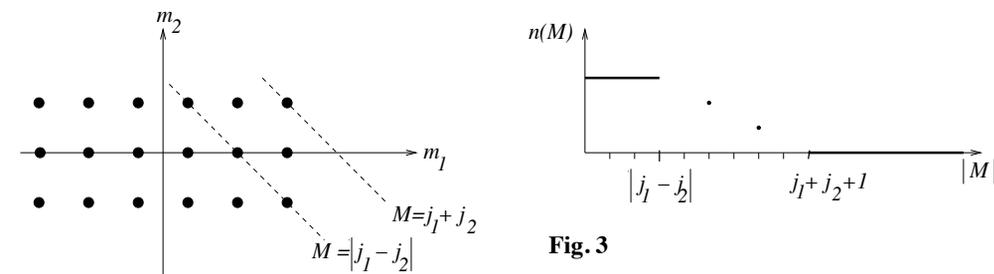


Fig. 3

En conclusion, nous venons de démontrer que les $(2j_1 + 1)(2j_2 + 1)$ vecteurs (0.74) (à j_1 et j_2 fixés) peuvent se réexprimer en fonction des vecteurs $|JM\rangle$ où

$$\begin{aligned} J &= |j_1 - j_2|, |j_1 - j_2| + 1, \dots, j_1 + j_2 \\ M &= -J, -J + 1, \dots, J. \end{aligned} \tag{0.82}$$

Noter qu'en définitive les multiplicités N_J valent 0 ou 1; c'est une particularité de $SU(2)$ que des multiplicités supérieures à 1 n'apparaissent pas dans la décomposition du produit de deux représentations "irréductibles", c'est-à-dire ici de spin fixé.

0.4.2 Coefficients de Clebsch-Gordan, symboles 3-j et 6-j ...

Le changement de base orthonormée $|j_1 m_1; j_2 m_2\rangle \rightarrow |(j_1 j_2) JM\rangle$ s'effectue à l'aide des coefficients de Clebsch-Gordan (C.G.) $\langle (j_1 j_2); JM | j_1 m_1; j_2 m_2 \rangle$ qui forment une matrice unitaire

$$|j_1 m_1; j_2 m_2\rangle = \sum_{J=|j_1-j_2|}^{j_1+j_2} \sum_{M=-J}^J \langle (j_1 j_2) JM | j_1 m_1; j_2 m_2 \rangle |(j_1 j_2) JM\rangle \tag{0.83}$$

$$|(j_1 j_2); JM\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} \langle (j_1 j_2) JM | j_1 m_1; j_2 m_2 \rangle^* |j_1 m_1; j_2 m_2\rangle. \tag{0.84}$$

Leur valeur dépend en fait d'un choix de phase relative entre les vecteurs (0.74) et (0.76); la convention habituelle est que pour chaque valeur de J , on choisit

$$\langle JM = J | j_1 m_1 = j_1; j_2 m_2 = J - j_1 \rangle \quad \text{real.} \tag{0.85}$$

Les autres vecteurs sont alors définis sans ambiguïté par (0.57) et on va montrer que tous les C.G. sont réels. Les C.G. satisfont des relations de récurrence conséquences de (0.57). Appliquant en effet J_{\pm} aux deux membres de (0.83), on obtient

$$\begin{aligned} \sqrt{J(J+1) - M(M \pm 1)} & \langle (j_1 j_2) J M | j_1 m_1; j_2 m_2 \rangle \\ & = \sqrt{j_1(j_1+1) - m_1(m_1 \pm 1)} \langle (j_1 j_2) J M \pm 1 | j_1 m_1 \pm 1; j_2 m_2 \rangle \\ & + \sqrt{j_2(j_2+1) - m_2(m_2 \pm 1)} \langle (j_1 j_2) J M \pm 1 | j_1 m_1; j_2 m_2 \pm 1 \rangle \end{aligned} \quad (0.86)$$

qui permet à l'aide de la normalisation $\sum_{m_1, m_2} |\langle j_1 m_1; j_2 m_2 | (j_1 j_2) J M \rangle|^2 = 1$ et de la convention (0.85) de déterminer tous les C.G. Comme annoncé, ils sont clairement tous réels.

Les C.G. du groupe SU(2), qui décrivent un changement de base orthonormée, satisfont des propriétés d'orthogonalité et de complétude

$$\sum_{m_1=-j_1}^{j_1} \langle j_1 m_1; j_2 m_2 | (j_1 j_2) J M \rangle \langle j_1 m_1; j_2 m_2 | (j_1 j_2) J' M' \rangle = \delta_{JJ'} \delta_{MM'} \quad \text{si } |j_1 - j_2| \leq J \leq j_1 + j_2 \quad (0.87)$$

$$\sum_{J=|j_1-j_2|}^{j_1+j_2} \langle j_1 m_1; j_2 m_2 | (j_1 j_2) J M \rangle \langle j_1 m'_1; j_2 m'_2 | (j_1 j_2) J M \rangle = \delta_{m_1 m'_1} \delta_{m_2 m'_2} \quad \text{si } |m_1| \leq j_1, |m_2| \leq j_2.$$

Noter que dans la première ligne, m_2 est fixé par la donnée de m_1 , à M donné; et que dans la deuxième, M est fixé en termes de m_1 et de m_2 . Chaque relation n'implique donc qu'une seule somme.

Plutôt que les coefficients de Clebsch-Gordan, on peut considérer un ensemble de coefficients équivalents, dits *symboles 3-j*. Ils sont définis par

$$\begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & -M \end{pmatrix} = \frac{(-1)^{j_1-j_2+M}}{\sqrt{2J+1}} \langle j_1 m_1; j_2 m_2 | (j_1 j_2) J M \rangle \quad (0.88)$$

et ont l'intérêt de jouir de propriétés de symétrie simples :

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$$

est invariant par permutation circulaire des trois colonnes et change par le signe $(-1)^{j_1+j_2+j_3}$ quand deux colonnes sont permutées ou quand on change les signes de m_1 , m_2 et m_3 . Le lecteur trouvera dans la littérature de nombreuses tables et formules explicites.

Contentons-nous de donner les valeurs pour les spins les plus bas

$$\frac{1}{2} \otimes \frac{1}{2} : \begin{aligned} |(\frac{1}{2}, \frac{1}{2})1, 1\rangle & = |(\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2})\rangle \\ |(\frac{1}{2}, \frac{1}{2})1, 0\rangle & = \frac{1}{\sqrt{2}} (|(\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2})\rangle + |(\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}, \frac{1}{2})\rangle) \\ |(\frac{1}{2}, \frac{1}{2})0, 0\rangle & = \frac{1}{\sqrt{2}} (|(\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2})\rangle - |(\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}, \frac{1}{2})\rangle) \\ |(\frac{1}{2}, \frac{1}{2})1, -1\rangle & = |(\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}, -\frac{1}{2})\rangle \end{aligned} \quad (0.89)$$

et

$$\begin{aligned}
 \frac{1}{2} \otimes 1 : \quad & |(\frac{1}{2}, 1) \frac{3}{2}, \frac{3}{2}\rangle = |\frac{1}{2}, \frac{1}{2}; 1, 1\rangle \\
 & |(\frac{1}{2}, 1) \frac{3}{2}, \frac{1}{2}\rangle = \frac{1}{\sqrt{3}} (\sqrt{2}|\frac{1}{2}, \frac{1}{2}; 1, 0\rangle + |\frac{1}{2}, -\frac{1}{2}; 1, 1\rangle) \\
 & |(\frac{1}{2}, 1) \frac{3}{2}, -\frac{1}{2}\rangle = \frac{1}{\sqrt{3}} (|\frac{1}{2}, \frac{1}{2}; 1, -1\rangle + \sqrt{2}|\frac{1}{2}, -\frac{1}{2}; 1, 0\rangle) \\
 & |(\frac{1}{2}, 1) \frac{3}{2}, -\frac{3}{2}\rangle = |\frac{1}{2}, -\frac{1}{2}; 1, -1\rangle \\
 & |(\frac{1}{2}, 1) \frac{1}{2}, \frac{1}{2}\rangle = \frac{1}{\sqrt{3}} (-|\frac{1}{2}, \frac{1}{2}; 1, 0\rangle + \sqrt{2}|\frac{1}{2}, -\frac{1}{2}; 1, 1\rangle) \\
 & |(\frac{1}{2}, 1) \frac{1}{2}, -\frac{1}{2}\rangle = \frac{1}{\sqrt{3}} (-\sqrt{2}|\frac{1}{2}, \frac{1}{2}; 1, -1\rangle + |\frac{1}{2}, -\frac{1}{2}; 1, 0\rangle)
 \end{aligned} \tag{0.90}$$

On note sur le cas $\frac{1}{2} \otimes \frac{1}{2}$ la propriété que les vecteurs de spin total $j = 1$ sont symétriques dans l'échange des deux spins, celui de spin 0 antisymétrique. La propriété est générale : dans la composition de deux représentations de spin $j_1 = j_2$, les vecteurs résultants de spin $j = 2j_1, 2j_1 - 2, \dots$ sont symétriques, ceux de spin $2j_1 - 1, 2j_1 - 3, \dots$ sont antisymétriques. Cela est apparent sur l'expression (0.88) ci-dessus, compte tenu des propriétés annoncées des symboles 3- j .

Dans le même ordre d'idées, soit le produit complètement antisymétrique de $2j + 1$ copies d'une représentation de spin j . On peut montrer que cette représentation est de spin 0 (exercice suivant). (Cela a une conséquence en physique atomique, dans le remplissage des couches électroniques : une couche complète a un moment orbital total et un spin total nuls donc aussi un moment angulaire total nul.)

Exercice. On considère le produit complètement antisymétrique de $N = 2j + 1$ représentations de spin j . Montrer que cette représentation est engendrée par le vecteur $\epsilon_{m_1 m_2 \dots m_N} |j m_1, j m_2, \dots, j m_N\rangle$, qu'il est invariant par l'action de SU(2) et donc que la représentation construite est celle de spin $J = 0$.

On introduit aussi les *symboles 6- j* qui décrivent les deux recombinaisons possibles de 3 représentations de spins j_1, j_2 et j_3

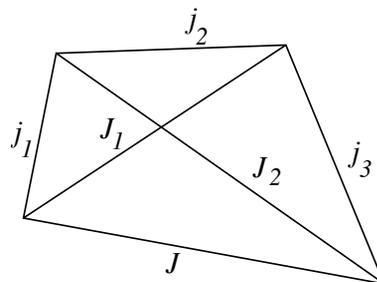


Fig. 4

$$\begin{aligned}
 |j_1 m_1; j_2 m_2; j_3 m_3\rangle &= \sum \langle (j_1 j_2) J_1 M_1 | j_1 m_1; j_2 m_2 \rangle \langle (J_1 j_3) J M | J_1 M_1; j_3 m_3 \rangle | (j_1 j_2) j_3; J M \rangle \\
 &= \sum \langle (j_2 j_3) J_2 M_2 | j_2 m_2; j_3 m_3 \rangle \langle (j_1 J_2) J' M' | j_1 m_1; J_2 M_2 \rangle | j_1 (j_2 j_3); J' M' \rangle
 \end{aligned} \tag{0.91}$$

selon que l'on compose d'abord j_1 et j_2 en J_1 puis J_1 et j_3 en J ou d'abord j_2 et j_3 en J_2 puis j_1 et J_2 en J' . La matrice de changement de base est notée

$$\langle j_1 (j_2 j_3); J M | (j_1 j_2) j_3; J' M' \rangle = \delta_{JJ'} \delta_{MM'} \sqrt{(2J_1 + 1)(2J_2 + 1)} (-1)^{j_1 + j_2 + j_3 + J} \begin{Bmatrix} j_1 & j_2 & J_1 \\ j_3 & J & J_2 \end{Bmatrix}. \tag{0.92}$$

et les $\left\{ \begin{matrix} j_1 & j_2 & J_1 \\ j_3 & J & J_2 \end{matrix} \right\}$ sont les symboles 6- j . On visualise l'opération d'addition des trois spins par un tétraèdre (cf. Fig. 4) dont les arêtes portent j_1, j_2, j_3, J_1, J_2 et J et le symbole est tel que deux spins portés par une paire d'arêtes opposées se trouvent dans la même colonne. Ces symboles sont tabulés dans la littérature.

0.5 Une application physique : l'isospin

Le groupe $SU(2)$ n'intervient pas en Physique qu'en tant que (relié au) groupe de rotation de l'espace euclidien. Illustrons une autre de ses apparitions par la symétrie d'isospin. Il existe dans la nature des particules élémentaires sujettes aux interactions fortes et de ce fait, appelées *hadrons*. Certaines de ces particules présentent des propriétés voisines, mais diffèrent par leur charge électrique. C'est le cas du proton et du neutron, de masses $938,28 \text{ MeV}/c^2$ et $939,57 \text{ MeV}/c^2$ respectivement, mais aussi du triplet de mésons pi, π^0 (masse $134,96 \text{ MeV}/c^2$) et π^\pm ($139,57 \text{ MeV}/c^2$), des quatre mésons K etc. Selon Heisenberg cela est la manifestation d'une symétrie brisée par les effets électromagnétiques. En l'absence d'interactions électromagnétiques, le proton et le neutron d'une part, les trois mésons π de l'autre seraient des particules de même nature, de même masse, différant seulement par un nombre quantique "interne", à la façon de deux électrons dotés de spins différents. En fait le groupe régissant cette symétrie est aussi $SU(2)$, mais un $SU(2)$ agissant dans un espace abstrait autre que l'espace usuel. On a donné le nom de *spin isotopique* ou *isospin* au nombre quantique correspondant. Pour résumer, la proposition est donc qu'il existe un groupe $SU(2)$ de symétrie de l'Hamiltonien des interactions fortes, et que les différentes particules sujettes à ces interactions forment des représentations de $SU(2)$: représentation d'isospin $I = \frac{1}{2}$ pour le nucléon (proton $I_z = +\frac{1}{2}$, neutron $I_z = -\frac{1}{2}$), isospin $I = 1$ pour les pions ($\pi^\pm : I_z = \pm 1$, $\pi^0 : I_z = 0$), etc. L'isospin est donc un "bon nombre quantique", conservé dans ces interactions. Ainsi le processus (virtuel) $N \rightarrow N + \pi$, (N pour nucléon) important en physique nucléaire, est compatible avec les règles d'addition des isospins ($\frac{1}{2} \otimes 1$ "contient" $\frac{1}{2}$). Les différentes réactions $N + \pi \rightarrow N + \pi$ autorisées par la conservation de la charge électrique

$$\begin{array}{lll}
 p + \pi^+ & \rightarrow & p + \pi^+ & I_z = \frac{3}{2} \\
 p + \pi^0 & \rightarrow & p + \pi^0 & I_z = \frac{1}{2} \\
 & & \rightarrow n + \pi^+ & \text{"} \\
 p + \pi^- & \rightarrow & p + \pi^- & I_z = -\frac{1}{2} \\
 & & \rightarrow n + \pi^0 & \text{"} \\
 n + \pi^- & \rightarrow & n + \pi^- & I_z = -\frac{3}{2}
 \end{array}$$

conservent aussi l'isospin total I et sa composante I_z mais l'hypothèse d'invariance par $SU(2)$ d'isospin nous apprend d'avantage. Les éléments de matrice de transition des deux réactions dans le canal $I_z = \frac{1}{2}$, par exemple, doivent être reliés par les règles d'addition de l'isospin. En inversant les relations (0.90), on obtient

$$\begin{aligned}
 |p, \pi^- \rangle &= \sqrt{\frac{1}{3}} |I = \frac{3}{2}, I_z = -\frac{1}{2} \rangle - \sqrt{\frac{2}{3}} |I = \frac{1}{2}, I_z = -\frac{1}{2} \rangle \\
 |n, \pi^0 \rangle &= \sqrt{\frac{2}{3}} |I = \frac{3}{2}, I_z = -\frac{1}{2} \rangle + \sqrt{\frac{1}{3}} |I = \frac{1}{2}, I_z = -\frac{1}{2} \rangle
 \end{aligned}$$

tandis que pour $I_z = 3/2$

$$|p, \pi^+ \rangle = |I = \frac{3}{2}, I_z = \frac{3}{2} \rangle .$$

L'invariance d'isospin implique que $\langle I I_z | \mathcal{T} | I' I'_z \rangle = \mathcal{T}_I \delta_{I I'} \delta_{I_z I'_z}$, : non seulement I et I_z sont conservés, mais l'amplitude ne dépend que de I , pas de I_z (comme on le justifiera plus tard

au chap. 2, par le lemme de Schur ou le théorème de Wigner-Eckart). En calculant alors les éléments de matrice de l'opérateur de transition \mathcal{T} entre ces différents états,

$$\begin{aligned} \langle p\pi^+ | \mathcal{T} | p\pi^+ \rangle &= \mathcal{T}_{3/2} \\ \langle p\pi^- | \mathcal{T} | p\pi^- \rangle &= \frac{1}{3} (\mathcal{T}_{3/2} + 2\mathcal{T}_{1/2}) \\ \langle n\pi^0 | \mathcal{T} | p\pi^- \rangle &= \frac{\sqrt{2}}{3} (\mathcal{T}_{3/2} - \mathcal{T}_{1/2}) \end{aligned} \quad (0.93)$$

on trouve que les amplitudes satisfont une relation

$$\sqrt{2} \langle n, \pi^0 | \mathcal{T} | p, \pi^- \rangle + \langle p, \pi^- | \mathcal{T} | p, \pi^- \rangle = \langle p, \pi^+ | \mathcal{T} | p, \pi^+ \rangle = \mathcal{T}_{3/2}$$

conséquence non triviale de l'invariance d'isospin, qui implique des inégalités triangulaires entre les modules carrés de ces amplitudes donc entre les sections efficaces de ces réactions

$$\begin{aligned} [\sqrt{\sigma(\pi^- p \rightarrow \pi^- p)} - \sqrt{2\sigma(\pi^- p \rightarrow \pi^0 n)}]^2 &\leq \sigma(\pi^+ p \rightarrow \pi^+ p) \leq \\ &\leq [\sqrt{\sigma(\pi^- p \rightarrow \pi^- p)} + \sqrt{2\sigma(\pi^- p \rightarrow \pi^0 n)}]^2 \end{aligned}$$

qui sont bien vérifiées expérimentalement. Mieux, on constate qu'à une énergie d'environ 180 MeV, les sections efficaces (proportionnelles aux carrés des amplitudes) sont dans les rapports

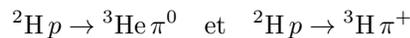
$$\sigma(\pi^+ p \rightarrow \pi^+ p) : \sigma(\pi^- p \rightarrow \pi^0 n) : \sigma(\pi^- p \rightarrow \pi^- p) = 9 : 2 : 1,$$

qui est ce qu'on obtiendrait à partir de (0.93) si on avait $\mathcal{T}_{\frac{1}{2}} = 0$. Cela indique qu'à cette énergie la diffusion dans le canal d'isospin 3/2 est prédominante et signale en fait l'existence d'un état intermédiaire, particule très instable ou "résonance", notée Δ , d'isospin 3/2 donc avec quatre états de charge

$$\Delta^{++}, \Delta^+, \Delta^0, \Delta^-$$

dont la contribution domine l'amplitude. Cette particule a un spin 3/2 et une masse $M(\Delta) \approx 1230 \text{ MeV}/c^2$.

Dans certains cas on peut parvenir à des prédictions plus précises. C'est le cas par exemple dans l'étude des réactions



impliquant des noyaux de deutérium ${}^2\text{H}$, de tritium ${}^3\text{H}$ et d'hélium ${}^3\text{He}$. A ces noyaux aussi on peut attribuer un isospin, 0 au deutéron qui est formé d'un proton et d'un neutron dans un état antisymétrique de leurs isospins (pour que la fonction d'onde de ces deux *fermions*, symétrique d'espace et de spin, soit antisymétrique), $I_z = -\frac{1}{2}$ à ${}^3\text{H}$ et $I_z = \frac{1}{2}$ à ${}^3\text{He}$ qui forment une représentation d'isospin $\frac{1}{2}$. Noter que dans tous les cas, la charge électrique est reliée à la composante I_z de l'isospin par la relation $Q = \frac{1}{2}\mathcal{B} + I_z$, avec \mathcal{B} la charge baryonique, égale ici au nombre de nucléons (protons ou neutrons). Montrer que le rapport des sections efficaces $\sigma({}^2\text{H}p \rightarrow {}^3\text{He}\pi^0)/\sigma({}^2\text{H}p \rightarrow {}^3\text{H}\pi^+)$ est $\frac{1}{2}$.

Remarque : l'invariance par $SU(2)$ d'isospin que nous venons de discuter est une symétrie des *interactions fortes*. Il existe aussi dans le cadre du Modèle Standard une notion d'"isospin faible", symétrie des *interactions électro-faibles*, on y reviendra au Chap. 5.

0.6 Représentations de SO(1,3) et SL(2,ℂ)

0.6.1 Petit rappel sur le groupe de Lorentz

L'espace de Minkowski est un espace \mathbb{R}^4 muni d'une métrique pseudo-euclidienne de signature $(+, -, -, -)$. Dans une base orthonormale avec des coordonnées $(x^0 = ct, x^1, x^2, x^3)$, la métrique est diagonale

$$g_{\mu\nu} = \text{diag}(1, -1, -1, -1)$$

et la norme carrée d'un 4-vecteur s'écrit

$$x \cdot x = x^\mu g_{\mu\nu} x^\nu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2.$$

Le groupe d'isométrie de cette forme quadratique, appelé O(1,3) ou *groupe de Lorentz* \mathcal{L} , est tel que

$$\Lambda \in \text{O}(1,3) \quad x' = \Lambda x \quad : \quad x' \cdot x' = \Lambda^\mu{}_\rho x^\rho g_{\mu\nu} \Lambda^\nu{}_\sigma x^\sigma = x^\rho g_{\rho\sigma} x^\sigma$$

c'est-à-dire

$$\Lambda^\mu{}_\rho g_{\mu\nu} \Lambda^\nu{}_\sigma = g_{\rho\sigma} \quad \text{or} \quad \Lambda^T g \Lambda = g. \quad (0.94)$$

Ces matrices pseudo-orthogonales satisfont $(\det \Lambda)^2 = 1$ et (en prenant l'élément de matrice 00 de (0.94)) $(\Lambda^0_0)^2 = 1 + \sum_{i=1}^3 (\Lambda^0_i)^2 \geq 1$ et donc $\mathcal{L} \equiv \text{O}(1,3)$ a quatre composantes connexes (or "nappes") selon que $\det \Lambda = \pm 1$ et $\Lambda^0_0 \geq 1$ ou ≤ -1 . Le sous-groupe des transformations *propres orthochrones* satisfaisant $\det \Lambda = 1$ et $\Lambda^0_0 \geq 1$ est noté \mathcal{L}_+^\uparrow . Toute transformation de \mathcal{L}_+^\uparrow peut être écrite comme le produit d'une rotation "ordinaire" de SO(3) et d'une "transformation spéciale de Lorentz" ou "boost".

Une différence majeure entre les groupes SO(3) et \mathcal{L}_+^\uparrow est que le premier est *compact* (l'ensemble de ses paramètres est borné et fermé, voir (0.2)), tandis que le second ne l'est pas : dans un "boost" dans la direction 1, disons, $x'_1 = \gamma(x_1 - vx_0/c)$, $x'_0 = \gamma(x_0 - vx_1/c)$, avec $\gamma = (1 - v^2/c^2)^{-\frac{1}{2}}$, la vitesse $|v| < c$ (inégalité stricte!) n'appartient pas à un domaine compact (ou encore, la variable de "rapidité" β , définie par $\cosh \beta = \gamma$ peut aller à l'infini). Cette compacité ou non a de très importantes implications sur la nature et les propriétés des représentations, comme nous allons voir.

Le *groupe de Poincaré* ou groupe de Lorentz inhomogène est engendré par les transformations de Lorentz $\Lambda \in \mathcal{L}$ et les translations d'espace-temps ; on peut noter (a, Λ) son élément générique avec une action sur un quadrivecteur x de l'espace de Minkowski et une loi de composition données par

$$\begin{aligned} (a, \Lambda) \quad : \quad x \mapsto x' &= \Lambda x + a \\ (a', \Lambda')(a, \Lambda) &= (a' + \Lambda' a, \Lambda' \Lambda) ; \end{aligned} \quad (0.95)$$

l'inverse de (a, Λ) est $(-\Lambda^{-1}a, \Lambda^{-1})$ (le vérifier!).

0.6.2 Algèbre de Lie des groupes de Lorentz et Poincaré

Une transformation infinitésimale de Poincaré s'écrit $(\alpha^\mu, \Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu)$. Prenant la forme infinitésimale de (0.94), on voit aisément que le tenseur $\omega_{\rho\nu} = \omega^\mu{}_\nu g_{\rho\mu}$ doit être antisymétrique : $\omega_{\nu\rho} + \omega_{\rho\nu} = 0$. Cela laisse 6 paramètres réels : le groupe de Lorentz est un groupe de dimension 6, et le groupe de Poincaré de dimension 10.

Pour déterminer l'algèbre de Lie des générateurs, procédons comme au § 0.2.3 : regardons l'algèbre de Lie engendrée par des opérateurs différentiels agissant sur les fonctions des coordonnées ; si $x'^\lambda = x^\lambda + \delta x^\lambda = x^\lambda + \alpha^\lambda + \omega^{\lambda\nu} x_\nu$, $\delta f(x) = f(x^\mu - \alpha^\lambda - \omega^{\lambda\nu} x_\nu) - f(x) = (I - i\alpha^\mu P_\mu - \frac{i}{2}\omega^{\mu\nu} J_{\mu\nu})f(x)$, (cf (0.45), donc

$$J_{\mu\nu} = i(x_\mu \partial_\nu - x_\nu \partial_\mu) \quad P_\mu = -i\partial_\mu \quad (0.96)$$

dont on calcule aisément les commutateurs

$$\begin{aligned} [J_{\mu\nu}, P_\rho] &= i(g_{\nu\rho} P_\mu - g_{\mu\rho} P_\nu) \\ [J_{\mu\nu}, J_{\rho\sigma}] &= i(g_{\nu\rho} J_{\mu\sigma} - g_{\mu\rho} J_{\nu\sigma} + g_{\mu\sigma} J_{\nu\rho} - g_{\nu\sigma} J_{\mu\rho}) \\ [P_\mu, P_\nu] &= 0 \end{aligned} \quad (0.97)$$

Noter la structure de ces relations : antisymétrie en $\mu \leftrightarrow \nu$ de la première, en $\mu \leftrightarrow \nu$, en $\rho \leftrightarrow \sigma$ et en $(\mu, \nu) \leftrightarrow (\rho, \sigma)$ de la seconde ; la première montre comment un vecteur (ici P_ρ) se transforme sous la transformation infinitésimale engendrée par $J_{\mu\nu}$, et la seconde a cette même forme dans les indices ρ et σ , exprimant que $J_{\rho\sigma}$ est un 2-tenseur.

Les générateurs qui commutent avec P_0 (qui est le générateur des translations de temps, donc le hamiltonien) sont les P_μ et les J_{ij} mais pas les J_{0j} : $i[P_0, J_{0j}] = P_j$.

Posons

$$J_{ij} = \epsilon_{ijk} J^k \quad K^i = J_{0i} . \quad (0.98)$$

On a alors

$$\begin{aligned} [J^i, J^j] &= i\epsilon_{ijk} J^k \\ [J^i, K^j] &= i\epsilon_{ijk} K^k \\ [K^i, K^j] &= -i\epsilon_{ijk} J^k \end{aligned} \quad (0.99)$$

et aussi

$$\begin{aligned} [J^i, P^j] &= i\epsilon_{ijk} P^k & [K^i, P^j] &= iP^0 \delta_{ij} \\ [J^i, P^0] &= 0 & [K^i, P^0] &= iP^i . \end{aligned} \quad (0.100)$$

N.B. Les deux premières des relations (0.99) et la première de (0.100) expriment bien, comme attendu, que $\mathbf{J} = \{J^j\}$, $\mathbf{K} = \{K^j\}$ et $\mathbf{P} = \{P^j\}$ se transforment comme des vecteurs sous l'action des rotations de \mathbb{R}^3 . Formons les combinaisons

$$M^j = \frac{1}{2}(J^j + iK^j) \quad N^j = \frac{1}{2}(J^j - iK^j) \quad (0.101)$$

elles satisfont

$$[M^i, M^j] = i\epsilon_{ijk} M^k$$

$$\begin{aligned} [N^i, N^j] &= i\epsilon_{ijk}N^k \\ [M^i, N^j] &= 0. \end{aligned} \quad (0.102)$$

On voit donc que, si on considère les combinaisons complexes M et N de ses générateurs, l'algèbre de Lie de $\mathcal{L} = O(1,3)$ est isomorphe à $su(2) \oplus su(2)$. L'introduction du $\pm i$, cependant, fait que les représentations *unitaires* de \mathcal{L} ne découlent pas simplement de celles de $SU(2) \times SU(2)$. Les représentations de dimension finie, *non unitaires*, de \mathcal{L} sont indexées par une paire (j_1, j_2) , entiers ou demi-entiers.

Exercice. Montrer que cette algèbre admet *deux* opérateurs de Casimir quadratiques indépendants, et les exprimer en termes de \mathbf{M} et \mathbf{N} d'abord, puis de \mathbf{J} et \mathbf{K} .

0.6.3 Groupes de recouvrement de \mathcal{L}_+^\uparrow et \mathcal{P}_+^\uparrow

De la même façon que de l'étude de $SO(3)$ on a été conduit (pour des raisons qui seront discutées aux chapitres 1 et 2) à celle de $SU(2)$, son "groupe de recouvrement", de même dans le cas du groupe de Lorentz, on est amené à étudier son groupe de recouvrement $SL(2,\mathbb{C})$.

Il existe une manière simple de voir comment $SL(2,\mathbb{C})$ et \mathcal{L}_+^\uparrow sont reliés, qui est une extension quadri-dimensionnelle de la méthode suivie au § 0.1.2. On utilise les matrices σ_μ constituées de $\sigma_0 = \mathbf{I}$ et des trois matrices de Pauli familières. Notons que l'on a

$$\text{tr } \sigma_\mu \sigma_\nu = 2\delta_{\mu\nu} \quad \sigma_\mu^2 = \mathbf{I} \quad \text{sans sommation sur l'indice } \mu.$$

À tout vecteur réel $x \in \mathbb{R}^4$, associons la matrice hermitique

$$X = x^\mu \sigma_\mu \quad x^\mu = \frac{1}{2} \text{tr}(X \sigma_\mu) \quad \det X = x^2 = (x^0)^2 - \mathbf{x}^2.$$

Une matrice $A \in SL(2, \mathbb{C})$ agit sur X selon

$$X \mapsto X' = AXA^\dagger$$

qui est bien hermitique et définit donc $x'^\mu = \frac{1}{2} \text{tr}(X' \sigma_\mu)$ réel, avec $\det X' = \det X$, donc $x^2 = x'^2$. C'est une transformation linéaire de \mathbb{R}^4 dans \mathbb{R}^4 qui préserve la norme minkovskienne x^2 , c'est donc une transformation de Lorentz, et on vérifie qu'elle est dans \mathcal{L}_+^\uparrow et que $A \rightarrow \Lambda$ est un homomorphisme de $SL(2, \mathbb{C})$ dans \mathcal{L}_+^\uparrow . On notera dans la suite $x' = A.x$ si $X' = AXA^\dagger$.

Cependant, les deux transformations A et $-A \in SL(2, \mathbb{C})$ donnent la même transformation de $\mathcal{L}_+^\uparrow : SL(2, \mathbb{C})$ est un recouvrement d'ordre 2 de \mathcal{L}_+^\uparrow . Pour le groupe de Poincaré, on raisonne de même, son recouvrement est le produit ("semi-direct") du groupe des translations par $SL(2, \mathbb{C})$. Si on note $\underline{a} := a^\mu \sigma_\mu$

$$(\underline{a}, A)(\underline{a}', A') = (\underline{a} + A\underline{a}'A^\dagger, AA')$$

(on parle aussi du "groupe $SL(2, \mathbb{C})$ inhomogène", ou $ISL(2, \mathbb{C})$).

0.6.4 Représentations irréductibles de dimension finie de $SL(2, \mathbb{C})$

La construction du § 0.3.3 fournit une représentation explicite de $GL(2, \mathbb{C})$ et donc de $SL(2, \mathbb{C})$.

Pour $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{C})$, (0.71) donne l'expression de $\mathcal{D}_{mm'}^j(A)$:

$$\mathcal{D}_{mm'}^j(A) = [(j+m)!(j-m)!(j+m')!(j-m')!]^{\frac{1}{2}} \sum_{\substack{n_1, n_2, n_3, n_4 \geq 0 \\ n_1+n_2=j+m; \quad n_3+n_4=j-m' \\ n_1+n_3=j+m; \quad n_2+n_4=j-m}} \frac{a^{n_1} b^{n_2} c^{n_3} d^{n_4}}{n_1! n_2! n_3! n_4!} \quad (0.71)$$

Noter que $\mathcal{D}^T(A) = \mathcal{D}(A^T)$ (car échanger $m \leftrightarrow m'$ équivaut à $n_2 \leftrightarrow n_3$, donc à $b \leftrightarrow c$) et $(\mathcal{D}(A))^* = \mathcal{D}(A^*)$ (car les coefficients numériques dans (0.71) sont réels) donc $\mathcal{D}^\dagger(A) = \mathcal{D}(A^\dagger)$.

Cette représentation est appelée $(j, 0)$, elle est de dimension $2j + 1$. Il en existe une autre de dimension $2j + 1$, non équivalente, notée $(0, j)$, c'est la représentation "contragrédiente conjuguée" (au sens du chap 2. § 2.1.4) $\mathcal{D}^j(A^{\dagger -1})$. Le remplacement de A par $A^{\dagger -1}$ s'interprète dans la construction du § 0.6.3 si au lieu d'associer $X = x^\mu \sigma_\mu = x^0 \sigma_0 + \mathbf{x} \cdot \boldsymbol{\sigma}$ à x , on lui associe $\tilde{X} = x^0 \sigma_0 - \mathbf{x} \cdot \boldsymbol{\sigma}$. On note que $\sigma_2(\sigma_i)^T \sigma_2 = -\sigma_i$ pour $i = 1, 2, 3$ donc $\tilde{X} = \sigma_2 X^T \sigma_2$. Pour la transformation $A : X \mapsto X' = AXA^\dagger$, on a

$$\tilde{X}' = \sigma_2 (X')^T \sigma_2 = \sigma_2 (AXA^\dagger)^T \sigma_2 = (\sigma_2 A^T \sigma_2)^\dagger \tilde{X} (\sigma_2 A^T \sigma_2).$$

Toute matrice A de $SL(2, \mathbb{C})$ peut elle-même s'écrire $A = a^\mu \sigma_\mu$, avec $(a^\mu) \in \mathbb{C}^4$, et comme $\det A = (a^0)^2 - \mathbf{a}^2 = 1$ (le "S" de $SL(2, \mathbb{C})$), on vérifie aisément que $A^{-1} = a^0 \sigma_0 - \mathbf{a} \cdot \boldsymbol{\sigma}$, donc

$$\sigma_2 A^T \sigma_2 = A^{-1}. \quad (0.103)$$

Finalement

$$X' = AXA^\dagger \iff \tilde{X}' = (A^{-1})^\dagger \tilde{X} A^{-1}. \quad (0.104)$$

Remarque. Les deux représentations $(j, 0)$ et $(0, j)$ sont inéquivalentes sur $SL(2, \mathbb{C})$, mais équivalentes sur $SU(2)$. En effet dans $SU(2)$, $A = U = (U^\dagger)^{-1}$.

Finalement, on démontre que toute représentation de dimension finie de $SL(2, \mathbb{C})$ est complètement réductible et peut donc s'écrire comme somme directe de représentations irréductibles. La représentation irréductible de dimension finie la plus générale de $SL(2, \mathbb{C})$ est notée (j_1, j_2) , avec j_1 et j_2 entiers ou demi-entiers ≥ 0 ; elle est définie par

$$(j_1, j_2) = (j_1, 0) \otimes (0, j_2).$$

Toutes ces représentations peuvent être obtenues à partir des représentations $(\frac{1}{2}, 0)$ et $(0, \frac{1}{2})$. En effet $(j_1, 0)$ et $(0, j_2)$ se construisent par produit tensoriel symétrisé des représentations $(\frac{1}{2}, 0)$ et $(0, \frac{1}{2})$, comme on l'a fait pour $SU(2)$. Seules les représentations (j_1, j_2) ayant j_1 et j_2 simultanément entiers ou demi-entiers fournissent de vraies représentations de \mathcal{L}_+^\dagger . Les autres sont des représentations à un signe près.

Exercice : montrer que la représentation $(0, j)$ est "équivalente" (à un changement de base près) à la complexe conjuguée de la représentation $(j, 0)$. (On pourra le montrer d'abord pour $j = \frac{1}{2}$ en se rappelant que $(A^{-1})^\dagger = \sigma_2 A^* \sigma_2$, puis pour les représentations de j quelconque obtenues par produit tensoriel d'ordre $2j$ à partir de $j = \frac{1}{2}$.)

Représentations spinorielles

Revenons aux deux représentations spinorielles $(\frac{1}{2}, 0)$ et $(0, \frac{1}{2})$. Ce sont des représentations de dimension 2 (spineurs à deux composantes). Il est traditionnel de noter les indices des composantes avec des indices “pointés” ou non pointés, pour la représentation $(0, \frac{1}{2})$ et la $(\frac{1}{2}, 0)$, respectivement. Pour $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{C})$

$$\begin{aligned} (\frac{1}{2}, 0) \quad \xi = (\xi^\alpha) &\mapsto \xi' = A\xi = \begin{pmatrix} a\xi^1 + b\xi^2 \\ c\xi^1 + d\xi^2 \end{pmatrix} \\ (0, \frac{1}{2}) \quad \xi = (\xi^{\dot{\alpha}}) &\mapsto \xi' = A^*\xi = \begin{pmatrix} a^*\xi^{\dot{1}} + b^*\xi^{\dot{2}} \\ c^*\xi^{\dot{1}} + d^*\xi^{\dot{2}} \end{pmatrix} \end{aligned} \quad (0.105)$$

On note que la forme alternée $(\xi, \eta) = \xi^1\eta^2 - \xi^2\eta^1 = \xi^T(i\sigma_2)\eta$ est invariante dans $(\frac{1}{2}, 0)$ (et aussi dans $(0, \frac{1}{2})$), ce qui découle à nouveau de (0.103)

$$(\sigma_2 A^T \sigma_2)A = A^{-1}A = \mathbf{I} \iff A^T(i\sigma_2)A = i\sigma_2 .$$

On peut donc utiliser cette forme pour abaisser les indices α (ou $\dot{\alpha}$). Ainsi

$$\begin{aligned} \text{dans } (\frac{1}{2}, 0) : (\xi, \eta) &= \xi_\alpha \eta^\alpha & \xi_2 &= \xi^1 & \xi_1 &= -\xi^2 \\ \text{dans } (0, \frac{1}{2}) : (\xi, \eta) &= \xi_{\dot{\alpha}} \eta^{\dot{\alpha}} & \xi_{\dot{2}} &= \xi^{\dot{1}} & \xi_{\dot{1}} &= -\xi^{\dot{2}} \end{aligned} \quad (0.106)$$

Représentation (j_1, j_2)

Les $\{\xi^{\alpha_1\alpha_2\cdots\alpha_{2j_1}}\beta_1\beta_2\cdots\beta_{2j_2}\}$ symétriques en $\alpha_1, \alpha_2, \dots, \alpha_{2j_1}$ et en $\beta_1, \beta_2, \dots, \beta_{2j_2}$, forment la représentation irréductible (j_1, j_2) . (On ne peut pas diminuer le rang en prenant des traces, le seul tenseur invariant étant la forme précédente *alternée*). La dimension de cette représentation est $(2j_1 + 1)(2j_2 + 1)$. Les représentations les plus usuelles rencontrées en théorie des champs sont $(0, 0)$, $(\frac{1}{2}, 0)$ et $(0, \frac{1}{2})$, $(\frac{1}{2}, \frac{1}{2})$. Cette dernière correspond aux 4-vecteurs, comme on l’a vu plus haut :

$$x \mapsto X = x^0\sigma_0 + \mathbf{x}\cdot\boldsymbol{\sigma} \xrightarrow{A \in SL(2,\mathbb{C})} X' = AXA^\dagger$$

c’est-à-dire

$$X = X^{\alpha\dot{\beta}} \rightarrow (X')^{\alpha\dot{\beta}} = A^{\alpha\alpha'} (A^{\dot{\beta}\dot{\beta}'})^* X^{\alpha'\dot{\beta}'},$$

ce qui montre que X se transforme bien selon la représentation $(\frac{1}{2}, \frac{1}{2})$.

Exercice. Montrer que les représentations $(1, 0)$ et $(0, 1)$, de dimension 3, décrivent des tenseurs $F^{\mu\nu}$ de rang 2 self-duaux ou anti-self-duaux, c’est-à-dire satisfaisant

$$F^{\mu\nu} = \pm \frac{i}{2} \epsilon^{\mu\nu\rho\sigma} F_{\rho\sigma},$$

où $\epsilon^{\mu\nu\rho\sigma}$ est le tenseur complètement antisymétrique à 4 indices, avec la convention que $\epsilon^{0123} = 1$, mais attention $\epsilon^{\mu\nu\rho\sigma} = -\epsilon_{\mu\nu\rho\sigma}$!

0.6.5 Représentations irréductibles unitaires du groupe de Poincaré. États à une particule.

Selon le théorème de Wigner qui sera discuté au chapitre 2, pour décrire l'action des transformations propres orthochrones de Lorentz ou de Poincaré sur les états d'une théorie quantique, nous avons besoin de représentations *unitaires* de ces groupes, ou plutôt de leurs "recouvrements universels" $SL(2, \mathbb{C})$ et $ISL(2, \mathbb{C})$. Comme on le verra plus bas (Chap. 2), les représentations unitaires (de classe L^2) du groupe non compact $SL(2, \mathbb{C})$ sont nécessairement de dimension infinie (à l'exception de la représentation triviale $(0, 0)$, qui décrit un état invariant par rotation et sous l'effet des boosts, c'est-à-dire le vide!).

Revenons aux relations de commutation de l'algèbre de Lie (0.97). On cherche un ensemble maximal d'opérateurs commutants. Les quatre P_μ commutent. Soit (p_μ) une valeur propre pour un vecteur propre des P_μ , état "à une particule", et on suppose que le vecteur propre noté $|p\rangle$ n'est indexé que par p^μ et par des indices discrets : c'est en effet le sens de "état à une particule", au contraire d'un état à deux particules qui dépendrait d'une impulsion relative, variable continue

$$P_\mu |p\rangle = p_\mu |p\rangle . \quad (0.107)$$

On considère aussi le tenseur de Pauli-Lubanski

$$W^\lambda = \frac{1}{2} \epsilon^{\lambda\mu\nu\rho} J_{\mu\nu} P_\rho \quad (0.108)$$

et on vérifie (exercice!) que (0.97) implique

$$\begin{aligned} [W_\mu, P_\nu] &= 0 \\ [W^\mu, W^\nu] &= -i \epsilon^{\mu\nu\rho\sigma} W_\rho P_\sigma \\ [J_{\mu\nu}, W_\lambda] &= i(g_{\nu\lambda} W_\mu - g_{\mu\lambda} W_\nu) . \end{aligned} \quad (0.109)$$

La dernière relation signifie que W est un 4-vecteur de Lorentz. On note aussi que $W.P = 0$ en raison de l'antisymétrie du tenseur ϵ . On montre enfin (le vérifier!) que $P^2 = P_\mu P^\mu$ et $W^2 = W_\mu W^\mu$ commutent avec tous les générateurs P et J : ce sont les opérateurs de Casimir de l'algèbre. Selon le lemme de Schur, (cf plus bas, chap. 2, § 2.1.6) ils sont dans toute représentation irréductible proportionnels à l'identité, autrement dit, leurs valeurs propres peuvent être utilisées pour indexer les représentations irréductibles.

En physique, on n'a en principe que deux types de représentations à considérer⁴ : les représentations où $P^2 > 0$ et celles où $P^2 = 0$, $W^2 = 0$. Leur construction détaillée sera effectuée dans le cours d'A. Bilal.

Bibliographie

La référence historique pour le physicien est le livre d'E. Wigner [Wi].

Pour une discussion détaillée des propriétés du groupe de Lorentz, voir le livre récent d'Éric Gourgoulhon, *Relativité restreinte. Des particules l'astrophysique*, (EDP Sciences / CNRS Éditions).

Pour une discussion détaillée du groupe des rotations, ainsi que de nombreuses formules et tables, se reporter à : J.-M. Normand, *A Lie group : Rotations in Quantum Mechanics*, North-Holland.

Pour une étude approfondie des représentations physiques des groupes de Lorentz et Poincaré, voir P. Moussa et R. Stora, *Angular analysis of elementary particle reactions*, dans *Analysis of scattering and decay*, édité par M. Nikolic, Gordon et Breach 1968.

4. ce qui ne veut pas dire qu'il n'existe pas d'autres représentations irréductibles ; par exemple les représentations "non physiques" où $P^2 = -M^2 < 0$

Problème

On considère deux représentations de spin $\frac{1}{2}$ du groupe $SU(2)$ et leur produit direct (ou tensoriel). On note $\mathbf{J}^{(1)}$ et $\mathbf{J}^{(2)}$ les générateurs infinitésimaux agissant dans chaque représentation, et $\mathbf{J} = \mathbf{J}^{(1)} + \mathbf{J}^{(2)}$ ceux agissant dans leur produit direct, cf. (0.75), (0.75').

1. Que peut-on dire des opérateurs $\mathbf{J}^{(1)2}$, $\mathbf{J}^{(2)2}$ et \mathbf{J}^2 et de leurs valeurs propres ?
2. Montrer que l'on peut exprimer $\mathbf{J}^{(1)} \cdot \mathbf{J}^{(2)}$ en termes de ces opérateurs et en déduire que les opérateurs

$$\frac{1}{4}(3\mathbf{I} + 4\mathbf{J}^{(1)} \cdot \mathbf{J}^{(2)}) \quad \text{et} \quad \frac{1}{4}(\mathbf{I} - 4\mathbf{J}^{(1)} \cdot \mathbf{J}^{(2)})$$

sont des projecteurs sur des espaces que l'on précisera.

3. En prenant en compte les symétries d'échange des vecteurs, que pouvez-vous dire de l'opérateur

$$\frac{1}{2}\mathbf{I} + 2\mathbf{J}^{(1)} \cdot \mathbf{J}^{(2)} ?$$

Appendice 0. Mesure et laplacien sur les sphères S^2 et S^3

On expose ici une méthode générale permettant de construire une mesure d'intégration à partir d'une géométrie riemannienne.

On considère une *variété riemannienne*, c'est-à-dire une variété (voir App. B du Chap. 1 pour quelques rappels) dotée d'une métrique :

$$ds^2 = g_{\alpha\beta} d\xi^\alpha d\xi^\beta \quad (0.110)$$

avec un tenseur métrique $g_{\alpha\beta}(\xi)$ dépendant *a priori* des coordonnées (locales) ξ^α , $\alpha = 1, \dots, n$; n est la dimension de la variété. Ce ds^2 doit être invariant par des changements (différentiables) des coordonnées, $\xi \rightarrow \xi'$, ce qui dicte la transformation du tenseur g

$$\xi \mapsto \xi', \quad g \mapsto g' \quad : \quad g'_{\alpha'\beta'} = \frac{\partial \xi^\alpha}{\partial \xi'^{\alpha'}} \frac{\partial \xi^\beta}{\partial \xi'^{\beta'}} g_{\alpha\beta}, \quad (0.111)$$

qui signifie que g est un tenseur de rang 2 covariant. Ce tenseur métrique est supposé non singulier, c'est-à-dire inversible, et le tenseur inverse est noté avec des indices supérieurs

$$g_{\alpha\beta} g^{\beta\gamma} = \delta_\alpha^\gamma. \quad (0.112)$$

Par ailleurs, son déterminant est traditionnellement noté g

$$g = \det(g_{\alpha\beta}). \quad (0.113)$$

Il y a alors une méthode générale pour construire un élément de volume sur la variété (c'est-à-dire une mesure d'intégration) et un laplacien, tous deux invariants par changement de coordonnées locales :

$$\begin{aligned} d\mu(\xi) &= \sqrt{g} \prod_{\alpha=1}^n d\xi^\alpha \\ \Delta &= \frac{1}{\sqrt{g}} \partial_\alpha \sqrt{g} g^{\alpha\beta} \partial_\beta \end{aligned} \quad (0.114)$$

où ∂_α est une notation abrégée pour l'opérateur différentiel $\frac{\partial}{\partial \xi^\alpha}$.

Exercice : vérifier que $d\mu(\xi)$ et Δ sont invariants par le changement de coordonnées $\xi \mapsto \xi'$.

Cela peut être appliqué dans de nombreux contextes et sera utilisé au Chap. 1 pour définir une mesure d'intégration sur les groupes de Lie compacts.

Appliquons-le ici à l'espace euclidien \mathbb{R}^n . En coordonnées sphériques, on a

$$ds^2 = dr^2 + r^2 d\Omega^2$$

où $d\Omega$ est une notation générique qui rassemble toutes les variables angulaires. Le tenseur métrique est donc de la forme générale $\begin{pmatrix} 1 & 0 \\ 0 & r^2 A \end{pmatrix}$ avec une matrice A $(n-1) \times (n-1)$ qui est indépendante de r et ne dépend que des variables angulaires. Ces dernières donnent lieu au laplacien sur la sphère unité S^{n-1} , noté $\Delta_{S^{n-1}}$; $\sqrt{g} = r^{n-1} \sqrt{\det A}$; et (0.114) nous dit que le laplacien sur \mathbb{R}^n a la forme générale

$$\Delta_{\mathbb{R}^n} = \frac{1}{r^{n-1}} \frac{\partial}{\partial r} r^{n-1} \frac{\partial}{\partial r} + \frac{1}{r^2} \Delta_{S^{n-1}} = \frac{\partial^2}{\partial r^2} + \frac{n-1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \Delta_{S^{n-1}}.$$

Écrivons les choses plus explicitement pour les sphères unités S^2 et S^3 . Considérons la sphère S^2 de rayon r fixé à 1 avec les coordonnées sphériques $0 \leq \theta \leq \pi$, $0 \leq \phi \leq 2\pi$ (Fig. 1). On a

$$\begin{aligned} ds^2 &= (d\theta^2 + \sin^2 \theta d\phi^2) \\ \sqrt{g} &= \sin \theta \\ d\mu(x) &= \sin \theta d\theta d\phi \\ \Delta_{S^2} &= \frac{1}{r^2} \left(\frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} \right). \end{aligned} \quad (0.115)$$

Les générateurs J_i s'écrivent

$$\begin{aligned} J_3 &= -i \frac{\partial}{\partial \phi} \\ J_1 &= -i \left[-\cos \phi \cotg \theta \frac{\partial}{\partial \phi} - \sin \phi \frac{\partial}{\partial \theta} \right] \\ J_2 &= -i \left[-\sin \phi \cotg \theta \frac{\partial}{\partial \phi} + \cos \phi \frac{\partial}{\partial \theta} \right] \end{aligned} \quad (0.116)$$

et on vérifie que $-\Delta_{S^2} = \vec{J}^2 = J_1^2 + J_2^2 + J_3^2$.

Pour la sphère S^3 on a des formules analogues. Dans la paramétrisation (0.12), on prendra par exemple

$$ds^2 = \frac{1}{2} \text{tr} dU dU^\dagger = \left(d\frac{\psi}{2} \right)^2 + \sin^2 \frac{\psi}{2} (d\theta^2 + \sin^2 \theta d\phi^2) \quad (0.117)$$

invariante par $U \rightarrow UV$, $U \rightarrow VU$ ou $U \rightarrow U^{-1}$ d'où la mesure invariante par ces mêmes transformations

$$d\mu(U) = \frac{1}{2} \left(\sin \frac{\psi}{2} \right)^2 \sin \theta d\psi d\theta d\phi. \quad (0.118)$$

Dans la paramétrisation des angles d'Euler,

$$U = e^{-i\alpha \frac{\sigma_3}{2}} e^{-i\beta \frac{\sigma_2}{2}} e^{-i\gamma \frac{\sigma_3}{2}} \quad (0.119)$$

on a

$$ds^2 = \frac{1}{2} \text{tr} dU dU^\dagger = \frac{1}{4} (d\alpha^2 + 2d\alpha d\gamma \cos \beta + d\gamma^2 + d\beta^2) \quad (0.120)$$

et avec $\sqrt{g} = \sin \beta$ on calcule

$$d\mu(U) = \frac{1}{8} \sin \beta d\alpha d\beta d\gamma \quad (0.121)$$

$$\Delta_{\text{sphère } S^3} = \frac{4}{\sin^2 \beta} \left[\frac{\partial^2}{\partial \alpha^2} + \frac{\partial^2}{\partial \gamma^2} + \frac{\partial^2}{\partial \alpha \partial \gamma} \right] + \frac{4}{\sin \beta} \frac{\partial}{\partial \beta} \sin \beta \frac{\partial}{\partial \sin \beta} . \quad (0.122)$$

Chapitre 1

Groupes. Groupes et algèbres de Lie

1.1 Généralités sur les groupes

1.1.1 Définitions de base et premiers exemples

On considère un groupe G , avec une opération notée selon les cas \cdot , \times ou $+$, un élément neutre e (ou 1 ou I ou 0), et un inverse g^{-1} (ou $-a$). Si l'opération est commutative, le groupe est dit *abélien*. Si le groupe est fini, c'est-à-dire a un nombre d'éléments fini, on appelle ce nombre l'*ordre* du groupe. On s'intéressera dans ce cours surtout à des groupes infinis, discrets ou continus.

Exemples (que le physicien peut rencontrer...)

1. Groupes finis

- le groupe cyclique \mathbb{Z}_p d'ordre p , considéré géométriquement comme le groupe d'invariance de rotation d'un cercle avec p points marqués équidistants, ou comme le groupe multiplicatif des racines p -ièmes de l'unité, $\{e^{2i\pi q/p}\}$, $q = 0, 1, \dots, p-1$, ou comme le groupe additif des entiers modulo p ;
- les groupes d'invariance de rotation et les groupes d'invariance de rotations *et* réflexions des solides réguliers ou des réseaux réguliers, d'une grande importance en physique des solides et en cristallographie;
- le groupe de permutation S_n de n objets, appelé aussi groupe symétrique, d'ordre $n!$; etc.

2. Groupes infinis discrets.

L'exemple le plus simple est le groupe additif \mathbb{Z} . Citons aussi les groupes de translations des réseaux réguliers, ou les *groupes d'espace* en cristallographie qui incluent toutes les isométries (rotations, translations, réflexions, leurs produits) laissant invariant un cristal...

Ou encore les groupes engendrés par les réflexions dans un nombre fini d'hyperplans de \mathbb{R}^n , qui sont finis ou infinis, selon l'arrangement de ces hyperplans, cf. les groupes de Weyl au Chapitre 4.

Un autre exemple important est le groupe modulaire $PSL(2, \mathbb{Z})$ des matrices $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ à coefficients entiers, de déterminant unité $ad - bc = 1$, où on identifie les matrices A et $-A$. Étant donné un réseau à 2 dimensions engendré dans le plan complexe par deux nombres complexes de rapport non réel ω_1 et ω_2 , ce groupe décrit les changements de base $(\omega_1, \omega_2)^T \rightarrow (\omega'_1, \omega'_2)^T = A(\omega_1, \omega_2)^T$ laissant invariant l'aire de la cellule élémentaire ($\Im(\omega_2\omega_1^*) = \Im(\omega'_2\omega_1'^*)$) et agissant sur $\tau = \omega_2/\omega_1$ selon $\tau \rightarrow (a\tau + b)/(c\tau + d)$. Ce groupe joue un rôle important en mathématiques dans l'étude des fonctions elliptiques, des formes modulaires, etc, et en physique dans l'étude des théories conformes et des théories de cordes...

Les groupes d'homotopie, que nous allons rencontrer bientôt, sont d'autres exemples de groupes discrets, finis ou infinis...

3. Groupes continus.

Nous n'aurons à faire qu'à des groupes de matrices de dimension finie, c'est-à-dire des sous-groupes des groupes linéaires $GL(n, \mathbb{R})$ ou $GL(n, \mathbb{C})$, pour un certain n . En particulier

- $U(n)$, groupe des matrices unitaires complexes, $UU^\dagger = I$, qui est le groupe d'invariance de la forme sesquilinéaire $(x, y) = \sum x^{*i}y^i$;
- $SU(n)$ son sous-groupe unimodulaire, des matrices unitaires de déterminant $\det U = 1$;
- $O(n)$ et $SO(n)$ sont les groupes orthogonaux laissant invariante la forme bilinéaire $\sum_{i=1}^n x_i y_i$. Les matrices de $SO(n)$ sont en outre de déterminant 1;
- $U(p, q)$, $SU(p, q)$, resp. $O(p, q)$, $SO(p, q)$, les groupes d'invariance d'une forme sesquilinéaire, resp. bilinéaire, de signature $((+)^p, (-)^q)$, (tel le groupe de Lorentz group). On considère le plus souvent les groupes $O(n, \mathbb{R})$, $SO(n, \mathbb{R})$ de matrices à coefficients réels mais les groupes $O(n, \mathbb{C})$, $SO(n, \mathbb{C})$ d'invariance de la même forme bilinéaire sur les complexes peuvent aussi jouer un rôle.
- $Sp(2n, \mathbb{R})$: Soit Z la matrice $2n \times 2n$ faite d'une diagonale de n blocs $i\sigma_2$: $Z = \text{diag} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, et considérons la forme bilinéaire antisymétrique

$$(X, Y) = X^T Z Y = \sum_{i=1}^n (x_{2i-1} y_{2i} - y_{2i-1} x_{2i}). \quad (1.1)$$

Le groupe symplectique $Sp(2n, \mathbb{R})$ est le groupe de matrices B réelles $2n \times 2n$ préservant cette forme $B^T Z B = Z$. La forme ci-dessus apparaît naturellement en mécanique hamiltonienne dans la 2-forme symplectique $\omega = \sum_{i=1}^n dp_i \wedge dq_i = \frac{1}{2} Z_{ij} d\xi_i \wedge d\xi_j$ avec les coordonnées $\xi = (p_1, q_1, p_2, \dots, q_n)$; ω est invariante par action de $Sp(2n, \mathbb{R})$ sur ξ . Pour $n = 1$, vérifier que $Sp(2, \mathbb{R}) = SL(2, \mathbb{R})$.

On peut aussi considérer le groupe symplectique complexe $Sp(2n, \mathbb{C})$. Un groupe relié, souvent noté $Sp(n)$ mais que je noterai $USp(n)$ pour éviter la confusion avec les précédents, est le *groupe symplectique unitaire*, groupe d'invariance d'une forme hermitienne quaternionique, $USp(n) = U(2n) \cap Sp(2n, \mathbb{C})$. Voir Appendice A.

- le groupe de “déplacements” dans \mathbb{R}^3 , – compositions de transformations de $O(3)$ et de translations –, et les groupes obtenus en lui ajoutant les dilatations, puis les inversions par rapport à un point;
- le groupe de *transformations conformes*, c'est-à-dire préservant les angles dans \mathbb{R}^n , voir le Problème à la fin de ce chapitre.
- le groupe de Galilée des transformations $\mathbf{x}' = O\mathbf{x} + \mathbf{v}t + \mathbf{x}_0$, $t' = t + t_0$, $O \in O(3)$;

- le groupe de Poincaré, dans lequel les translations sont adjointes au groupe de Lorentz $O(1,3)$,
- etc etc.

1.1.2 Classes de conjugaison d'un groupe

On définit sur un groupe G la relation d'équivalence suivante :

$$a \sim b \text{ ssi } \exists g \in G : a = g.b.g^{-1} \quad (1.2)$$

et on dit alors que les éléments a et b sont *conjugués*.

Les classes d'équivalence qui en découlent réalisent une partition de G , puisque tout élément appartient à une classe et une seule. Noter que l'élément neutre constitue à lui seul une classe. Pour un groupe fini, les différentes classes ont en général des ordres différents. Par exemple, la classe de l'identité e ne contient que le seul élément e .

On a déjà noté (Chap. 0) que dans le groupe $SO(3)$, une classe de conjugaison est caractérisée par l'angle de rotation ψ (autour d'un vecteur unitaire \mathbf{n}). Mais cette notion est aussi familière dans le cas du groupe $U(n)$, où une classe est caractérisée par un n -tuple non ordonné de valeurs propres $(e^{i\alpha_1}, \dots, e^{i\alpha_n})$. La notion de classe joue un rôle important dans la discussion des représentations des groupes et sera abondamment illustrée par la suite.

Pour le groupe symétrique S_n , que sont ces classes de conjugaison ? On peut aisément décomposer toute permutation σ de S_n en un produit de cycles (permutations cycliques) portant sur des éléments disjoints. (Pour s'en convaincre, on construit le cycle $(1, \sigma(1), \sigma^2(1), \dots)$, puis une fois revenu en 1, on construit un autre cycle à partir d'un nombre pas encore atteint, etc.). Au final, si σ est fait de p_1 cycles de longueur 1, de p_2 de longueur 2, etc, avec $\sum ip_i = n$, on écrit $\sigma \in [1^{p_1}2^{p_2}\dots]$, et on démontre que cette décomposition en cycles caractérise les classes de conjugaison : deux permutations sont conjuguées ssi elles ont la même décomposition en cycles.

1.1.3 Sous-groupes

La notion de sous-groupe, sous-ensemble d'un groupe lui-même doté de la structure de groupe, est familière. Le sous-groupe est *propre* s'il n'est pas identique à G . Si H est un sous-groupe, pour tout $a \in G$, l'ensemble $a^{-1}.H.a$ des éléments de la forme $a^{-1}.h.a$, $h \in H$ forme aussi un sous-groupe, dit *sous-groupe conjugué* de H .

Des exemples de sous-groupes particuliers sont donnés par :

- le centre Z :

Soit G un groupe. On appelle *centre* de G l'ensemble Z des éléments qui commutent avec tous les éléments de G :

$$Z = \{a \mid \forall g \in G, a.g = g.a\} \quad (1.3)$$

Z est un sous-groupe de G , propre si G est non-abélien. Exemples : le centre du groupe $GL(2, \mathbb{R})$ des matrices régulières 2×2 est l'ensemble des matrices multiples de I ; le centre de $SU(2)$ est le groupe \mathbb{Z}_2 des matrices $\pm I$ (le vérifier par le calcul direct).

— le centralisateur d'un élément a :

Le *centralisateur* (ou *commutant*) d'un élément a fixé de G est l'ensemble des éléments de G qui commutent avec a .

$$Z_a = \{g \in G \mid a.g = g.a\} \quad (1.4)$$

Le commutant Z_a n'est jamais vide : il contient au moins le sous-groupe engendré par a . Le centre Z est l'intersection de tous les commutants. Exemple : dans le groupe $GL(2, \mathbb{R})$, le commutant de la matrice de Pauli $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ est le groupe abélien des matrices de la forme $a\mathbf{I} + b\sigma_1$, $a^2 - b^2 \neq 0$.

— Plus généralement, étant donnée une partie S d'un groupe G , on définit son *centralisateur* $Z(S)$ et son *normalisateur* $N(S)$ comme les sous-groupes commutant respectivement individuellement avec tout élément de S ou globalement avec S tout entier

$$Z(S) = \{y : \forall s \in S \quad y.s = s.y\} \quad (1.5)$$

$$N(S) = \{x : x^{-1}.S.x = S\} . \quad (1.6)$$

1.1.4 Homomorphisme d'un groupe G dans un groupe G'

Un *homomorphisme* d'un groupe G dans un groupe G' est une application ρ de G dans G' qui respecte la loi de composition :

$$\forall g, h \in G, \quad \rho(g.h) = \rho(g).\rho(h) \quad (1.7)$$

En particulier, à l'élément neutre de G correspond par ρ celui de G' , à l'inverse de g correspond celui de $g' = \rho(g) : \rho(g^{-1}) = (\rho(g))^{-1}$.

Un exemple d'homomorphisme que nous allons particulièrement étudier est celui d'une représentation linéaire de groupe, dont la définition a été donnée au chapitre 0 et sur laquelle on va revenir au chap. 2.

Le *noyau* de l'homomorphisme noté $\ker \rho$ ("Kern" en allemand, "kernel" en anglais) est l'ensemble des *antécédents* (ou *préimages*) de l'élément neutre e' de G' : $\ker \rho = \{x \in G : \rho(x) = e'\}$. C'est un sous-groupe de G .

Par exemple, la parité (ou signature) d'une permutation de S_n définit un homomorphisme de S_n dans \mathbb{Z}_2 . Son noyau est constitué des permutations paires : c'est le *groupe alterné* A_n d'ordre $n!/2$.

1.1.5 Classes par rapport à un sous-groupe

Soit H un sous-groupe d'un groupe G . On définit la relation entre éléments de G :

$$g \sim g' \iff g.g'^{-1} \in H, \quad (1.8)$$

ce qu'on peut encore écrire comme

$$g \sim g' \iff \exists h \in H : g = h.g' \quad \text{ou encore} \quad g \in H.g' . \quad (1.9)$$

C'est une relation d'équivalence (le vérifier), dite *équivalence à droite*. On peut définir de la même façon une *équivalence à gauche* par

$$g \sim_L g' \iff g^{-1}.g' \in H \iff g \in g'.H. \quad (1.10)$$

La relation (disons à droite) définit des classes d'équivalence qui donnent une partition de G ; si g_j est un représentant de la classe j , on peut noter cette dernière $H.g_j$. (Les anglophones utilisent le terme "right-coset" pour cette classe). Les éléments de H forment à eux-seuls une classe. On note G/H l'ensemble quotient, c'est-à-dire l'ensemble des classes d'équivalence. Son cardinal (le nombre de classes) est appelé l'*indice* de H dans G et noté $|G : H|$. Par exemple, le groupe (additif) $H = 2\mathbb{Z}$ des entiers pairs est d'indice fini égal à 2 dans $G = \mathbb{Z}$. En revanche, \mathbb{Z} est d'indice infini dans \mathbb{R} .

Si H est d'ordre fini $|H|$, toutes les classes ont $|H|$ éléments, et si G est lui-même d'ordre fini $|G|$, il est partitionné en $|G : H| = |G|/|H|$ classes, et on obtient comme corollaire le théorème de Lagrange : l'ordre $|H|$ de tout sous-groupe H divise celui de G , et l'indice $|G : H| = |G|/|H|$ est l'ordre de l'ensemble quotient G/H .

L'équivalence à gauche donne en général une partition différente, mais de même indice. Par exemple, le groupe S_3 possède un sous-groupe \mathbb{Z}_2 engendré par la permutation des deux éléments 1 et 2. Exercice : vérifier que les classes à gauche et à droite ne coïncident pas.

1.1.6 Sous-groupe invariant

Soit G un groupe, H un sous-groupe de G . H est un *sous-groupe invariant* (on dit aussi *normal*) si l'une des propriétés équivalentes suivantes est vraie

- $\forall g \in G, \forall h \in H, ghg^{-1} \in H$;
- les classes à gauche et à droite coïncident ;
- H est égal à tous ses conjugués, $\forall g \in G, gHg^{-1} = H$.

Exercice : vérifier l'équivalence entre ces trois définitions.

La propriété importante à retenir est la suivante :

- Si H est un sous-groupe invariant de G , on peut munir l'ensemble quotient G/H de la structure de groupe.

Noter qu'en général on ne peut pas considérer le *groupe quotient* G/H comme un sous-groupe de G .

Esquissons la démonstration. Si $g_1 \sim g'_1$ et $g_2 \sim g'_2$, $\exists h_1, h_2 \in H : g_1 = h_1.g'_1, g_2 = g'_2.h_2$, donc $g_1.g_2 = h_1.(g'_1.g'_2).h_2$ c'est-à-dire $g_1.g_2 \sim g'_1.g'_2$ et $g_1^{-1} = g'^{-1}_1.h_1^{-1} \sim g'^{-1}_1$. La relation d'équivalence est compatible avec les opérations de produit et de passage à l'inverse. Si $[g_1]$ et $[g_2]$ sont deux classes, on définit leur produit $[g_1].[g_2] = [g_1.g_2]$ où au membre de droite on peut prendre tout représentant g_1 de $[g_1]$ et g_2 de $[g_2]$; de même pour l'inverse. Donc la structure de groupe passe donc au quotient, ensemble des classes. La classe constituée de H est l'élément neutre du quotient.

Exemple de sous-groupe invariant : Le *noyau* d'un homomorphisme ρ d'un groupe G dans un groupe G' est un sous-groupe invariant ; montrer que son groupe quotient est un groupe isomorphe à l'image $\rho(G) \subset G'$ de G par ρ .

1.1.7 Groupe simple, groupe semi-simple

Un groupe est *simple* s'il n'a pas de sous-groupe invariant non trivial (c'est-à-dire différent de $\{e\}$ et de G tout entier). Un groupe est *semi-simple* s'il n'a pas de sous-groupe invariant abélien non trivial. Tout groupe simple est évidemment semi-simple.

Cette notion est importante dans l'étude des représentations et la classification des groupes.

Exemples : Le groupe des rotations à deux dimensions n'est pas simple, ni même semi-simple (pourquoi?). Le groupe $SO(3)$ est simple (preuve non triviale, voir plus bas, §1.2.2). Le groupe $SU(2)$ n'est ni simple, ni semi-simple, il contient en effet le sous-groupe invariant $\mathbb{Z}_2 = \{I, -I\}$. Le groupe S_n n'est pas simple, pour $n > 2$ (pourquoi?).

Produits direct, semi-direct

Considérons deux groupes G_1 et G_2 et leur *produit direct* $G = G_1 \times G_2$: c'est l'ensemble des paires (g_1, g_2) doté du produit naturel $(g'_1, g'_2) \cdot (g_1, g_2) = (g'_1 \cdot g_1, g'_2 \cdot g_2)$. À l'évidence ses sous-groupes $\{(g_1, e)\} \simeq G_1$ et $\{(e, g_2)\} \simeq G_2$ sont des sous-groupes invariants, et G n'est pas simple.

Une construction plus subtile fait appel au groupe d'automorphismes de G_1 noté $\text{Aut}(G_1)$: c'est le groupe des bijections β de G_1 dans lui-même qui respecte le produit de G_1 (homomorphisme de groupe) : $\beta(g'_1 \cdot g_1) = \beta(g'_1) \beta(g_1)$. Supposons donné un homomorphisme de groupe φ d'un autre groupe G_2 dans $\text{Aut}(G_1)$: $\forall g_2 \in G_2, \varphi(g_2) \in \text{Aut}(G_1)$. On définit alors sur les paires (g_1, g_2) le produit suivant

$$(g'_1, g'_2) \cdot (g_1, g_2) = (g'_1 \cdot \varphi(g'_2) g_1, g'_2 \cdot g_2).$$

Exercice : montrer que ceci une structure de groupe sur l'ensemble de ces paires : c'est le produit semi-direct de f G_1 et G_2 (pour un φ donné) et il est noté $G_1 \rtimes_{\varphi} G_2$. Vérifier que le sous-groupe $\{(g_1, e)\} \simeq G_1$ est un sous-groupe invariant de G .

Exemples : le groupe des déplacements préservant l'orientation, engendré par les translations et les rotations de l'espace euclidien \mathbb{R}^n , est le produit semi-direct $\mathbb{R}^n \rtimes SO(n)$, avec $(\vec{a}', R')(\vec{a}, R) = (\vec{a}' + R'\vec{a}, R'R)$. De même le groupe de Poincaré dans l'espace de Minkowski est le produit semi-direct $\mathbb{R}^4 \rtimes \mathcal{L}$.

1.2 Groupes continus. Propriétés topologiques. Groupes de Lie

Un groupe continu (ou encore *groupe topologique*) est un espace topologique (donc doté d'une base de voisinages permettant de définir les notions de continuité etc¹) muni d'une structure de groupe, telle que les opérations de groupe $(g, h) \mapsto g \cdot h$ et $g \mapsto g^{-1}$ soient des fonctions continues. Autrement dit, si g' est proche (au sens de la topologie de G) de g et h' de h , alors $g' \cdot h'$ est proche de $g \cdot h$ et g'^{-1} est proche de g^{-1} .

Les groupes de matrices présentés plus haut entrent bien dans cette classe de groupes topologiques, mais aussi des groupes "de dimension infinie" comme le groupe des difféomorphismes invoqué en Relativité Générale, ou les groupes de jauge des théories de jauge.

Commençons par étudier quelques propriétés topologiques de tels groupes continus.

1. Voir l'Appendice B pour quelques rappels de vocabulaire. . .

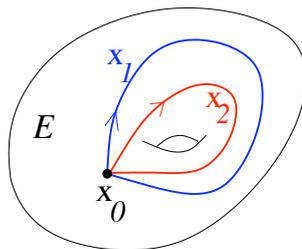


FIGURE 1.1 – Les lacets x_1 et x_2 sont homotopes. Mais aucun d’eux n’est homotope au lacet “trivial” qui reste en x_0 . L’espace n’est pas simplement connexe.

1.2.1 Connexité

Un groupe peut être ou non connexe. Si G n’est pas connexe, la composante connexe de l’identité est un sous-groupe invariant.

On peut s’intéresser à la propriété de connexité au sens topologique général (un espace E est connexe si ses seuls sous-espaces à la fois ouverts et fermés sont E et \emptyset), mais c’est surtout la connexité par arcs qui nous concernera (pour toute paire de points, il existe un chemin continu les joignant). Démontrer que la composante connexe de l’identité est un sous-groupe invariant dans l’une et l’autre définition. Réf. [K-S, Po].

Exemples. $O(3)$ est disconnexe et la composante connexe de l’identité est $SO(3)$; pour le groupe de Lorentz $\mathcal{L}=O(1,3)$ on a défini sa composante propre orthochrone \mathcal{L}_+^\uparrow , les autres “nappes” s’en déduisant via la parité P , le renversement du temps T et leur produit PT ...

1.2.2 Simple connexité. Groupe d’homotopie. Recouvrement universel

La notion de *simple connexité* ne doit pas être confondue avec la précédente. Discutons-la d’abord dans le cadre d’un espace topologique arbitraire E avant de nous spécialiser au cas d’un groupe.

On considère les chemins fermés à extrémité fixée x_0 ou *lacets* tracés dans l’espace E , c’est-à-dire les applications continues $x(t)$ de $[0, 1]$ dans E telles que $x(0) = x(1) = x_0$. Étant donnés deux tels chemins $x_1(\cdot)$ et $x_2(\cdot)$ de x_0 à x_0 , peut-on les déformer continûment l’un en l’autre? Autrement dit, existe-t-il une fonction continue $f(t, \xi)$ de deux variables $t, \xi \in [0, 1]$, à valeurs dans E , telle que

$$\begin{aligned} \forall \xi \in [0, 1] \quad f(0, \xi) &= f(1, \xi) = x_0 & : \text{trajectoires fermées} & \quad (1.11) \\ \forall t \in [0, 1] \quad f(t, 0) &= x_1(t) \quad f(t, 1) = x_2(t) & : \text{interpolation} & \end{aligned}$$

Si c’est le cas, on dit que les lacets x_1 et x_2 sont *homotopes* (c’est une relation d’équivalence), ou encore qu’ils appartiennent à la même *classe d’homotopie*, voir Fig. 1.1.

On peut aussi *composer* les chemins : Si $x_1(\cdot)$ et $x_2(\cdot)$ sont deux lacets de x_0 à x_0 , le chemin $x_2 \circ x_1$ va aussi de x_0 à x_0 en parcourant d’abord x_1 puis x_2 . Le lacet inverse de $x_1(\cdot)$ pour cette composition est le lacet parcouru en sens inverse : $x_1^{-1}(t) := x_1(1 - t)$. La composition et le passage à l’inverse sont compatibles avec l’homotopie : si $x_1 \sim x_1'$ et $x_2 \sim x_2'$,

alors $x_2 \circ x_1 \sim x'_2 \circ x'_1$ et $x_1^{-1} \sim x'_1{}^{-1}$. Ces opérations passent donc aux classes, ce qui munit l'ensemble des classes d'homotopie d'une structure de groupe pour cette composition, c'est la *groupe d'homotopie* $\pi_1(E, x_0)$. Ainsi, un représentant de la classe identité est fourni par le lacet "trivial", $x(t) = x_0, \forall t$. On montre enfin que les groupes relatifs à des extrémités x_0 différentes sont isomorphes (dans un espace connexe); par exemple dans le cas d'un groupe connexe, on peut se ramener au choix du point de base à l'identité $x_0 = e$. On parle donc **du** groupe d'homotopie (ou groupe fondamental) $\pi_1(E)$. Pour plus de détails, voir par exemple [Po], [DNF].

Si tous les lacets de x_0 à x_0 peuvent être contractés en le lacet trivial $\{x_0\}$, on dit que E est *simplement connexe*. Dans le cas contraire, on démontre, et nous admettrons, que l'on peut construire un espace \tilde{E} , dit *espace de recouvrement universel* de E , tel que \tilde{E} est simplement connexe et que **localement**, E et \tilde{E} sont *homéomorphes*. Cela signifie qu'il existe une application continue surjective p de \tilde{E} dans G tel que tout point x de \tilde{E} ait un voisinage V_x et que $V_x \mapsto p(V_x)$ soit un homéomorphisme, c'est-à-dire une application bijective et bicontinue². L'espace \tilde{E} de recouvrement universel de x est unique (à un homéomorphisme près).

Dans le cas qui nous occupe où $E = G$ est un groupe topologique, on montre que \tilde{G} est lui-même un groupe et que de plus, l'application p est un homomorphisme de \tilde{G} dans G ([Po], § 51). Son noyau qui est un sous-groupe invariant n'est autre que le groupe d'homotopie $\pi_1(G)$. Le groupe quotient est isomorphe à G

$$\tilde{G}/\pi_1(G) \simeq G, \quad (1.12)$$

(selon une propriété générale du groupe quotient par le noyau d'un homomorphisme, cf. § 1.1.6).

On peut construire le groupe de recouvrement universel \tilde{G} en considérant les chemins qui joignent l'identité e à un point g et leurs classes d'équivalence par déformation continue **à extrémités fixes**. \tilde{G} est l'ensemble de ces classes d'équivalence. C'est un groupe pour la multiplication des chemins définie comme suit : si deux chemins $g_1(t)$ et $g_2(t)$ joignent e à g_1 et à g_2 respectivement, le chemin $g_1(t).g_2(t)$ joint e à $g_1.g_2$. Cette loi de composition est compatible avec l'équivalence et munit \tilde{G} d'une structure de groupe et on montre que \tilde{G} est simplement connexe (cf. [Po] § 51). La projection p de \tilde{G} dans G associe à toute classe de chemins leur extrémité commune. On vérifie que c'est bien un homéomorphisme local et un homomorphisme de groupes, et que son noyau est le groupe d'homotopie $\pi_1(G)$.

Exemple : Le groupe des phases $G = U(1)$, vu comme le cercle unité S^1 , n'est pas simplement connexe : un chemin de l'identité 1 à 1 peut faire un nombre arbitraire de fois le tour du cercle et ce nombre de tours (positif ou négatif) distingue les différentes classes d'homotopie : le groupe d'homotopie est $\pi_1(U(1)) = \mathbb{Z}$. Le groupe \tilde{G} n'est autre que le groupe additif \mathbb{R} qu'on peut visualiser comme une hélice au-dessus du cercle $U(1)$. Le quotient est $\mathbb{R}/\mathbb{Z} \simeq U(1)$, ce qu'il faut rapprocher du fait qu'un point de $U(1)$, c'est-à-dire un angle, est un nombre réel modulo un multiple entier de 2π . On peut dire encore $\pi_1(S^1) = \mathbb{Z}$. Plus généralement on se convainc que pour les sphères, $\pi_1(S^n)$ est trivial (tous les lacets sont contractibles) dès que $n > 1$ ³ . .

Autre exemple fondamental : Le groupe des rotations $SO(3)$ n'est pas simplement connexe, comme cela a été pressenti au Chapitre 0. Pour nous en convaincre, visualisons la rotation $R_{\mathbf{n}}(\psi)$

2. "bicontinue" signifie que l'application et son inverse sont continues.

3. En particulier $\pi_1(S^2) = 0$ et "you cannot lasso a basketball" selon la formule imagée de S. Coleman!

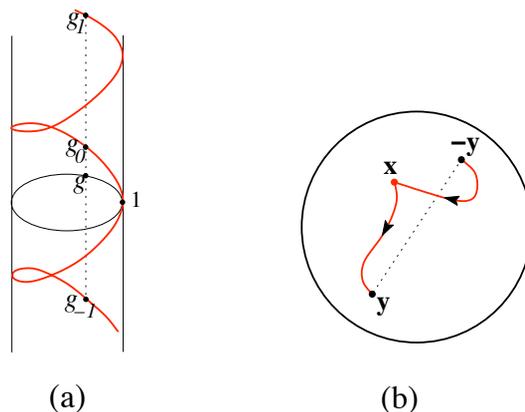


FIGURE 1.2 – (a) Le groupe $U(1)$, identifié au cercle et son groupe de recouvrement universel \mathbb{R} , identifié à l’hélice. Un élément $g \in U(1)$ se relève en des points $\dots, g_{-1}, g_0, g_1, \dots$ sur l’hélice. (b) Dans la boule B^3 représentant $SO(3)$, les points \mathbf{y} et $-\mathbf{y}$ de la surface sont identifiés. Un chemin allant de \mathbf{x} à \mathbf{x} via \mathbf{y} et $-\mathbf{y}$ est donc fermé et non contractible : $SO(3)$ est non simplement connexe.

par le point $\mathbf{x} = \tan \frac{\psi}{4} \mathbf{n}$ d’un espace \mathbb{R}^3 auxiliaire ; ces points sont tous dans la boule B^3 de rayon 1, avec la rotation identité au centre et les rotations d’angle π sur la surface de la sphère, mais en raison de $R_{\mathbf{n}}(\pi) = R_{-\mathbf{n}}(\pi)$, il faut identifier les points de la sphère diamétralement opposés. Il s’ensuit qu’il existe dans $SO(3)$ des courbes fermées non contractibles : une courbe de \mathbf{x} à \mathbf{x} passant par deux points diamétralement opposés sur la sphère S^2 doit être considérée comme fermée mais n’est pas contractible (Fig. 2). Il existe deux classes de chemins fermés non homotopes et le groupe $SO(3)$ est “doublement connexe” : son groupe d’homotopie est $\pi_1(SO(3)) = \mathbb{Z}_2$. En fait, nous connaissons déjà le groupe de recouvrement universel de $SO(3)$: c’est le groupe $SU(2)$, dont on a montré qu’il était homéomorphe à la sphère S^3 , donc simplement connexe, et qu’il existait un homomorphisme l’envoyant dans $SO(3)$, selon $\pm U_{\mathbf{n}}(\psi) = \pm(\cos \frac{\psi}{2} - i \sin \frac{\psi}{2} \sigma \cdot \mathbf{n}) \mapsto R_{\mathbf{n}}(\psi)$, cf. Chapitre 0, § 1.2.

Cette propriété de $SO(3)$ d’être non simplement connexe peut être illustrée par différentes expériences de salon, dont l’interprétation précise n’est pas toujours évidente, telles “la ceinture de Dirac” et “l’assiette de Feynman”, voir <http://gregegan.customer.netspace.net.au/APPLETS/21/21.html> et <http://www.math.utah.edu/~palais/links.html> pour des animations, et V. Stojanoska et O. Stoytchev, *Mathematical Magazine*, **81**, 2008, 345-357, pour une discussion détaillée impliquant le groupe des tresses.

Cette même visualisation des rotations par l’intérieur de la boule unité permet de comprendre l’assertion faite plus haut que le groupe $SO(3)$ est simple. Supposons qu’il ne le soit pas, et soit $R = R_{\mathbf{n}}(\psi)$ un élément d’un sous-groupe invariant de $SO(3)$, qui contient aussi tous les conjugués de R (par définition d’un sous-groupe invariant). Ces conjugués sont visualisés par les points de la sphère de rayon $\tan \psi/4$. Le sous-groupe invariant contenant $R_{\mathbf{n}}(\psi)$ et des points arbitrairement proches de son inverse $R_{-\mathbf{n}}(\psi)$ contient des points arbitrairement proches de l’identité, qui par conjugaison, remplissent une petite boule au voisinage de l’identité. Il reste à montrer que le produit de tels éléments permet de remplir toute la boule, c’est-à-dire que le sous-groupe invariant ne peut être que le groupe $SO(3)$ tout entier ; ceci est en fait vrai pour tout groupe de Lie connexe, comme on le verra plus bas.

Autres exemples : les groupes classiques. On démontre que

- les groupes $SU(n)$ sont tous simplement connexes, pour tout n , tandis que $\pi_1(U(n)) = \mathbb{Z}$;
- pour le groupe $SO(2) \cong U(1)$, on a vu que $\pi_1(SO(2)) = \mathbb{Z}$;

• pour tout $n > 2$, $SO(n)$ est doublement connexe, $\pi_1(SO(n)) = \mathbb{Z}_2$, et on appelle $Spin(n)$ son groupe de recouvrement universel. Donc $Spin(3) = SU(2)$.

La notion d'homotopie, c'est-à-dire de déformation continue, qu'on vient d'appliquer à des lacets, c'est-à-dire à des applications de S^1 dans une variété \mathcal{V} (un groupe G ici), peut s'étendre à des applications d'une sphère S^n dans \mathcal{V} . Même si la composition de telles applications est moins aisée à visualiser, elle peut être définie et est à nouveau compatible avec l'homotopie, ce qui conduit à la définition du groupe d'homotopie $\pi_n(\mathcal{V})$. Par exemple $\pi_n(S^n) = \mathbb{Z}$. Voir [DNF] pour plus de détails et des calculs de ces groupes π_n . Cette notion est importante pour le physicien pour décrire des défauts topologiques, solitons, instantons, monopoles, etc. Voir par exemple sur la Fig. 1.3 deux configurations de *vortex* ou d'*anti-vortex* de vecteurs unitaires à deux dimensions, de nombre d'enroulement (ou *vorticité*) respectif ± 1 .

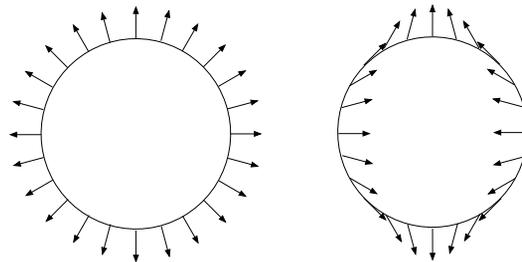


FIGURE 1.3 – Deux configurations de vecteurs unitaires réalisant des applications $S^1 \rightarrow S^1$ homotopiquement non triviales. Ce sont respectivement les *vortex* et *anti-vortex* du modèle XY de mécanique statistique, voir par exemple <http://www.ibiblio.org/e-notes/Perc/xy.htm> pour plus de détails et de belles figures.

1.2.3 Groupes compacts et non compacts

Si le domaine \mathcal{D} dans lequel vivent les paramètres du groupe G est compact, on dit que G est un *groupe compact*.

Rappelons la définition et quelques-unes des nombreuses propriétés d'un espace compact E . Un e.t. séparé E est compact si, étant donné un recouvrement de E par un ensemble d'ouverts U_i , E peut être recouvert par un nombre fini d'entre eux. Alors, toute suite infinie y admet un sous-suite convergente (généralisation du théorème de Bolzano–Weierstrass). Toute fonction réelle continue y est bornée, etc. Pour un domaine \mathcal{D} de \mathbb{R}^d , la propriété de compacité équivaut à la propriété de \mathcal{D} d'être fermé et borné.

Exemples. Les groupes de matrices unitaires $U(n)$ et leurs sous-groupes $SU(n)$, $O(n)$, $SO(n)$, $USp(n/2)$ (n pair), sont compacts. Les groupes $SL(n, \mathbb{R})$ ou $SL(n, \mathbb{C})$, $Sp(n, \mathbb{R})$ ou $Sp(n, \mathbb{C})$, le groupe de translation dans \mathbb{R}^n , le groupe de Galilée, les groupes de Lorentz et Poincaré ne le sont pas, pourquoi ?

1.2.4 Mesure invariante de Haar

Quand on traite d'un groupe *fini*, on est souvent amené à considérer des sommes sur tous les éléments du groupe et à utiliser le “lemme de réarrangement”, qui consiste à écrire

$$\forall g' \in G \quad \sum_{g \in G} f(g'g) = \sum_{h=g'g \in G} f(g'g) = \sum_{g \in G} f(g) ,$$

(invariance à gauche), la même chose avec $g'g$ changé en gg' (invariance à droite), et aussi

$$\sum_{g \in G} f(g^{-1}) = \sum_{g^{-1} \in G} f(g^{-1}) = \sum_{g \in G} f(g) .$$

On aimerait pouvoir effectuer de telles opérations dans le cas d'un groupe continu, la somme finie étant remplacée par une intégrale, finie et dotée des mêmes invariances. Cela nécessite de pouvoir disposer d'une mesure d'intégration invariante à gauche et à droite

$$d\mu(g) = d\mu(g'.g) = d\mu(g.g') = d\mu(g^{-1})$$

telle que $\int d\mu(g)f(g)$ soit finie pour toute fonction f continue.

On démontre que

- si le groupe est compact, une telle mesure existe et est unique à une normalisation près.

C'est la *mesure de Haar*.

Par exemple, pour le groupe unitaire $U(n)$, on peut construire la mesure de Haar explicitement. On peut utiliser la méthode proposée au chapitre 0, Appendice 0 : on définit d'abord une métrique sur $U(n)$ en écrivant $ds^2 = \text{tr } dU.dU^\dagger$ dans la paramétrisation de son choix ; cette métrique est bien invariante par $U \rightarrow UU'$ ou $U \rightarrow U'U$ et par $U \rightarrow U^{-1} = U^\dagger$; la mesure $d\mu(U)$ qu'on en tire a les mêmes propriétés. On trouvera dans l'Appendice C le calcul explicite de cette mesure pour $SU(2)$ et $U(n)$, et plus de détails en TD.

Inversement si le groupe n'est pas compact, les mesures invariantes à gauche et à droite peuvent exister, elles peuvent même coïncider (groupes non compacts abéliens ou semi-simples), mais leur intégrale sur le groupe diverge.

Ainsi, si G est localement compact, (c'est-à-dire tout point a une base de voisinages compacts), on démontre qu'il existe une mesure invariante à gauche, unique à une constante près. Il existe aussi une mesure invariante à droite, mais elles peuvent ne pas coïncider. Exemple

$$G = \left\{ \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \mid x, y \in \mathbb{R}, y > 0 \right\}$$

on vérifie aisément que $d\mu_L(g) = y^{-2}dxdy$, $d\mu_R(g) = y^{-1}dxdy$ sont les mesures invariantes à gauche et à droite, respectivement, et que leurs intégrales divergent. Réf. [Bu].

1.2.5 Groupes de Lie

En imposant davantage de structure à un groupe continu, nous sommes amenés à la notion de groupe de Lie.

Selon la définition la plus usuelle, un groupe de Lie est un espace topologique muni d'une loi de groupe, (un groupe topologique), qui en outre est une variété différentiable et qui est tel que les lois de composition et de passage à l'inverse $G \times G \rightarrow G$ et $G \rightarrow G$ soient des fonctions infiniment différentiables. On impose parfois que ce soit des fonctions analytiques réelles, c'est-à-dire des fonctions dont le développement de Taylor converge vers la fonction considérée. Le fait que l'une et l'autre de ces deux propriétés se trouvent dans la littérature laisse présager que la plus faible (différentiabilité) implique la plus forte (analyticité). En fait, selon un théorème très puissant de Montgomery et Zippen (1955), des hypothèses beaucoup plus faibles suffisent à assurer la propriété de groupe de Lie. Un groupe topologique connexe qui est localement homéomorphe à \mathbb{R}^d , *d fini*, est un groupe de Lie. Autrement dit, l'existence de coordonnées locales (en nombre fini) et les propriétés de groupe topologique

(la continuité des opérations de groupe) suffisent à entraîner les propriétés d'analyticité!⁴ Ceci laisse entrevoir que la structure de groupe de Lie est très puissante et très rigide. Cela dit, il existe des groupes de Lie de dimension infinie.

Pour ne pas rentrer dans une discussion mathématique inutile pour nos besoins, nous nous restreindrons à des groupes continus de matrices de taille finie. Pour un tel groupe, les éléments de matrices de $g \in G$ dépendent de façon continue de paramètres réels $(\xi^1, \xi^2, \dots, \xi^d) \in \mathcal{D} \subset \mathbb{R}^d$, et dans l'opération du groupe $g(\xi'') = g(\xi') \cdot g(\xi)$, ou dans l'inverse $g(\xi)^{-1} = g(\xi'')$, les ξ''^i sont fonctions continues et en fait analytiques des ξ^j (et des ξ'^j). Un tel groupe est appelé *groupe de Lie*, et d est appelé sa *dimension*.

Pour être plus précis, dans l'esprit de la géométrie différentielle, il faut en général introduire plusieurs domaines \mathcal{D}_j , avec des fonctions de recollement continues, et en fait analytiques, etc.

Exemples : tous les groupes de matrices présentés plus haut sont des groupes de Lie. Vérifier que la dimension de $U(n)$ est n^2 , celle de $SU(n)$ est $n^2 - 1$, celle de $O(n)$ ou $SO(n)$ est $n(n-1)/2$. Quelle est celle de $Sp(2n, \mathbb{R})$? du groupe de Galilée dans \mathbb{R}^3 ? des groupes de Lorentz et de Poincaré? Montrer que $\dim(Sp(2n, \mathbb{R})) = \dim(USp(n)) = \dim(SO(2n+1))$. Nous verrons plus bas au Chap. 3 que cela n'est pas un accident.

Dans l'étude d'un groupe de Lie et de ses représentations, on est conduit à se livrer à une double étude : d'une part une étude locale de son espace tangent au voisinage de l'identité (son algèbre de Lie), et d'autre part, une étude globale sur la topologie du groupe, information que ne révèle pas l'étude locale.

1.3 Étude locale d'un groupe de Lie. Algèbre de Lie

1.3.1 Algèbres et algèbres de Lie. Définitions

On rappelle d'abord la définition d'une algèbre.

Une *algèbre* est un espace vectoriel sur un corps (en pratique, toujours \mathbb{R} ou \mathbb{C} pour nous), doté d'un produit noté $X * Y$, (pas nécessairement associatif), bilinéaire en X et Y

$$\begin{aligned} (\lambda_1 X_1 + \lambda_2 X_2) * Y &= \lambda_1 X_1 * Y + \lambda_2 X_2 * Y \\ X * (\mu_1 Y_1 + \mu_2 Y_2) &= \mu_1 X * Y_1 + \mu_2 X * Y_2 . \end{aligned} \quad (1.13)$$

Exemples : l'ensemble des matrices $n \times n$ à coefficients réels ou complexes, $M(n, \mathbb{R})$ ou $M(n, \mathbb{C})$, est une algèbre associative pour le produit matriciel usuel ; l'ensemble des vecteurs de \mathbb{R}^3 est une algèbre (non associative!) pour le produit vectoriel (noté \wedge dans la littérature française et \times dans l'anglo-saxonne).

Une *algèbre de Lie* est une algèbre dont le produit noté $[X, Y]$ et appelé *crochet de Lie* a la

4. Pour un exemple élémentaire d'un tel phénomène, considérer une fonction f d'une variable réelle satisfaisant $f(x)f(y) = f(x+y)$. Sous la seule hypothèse que f est continue, démontrer que $f(x) = \exp kx$, donc qu'elle est analytique!

propriété supplémentaire d'être antisymétrique et de satisfaire l'identité de Jacobi

$$\begin{aligned} [X, Y] &= -[Y, X] \\ [X_1, [X_2, X_3]] + [X_2, [X_3, X_1]] + [X_3, [X_1, X_2]] &= 0 . \end{aligned} \quad (1.14)$$

Exemples : Toute algèbre associative pour un produit noté $*$, en particulier toute algèbre de matrices, est une algèbre de Lie pour le crochet de Lie défini par le commutateur

$$[X, Y] = X * Y - Y * X .$$

Les propriétés de bilinéarité et d'antisymétrie sont évidentes, et l'identité de Jacobi est vérifiée au prix d'une ligne de calcul. Autre exemple : l'algèbre des vecteurs de \mathbb{R}^3 pour le produit vectoriel mentionné précédemment est en fait une algèbre de Lie, l'identité de Jacobi est aisément vérifiée compte tenu de la formule connue du "double produit vectoriel", $\mathbf{u} \wedge (\mathbf{v} \wedge \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$. Si on écrit $(\mathbf{v} \wedge \mathbf{w})_i = \epsilon_{ijk}v_jw_k$ en termes du tenseur complètement antisymétrique ϵ , l'identité de Jacobi est celle rencontrée en (0.27).

1.3.2 Espace tangent d'un groupe de Lie G

Soit G un groupe de Lie. On considère un *sous-groupe à un paramètre* $g(t)$, où t est un paramètre réel prenant ses valeurs dans un voisinage de 0, avec $g(0) = e$; autrement dit, il s'agit d'une courbe (supposée différentiable) dans G passant par l'origine, et on suppose que (toujours au voisinage de 0),

$$g(t_1)g(t_2) = g(t_1 + t_2) \quad g^{-1}(t) = g(-t) . \quad (1.15)$$

La loi de composition de ce sous-groupe équivaut donc localement à l'addition des paramètres t ; autrement dit, localement, ce groupe à un paramètre est isomorphe au groupe abélien \mathbb{R} . Il est donc naturel de différencier

$$g(t + \delta t) = g(t)g(\delta t) \quad \Leftrightarrow \quad g^{-1}(t)g(t + \delta t) = g(\delta t) . \quad (1.16)$$

Puisque nous avons choisi de nous restreindre à des groupes de matrices, (avec $e \equiv I$, la matrice identité), nous pouvons écrire l'application linéaire tangente sous la forme

$$g(\delta t) = I + \delta t X + \dots$$

ce qui définit un vecteur X dans l'espace tangent. On écrit encore

$$X = \left. \frac{d}{dt} g(t) \right|_{t=0} , \quad (1.17)$$

c'est le vecteur vitesse en $t = 0$ (ou en $g = e$) le long de la courbe. L'équation (1.16) se réécrit donc

$$g'(t) = g(t)X . \quad (1.18)$$

Comme il est habituel en géométrie des variétés, (cf Appendice B.3), l'espace tangent $T_e G$ en e au groupe G , que nous noterons désormais \mathfrak{g} , est l'espace vectoriel engendré par les vecteurs X tangents à tous les sous-groupes à un paramètre (= tous les vecteurs vitesse). Si on a choisi dans G des coordonnées ξ^α au voisinage de e , un vecteur tangent est un opérateur différentiel $X = X^\alpha \frac{\partial}{\partial \xi^\alpha}$. La dimension de \mathfrak{g} (comme espace vectoriel) égale celle du groupe G , définie comme le nombre de paramètres (réels), $\dim \mathfrak{g} = \dim G$.

Dans le cas auquel nous nous restreignons d'un groupe $G \subset GL(n, \mathbb{R})$, $X \in \mathfrak{g} \subset M(n, \mathbb{R})$, l'ensemble des matrices réelles $n \times n$, et on peut effectuer tous les calculs dans cette algèbre. En particulier, on peut intégrer (1.18) selon

$$g(t) = \exp tX = \sum_{n=0}^{\infty} \frac{t^n}{n!} X^n, \quad (1.19)$$

une somme toujours convergente. (En fait on peut se passer de l'hypothèse que le groupe est un groupe matriciel, à condition de donner un sens à l'application \exp de \mathfrak{g} dans G , application dotée des propriétés usuelles de l'exponentielle, cf Appendice B.4.)

1.3.3 Relations entre l'espace tangent \mathfrak{g} et le groupe G

1. Si G est le groupe linéaire $GL(n, \mathbb{R})$, \mathfrak{g} est l'ensemble des matrices réelles $n \times n$, noté $M(n, \mathbb{R})$. Si G est le groupe de matrices unitaires $U(n)$, \mathfrak{g} est l'ensemble des matrices antihermitiennes $n \times n$. Elles sont en outre de trace nulle si $G = SU(n)$. De même, pour le groupe orthogonal $O(n)$, \mathfrak{g} est constitué des matrices antisymétriques, et donc de trace nulle.

Pour le groupe symplectique $G = USp(n)$, \mathfrak{g} est engendré par les matrices quaternioniques "antiselfduales", cf. Appendice A.

Dans chacun de ces cas, vérifier que la propriété caractéristique (anti-hermiticité, antisymétrie, trace nulle, ...) est préservée par le commutateur, ce qui fait bien de \mathfrak{g} une algèbre de Lie.

2. L'application exponentielle joue un rôle important dans la reconstruction du groupe de Lie G à partir de son espace tangent \mathfrak{g} . On démontre, et nous admettrons,

- que l'application $X \in \mathfrak{g} \mapsto e^X \in G$ est bijective au voisinage de l'identité ;
- qu'elle est surjective (= elle atteint tout élément de G) si G est connexe et compact ;
- qu'elle est injective (un $g \in G$ n'a qu'un seul antécédent) seulement si G est simplement connexe. Un exemple de non injectivité est fourni par $G = U(1)$, pour lequel $\mathfrak{g} = i\mathbb{R}$ et tous les $i(x + 2\pi k)$, $k \in \mathbb{Z}$ ont la même image par \exp . La réciproque est en général fautive : par exemple, dans $SU(2)$ qui est simplement connexe, si \mathbf{n} est de norme 1, $e^{i\pi \mathbf{n} \cdot \boldsymbol{\sigma}} = -I$, donc tous les éléments $i\pi \mathbf{n} \cdot \boldsymbol{\sigma}$ de $\mathfrak{g} = \mathfrak{su}(2)$ ont même image !

★ Exemple de groupe non compact pour lequel l'application \exp n'est pas surjective : $G = SL(2, \mathbb{R})$, pour lequel $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{R})$, ensemble des matrices réelles de trace nulle. Pour toute matrice $A \in \mathfrak{g}$, donc de trace nulle, montrer en utilisant son équation caractéristique que $\text{tr } A^{2n+1} = 0$, $\text{tr } A^{2n} = 2(-\det A)^n$, donc $\text{tr } e^A = 2 \cosh \sqrt{-\det A} \geq -2$. Cependant il existe dans G des matrices de trace < -2 , par exemple $\text{diag}(-2, -\frac{1}{2})$.

★ Pour un groupe non compact, l'application exponentielle est aussi utile. On démontre que tout élément d'un groupe de Lie matriciel peut s'écrire comme le produit d'un nombre fini d'exponentielles d'éléments de son algèbre de Lie. [Cornwell p 151].

★ On a encore $\det e^X = e^{\text{tr } X}$, une propriété qu'on établit aisément si X appartient à l'ensemble des matrices diagonalisables. Ces dernières étant denses dans $M(d, \mathbb{R})$, la propriété est vraie en général.

1.3.4 L'espace tangent comme algèbre de Lie

On va maintenant montrer que l'espace tangent \mathfrak{g} en $e \equiv I$ au groupe de Lie G est muni d'une structure d'algèbre de Lie. Étant donnés deux groupes à un paramètre engendrés par deux éléments distincts X et Y de \mathfrak{g} , nous mesurons leur défaut de commutativité en formant leur *commutateur* (dans un sens différent du sens usuel!) $g = e^{tX}e^{uY}e^{-tX}e^{-uY}$; pour $t \sim u$ petits, ce g est proche de l'identité, donc s'écrit $g = \exp Z$, $Z \in \mathfrak{g}$. Calculons Z au premier ordre non trivial

$$\begin{aligned} e^{tX}e^{uY}e^{-tX}e^{-uY} &= (I + tX + \frac{1}{2}t^2X^2)(I + uY + \frac{1}{2}u^2Y^2)(I - tX + \frac{1}{2}t^2X^2)(I - uY + \frac{1}{2}u^2Y^2) \\ &= I + (XY - YX)tu + O(t^3). \end{aligned} \quad (1.20)$$

On a effectué le calcul dans l'algèbre (associative) des matrices, l'élément neutre a été noté I . Tous les termes négligés sont du 3ème ordre puisque $t \sim u$. À l'ordre 2, on voit donc apparaître le commutateur $XY - YX$ au sens habituel, c'est-à-dire le crochet de Lie des matrices X et Y . En général, pour un groupe de Lie quelconque, on définit le crochet par

$$e^{tX}e^{uY}e^{-tX}e^{-uY} = e^Z \quad , \quad Z = tu[X, Y] + O(t^3) \quad (1.21)$$

et on démontre que ce crochet a les propriétés (1.14) d'un crochet de Lie.

Ce résultat fondamental découle d'une discussion détaillée de la forme locale des opérations dans un groupe de Lie ("équations de Lie", voir par exemple [OR]).

• Application adjointe dans l'algèbre de Lie \mathfrak{g} . Formule de Baker-Campbell-Hausdorff

Introduisons une notation commode. Pour tout $X \in \mathfrak{g}$, soit $\text{ad } X$ l'opérateur linéaire dans l'algèbre de Lie défini par

$$Y \mapsto (\text{ad } X)Y := [X, Y] \quad , \quad (1.22)$$

et donc

$$(\text{ad}^p X)Y = [X, [X, \dots [X, Y] \dots]]$$

avec p crochets (commutateurs).

Étant donnés deux éléments X et Y de \mathfrak{g} , e^X et e^Y les éléments de G qu'ils engendrent, existe-t-il $Z \in \mathfrak{g}$ tel que $e^Xe^Y = e^Z$? La réponse est positive, au moins pour X et Y suffisamment petits.

Notons d'abord que si $[X, Y] = 0$, les règles du calcul ordinaire s'appliquent et $Z = X + Y$. En général, la *formule de Baker-Campbell-Hausdorff*, que nous admettrons, donne une expression explicite de Z .

$$\begin{aligned} e^Xe^Y &= e^Z \\ Z &= X + \int_0^1 dt \psi(\exp \text{ad } X \exp t \text{ad } Y)Y \end{aligned} \quad (1.23)$$

où $\psi(\cdot)$ est la fonction

$$\psi(u) = \frac{u \ln u}{u-1} = 1 + \frac{1}{2}(u-1) - \frac{1}{6}(u-1)^2 + \dots, \quad (1.24)$$

régulière en $u = 1$. Explicitement, les premiers termes du développement en puissances de X et Y s'écrivent

$$Z = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}([X, [X, Y]] + [Y, [Y, X]]) + \dots \quad (1.25)$$

La formule admet des cas particuliers intéressants à connaître. Ainsi si X et Y commutent avec $[X, Y]$, on a simplement

$$e^X e^Y = e^{X+Y+\frac{1}{2}[X,Y]} = e^{X+Y} e^{\frac{1}{2}[X,Y]}, \quad (1.26)$$

formule qu'on démontre en utilisant l'identité vraie en général

$$e^X Y e^{-X} = \sum_0^{\infty} \frac{1}{n!} \text{ad}^n X Y \quad (1.27)$$

(qui n'est autre que le développement de Taylor à $t = 0$ de $e^{tX} Y e^{-tX}$ évalué en $t = 1$) et en écrivant et en résolvant l'équation différentielle satisfaite par $f(t) = e^{tX} e^{tY}$, $f(0) = 1$

$$f'(t) = (X + e^{tX} Y e^{-tX}) f(t) \quad (1.28)$$

$$= (X + Y + t[X, Y]) f(t). \quad (1.29)$$

Par ailleurs, au premier ordre en Y , on peut remplacer l'argument de ψ dans (1.23) par $\exp \text{ad} X$ et on voit qu'on a

$$Z = X + \sum_{n=0}^{\infty} \frac{B_n}{n!} (-1)^n (\text{ad} X)^n Y + O(Y^2) \quad (1.30)$$

où les B_n sont les nombres de Bernoulli : $\frac{t}{e^t-1} = \sum_0 B_n \frac{t^n}{n!}$, $B_0 = 1, B_2 = \frac{1}{6}, B_4 = -\frac{1}{30}$ et en dehors de $B_1 = -\frac{1}{2}$, tous les B d'indice impair sont nuls. Toujours au premier ordre en Y , on a encore

$$e^{X+Y} = e^X + \int_0^1 dt e^{tX} Y e^{(1-t)X} + O(Y^2)$$

qu'on obtient en écrivant et en intégrant l'équation différentielle satisfaite par $F(t) = \exp t(X+Y) \cdot \exp -tX$.

La convergence des expressions peut se démontrer pour X et Y assez petits. Bien noter que cette formule de BCH ne fait appel qu'à l'application ad dans l'algèbre de Lie, et non à la multiplication ordinaire des matrices de $\text{GL}(d, \mathbb{R})$. C'est cela qui lui donne un caractère canonique et universel.

1.3.5 Un exemple explicite : l'algèbre de Lie $\mathfrak{so}(n)$ de $\text{SO}(n)$

De la définition des éléments de \mathfrak{g} comme vecteurs tangents à G en $e \equiv I$, ou encore de la construction de sous-groupes à un paramètre associés à chaque $X \in \mathfrak{g}$, il découle l'interprétation de X comme "générateur infinitésimal" du groupe G . Le calcul concret de l'algèbre de Lie d'un groupe de Lie donné G peut s'effectuer de diverses façons, selon la manière dont on définit ou représente le groupe.

Si on a une paramétrisation explicite des éléments de G en termes de d paramètres réels, les générateurs infinitésimaux s'obtiennent par différentiation par rapport à ces paramètres. Voir au chapitre 0, le cas explicite de $SO(3)$ ou $SU(2)$ traité de cette façon.

Si le groupe a été défini comme groupe d'invariance d'une forme quadratique dans des variables x , on peut en tirer une expression des générateurs infinitésimaux comme opérateurs différentiels en x . Illustrons cela sur le groupe $O(n)$, groupe d'invariance de la forme $\sum_{i=1}^n x_i^2$. La transformation linéaire la plus générale laissant cette forme invariante s'écrit $x \rightarrow x' = O x$, avec O orthogonale. Sous forme infinitésimale, $O = I + \omega$, et $\omega = -\omega^T$ est une matrice antisymétrique arbitraire. Une transformation infinitésimale de la forme $\delta x^i = \omega^i_j x^j$ peut encore s'écrire

$$\delta x^i = \omega^i_j x^j = -\frac{1}{2} \omega^{kl} J_{kl} x^i \quad (1.31)$$

$$J_{kl} = x^k \partial_l - x^l \partial_k \quad : \quad J_{kl} x^i = x^k \delta_{il} - x^l \delta_{ik} \quad (1.32)$$

(notons que nous nous autorisons à monter et descendre librement les indices, ce qui est justifié avec la métrique de signature $(+)^n$). On dispose ainsi d'une représentation explicite des générateurs infinitésimaux de l'algèbre $so(n)$. C'est alors un calcul simple de calculer les relations de commutation⁵

$$[J_{ij}, J_{kl}] = \delta_{il} J_{jk} - \delta_{ik} J_{jl} - \delta_{jl} J_{ik} + \delta_{jk} J_{il} . \quad (1.33)$$

(Autrement dit, les seuls commutateurs non nuls sont de la forme $[J_{ij}, J_{ik}] = -J_{jk}$ pour tout triplet $i \neq j \neq k \neq i$, et tous ceux qui s'en déduisent par antisymétrie dans les indices.)

On peut enfin procéder autrement, en utilisant une base des matrices de l'algèbre de Lie, considérée comme ensemble des matrices antisymétriques $n \times n$. Une telle base est donnée par des matrices A_{ij} indexées par des paires d'indices $1 \leq i < j \leq n$, d'éléments de matrice

$$(A_{ij})^k_l = \delta_{ik} \delta_{jl} - \delta_{il} \delta_{jk} .$$

Autrement dit, la matrice A_{ij} n'a que deux éléments non nuls (et opposés), à l'intersection de la ligne i et de la colonne j et vice versa. Vérifier que ces matrices A_{ij} ont les relations de commutation données par (1.33).

Exercice : répéter cette discussion et le calcul des relations de commutation pour le groupe $SO(p, q)$ d'invariance de la forme $\sum_{i=1}^p x_i^2 - \sum_{i=p+1}^{p+q} x_i^2$. On introduira le tenseur métrique $g = \text{diag}((+1)^p, (-1)^q)$.

Application physique. Courants de Noether du "modèle $O(n)$ "

Une théorie des champs très souvent étudiée (cf. cours de F. David et Chap. 4) est le modèle $O(n)$. Son lagrangien, écrit ici dans la version euclidienne de la théorie et pour un champ bosonique $\phi = \{\phi^k\}$ réel à n composantes,

$$\mathcal{L} = \frac{1}{2} (\partial \phi)^2 + \frac{1}{2} m^2 \phi^2 + \frac{\lambda}{4} (\phi^2)^2 \quad (1.34)$$

est invariant sous l'effet des rotations de $O(n)$. Les courants de Noether se calculent en considérant des transformations infinitésimales du type précédent $\delta \phi = \sum_{i < j} \delta \omega^{ij} A_{ij} \phi$, ou, en composantes, $\delta \phi^k = \sum_{i < j} \delta \omega^{ij} (A_{ij})^k_l \phi^l$, donc (à un possible facteur près) $j_\mu^{(ij)} = \frac{\partial \mathcal{L}}{\partial \partial^\mu \phi^k} (A_{ij})^k_l \phi^l =$

5. Noter que par rapport au calcul mené pour le groupe $O(1,3)$ au chapitre 0, § 0.6.2, nous avons changé nos conventions et adopté ici des générateurs infinitésimaux antihermitiens.

$\partial_\mu \phi_k (A_{ij})^k_l \phi^l$. Vérifier en utilisant l'antisymétrie des matrices A et l'équation d'Euler-Lagrange que ces courants sont bien de divergence nulle, ce qui implique la conservation de $\dim \mathfrak{so}(n) = \frac{1}{2}n(n-1)$ "charges".

1.3.6 Un exemple de dimension infinie : l'algèbre de Virasoro

Dans ces notes nous avons convenu de nous restreindre à des groupes et algèbres de Lie de dimension finie. Donnons ici un exemple de dimension infinie. On s'intéresse aux difféomorphismes $z \mapsto z' = f(z)$ où f est une fonction analytique (holomorphe) de son argument sauf en 0 et à l'infini. (On parle aussi des "difféomorphismes du cercle".) C'est à l'évidence un groupe et une variété de dimension infinie, et cela se manifeste dans son algèbre des difféomorphismes infinitésimaux $z \mapsto z' = z + \epsilon(z)$, engendrés par les opérateurs différentiels ℓ_n

$$\ell_n = -z^{n+1} \frac{\partial}{\partial z}, \quad n \in \mathbb{Z} \quad (1.35)$$

qui satisfont

$$[\ell_n, \ell_m] = (n-m)\ell_{n+m}. \quad (1.36)$$

comme un calcul immédiat le montre. Cette algèbre est l'*algèbre de Witt*. C'est sous la forme de son *extension centrale* (cf. chap. 2), où on lui ajoute un générateur c supplémentaire "central", c'est-à-dire commutant avec tous les générateurs, que cette algèbre, dite alors *algèbre de Virasoro*, est la plus intéressante. Appelons maintenant L_n et c les générateurs

$$[L_n, L_m] = (n-m)L_{n+m} + \frac{c}{12}n(n^2-1)\delta_{n,-m} \quad [c, L_n] = 0. \quad (1.37)$$

(On peut penser aux L_n comme réalisant dans une théorie quantique des champs les opérateurs ℓ_n , le terme c résultant d'effets quantiques...)

Vérifier que l'identité de Jacobi est bien satisfaite par cette algèbre. On montre que c'est l'extension centrale la plus générale de (1.36) respectant l'identité de Jacobi. Montrer que la sous-algèbre engendrée par $L_{\pm 1}$, L_0 n'est pas affectée par le terme central. Quelle est l'interprétation géométrique des transformations correspondantes ?

L'algèbre de Virasoro joue un rôle central dans la construction des théories de champs invariantes conformes et dans leur application à la physique des phénomènes critiques bidimensionnels et à la théorie des cordes... Plus de détails dans [DFMS].

1.4 Relations entre les propriétés de \mathfrak{g} et de G

Certaines propriétés du groupe G se traduisent sur son algèbre de Lie \mathfrak{g} .

1.4.1 Simplicité, semi-simplicité

Considérons la version infinitésimale de la notion de sous-groupe invariant. Un *idéal* (on dit aussi une sous-algèbre invariante) dans une algèbre (de Lie) \mathfrak{g} est un sous-espace \mathfrak{I} de \mathfrak{g} stable par multiplication (au sens du crochet de Lie) par un élément quelconque de \mathfrak{g} , c'est-à-dire tel que $[\mathfrak{I}, \mathfrak{g}] \subset \mathfrak{I}$. L'idéal est dit *abélien* si $[\mathfrak{I}, \mathfrak{I}] = \{0\}$.

Une algèbre de Lie \mathfrak{g} est *simple* si \mathfrak{g} n'a pas d'autre idéal que $\{0\}$. Elle est *semi-simple* si \mathfrak{g} n'a pas d'autre idéal abélien que $\{0\}$.

Exemple. Considérons l'algèbre de Lie de $\mathrm{SO}(4)$, notée $\mathfrak{so}(4)$, cf les formules données en (1.33) pour $\mathfrak{so}(n)$. On vérifie aisément que les combinaisons

$$A_1 := \frac{1}{2}(J_{12} - J_{34}), \quad A_2 := \frac{1}{2}(J_{13} + J_{24}), \quad A_3 := \frac{1}{2}(J_{14} - J_{23})$$

commutent avec

$$B_1 := \frac{1}{2}(J_{12} + J_{34}), \quad B_2 = \frac{1}{2}(-J_{13} + J_{24}), \quad B_3 := \frac{1}{2}(J_{14} + J_{23})$$

et que

$$[A_i, A_j] = \epsilon_{ijk} A_k \quad [B_i, B_j] = \epsilon_{ijk} B_k, \quad [A_i, B_j] = 0$$

où on reconnaît deux copies commutantes de l'algèbre $\mathfrak{so}(3)$. On écrit $\mathfrak{so}(4) = \mathfrak{so}(3) \oplus \mathfrak{so}(3)$. À l'évidence l'algèbre $\mathfrak{so}(4)$ n'est pas simple, mais elle est semi-simple.

Bien noter la différence entre ce cas de $\mathfrak{so}(4)$ et le cas de l'algèbre $\mathfrak{so}(1,3)$ étudiée au chapitre 0, § 0.6.2. Là, la signature indéfinie nous a obligés à complexifier l'algèbre pour "découpler" les deux copies de l'algèbre $\mathfrak{so}(3)$.

On a les relations suivantes

$$\begin{aligned} G \text{ simple} &\implies \mathfrak{g} \text{ simple} \\ G \text{ semi-simple} &\implies \mathfrak{g} \text{ semi-simple} \end{aligned}$$

mais la réciproque n'est pas vraie! Plusieurs groupes de Lie différents peuvent en effet avoir la même algèbre de Lie, tels $\text{SO}(3)$ qui est simple, et $\text{SU}(2)$ qui n'est pas semi-simple, comme on l'a vu plus haut au § 1.1.7.⁶

1.4.2 Compacité. Complexification

Une algèbre de Lie semi-simple est dite compacte si elle est l'algèbre de Lie d'un groupe de Lie compact.

A ce stade, cette définition semble non intrinsèque à l'algèbre, et liée au groupe de Lie dont elle est issue. On verra plus bas une condition (critère de Cartan) qui permet de s'affranchir de cette relation.

Il faut aussi examiner la notion de complexification. Plusieurs groupes distincts peuvent avoir des algèbres de Lie différentes mais qui deviennent isomorphes si on autorise la complexification des paramètres. Par exemple les groupes $\text{O}(3)$ et $\text{O}(2,1)$, l'un compact, l'autre non, ont pour algèbres de Lie

$$\begin{aligned} \mathfrak{o}(3) &\left\{ \begin{array}{l} X_1 = z\partial_y - y\partial_z \\ X_2 = x\partial_z - z\partial_x \\ X_3 = y\partial_x - x\partial_y \end{array} \right. \quad [X_1, X_2] = y\partial_x - x\partial_y = X_3 \text{ etc} \\ \mathfrak{o}(2,1) &\left\{ \begin{array}{l} \tilde{X}_1 = z\partial_y + y\partial_z \\ \tilde{X}_2 = x\partial_z + z\partial_x \\ \tilde{X}_3 = y\partial_x - x\partial_y \end{array} \right. \quad \begin{array}{l} [\tilde{X}_1, \tilde{X}_2] = y\partial_x - x\partial_y = \tilde{X}_3 \\ [\tilde{X}_2, \tilde{X}_3] = -z\partial_y - y\partial_z = -\tilde{X}_1 \\ [\tilde{X}_3, \tilde{X}_1] = -x\partial_z - z\partial_x = -\tilde{X}_2 \end{array} \end{array} \quad (1.38)$$

qui ne sont pas isomorphes sur les réels, mais $i\tilde{X}_1$, $i\tilde{X}_2$ et $-\tilde{X}_3$ vérifient l'algèbre $\mathfrak{o}(3)$. On dit que les algèbres $\mathfrak{o}(3)$ et $\mathfrak{o}(2,1)$ ont la même complexifiée, et qu'elles en sont des formes réelles, mais seule la forme réelle $\mathfrak{o}(3)$ (ou $\mathfrak{so}(3) = \mathfrak{su}(2)$) de cette complexifiée est compacte.

6. Attention! Certains auteurs appellent "simple" tout groupe de Lie dont l'algèbre de Lie est simple. Cela revient à faire une distinction entre le concept de groupe simple et groupe de Lie simple. Ce dernier est tel qu'il ne possède pas de sous-groupe *de Lie* invariant non trivial. Le groupe de Lie $\text{SU}(2)$ est simple au sens des groupes de Lie, mais pas simple au sens général des groupes (il a un sous-groupe invariant \mathbb{Z}_2 qui n'est pas de Lie) ...

Cette complexifiée n'est autre que l'algèbre $\mathfrak{sl}(2, \mathbb{C})$, dont $\mathfrak{sl}(2, \mathbb{R})$ est aussi une forme réelle non compacte. (Voir Exercice B et les TD).

Les algèbres $\mathfrak{so}(4)$ et $\mathfrak{so}(1,3)$ étudiées plus haut et au Chap. 0 offrent un autre exemple de deux algèbres, qui sont deux formes réelles non isomorphes de la même complexifiée. Autre exemple, $\mathfrak{sp}(2n, \mathbb{R})$ et $\mathfrak{usp}(n)$. (Voir Appendice A).

De façon générale, on démontre ([FH] p. 130) que

- toute algèbre de Lie complexe semi-simple a une unique forme réelle compacte.

En résumé, les propriétés topologiques locales d'un groupe de Lie sont transcrites dans son algèbre de Lie. L'algèbre de Lie ne capte cependant pas les propriétés topologiques globales du groupe (connexité, simple-connexité, ...), comme nous le discutons maintenant.

1.4.3 Connexité, simple-connexité

– Si G n'est pas connexe et G' est le sous-groupe composante connexe de l'identité, les algèbres de Lie de G et G' coïncident $\mathfrak{g} = \mathfrak{g}'$. Par exemple, $\mathfrak{o}(3) = \mathfrak{so}(3)$.

– Si G n'est pas simplement connexe, soit \tilde{G} son groupe de recouvrement universel. G et \tilde{G} étant localement isomorphes, ils ont mêmes algèbres de Lie. Exemples $U(1)$ et \mathbb{R} ; $SO(3)$ et $SU(2)$; $SO(1,3)$ et $SL(2, \mathbb{C})$.

Pour résumer :

Étant donné un groupe de Lie G , on a construit son algèbre de Lie. Réciproquement, un théorème de Lie affirme que toute algèbre de Lie de dimension finie est l'algèbre de Lie d'un certain groupe de Lie [Ki-Jr, p.34]. Plus précisément, à toute algèbre de Lie \mathfrak{g} correspond un unique groupe de Lie G connexe et simplement connexe dont \mathfrak{g} est l'algèbre de Lie. Tout autre groupe de Lie G' connexe ayant \mathfrak{g} comme algèbre de Lie est de la forme $G' = G/H$ où H est un sous-groupe invariant fini ou discret de G . Cela est en accord avec ce nous avons vu plus haut : si G est le groupe de recouvrement de G' , $G' = G/\pi_1(G')$. Par exemple $U(1) = \mathbb{R}/\mathbb{Z}$, $SO(3) = SU(2)/\mathbb{Z}_2$. Si G' n'est pas connexe, la propriété précédente s'applique à la composante connexe de l'identité.

1.4.4 Constantes de structure. Forme de Killing. Critères de Cartan

Choissant une base $\{t_\alpha\}$ dans l'algèbre de \mathfrak{g} de dimension d , tout élément X s'écrit $X = \sum_{\alpha=1}^d x^\alpha t_\alpha$. Nous définissons les *constantes de structure* de \mathfrak{g} (dans cette base) par

$$[t_\alpha, t_\beta] = C_{\alpha\beta}^\gamma t_\gamma, \quad (1.39)$$

qui sont évidemment antisymétriques dans leurs deux indices inférieurs, $C_{\alpha\beta}^\gamma = -C_{\beta\alpha}^\gamma$. Pour l'opérateur ad défini en (1.22), on a donc

$$\text{ad } X Y = [X, Y] = \sum x^\alpha y^\beta C_{\alpha\beta}^\gamma t_\gamma.$$

Considérons alors l'opérateur linéaire $\text{ad } X \text{ad } Y$

$$\text{ad } X \text{ad } Y Z = [X, [Y, Z]] = C_{\alpha\delta}^\epsilon C_{\beta\gamma}^\delta x^\alpha y^\beta z^\gamma t_\epsilon.$$

Exercices (faciles!) : montrer que l'identité de Jacobi est équivalente à l'identité

$$\sum_{\delta} (C_{\alpha\delta}^{\epsilon} C_{\beta\gamma}^{\delta} + C_{\beta\delta}^{\epsilon} C_{\gamma\alpha}^{\delta} + C_{\gamma\delta}^{\epsilon} C_{\alpha\beta}^{\delta}) = 0 \quad (1.40)$$

(bien noter la structure : permutation cyclique sur les trois indices α, β, γ à ϵ fixe et δ sommé) ; et montrer que cette identité s'exprime encore comme

$$[\text{ad } X, \text{ad } Y]Z = \text{ad } [X, Y]Z . \quad (1.41)$$

En prenant la trace de cet opérateur linéaire $\text{ad } X \text{ad } Y$, on définit la *forme de Killing*

$$(X, Y) := \text{tr}(\text{ad } X \text{ad } Y) = \sum_{\gamma, \delta} C_{\alpha\delta}^{\gamma} C_{\beta\gamma}^{\delta} x^{\alpha} y^{\beta} =: g_{\alpha\beta} x^{\alpha} y^{\beta} , \quad (1.42)$$

qui est une forme bilinéaire symétrique (un produit scalaire) sur les vecteurs de l'algèbre de Lie. Autrement dit, le tenseur symétrique $g_{\alpha\beta}$ est donné par

$$g_{\alpha\beta} = \sum_{\gamma, \delta} C_{\alpha\delta}^{\gamma} C_{\beta\gamma}^{\delta} = \text{tr}(\text{ad } t_{\alpha} \text{ad } t_{\beta}) .$$

(La symétrie en α, β est manifeste sur la 1ère expression, elle résulte de la cyclicité de la trace dans la 2ème.)

Noter que cette forme de Killing est invariante sous l'action de tout $\text{ad } Z$:

$$\forall X, Y, Z \in \mathfrak{g} \quad (\text{ad } Z X, Y) + (X, \text{ad } Z Y) = ([Z, X], Y) + (X, [Z, Y]) = 0 \quad (1.43)$$

(penser à l'action de $\text{ad } Z$ comme celle d'un générateur infinitésimal agissant à la manière d'une dérivation, soit sur le premier terme, soit sur le second). En effet par (1.41), le premier terme vaut $\text{tr}(\text{ad } Z \text{ad } X \text{ad } Y - \text{ad } X \text{ad } Z \text{ad } Y)$ et le second $\text{tr}(\text{ad } X \text{ad } Z \text{ad } Y - \text{ad } X \text{ad } Y \text{ad } Z)$, et ils sont opposés grâce à la cyclicité de la trace. On démontre que dans une algèbre de Lie simple, une forme bilinéaire invariante est nécessairement multiple de la forme de Killing.

On peut alors utiliser le tenseur $g_{\alpha\beta}$ pour abaisser le 3ème indice de $C_{\alpha\beta}^{\gamma}$, définissant ainsi

$$C_{\alpha\beta\gamma} := C_{\alpha\beta}^{\delta} g_{\gamma\delta} = C_{\alpha\beta}^{\delta} C_{\gamma\epsilon}^{\kappa} C_{\delta\kappa}^{\epsilon}$$

Montrons alors que ce $C_{\alpha\beta\gamma}$ est complètement antisymétrique en α, β, γ . Compte tenu de l'antisymétrie en α, β déjà connue, il suffit d'établir que $C_{\alpha\beta\gamma}$ est invariant par permutation cyclique. Cela découle de (1.43) qui peut être écrite sous une forme plus symétrique selon

$$(X, [Y, Z]) = (Y, [Z, X]) = (Z, [X, Y]) = C_{\alpha\beta\gamma} x^{\alpha} y^{\beta} z^{\gamma} = C_{\beta\gamma\alpha} y^{\beta} z^{\gamma} x^{\alpha} = C_{\gamma\alpha\beta} z^{\gamma} x^{\alpha} y^{\beta} , \quad (1.44)$$

ce qui établit la propriété annoncée.

Un théorème très remarquable d'E. Cartan affirme :

- (i) Une algèbre de Lie est semi-simple ssi la forme de Killing est non-dégénérée, c'est-à-dire $\det g \neq 0$.
- (ii) Une algèbre de Lie semi-simple réelle est compacte ssi la forme de Killing g est définie

négative.

Ce sont les *critères de Cartan*.

Dans un sens, la propriété (i) est aisée à établir. Supposons que \mathfrak{g} n'est pas semi-simple et montrons que $\det g = 0$. Soit \mathfrak{J} un idéal de \mathfrak{g} , choisissons une base de \mathfrak{g} faite d'une base de \mathfrak{J} , $\{t_i\}$, $i = 1, \dots, r$, complétée par t_a , $a = r + 1, \dots, d$. Calculons, pour $1 \leq i, j \leq r$, $g_{ij} = \sum_{\alpha\beta} C_{i,\alpha}^\beta C_{j,\beta}^\alpha$. La propriété d'idéal nous dit que α et β sont eux-mêmes entre 1 et r , $g_{ij} = \sum_{1 \leq k, l \leq r} C_{i,k}^l C_{j,l}^k$. donc la restriction de la forme de Killing de \mathfrak{g} à \mathfrak{J} est la forme de Killing de \mathfrak{J} . Si en outre, on suppose l'idéal abélien, $g_{ij} = 0$ et $g_{ia} = 0$. La forme est clairement dégénérée ($\det g = 0$). La réciproque, $\det g = 0 \Rightarrow \mathfrak{g}$ non semi-simple, est un peu plus délicate à établir.

De même, la propriété (ii) est assez aisée à établir dans le sens compacité \Rightarrow forme définie négative. Partons d'une forme bilinéaire symétrique définie positive arbitraire ; par exemple dans une base $\{t_\alpha\}$ donnée, considérons $\langle X, Y \rangle = \sum x^\alpha y^\beta$. Pour un groupe G compact, on peut rendre cette forme invariante en moyennant sur G : $\varphi(X, Y) := \int d\mu(g) \langle gXg^{-1}, gYg^{-1} \rangle$. Elle est invariante $\varphi(gXg^{-1}, gYg^{-1}) = \varphi(X, Y)$, soit sous forme infinitésimale, $\varphi([Z, X], Y) + \varphi(X, [Z, Y]) = 0$, (cf (1.43)). Elle est aussi définie positive. Soit e_α une base qui la diagonalise, $\varphi(e_\alpha, e_\beta) = \delta_{\alpha\beta}$. Calculons dans cette base la matrice de l'opérateur $\text{ad } X$ et montrons qu'elle est antisymétrique, $(\text{ad } X)_{\alpha\beta} = -(\text{ad } X)_{\beta\alpha}$:

$$(\text{ad } X)_{\alpha\beta} = \varphi(e_\alpha, [X, e_\beta]) = -\varphi(e_\beta, [X, e_\alpha]) = -(\text{ad } X)_{\beta\alpha} .$$

Donc la forme de Killing

$$(X, X) = \text{tr}(\text{ad } X \text{ad } X) = \sum_{\alpha,\beta} (\text{ad } X)_{\alpha\beta} (\text{ad } X)_{\beta\alpha} = -\sum_{\alpha,\beta} ((\text{ad } X)_{\alpha\beta})^2 \leq 0$$

est semi-définie négative, et si l'algèbre est semi-simple, définie négative, q.e.d.

Exemple. Le cas de $\text{SO}(3)$ ou $\text{SU}(2)$ est bien familier. Les constantes de structure sont données par le tenseur complètement antisymétrique $C_{\alpha\beta\gamma} = \epsilon_{\alpha\beta\gamma}$. La forme de Killing est $g_{\alpha\beta} = -2\delta_{\alpha\beta}$. Exercice : calculer la forme de Killing pour l'algèbre $\mathfrak{so}(2, 1)$ (voir Exercice B).

Enfin un dernier théorème important (toujours de Cartan!) énonce que

- Toute algèbre de Lie semi-simple \mathfrak{g} est somme directe d'algèbres de Lie simples \mathfrak{g}_i

$$\mathfrak{g} = \bigoplus_i \mathfrak{g}_i .$$

Ceci est une conséquence simple de (1.44). Considérons une algèbre semi-simple \mathfrak{g} ayant un idéal \mathfrak{J} et appelons \mathfrak{C} le complément de \mathfrak{J} par rapport à la forme de Killing, c'est-à-dire $(\mathfrak{J}, \mathfrak{C}) = 0$. Par (1.44), $([\mathfrak{C}, \mathfrak{J}], \mathfrak{J}) = (\mathfrak{C}, [\mathfrak{J}, \mathfrak{J}]) = (\mathfrak{C}, \mathfrak{J}) = 0$ (puisque \mathfrak{J} est une sous-algèbre), et $([\mathfrak{C}, \mathfrak{J}], \mathfrak{C}) = (\mathfrak{J}, \mathfrak{C}) = 0$ (puisque \mathfrak{J} est un idéal), donc $[\mathfrak{C}, \mathfrak{J}]$, orthogonal à tout \mathfrak{g} pour la forme de Killing non dégénérée, s'annule, $[\mathfrak{C}, \mathfrak{J}] = 0$, ce qui signifie que $\mathfrak{g} = \mathfrak{J} \oplus \mathfrak{C}$. En itérant l'argument sur \mathfrak{C} , on obtient la propriété annoncée.

Ces propriétés ont été mises à profit par Cartan pour classifier les algèbres de Lie simples complexes ou réelles. Nous reviendrons au Chapitre 3 sur cette classification.

1.4.5 Opérateur(s) de Casimir

Avec les notations précédentes, étant données une algèbre \mathfrak{g} semi-simple, donc dotée d'une forme de Killing g inversible, et une base $\{t_\alpha\}$ de \mathfrak{g} , on définit

$$C_2 = \sum_{\alpha,\beta} g^{\alpha\beta} t_\alpha t_\beta \tag{1.45}$$

où $g^{\alpha\beta}$ est l'inverse de $g_{\alpha\beta}$, c'est-à-dire $g_{\alpha\gamma} g^{\gamma\beta} = \delta_\alpha^\beta$.

Formellement, cette combinaison des t qui ne fait pas appel au crochet de Lie ne vit pas dans l'algèbre de Lie mais dans son *algèbre enveloppante universelle* $U\mathfrak{g}$, définie comme l'algèbre associative des polynômes dans les éléments de \mathfrak{g} . Ici, puisque nous nous sommes restreints à $\mathfrak{g} \subset M(n, \mathbb{R})$, $U\mathfrak{g}$ peut aussi être considérée comme une sous-algèbre de $M(n, \mathbb{R})$.

Montrons que C_2 a un crochet (commutateur) nul avec tout t_γ donc avec tout élément de \mathfrak{g} . C'est l'*opérateur de Casimir* quadratique.

$$\begin{aligned}
[C_2, t_\gamma] &= \sum_{\alpha, \beta} g^{\alpha\beta} [t_\alpha t_\beta, t_\gamma] \\
&= \sum_{\alpha, \beta} g^{\alpha\beta} (t_\alpha [t_\beta, t_\gamma] + [t_\alpha, t_\gamma] t_\beta) \\
&= \sum_{\alpha, \beta, \delta} g^{\alpha\beta} C_{\beta\gamma}^\delta (t_\alpha t_\delta + t_\delta t_\alpha) \\
&= \sum_{\alpha, \beta, \delta, \kappa} g^{\alpha\beta} g^{\delta\kappa} C_{\beta\gamma\kappa} (t_\alpha t_\delta + t_\delta t_\alpha).
\end{aligned} \tag{1.46}$$

Le terme $\sum_{\beta\kappa} g^{\alpha\beta} g^{\delta\kappa} C_{\beta\gamma\kappa}$ est antisymétrique en $\alpha \leftrightarrow \delta$, tandis que le terme entre parenthèses est symétrique. La somme s'annule donc, qed.

On démontre que dans une algèbre de Lie simple, (plus précisément dans son algèbre enveloppante universelle), une expression *quadratique* dans les t qui commute avec tous les t est proportionnelle à l'opérateur de Casimir C_2 . Autrement dit, l'opérateur de Casimir quadratique est unique à un facteur près.

Exemple. Dans l'algèbre $\mathfrak{so}(3) \cong \mathfrak{su}(2)$, l'opérateur de Casimir C_2 est (à un signe près) \mathbf{J}^2 , qui, comme chacun sait, commute avec les générateurs infinitésimaux J^i de l'algèbre. Dans une algèbre non simple, il y a autant d'opérateurs de Casimir quadratiques qu'il y a de composantes simples, voir par exemple les deux opérateurs de Casimir \mathbf{J}^2 et \mathbf{K}^2 dans l'algèbre du groupe de Lorentz complexifiée (see Chap. 0 § 0.6.2); ou P^2 et W^2 dans l'algèbre de Poincaré (non semi-simple), see Chap. 0, § 0.6.5.

Il peut exister d'autres opérateurs de Casimir de degré plus élevé. Vérifier ainsi que

$$C_r = g^{\alpha_1 \alpha'_1} g^{\alpha_2 \alpha'_2} \dots g^{\alpha_r \alpha'_r} C_{\alpha_1 \beta_1}^{\beta_2} C_{\alpha_2 \beta_2}^{\beta_3} \dots C_{\alpha_r \beta_r}^{\beta_1} t_{\alpha'_1} t_{\alpha'_2} \dots t_{\alpha'_r} \tag{1.47}$$

a un crochet nul avec tout t_γ . Que vaut C_3 dans $\mathfrak{su}(2)$? Voir Bourbaki [Bo] pour la discussion de ces opérateurs de Casimir généraux. Voir aussi l'exercice C ci-dessous.

Si on se rappelle que les générateurs infinitésimaux (vecteurs de l'algèbre de Lie) s'interprètent comme des opérateurs différentiels dans les coordonnées sur le groupe, on conçoit que les opérateurs de Casimir fournissent des opérateurs différentiels invariants (puisque commutant avec les générateurs infinitésimaux). En particulier, l'opérateur de Casimir quadratique correspond à un laplacien sur le groupe (see Chap. 0, § 0.2.3 pour le cas de $\text{SO}(3)$).

Ces opérateurs de Casimir vont jouer un rôle important dans l'étude des représentations des groupes.

Bibliographie sommaire

Ouvrages mathématiques

[Bo] N. Bourbaki, *Groupes et Algèbres de Lie*, Chap. 1-9, Hermann 1960-1983.

[Bu] D. Bump, *Lie groups*, Series “Graduate Texts in Mathematics”, vol. **225**, Springer 2004.

[Ch] C. Chevalley, *Theory of Lie groups*, Princeton University Press.

[D] J. Dieudonné, *Éléments d’analyse*, Gauthier-Villars, en particulier tomes 5-7 (très complets mais difficiles!).

[DNF] B. Doubrovine, S. Novikov et A. Fomenko, *Géométrie contemporaine*, Éditions de Moscou, en particulier, les §14, 23 et 24 du volume 1 et les chapitres 1 (variétés) et 4 et 5 (homotopie) du volume 2. La version française étant épuisée, voici les références de la version anglaise : Dubrovin, B. A., Fomenko, A. T., Novikov, S. P. *Modern geometry—methods and applications*. Part I. The geometry of surfaces, transformation groups, and fields. Graduate Texts in Mathematics, 93. Springer-Verlag, New York, 1992. Part II. The geometry and topology of manifolds. Graduate Texts in Mathematics, 104. Springer-Verlag, New York, 1985. (a third volume deals with homology...)

[Ki-Jr] A. Kirillov Jr, *An Introduction to Lie groups and Lie algebras*, (Cambridge Studies in Advanced Mathematics), Cambridge Univ. Pr., 2008.

[Po] L.S. Pontryagin, *Topological Groups*, Gordon and Breach, 1966.

[W] H. Weyl, *Classical groups*, Princeton University Press

Un ouvrage récent écrit par une mathématicienne, mais avec un contenu et dans un esprit proches du présent cours, est celui de

[K-S] Y. Kosmann-Schwarzbach, *Groupes et symétries, Groupes finis, groupes et algèbres de Lie, représentations*, Les Éditions de l’École Polytechnique, 2006.

Théorie des groupes pour physiciens

[Wi] E. Wigner, *Group Theory and its Applications to Quantum Mechanics*. Academ. Pr. 1959

[Co] J.F. Cornwell, *Group theory in physics. An introduction*, Academic Pr. contient beaucoup d’information mais utilise une terminologie parfois différente du reste de la littérature...

[Gi] R. Gilmore, *Lie groups, Lie algebras and some of their applications*, Wiley

[Ha] M. Hamermesh, *Group theory and its applications to physical problems*, Addison-Wesley

[Itz] C. Itzykson, *Notes de cours pour l’Ecole de Physique Mathématique de l’Université de Toulouse* (Saclay report (in French), September 1974)

[OR] L. O’ Raifeartaigh, *Group structure of gauge theories*, Cambridge Univ. Pr. 1986.

Voir aussi plusieurs cours de théorie des groupes par et pour des physiciens disponibles sur le serveur du CCSD, <http://cel.ccsd.cnrs.fr/>, entre autres

J.-B. Z., *Introduction à la théorie des groupes et de leurs représentations*, (Notes de cours au Magistère MIP 1994), qui met plutôt l’accent sur les groupes finis.

Appendix A. Corps des quaternions et groupes symplectiques

A.1 Quaternions

L'ensemble des quaternions est l'algèbre engendrée par 4 éléments 1, e_i , $i = 1, 2, 3$,

$$q = q^{(0)}1 + q^{(1)}e_1 + q^{(2)}e_2 + q^{(3)}e_3 \quad q^{(\cdot)} \in \mathbb{C} \quad (\text{A.1})$$

dotée de la multiplication $e_i^2 = e_1e_2e_3 = -1$, d'où il découle que

$$e_1e_2 = -e_2e_1 = e_3$$

et ses permutations cycliques. On peut représenter les e_i en termes des matrices de Pauli : $e_i \mapsto -i\sigma_i$.

Le conjugué d'un quaternion q est le quaternion

$$\bar{q} = q^{(0)}1 - q^{(1)}e_1 - q^{(2)}e_2 - q^{(3)}e_3 . \quad (\text{A.2})$$

à ne pas confondre avec son complexe conjugué

$$q^* = q^{(0)*}1 + q^{(1)*}e_1 + q^{(2)*}e_2 + q^{(3)*}e_3 . \quad (\text{A.3})$$

Noter que $q\bar{q} := |q|^2 = |q^{(0)}|^2 + |q^{(1)}|^2 + |q^{(2)}|^2 + |q^{(3)}|^2$, la norme carrée du quaternion, et donc $q^{-1} = \bar{q}/|q|^2$ si sa norme est non nulle.

On définit encore le conjugué hermitique de q

$$q^\dagger = \bar{q}^* = q^{(0)*}1 - q^{(1)*}e_1 - q^{(2)*}e_2 - q^{(3)*}e_3 \quad (\text{A.4})$$

(en accord avec le fait que les matrices de Pauli sont hermitiennes).

Noter que la conjugaison et la conjugaison hermitique renversent l'ordre des facteurs

$$\overline{(q_1q_2)} = \bar{q}_2\bar{q}_1 \quad (q_1q_2)^\dagger = q_2^\dagger q_1^\dagger . \quad (\text{A.5})$$

Un quaternion *réel* est un quaternion de la forme (A.1) avec $q^{(\mu)} \in \mathbb{R}$, donc identique à son complexe conjugué.

L'ensemble des quaternions réels forment un corps, qui est aussi un espace de dimension 4 sur \mathbb{R} . Il est désigné par \mathbb{H} (H comme Hamilton).

A.2 Matrices de quaternions

On considère des matrices Q d'éléments quaternioniques $(Q)_{ij} = q_{ij}$, ou $Q = (q_{ij})$. On peut appliquer à Q les conjugaisons définies plus haut. En outre, on peut transposer Q . L'hermitique conjugué de Q est défini par

$$(Q^\dagger)_{ij} = q_{ji}^\dagger . \quad (\text{A.6})$$

Le *dual* Q^R d'une matrice quaternionique Q est la matrice

$$(Q^R)_{ij} = \bar{q}_{ji} . \quad (\text{A.7})$$

(Cela joue pour les matrices quaternioniques le même rôle que le conjugué hermitique pour des matrices complexes.) Une matrice quaternionique est donc dite *self-duale* si

$$Q^R = Q = (q_{ij}) = (\bar{q}_{ji}) , \quad (\text{A.8})$$

elle est *quaternionique réelle* si

$$Q^R = Q^\dagger \quad \text{donc} \quad q_{ij} = q_{ij}^* , \quad (\text{A.9})$$

donc si ses éléments sont des quaternions réels.

A.3 Groupes symplectiques $\mathrm{Sp}(2n, \mathbb{R})$ et $\mathrm{USp}(n)$, algèbres de Lie $\mathfrak{sp}(2n)$ et $\mathfrak{usp}(n)$

Soit la matrice $2n \times 2n$

$$S = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} \quad (\text{A.10})$$

où I_n est la matrice identité en dimension n , et la forme bilinéaire alternée (“skew-symmetric” en anglais) associée

$$(X, Y) = X^T S Y = \sum_{i=1}^n (x_i y_{i+n} - y_i x_{i+n}) . \quad (\text{A.11})$$

Le groupe symplectique $\mathrm{Sp}(2n, \mathbb{R})$ est le groupe de matrices réelles $2n \times 2n$ préservant cette forme

$$B^T S B = S . \quad (\text{A.12})$$

Dans la base où $X^T = (x_1, x_{n+1}, x_2, x_{n+2}, \dots)$, la matrice $S = \mathrm{diag} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \mathrm{diag}(-e_2)$ en termes quaternioniques, et le groupe symplectique est alors engendré par des matrices quaternioniques $n \times n$ Q satisfaisant $Q^R \cdot Q = I$, (le vérifier !); cependant, la matrice B étant réelle, les éléments de Q sont tels que les $q_{ij}^{(\alpha)}$ sont réels pour $\alpha = 0, 2$ et imaginaires purs pour $\alpha = 1, 3$. Ce groupe n’est pas compact. Son algèbre de Lie $\mathfrak{sp}(2n, \mathbb{R})$ est engendrée par les matrices réelles A telles que $A^T S + S A = 0$. La dimension de ce groupe ou de son algèbre de Lie est $n(2n + 1)$. Pour $n = 1$, $\mathrm{Sp}(2, \mathbb{R}) = \mathrm{SL}(2, \mathbb{R})$.

Un groupe relié est le groupe $\mathrm{USp}(n)$ engendré par les matrices $n \times n$ unitaires et quaternioniques réelles $Q^R = Q^\dagger = Q^{-1}$. C’est le groupe d’invariance de la forme hermitienne quaternionique $\sum \bar{x}_i y_i$, $x, y \in \mathbb{H}^n$. Il est compact car c’est un sous groupe de $\mathrm{U}(2n)$. Son algèbre de Lie $\mathfrak{usp}(n)$ est engendrée par les matrices quaternioniques réelles antiselfduales $A = -A^R = -A^\dagger$ (le vérifier). Elle a aussi $n(2n + 1)$ pour dimension. Pour $n = 1$, $\mathrm{USp}(1) = \mathrm{SU}(2)$.

En exprimant la condition sur les matrices A de $\mathfrak{sp}(n, \mathbb{R})$ en termes de quaternions, on constate que les deux algèbres $\mathfrak{sp}(2n, \mathbb{R})$ et $\mathfrak{usp}(n)$ ont la même algèbre complexifiée, qui n’est autre que $\mathfrak{sp}(2n, \mathbb{C})$. Seule $\mathfrak{usp}(n)$ est compacte.

Appendix B. Rappels et compléments de topologie et de géométrie différentielle

B.1 Petit lexique de quelques concepts de topologie utilisés dans ces notes

Espace topologique : ensemble E doté d’une collection de sous-ensembles, dits *ouverts*, avec la propriété que l’union d’ouverts ou l’intersection d’un nombre fini d’ouverts est un ouvert, et que E et \emptyset sont des ouverts.

Fermé de E : complémentaire d’un ouvert dans E .

Voisinage d’un point x : sous-ensemble de E contenant un ouvert contenant x . Soit $\mathcal{V}(x)$ l’ensemble des voisinages de x .

Un espace topologique est *séparé* (ou de Hausdorff) si deux points distincts possèdent deux voisinages distincts. On supposera toujours cette propriété satisfaite dans ces notes.

Base de voisinages $\mathcal{B}(x)$ d’un point x : sous-ensemble de $\mathcal{V}(x)$ tel que tout $V \in \mathcal{V}(x)$ contient un $W \in \mathcal{B}(x)$. (Intuitivement, une base est constituée de “suffisamment” de voisinages.)

Espace compact E : espace topologique (séparé) tel que de tout recouvrement de E par des ouverts, on peut extraire un recouvrement fini ;

Conséquences : si E est compact,

- toute suite infinie de points de E admet un point d'accumulation ;
- si $f : E \mapsto F$ est continue, $f(E)$ est compact ;
- toute fonction continue sur E est bornée.

Si E est un sous-espace de \mathbb{R}^n , E compact $\Leftrightarrow E$ borné et fermé (théorème de Heine–Borel).

Espace localement compact : espace (séparé) dont tout point a au moins un voisinage compact. Exemples : \mathbb{R} n'est pas compact mais localement compact ; \mathbb{Q} n'est ni compact ni localement compact.

B.2 Notion de variété

Une *variété* M (manifold en anglais) de dimension n est un espace qui localement, au voisinage de chaque point, “ressemble” à \mathbb{R}^n ou \mathbb{C}^n . Des contre-exemples sont fournis par deux droites sécantes, ou par $\text{---}\bigcirc$. Plus précisément, il existe une base de voisinages U_i couvrant M , avec des *cartes* f_i , fonctions inversibles et bicontinues (homéomorphismes) entre U_i et un ouvert de \mathbb{R}^n : $f_i(U_i) \subset \mathbb{R}^n$. Soit m un point de M , $m \in U_i$, et $f_i(m) = (x^1, x^2, \dots, x^n)$ son image dans \mathbb{R}^n : (x^1, x^2, \dots, x^n) sont les *coordonnées locales* de m , elles dépendent de la carte. Il est fondamental de savoir changer de carte (de système de coordonnées). On dit que la variété est différentiable de classe C^k si pour toute paire d'ouverts U_i et U_j d'intersection non nulle, $f_j \circ f_i^{-1}$ qui applique $f_i(U_i \cap U_j) \subset \mathbb{R}^n$ sur $f_j(U_i \cap U_j)$ est de classe C^k .

Exemple : la sphère S^2 est une variété analytique de dimension 2. On peut choisir comme deux ouverts la sphère privée de son pôle Nord, resp. Sud, avec comme carte la projection stéréographique (cf Problème ci-dessous) à partir de ce pôle.

Une *variété riemannienne* est une variété différentiable réelle dont l'espace tangent est muni d'un produit scalaire défini positif. Si ce produit scalaire est seulement supposé non dégénéré et de signature $((+1)^p, (-1)^{n-p})$, la variété est dite *pseudo-riemannienne*. Dans des coordonnées locales x^i , un vecteur tangent (voir ci-dessous § B.3) s'écrit $X = X^i \frac{\partial}{\partial x^i}$, et ce produit scalaire et l'élément de longueur carrée sont donnés par le tenseur métrique g

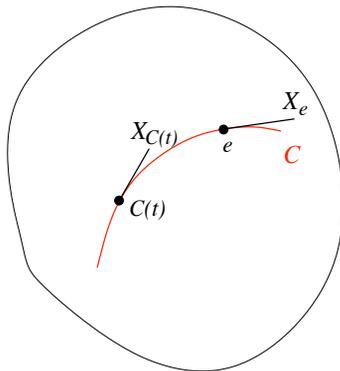
$$(X, Y) = g_{ij} X^i Y^j, \quad ds^2 = g_{ij} dx^i dx^j. \quad (\text{B.1})$$

B.3 Espace tangent

En géométrie différentielle, un vecteur tangent X à une variété M en un point x_0 est un opérateur différentiel linéaire, du premier ordre dans les dérivées en x_0 , agissant sur les fonctions f sur M . En coordonnées locales x^i ,

$$X : f(x) \mapsto \sum_i X^i \frac{\partial}{\partial x^i} \Big|_{x_0} f(x)$$

et par changement de coordonnées $\{x^i\} \rightarrow \{y^j\}$, ces opérateurs se transforment par la matrice jacobienne $\frac{\partial}{\partial y^j} = \sum_i \frac{\partial x^i}{\partial y^j} \Big|_{x_0} \frac{\partial}{\partial x^i}$ avec la transformation des $X^i \rightarrow Y^j$ qui en découle.

FIGURE 1.4 – Le champ de vecteurs tangents à la courbe $C(t)$ est un champ invariant à gauche

Vecteur tangent à une courbe : si une courbe $C(t)$ passe par le point x_0 en $t = 0$, on peut dériver une fonction f le long de cette courbe

$$f \mapsto \left. \frac{df(C(t))}{dt} \right|_{t=0} .$$

Cela définit le vecteur tangent à la courbe C au point x_0 , appelé aussi vecteur vitesse et noté $C'(t)|_{t=0} = C'(0)$.

L'espace tangent à M en x_0 , noté $T_{x_0}M$, est l'espace vectoriel engendré par les vecteurs vitesses à toutes les courbes passant par x_0 . L'espace $T_{x_0}M$ a pour base les $\left. \frac{\partial}{\partial x^i} \right|_{x_0}$: il a la même dimension que M .

On appelle *champ de vecteurs* sur une variété M la donnée en tout point $x \in M$ d'un vecteur X_x tangent à M en x .

B.4 Groupe de Lie. Application exponentielle

Soit un groupe G , e son identité. Soit $C(t)$ une courbe passant par $C(0) = e$, et soit $X_e = (C'(t))_{t=0}$ son vecteur vitesse en e . Pour $g \in G$, on définit la *translatée à gauche* $g.C(t)$ de C par g . Son vecteur vitesse en g , $X_g = (g.C(t))'_{t=0}$, est appelé *vecteur translaté à gauche* de X_e . Le champ de vecteurs $g \mapsto X_g$ est dit invariant à gauche, ce qui se justifie, puisque c'est l'ensemble des vecteurs translatés à gauche de X_e . L'espace tangent en e et l'espace des champs de vecteurs invariants sont donc isomorphes, on les note \mathfrak{g} .

Inversement, étant donné X_e un vecteur tangent à G en e , on note

$$C(t) = \exp tX_e \tag{B.2}$$

l'unique solution de l'équation différentielle

$$C'(t) = X_{C(t)} \tag{B.3}$$

qui exprime que la courbe $C(t)$ est tangente en chacun de ses points au champ de vecteurs invariant à gauche, équation complétée par la condition initiale $C(0) = e$. Cette équation différentielle du premier ordre a en effet une solution déterminée à une constante (dans le groupe) près, constante qui est fixée de façon unique par la condition initiale.

La fonction définie par (B.2) satisfait la propriété (1.15). En effet, si $C(t)$ satisfait (B.3), $C(t+t')$ la satisfait aussi et diffère donc de $C(t)$ par une constante, $C(t+t') = k.C(t)$, (k dans le groupe), constante qui est fixée en prenant $t = 0$, $C(t') = k$, donc $C(t+t') = C(t')C(t)$ et $C(-t) = C(t)^{-1}$, qed.

Dans le cas des groupes matriciels considéré dans ce cours, cette fonction \exp s'identifie bien sûr à la fonction exponentielle définie par son développement de Taylor (1.19).

Appendix C. Mesures invariantes sur SU(2) et sur U(n)

Le groupe SU(2) isomorphe à une sphère est compact et on peut donc intégrer une fonction sur ce groupe avec une grande variété de mesures d'intégration $d\mu(g)$. La mesure invariante, c'est-à-dire telle que $d\mu(g.g_1) = d\mu(g_1.g) = d\mu(g^{-1}) = d\mu(g)$, est, elle, unique à un facteur près.

Une manière possible de trouver cette mesure est de considérer la transformation $U \rightarrow U' = U.V$ où U, V et donc U' sont unitaires de la forme (0.10) (c'est-à-dire $U = u_0 I - \mathbf{u} \cdot \boldsymbol{\sigma}$, $u \in S^3$ etc) ; si on relâche momentanément la condition que $u_0^2 + \mathbf{u}^2 = 1$ (mais qu'on garde $v_0^2 + \mathbf{v}^2 = 1$), cela définit une transformation linéaire $u \rightarrow u'$ qui conserve la norme $\det U = u_0^2 + \mathbf{u}^2 = u_0'^2 + \mathbf{u}'^2 = \det U'$. C'est donc une isométrie de l'espace \mathbb{R}^4 qui préserve la mesure naturelle $d^4 u \delta(u^2 - 1)$ sur la sphère unité S^3 d'équation $\det U = 1$. En d'autres termes, cette mesure sur la sphère S^3 fournit une mesure invariante à droite : $d\mu(U) = d\mu(U.V)$. On démontrerait de la même façon que cette mesure est aussi invariante à gauche : $d\mu(U) = d\mu(V.U)$. Cette mesure est aussi invariante par $U \rightarrow U^{-1}$, car l'inversion dans SU(2) est la restriction à S^3 de la transformation orthogonale $u_0 \rightarrow u_0$, $\mathbf{u} \rightarrow -\mathbf{u}$ de \mathbb{R}^4 , qui préserve bien sûr la mesure naturelle sur S^3 :

$$d\mu(U) = d\mu(UV) = d\mu(VU) = d\mu(U^{-1}) .$$

La forme explicite de la mesure dépend de la paramétrisation utilisée. Si on adopte la direction \mathbf{n} (ou ses deux angles polaires θ et ϕ) et l'angle de rotation ψ , on prendra

$$d\mu(U) = \frac{1}{2} \sin^2 \frac{\psi}{2} \sin \theta d\psi d\theta d\phi \quad (\text{C.1})$$

normalisée pour SU(2) à

$$v(\text{SU}(2)) = \int_{\text{SU}(2)} d\mu(U) = \frac{1}{2} \int_0^\pi d\theta \sin \theta \int_0^{2\pi} d\phi \int_0^{2\pi} d\psi \sin^2 \frac{\psi}{2} = 2\pi^2 \quad (\text{C.2})$$

qui est l'"aire" de la sphère unité S^3 et le "volume" de SU(2). Pour SO(3) où l'angle ψ est restreint à $(0, \pi)$, on a plutôt $v(\text{SO}(3)) = \int_{\text{SO}(3)} d\mu(g) = \pi^2$.

On obtient l'expression dans tout autre système de coordonnées, par exemple les angles d'Euler, en calculant le jacobien adéquat,

$$d\mu(U) = \frac{1}{8} \sin \beta d\alpha d\beta d\gamma . \quad (\text{C.3})$$

(Noter que $0 \leq \gamma \leq 4\pi$ pour SU(2), tandis que $0 \leq \alpha \leq 2\pi$ et $0 \leq \beta \leq \pi$).

Comparer ces formules avec celles trouvées par une autre méthode au Chapitre 0, App. 0.

• Cas de U(n).

Examinons finalement rapidement le cas du groupe U(n) en lui appliquant la méthode du Chap. 0, App. 0. Toute matrice unitaire $U \in \text{U}(n)$ peut se diagonaliser sous la forme

$$U = V \Lambda V^\dagger , \quad (\text{C.4})$$

où $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ et les λ_i sont en fait des phases $\lambda_j = e^{i\alpha_j}$. Les λ_i peuvent être considérées comme des variables "radiales", tandis que V représente les variables "angulaires". Noter que V doit être restreint à ne pas commuter avec la matrice diagonale Λ . Supposant cette dernière générique, avec des valeurs propres λ_i

toutes distinctes, V vit dans $U(n)/U(1)^n$. La métrique naturelle, invariante par $U \mapsto U'U$ ou $\mapsto UU'$, s'écrit $\text{tr}(dUdU^\dagger)$. Or $dU = V(d\Lambda + [dX, \Lambda])V^\dagger$, où $dX := V^\dagger dV$ est antihermitienne (et sans termes diagonaux, pourquoi?). On a donc $\text{tr}(dUdU^\dagger) = \sum_i |d\alpha_i|^2 + 2 \sum_{i < j} |dX_{ij}|^2 |\lambda_i - \lambda_j|^2$ ce qui définit le tenseur métrique $g_{\alpha\beta}$ dans les coordonnées $\xi^\alpha = (\alpha_i, \Re X_{ij}, \Im X_{ij})$ et détermine la mesure d'intégration

$$d\mu(U) = \sqrt{\det g} \prod d\xi^\alpha = \text{const.} |\Delta(e^{i\alpha})|^2 \prod d\alpha_i d\mu(V). \quad (\text{C.5})$$

Ici $\Delta(\lambda)$ est le *déterminant de Vandermonde*

$$\Delta(\lambda) := \prod_{i < j} (\lambda_i - \lambda_j) = \begin{vmatrix} \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \\ \vdots & & & \vdots \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ 1 & 1 & \dots & 1 \end{vmatrix}. \quad (\text{C.6})$$

La partie "radiale" de la mesure d'intégration est donc donnée par $|\Delta(e^{i\alpha})|^2 \prod d\alpha_i$ à un facteur constant près, soit encore

$$d\mu(U) = \text{const.} \prod_{i < j} \sin^2 \left(\frac{\alpha_i - \alpha_j}{2} \right) \prod d\alpha_i \times \text{partie angulaire}. \quad (\text{C.7})$$

Noter que cette partie radiale de la mesure est suffisante si on a à intégrer sur le groupe une fonction de U invariante par $U \rightarrow VUV^\dagger$, $V \in U(n)$. Par exemple $\int d\mu(U) \text{tr} P(U)$, avec P un polynôme.

Exercices et problème du chapitre 1

A. Action d'un groupe dans un ensemble.

Soit E un ensemble, G un groupe. On dit que le groupe G agit dans l'ensemble E s'il existe un homomorphisme β de G dans le groupe des bijections de E dans lui-même.

1. Écrire précisément les conditions requises.

On définit alors l'*orbite* $O(x)$ d'un point $x \in E$ comme l'ensemble des images $\beta(g)x$ pour $g \in G$.

2. Montrer que l'appartenance à une même orbite est une relation d'équivalence.

3. Exemple : action du groupe $O(n)$ sur l'espace \mathbb{R}^n . Que sont les orbites ?

4. Un espace est homogène s'il n'a qu'une seule orbite. Exemple trivial : \mathbb{R}^n sous l'action des translations.

Plus généralement, qu'en est-il de l'action à gauche de G sur lui-même, avec $E = G$? Donner d'autres exemples d'espaces homogènes pour $G = O(3)$ ou $\mathcal{L} = O(1,3)$.

5. On définit aussi le *groupe d'isotropie*, (appelé aussi *stabilisateur*, ou, par les physiciens, *petit groupe*) $S(x)$ de l'élément $x \in E$: c'est le sous-groupe de G laissant x invariant :

$$S(x) = \{g \in G | \beta(g)x = x\} .$$

Montrer que si x et y appartiennent à la même orbite, leurs groupes d'isotropie sont conjugués. Quel est le groupe d'isotropie d'un point $x \in \mathbb{R}^n$ sous l'action de $SO(n)$? d'un vecteur de genre temps p dans l'espace de Minkowski ?

Le stabilisateur $S(x)$ est-il un sous-groupe invariant ?

6. Montrer qu'il existe une bijection entre les points de l'orbite $O(x)$ et l'ensemble quotient $G/S(x)$. Pour un groupe fini G , en déduire une relation entre les ordres de G , de $O(x)$ et de $S(x)$. Cet ensemble $G/S(x)$ est-il un espace homogène pour l'action de G ?

Le sujet du chapitre 2 porte sur le cas particulier où E est un espace vectoriel avec comme bijections les transformations linéaires du groupe $GL(E)$: on parle alors de représentations du groupe G dans E .

B. Groupes et algèbres de Lie de dimension 3

1. Rappeler la définition du groupe $SU(1,1)$. Quelle est sa dimension ?

2. Quelle équation définit son algèbre de Lie ? Quelle conséquence cela implique-t-il sur les éléments de matrices de $X \in \mathfrak{su}(1,1)$? Montrer qu'on peut écrire une base de $\mathfrak{su}(1,1)$ en termes des 3 matrices de Pauli et en calculer les relations de commutation. Cette algèbre est-elle isomorphe à l'algèbre de $\mathfrak{so}(3)$?

3. On considère maintenant le groupe linéaire réel $SL(2, \mathbb{R})$. Quelle est sa définition ? Comment son algèbre de Lie est-elle définie ? En donner une base en termes de matrices de Pauli.

4. Montrer l'isomorphisme des deux algèbres $\mathfrak{su}(1,1)$ et $\mathfrak{sl}(2, \mathbb{R})$.

5. Mêmes questions avec l'algèbre $\mathfrak{so}(2,1)$: définition, dimension, relations de commutation, isomorphisme avec l'une des précédente ?

6. En utilisant les critères de Cartan, discuter la semi-simplicité et la compacité de ces différentes algèbres. Quelle est leur complexifiée et leur relation avec $\mathfrak{su}(2)$?

(Pour la relation géométrique entre les groupes $SU(1,1)$, $SL(2, \mathbb{R})$ et $SO(1,2)$), cf. le §13, vol. 1 de [DNF].

C. Opérateurs de Casimir de $\mathfrak{u}(n)$.

1. Montrer que les n^2 matrices $t_{(ij)}$, $1 \leq i, j \leq n$, matrices $n \times n$ d'éléments $(t_{(ij)})_{ab} = \delta_{ia}\delta_{jb}$ forment une base de l'algèbre $\mathfrak{u}(n)$. En calculer les relations de commutation et les constantes de structure.

2. Calculer la forme de Killing dans cette base et en vérifier les propriétés liées aux critères de Cartan.

3. Montrer que les éléments de l'algèbre enveloppante $C^{(r)} = \sum_{1 \leq i_1, i_2, \dots, i_r \leq n} t_{(i_1 i_2)} t_{(i_2 i_3)} \dots t_{(i_r i_1)}$ commutent avec tous les $t_{(ij)}$ et sont donc des opérateurs de Casimir de degré r .

4. Comment modifier ce qui précède pour l'algèbre $\mathfrak{su}(n)$? ([Bu], chap 10)

Problème : Transformations conformes

I-1. On rappelle que dans une théorie (classique) des champs locale invariante par translations, on sait définir un tenseur énergie-impulsion $\Theta_{\mu\nu}(x)$ tel que

- sous l'effet d'un changement de coordonnées infinitésimal $x^\mu \rightarrow x'^\mu = x^\mu + a^\mu(x)$, l'action subit une variation

$$\delta S = \int d^d x (\partial_\mu a_\nu) \Theta^{\mu\nu}(x); \quad (1.49)$$

- $\Theta_{\mu\nu}$ est conservé : $\partial_\mu \Theta^{\mu\nu}(x) = 0$;
- on suppose $\Theta_{\mu\nu}$ symétrique en μ, ν .

Montrer que si Θ est en outre de trace nulle, $\Theta_\mu^\mu = 0$, l'action est aussi invariante sous l'effet des dilatations, $x^\mu \rightarrow x'^\mu = (1 + \delta\lambda)x^\mu$.

2. Dans un espace riemannien ou pseudo-riemannien de dimension d , doté d'une métrique $g_{\mu\nu}(x)$ de signature $\{(+1)^p, (-1)^{d-p}\}$, on appelle transformation conforme une transformation des coordonnées $x^\mu \rightarrow x'^\mu$ qui dilate localement les longueurs

$$ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu \rightarrow ds'^2 = g_{\mu\nu}(x') dx'^\mu dx'^\nu = \alpha(x) ds^2 \quad (1.50)$$

a) Écrire la forme infinitésimale de cette condition, quand $x^\mu \rightarrow x'^\mu = x^\mu + a^\mu(x)$. (On reliera le paramètre de dilatation $1 + \delta\alpha$ à a^μ en prenant une trace adéquate.)

b) Montrer que pour un espace euclidien ou pseudo-euclidien de métrique $g_{\mu\nu} = \text{diag}\{(+1)^p, (-1)^{d-p}\}$, la condition se ramène à

$$\partial_\mu a_\nu + \partial_\nu a_\mu = \frac{2}{d} g_{\mu\nu} \partial_\rho a^\rho. \quad (1.51)$$

3. Montrer en utilisant (1.49), (1.51) que sous les conditions du 1. et du 2.b, toute théorie invariante par translations, rotations et dilatations l'est aussi sous l'effet des transformations conformes.

4. On va maintenant étudier les conséquences de (1.51). On pose $D := \frac{1}{d} \partial_\rho a^\rho$.

- a) En dérivant (1.51) par rapport à x^ν , montrer que

$$\partial^2 a_\mu = (2 - d) \partial_\mu D. \quad (1.52)$$

- b) En dérivant (1.52) par rapport à x^μ , montrer qu'en dimension $d > 1$, D est une fonction harmonique : $\partial^2 D = 0$.

- c) On suppose dans la suite que $d \geq 2$. En dérivant (1.52) par rapport à x^ν , en symétrisant en μ et ν et en utilisant (1.51), montrer que si $d > 2$, alors $\partial_\mu \partial_\nu D = 0$. En conclure qu'il existe une constante h et un vecteur constant k tel que $D = k_\mu x^\mu + h$.

- d) En dérivant (1.51) par rapport à x^σ et en antisymétrisant en ν et σ , montrer que

$$\partial_\mu (\partial_\sigma a_\nu - \partial_\nu a_\sigma) = 2(g_{\mu\nu} k_\sigma - g_{\mu\sigma} k_\nu) = \partial_\mu (2k_\sigma x_\nu - 2k_\nu x_\sigma). \quad (1.53)$$

- e) En déduire qu'il existe un tenseur antisymétrique constant $l_{\sigma\nu}$ tel que

$$\partial_\sigma a_\nu - \partial_\nu a_\sigma = (2k_\sigma x_\nu - 2k_\nu x_\sigma) + 2l_{\sigma\nu}, \quad (1.54)$$

qui combiné avec (1.51) donne

$$\partial_\sigma a_\nu = x_\nu k_\sigma - x_\sigma k_\nu + l_{\sigma\nu} + g_{\nu\sigma} k_\rho x^\rho + h g_{\nu\sigma}.$$

- f) En conclure que l'expression générale d'une transformation conforme infinitésimale en dimension $d > 2$ s'écrit

$$a_\nu = k_\sigma x^\sigma x_\nu - \frac{1}{2} x_\sigma x^\sigma k_\nu + l_{\sigma\nu} x^\sigma + h x_\nu + c_\nu \quad (1.55)$$

avec c un vecteur constant⁷. De combien de paramètres réels indépendants dépend une telle transformation en dimension d ?

7. Ce joli raisonnement est dû à Michel Bauer.

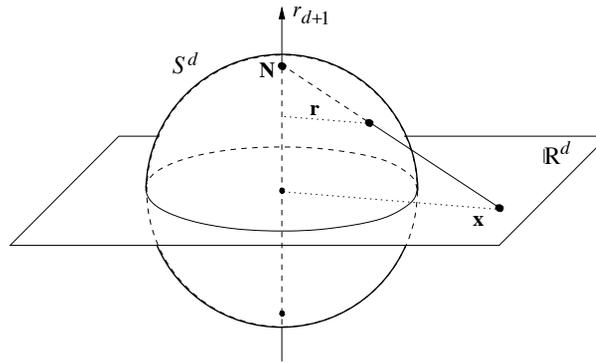


FIGURE 1.5 – Projection stéréographique depuis le pôle Nord

II-1. On apprend que dans l'espace (pseudo-)euclidien de dimension $d > 2$, les transformations conformes sont engendrées par les translations, les rotations, les dilatations et “les transformations conformes spéciales”, obtenues en composant une inversion $x^\mu \rightarrow x^\mu/x^2$, une translation et à nouveau une inversion. Écrire la forme finie puis la forme infinitésimale de ces transformations conformes spéciales, et vérifier que le résultat est bien en accord avec (1.55), ce qui justifie l’assertion précédente.

2. Écrire l’expression des générateurs infinitésimaux P_μ des translations, $J_{\mu\nu}$ des rotations, D des dilatations et K_μ des transformations spéciales, comme opérateurs différentiels en x .

3. Écrire avec le minimum de calculs les relations de commutation de ces générateurs (on utilisera les résultats déjà connus sur les générateurs P_μ et $J_{\mu\nu}$ et on tirera profit de l’homogénéité et de la définition des transformations conformes spéciales pour réduire le seul calcul non trivial à celui de $[K_\mu, P_\nu]$). Vérifier que des relations de commutation se ferment bien sur les générateurs P, J, D et K .

4. Quelle est la dimension du groupe conforme dans l’espace euclidien \mathbb{R}^d ?

III-1. Pour mieux comprendre la nature du groupe conforme, on applique l’espace \mathbb{R}^d , complété du point à l’infini et doté de sa métrique $\mathbf{x}^2 = x_1^2 + \dots + x_d^2$, sur la sphère S^d . Cette sphère est définie par l’équation $\mathbf{r}^2 + r_{d+1}^2 = 1$ dans l’espace \mathbb{R}^{d+1} , et l’application est réalisée grâce à la projection stéréographique à partir du “pôle Nord” $\mathbf{r} = 0, r_{d+1} = 1$ (voir Fig. 1.5). Montrer que l’on a

$$\mathbf{r} = \frac{2\mathbf{x}}{\mathbf{x}^2 + 1} \quad r_{d+1} = \frac{\mathbf{x}^2 - 1}{\mathbf{x}^2 + 1} .$$

Quelle est l’image du point à l’infini ? Quel est l’effet de l’inversion dans \mathbb{R}^d sur le point $r = (\mathbf{r}, r_{d+1}) \in S^d$?

2. La sphère précédente est à son tour considérée comme la section du cône de lumière \mathcal{C} dans l’espace de Minkowski $\mathcal{M}_{1,d+1}$ de métrique $z_0^2 - \mathbf{z}^2 - z_{d+1}^2 = 0$ par l’hyperplan $z_0 = 1$. Montrer que de cette façon on a une correspondance biunivoque entre les points de $\mathbb{R}^d \cup \{\infty\}$ et les rayons du cône de lumière (c’est-à-dire les vecteurs à une dilatation près) et que l’expression de $\mathbf{x} \in \mathbb{R}^d$ en fonction de $z = (z_0, \mathbf{z}, z_{d+1}) \in \mathcal{C}$ est

$$\mathbf{x} = \frac{\mathbf{z}}{z_0 - z_{d+1}} .$$

3. On va montrer maintenant que l’action du groupe conforme dans \mathbb{R}^d découle de transformations linéaires dans $\mathcal{M}_{1,d+1}$ préservant le cône de lumière. Sans aucun calcul, montrer que ces transformations doivent alors appartenir au groupe de Lorentz de $\mathcal{M}_{1,d+1}$, soit $O(1, d + 1)$.

a) Identifier les transformations linéaires de z correspondant aux rotations de \mathbf{x} dans \mathbb{R}^d . Montrer que les dilatations de x correspondent à des “boosts” de rapidité β dans le plan (z_0, z_{d+1}) , en donnant la relation entre le paramètre de dilatation et la rapidité.

b) On considère ensuite les transformations de $O(1, d + 1)$ qui préservent $z_0 - z_{d+1}$. Écrire la matrice T_a d’une telle transformation infinitésimale agissant sur les coordonnées $(z_0, \mathbf{z}, z_{d+1})$ telle que $\delta\mathbf{z} = \mathbf{a}(z_0 - z_{d+1})$

(au premier ordre en \mathbf{a}). A quelle transformation de $\mathbf{x} \in \mathbb{R}^d$ correspond-elle? Calculer par exponentiation de T_a la matrice d'une transformation finie (on pourra par exemple calculer les premières puissances $T_a^2, T_a^3 \dots$).

c) Quelle est enfin l'interprétation de l'inversion de \mathbb{R}^d dans le groupe de Lorentz de $\mathcal{M}_{1,d+1}$? Que dire des transformations conformes spéciales? Quelle est la dimension du groupe $O(1, d+1)$? Qu'en conclure sur la relation entre le groupe de Lorentz dans l'espace de Minkowski $\mathcal{M}_{1,d+1}$ et le groupe conforme dans \mathbb{R}^d ?

IV. Question subsidiaire : Connaissez-vous des transformations conformes de l'espace \mathbb{R}^2 autres que celles mentionnées au **II.1**?

Chapitre 2

Représentations linéaires des groupes

La question de l'action d'un groupe dans un ensemble a déjà été évoquée au chapitre précédent (Exercice A et TD). On va s'intéresser maintenant plus particulièrement à l'action *linéaire* d'un groupe dans un espace vectoriel. Cette situation est rencontrée fréquemment en géométrie et en physique (mécanique quantique, mécanique statistique, théorie des champs, ...). Il faut cependant garder à l'esprit que d'autres actions de groupe peuvent aussi avoir un intérêt physique : ainsi le groupe des rotations $SO(n)$ agit sur la sphère S^{n-1} de façon non linéaire, et cela apparaît par exemple dans des modèles de ferromagnétisme et des théories de champs dites modèles σ non linéaires, cf cours de F. David.

2.1 Définitions et propriétés générales.

2.1.1 Définitions de base

On dit qu'un groupe G est représenté dans un espace vectoriel E (sur un corps qui pour nous sera toujours \mathbb{R} ou \mathbb{C}), ou encore, que E porte une représentation de G , si on a un homomorphisme D du groupe G dans le groupe des transformations linéaires $GL(E)$:

$$\begin{aligned}\forall g \in G \quad g &\mapsto D(g) \in GL(E) \\ \forall g, g' \in G \quad D(g.g') &= D(g).D(g') \\ D(e) &= I \\ \forall g \in G \quad D(g^{-1}) &= (D(g))^{-1}\end{aligned}\tag{2.1}$$

où I désigne l'opérateur identité dans $GL(E)$. Si l'espace de représentation est de dimension p , la représentation est dite elle-même de dimension p . La représentation qui à tout $g \in G$ associe 1 (considéré comme $\in GL(\mathbb{R})$) est appelée *triviale* ou *représentation identité*; elle est de dimension 1.

Si G est un groupe topologique, resp. un groupe de Lie, on demandera aussi à l'application $g \mapsto D(g)$ d'être continue, resp. différentiable. Dans la suite de ces notes, on supposera toujours ces conditions satisfaites.

La représentation est dite *fidèle* si $\ker D = \{e\}$, ou encore si $D(g) = D(g') \Leftrightarrow g = g'$. Sinon, le noyau de l'homomorphisme est un sous-groupe invariant H , et la représentation du groupe G/H dans E est fidèle (le vérifier). En conséquence, toute représentation non triviale d'un groupe simple est fidèle. Inversement, si G a un sous-groupe invariant H , toute représentation de G/H fournit une représentation dégénérée (= non fidèle) de G .

Si E est de dimension finie p , on peut choisir une base e_i , $i = 1, \dots, p$, et associer à tout $g \in G$ la matrice représentative de $D(g)$, notée avec une lettre calligraphiée

$$D(g)e_j = e_i \mathcal{D}_{ij}(g) \quad (2.2)$$

avec, comme (presque) toujours dans ces notes, la convention de sommation sur les indices répétés. La disposition des indices (i : indice de ligne, j indice de colonne) est dictée par la loi (2.1). En effet, on a bien

$$\begin{aligned} D(g.g')e_k &= e_i \mathcal{D}_{ik}(g.g') \\ &= D(g)(D(g')e_k) = D(g)e_j \mathcal{D}_{jk}(g') \\ &= e_i \mathcal{D}_{ij}(g) \mathcal{D}_{jk}(g') \\ \text{hence } \mathcal{D}_{ik}(g.g') &= \mathcal{D}_{ij}(g) \mathcal{D}_{jk}(g') . \end{aligned} \quad (2.3)$$

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

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (2.4)$$

qui décrivent bien les rotations d'angle θ autour de l'origine.

Le groupe $SU(3)$ est défini comme l'ensemble des matrices U unitaires, 3×3 et de déterminant 1. Ces matrices forment elles-mêmes une représentation de $SU(3)$, c'est la "représentation de définition". Montrer que les matrices U^* complexes conjuguées forment aussi une représentation de $SU(3)$.

De quel groupe les matrices $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ forment-elles une représentation ?

2.1.2 Représentations équivalentes. Caractères

Soient D et D' deux représentations de G dans des espaces E et E' , supposons qu'il existe un opérateur linéaire V de E dans E' tel que

$$\forall g \in G \quad VD(g) = D'(g)V . \quad (2.5)$$

Un tel V est dit *opérateur d'entrelacement* ("intertwiner" en anglais). Si V est inversible (et donc E et E' ont même dimension, si elle est finie), on dit que les représentations D et D' sont *équivalentes*. (C'est une relation d'équivalence entre représentations!).

Dans le cas de dimension finie, où on identifie E et E' , on voit que les matrices représentatives de D et D' sont reliées par une transformation de similitude et peuvent être considérées comme différant par un changement de base. Il n'y a donc pas lieu de distinguer fondamentalement

deux représentations équivalentes, et en “théorie de la représentation” on va s’attacher à étudier les représentations inéquivalentes.

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

$$\chi(g) = \text{tr } D(g) . \quad (2.6)$$

C’est une fonction de G dans \mathbb{R} ou \mathbb{C} qui satisfait les propriétés suivantes (les vérifier) :

- Le caractère est indépendant du choix de base dans E .
- Deux représentations équivalentes ont le même caractère.
- Le caractère prend la même valeur pour les différents éléments d’une même classe¹ de G : on dit que le caractère est une *fonction de classe* : $\chi(g) = \chi(hgh^{-1})$.

La réciproque de cette dernière propriété, à savoir une fonction de classe peut-elle s’exprimer en termes des caractères, est vraie pour tout groupe fini ; elle l’est aussi pour tout groupe de Lie compact et toute fonction continue (ou de la classe de carré sommable) sur G : c’est l’objet du théorème de Peter–Weyl, voir plus bas § 2.3.1.

On notera encore que le caractère, évalué pour l’élément identité du groupe, fournit la dimension de la représentation

$$\chi(e) = \dim D . \quad (2.7)$$

2.1.3 Représentations réductibles et irréductibles

Un autre type de redondance est lié à la somme directe des représentations. Supposons qu’on a construit deux représentations D_1 et D_2 de G dans deux espaces E_1 et E_2 . On peut alors construire une représentation dans l’espace somme directe $E = E_1 \oplus E_2$ et la représentation est dite *somme directe des représentations* D_1 et D_2 et notée $D_1 \oplus D_2$. (Rappelons que tout vecteur de $E_1 \oplus E_2$ peut s’écrire de façon unique comme combinaison linéaire d’un vecteur de E_1 et d’un vecteur de E_2). Les sous-espaces E_1 et E_2 de E sont bien sûr laissés invariants par l’action of $D_1 \oplus D_2$.

Inversement, si une représentation de G dans un espace E laisse invariant un sous-espace de E , elle est dite *réductible*. Dans le cas contraire, elle est *irréductible*. Si D est réductible et laisse le sous-espace E_1 invariant, et aussi son sous-espace supplémentaire E_2 , on dit que la représentation est *complètement réductible* (on dit aussi *décomposable*) ; on peut alors considérer E comme somme directe de E_1 et E_2 et la représentation comme somme directe des représentations dans E_1 et E_2 .

Quand on s’intéresse à un groupe topologique, resp. de Lie, il est préférable d’ajouter dans la définition de réductibilité d’une représentation la condition que le sous-espace invariant est fermé, ou une condition de même nature, en accord avec la topologie. Cela sera considéré comme implicite dans la suite.

Si E est de dimension finie, ceci se traduit simplement sur les matrices de la représentation qui prennent la forme suivante (dans une base adaptée à la décomposition !) en termes de blocs

1. Les classes du groupe G dont il s’agit ici sont celles résultant de la relation d’équivalence dans G : $g \sim g' \Leftrightarrow \exists h \in G : g' = hgh^{-1}$. Ne pas confondre ces classes avec celles (“cosets”) liées à un sous-groupe.

de dimensions $\dim E_1$ et $\dim E_2$

$$\forall g \in G \quad \mathcal{D}(g) = \begin{pmatrix} \mathcal{D}_1(g) & 0 \\ 0 & \mathcal{D}_2(g) \end{pmatrix}. \quad (2.8)$$

Si la représentation est réductible sans être complètement réductible, (*représentation indécomposable*), sa matrice prend la forme, dans une base constituée par une base de E_1 et d'une base d'un sous-espace supplémentaire

$$\mathcal{D}(g) = \begin{pmatrix} \mathcal{D}_1(g) & \mathcal{D}'(g) \\ 0 & \mathcal{D}_2(g) \end{pmatrix}. \quad (2.9)$$

C'est le cas des représentations du groupe des translations à une dimension. La représentation

$$\mathcal{D}(a) = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \quad (2.10)$$

est réductible, puisqu'elle laisse invariants les vecteurs $(X, 0)$ mais n'a pas de sous-espace supplémentaire invariant.

En revanche, si la représentation réductible de G dans E laisse invariant le sous-espace E_1 , il existe une représentation dans le sous-espace $E_2 = E/E_1$. Dans les notations de l'équ. (2.9), sa matrice représentative est $\mathcal{D}_2(g)$.

Il faut encore souligner l'importance du corps de base dans la discussion de l'irréductibilité. C'est ainsi que la représentation (2.4) qui est irréductible sur un espace vectoriel sur \mathbb{R} ne l'est pas sur \mathbb{C} : on peut au prix d'un changement de base la récrire comme

$$\begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix}. \quad (2.11)$$

2.1.4 Représentations conjuguée et contragrédiente

Etant donnée une représentation D , soit \mathcal{D} sa matrice dans une certaine base, les matrices \mathcal{D}^* complexes conjuguées forment une autre représentation D^* dite conjuguée, puisqu'elles satisfont bien (2.3)

$$\mathcal{D}_{ik}^*(g.g') = \mathcal{D}_{ij}^*(g)\mathcal{D}_{jk}^*(g').$$

La représentation D est dite *réelle* s'il existe une base où $\mathcal{D} = \mathcal{D}^*$. Cela implique que son caractère χ est réel. Réciproquement si χ est réel, la représentation D est équivalente à sa conjuguée D^* ². Si les représentations D et D^* sont équivalentes mais qu'il n'existe pas de base où $\mathcal{D} = \mathcal{D}^*$, les représentations sont dites *pseudoréelles*. (C'est par exemple le cas de la représentation de spin $\frac{1}{2}$ de $SU(2)$.) Voir le problème III à la fin du chapitre pour une définition plus canonique de cette notion de représentation réelle ou pseudoréelle.

2. Ceci est vrai au moins pour les représentations irréductibles des groupes finis ou continus compacts pour lesquelles on verra plus bas (§ 2.3) que deux représentations irréductibles non équivalentes ne peuvent avoir le même caractère.

Ce concept joue un rôle dans le contrôle de l’“anomalie chirale non-singlet” dans les théories de jauge : si les fermions appartiennent à une représentation réelle ou pseudoréelle du groupe de jauge, ils n’induisent pas d’anomalie. Voir le chapitre 5 de ce cours.

La représentation *contragrédiente* de D est définie quant à elle par

$$\bar{D}(g) = D^{-1T}(g)$$

c’est-à-dire $\bar{D}_{ij}(g) = \mathcal{D}_{ji}(g^{-1})$, qui satisfait bien aussi (2.3). Pour une représentation unitaire, cf alinéa suivant, $\bar{D}_{ij}(g) = \mathcal{D}_{ij}^*(g)$, la contragrédiente s’identifie à la conjuguée. Les représentations D , D^* et \bar{D} sont simultanément réductibles ou irréductibles.

2.1.5 Représentations unitaires

Supposons que l’espace vectoriel E est “préhilbertien”, c’est-à-dire possède un produit scalaire (forme $J(x, y) = \langle x|y \rangle = \langle y|x \rangle^*$ bilinéaire symétrique sur \mathbb{R} , ou sesquilinéaire sur \mathbb{C}), tel que la norme est définie positive : $x \neq 0 \Rightarrow \langle x|x \rangle > 0$. Si la dimension de E est finie, on peut trouver une base orthonormale où la matrice de J se réduit à I et y définir des opérateurs unitaires U tels que $U^\dagger U = I$. Si on est dans un espace de dimension infinie, espace que nous supposerons préhilbertien séparable³, on sait qu’on peut trouver une base orthonormale dénombrable, donc indexée par un indice discret. Une représentation de G dans E est dite *unitaire* si pour tout $g \in G$, l’opérateur $D(g)$ est unitaire. On a donc pour tous $g \in G$ et $x, y \in E$

$$\langle x|y \rangle = \langle D(g)x|D(g)y \rangle \quad (2.12)$$

$$\text{donc } D(g)^\dagger D(g) = I \quad (2.13)$$

$$\text{et } D(g^{-1}) = D^{-1}(g) = D^\dagger(g) . \quad (2.14)$$

On a les propriétés importantes suivantes :

(i) *Toute représentation unitaire réductible est complètement réductible* (théorème de Maschke).

En effet soit E_1 un sous-espace invariant, $E_2 = (E_1)_\perp$ son espace supplémentaire est invariant puisque pour tous $g \in G$, $x \in E_1$ et $y \in E_2$ on a

$$\langle x|D(g)y \rangle = \langle D(g^{-1})x|y \rangle = 0 \quad (2.15)$$

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

(ii) *Toute représentation d’un groupe fini ou d’un groupe compact dans un espace doté d’un produit scalaire est “unitarisable”, c’est-à-dire équivalente à une représentation unitaire.*

Preuve. Considérons d’abord le cas d’un groupe fini et formons

$$Q = \sum_{g' \in G} D^\dagger(g')D(g') \quad (2.16)$$

3. Un espace est dit *séparable* s’il contient un sous-ensemble dénombrable dense.

qui satisfait

$$D^\dagger(g)QD(g) = \sum_{g' \in G} D^\dagger(g'.g)D(g'.g) = Q \quad (2.17)$$

où on a remplacé $\sum_{g'}$ par $\sum_{g'.g}$ (“lemme de réarrangement”), (cf § 1.2.4) L’opérateur auto-adjoint Q est défini positif, (pourquoi?), on peut donc l’écrire sous la forme

$$Q = V^\dagger V \quad (2.18)$$

avec V inversible. (Par exemple, la diagonalisation de l’opérateur auto-adjoint Q par un opérateur unitaire, $Q = U\Lambda^2U^\dagger$, avec Λ diagonal réel, permet d’en extraire la “racine carrée” $V = U\Lambda U^\dagger$.) L’entrelaceur V définit une représentation D' équivalente à D et unitaire :

$$\begin{aligned} D'(g) &= VD(g)V^{-1} \\ D'^\dagger(g)D'(g) &= V^{\dagger-1}D^\dagger(g)V^\dagger VD(g)V^{-1} \\ &= V^{\dagger-1}D^\dagger(g)QD(g)V^{-1} \stackrel{(2.17)}{=} V^{\dagger-1}QV^{-1} = I. \end{aligned} \quad (2.19)$$

Dans le cas d’un groupe continu compact, l’existence de la mesure invariante de Haar (cf § 1.2.4) permet de répéter l’argument avec $Q = \int d\mu(g')D^\dagger(g')D(g')$. cqfd

Comme corollaire des deux propriétés précédentes, toute représentation réductible d’un groupe fini ou d’un groupe compact sur un espace préhilbertien est (équivalente à) une représentation unitaire et complètement réductible. Il s’agit donc pour nous de construire et de classer les représentations unitaires irréductibles. On va montrer plus bas que, pour un groupe fini ou compact, ces représentations irréductibles sont de dimension finie.

Contre-exemple dans le cas d’un groupe non compact : les matrices $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ forment une représentation indécomposable (non complètement réductible).

2.1.6 Lemme de Schur

Soit deux représentations irréductibles D dans E et D' dans E' et un opérateur d’entrelacement entre elles, comme défini en (2.5). On a alors l’important

Lemme de Schur : ou bien $V = 0$, ou bien V est une bijection et les représentations sont équivalentes.

Preuve : Supposons $V \neq 0$. Alors $VD(g) = D'(g)V$ implique que le noyau de V est un sous-espace de E invariant par D ; par l’hypothèse d’irréductibilité, il se réduit donc à 0 (il ne peut être égal à E tout entier sans quoi V serait nul). De même, l’image de V est un sous-espace de E' invariant par D' , il ne peut être nul et est donc identique à E' . Des théorèmes classiques sur les applications linéaires entre espaces vectoriels, il découle que V est une bijection de E dans E' et que les représentations sont donc équivalentes. c.q.f.d.

N.B. Si les deux représentations ne sont pas irréductibles, c’est bien sûr faux en général. Un contre-exemple est fourni par la représentation (2.10) qui commute avec les matrices $V = \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix}$.

Corollaire 1. *Tout opérateur d'entrelacement d'une représentation irréductible sur le corps \mathbb{C} avec elle-même, c'est-à-dire tout opérateur commutant avec tous les représentants du groupe, est un multiple de l'identité.*

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

Corollaire 2. *Une représentation irréductible sur \mathbb{C} d'un groupe abélien est nécessairement de dimension 1.*

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

Insistons sur l'importance du caractère algébriquement clos de \mathbb{C} , par opposition à \mathbb{R} , dans ces deux corollaires. La représentation sur \mathbb{R} du groupe $SO(2)$ par les matrices $\mathcal{D}(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ vient fournir des contrexemples aux deux propositions précédentes : toute matrice $\mathcal{D}(\alpha)$ commute avec $\mathcal{D}(\theta)$ mais n'a pas de valeur propre réelle (si $\alpha \neq 0, \pi$) et la représentation est irréductible sur \mathbb{R} , quoique de dimension deux.

Application du Corollaire 1 : dans l'algèbre de Lie d'un groupe de Lie, les opérateurs de Casimir quadratiques définis à la fin du Chap. 1 commutent avec tous les générateurs infinitésimaux et donc avec tous les éléments du groupe. Anticipant un peu sur la discussion à venir des représentations d'une algèbre de Lie, dans une représentation unitaire on peut choisir ces opérateurs de Casimir hermitiens, donc diagonalisables, ce qui permet d'appliquer le raisonnement du Corollaire 1 : dans toute représentation irréductible, ils sont multiples de l'identité. Ainsi pour $SU(2)$, $\mathbf{J}^2 = j(j+1)I$ dans la représentation de spin j .

2.1.7 Produit tensoriel de représentations ; décomposition de Clebsch-Gordan

Produit tensoriel de représentations

Une méthode couramment utilisée pour construire des représentations irréductibles d'un groupe donné consiste à construire le produit tensoriel de représentations connues et à le décomposer en représentations irréductibles. C'est aussi la situation qu'on rencontre en Mécanique Quantique, quand on connaît la transformation des composantes d'un système et qu'on étudie comment le système composé se transforme (système de deux particules de spin j_1 et j_2 par exemple).

Soient E_1 et E_2 deux espaces vectoriels portant des représentations D_1 et D_2 d'un groupe G . L'espace produit tensoriel⁴ $E = E_1 \otimes E_2$ est l'espace engendré par les combinaisons linéaires de "produits" (tensoriels) d'un élément de E_1 et d'un élément de E_2 : $z = \sum_i x^{(i)} \otimes y^{(i)}$. L'espace E porte lui même une représentation, notée $D = D_1 \otimes D_2$, *produit tensoriel* (ou *produit direct*) des représentations D_1 et D_2 . (Voir au Chap. 0 l'exemple du groupe $SU(2)$). Sur l'élément z ci-dessus

$$D(g)z = \sum_i D_1(g)x^{(i)} \otimes D_2(g)y^{(i)}. \quad (2.20)$$

4. On trouvera à l'Appendice D un petit rappel sur les produits tensoriels et les tenseurs.

On vérifie immédiatement que le caractère de la représentation D est le produit des caractères χ_1 et χ_2 de D_1 et D_2

$$\chi(g) = \chi_1(g)\chi_2(g) \quad (2.21)$$

En particulier en évaluant cette relation pour $g = e$, on a pour des représentations de dimension finie

$$\dim D = \dim(E_1 \otimes E_2) = \dim E_1 \cdot \dim E_2 = \dim D_1 \cdot \dim D_2 \quad (2.22)$$

comme il est bien connu pour un produit tensoriel.

Décomposition de Clebsch-Gordan

La représentation produit direct de deux représentations irréductibles D et D' n'est en général pas irréductible. Si elle est complètement réductible (comme c'est le cas pour les représentations unitaires qui vont nous intéresser au premier chef), on effectue la *décomposition de Clebsch-Gordan* en représentations irréductibles

$$D \otimes D' = \bigoplus_j D_j \quad (2.23)$$

où au second membre apparaissent un certain nombre de représentations irréductibles D_1, \dots . La notation \bigoplus_j recouvre des situations très variées : sommation sur un ensemble fini (groupes finis), sur un sous-ensemble fini d'un ensemble a priori infini mais discret (groupes compacts) ou sur des variables éventuellement continues (groupes non compacts).

Si G est fini ou compact et si ses représentations irréductibles inéquivalentes ont été classées et indexées : $D^{(\rho)}$, on peut préférer à (2.23) une autre écriture qui indique lesquelles de ces représentations inéquivalentes apparaissent, *et avec quelle multiplicité*

$$D \otimes D' = \bigoplus_{\rho} m_{\rho} D^{(\rho)} . \quad (2.24)$$

Une écriture plus correcte serait $E \otimes E' = \bigoplus_{\rho} F_{\rho} \otimes E^{(\rho)}$ où F_{ρ} est un espace vectoriel de dimension m_{ρ} , l'«espace de multiplicité».

Les entiers m_{ρ} sont non négatifs. Les équations (2.23) et (2.24) impliquent des règles simples sur les caractères et les dimensions

$$\chi_D \cdot \chi_{D'} = \sum_j \chi_j = \sum_{\rho} m_{\rho} \chi^{(\rho)} \quad (2.25)$$

$$\dim D \cdot \dim D' = \sum_j \dim D_j = \sum_{\rho} m_{\rho} \dim D^{(\rho)} . \quad (2.26)$$

Exemple : le produit tensoriel de deux copies de l'espace euclidien de dimension 3 ne forme pas une représentation irréductible du groupe des rotations $SO(3)$. Cet espace est engendré par les produits tensoriels de vecteurs \vec{x} et \vec{y} et on sait construire le produit scalaire $\vec{x} \cdot \vec{y}$ qui est invariant par le groupe (représentation triviale), un tenseur antisymétrique à deux indices

$$A_{ij} = x_i y_j - x_j y_i$$

qui se transforme comme une représentation irréductible de dimension 3 (de spin 1),⁵ et un tenseur symétrique de trace nulle

$$S_{ij} = x_i y_j + x_j y_i - \frac{2}{3} \delta_{ij} \vec{x} \cdot \vec{y}$$

qui se transforme selon une représentation irréductible de dimension 5 (spin 2); on peut donc décomposer tout tenseur $x_i y_j$ selon

$$x_i y_j = \frac{1}{3} \delta_{ij} \vec{x} \cdot \vec{y} + \frac{1}{2} A_{ij} + \frac{1}{2} S_{ij};$$

le total des dimensions est bien sûr de $9 = 3 \times 3 = 1 + 3 + 5$ et en repérant dans ce cas simple les représentations par leur dimension, on écrit

$$D^{(3)} \otimes D^{(3)} = D^{(1)} \oplus D^{(3)} \oplus D^{(5)}. \quad (2.27)$$

Ou encore, dans la notation en “spins”

$$(1) \otimes (1) = (0) \oplus (1) \oplus (2)$$

où on reconnaît bien sûr les règles familières d’“addition du moment angulaire” (voir Chap. 0)

$$(j) \otimes (j') = \bigoplus_{j''=|j-j'|}^{j+j'} (j''). \quad (2.28)$$

En itérant, on trouve de même

$$D^{(3)} \otimes D^{(3)} \otimes D^{(3)} = D^{(1)} \oplus 3D^{(3)} \oplus 2D^{(5)} \oplus D^{(7)}, \quad (2.29)$$

avec cette fois des multiplicités.

Invariants.

Un problème rencontré fréquemment consiste à compter le nombre d’invariants (par l’action d’un groupe G) indépendants, obtenus en “combinant” (par produit tensoriel) des quantités ayant des transformations prescrites (c’est-à-dire se transformant selon des représentations données de G). C’est précisément l’information contenue dans les décompositions en représentations irréductibles comme (2.24, 2.27, 2.29), où la multiplicité de la représentation identité fournit bien ce nombre d’invariants dans le produit tensoriel des représentations considérées. Exercice : interpréter en termes d’invariants géométriques classiques les multiplicités $m_0 = 1, 1, 3$ de la représentation identité apparaissant dans les représentations $(1) \otimes (1)$, $(1) \otimes (1) \otimes (1)$, $(1) \otimes (1) \otimes (1) \otimes (1)$ de $SO(3)$. On fera grand usage de ce type de considérations au Chap 4 en discutant les amplitudes invariantes par le groupe $SU(3)$. Voir aussi le Problème II à la fin de ce chapitre.

5. (un tel tenseur est “dual” d’un vecteur : $A_{ij} = \epsilon_{ijk} z_k$, $z = x \wedge y$.)

Coefficients de Clebsch-Gordan

La formule (2.23) décrit comment dans une transformation du groupe les matrices de représentation se décomposent en représentations irréductibles. Il est aussi souvent important de savoir comment les vecteurs des représentations concernées se décomposent. Soit $e_\alpha^{(\rho)}$, $\alpha = 1, \dots, \dim D^{(\rho)}$, une base de vecteurs de la représentation ρ . On cherche à développer le produit de deux tels objets, soit $e_\alpha^{(\rho)} \otimes e_\beta^{(\sigma)}$, sur des $e_\gamma^{(\tau)}$. Comme la représentation τ peut intervenir un nombre m_τ de fois, il convient d'introduire un indice supplémentaire $i = 1, \dots, m_\tau$. On écrira

$$e_\alpha^{(\rho)} \otimes e_\beta^{(\sigma)} = \sum_{\tau, \gamma, i} C_{\rho, \alpha; \sigma, \beta | \tau_i, \gamma} e_\gamma^{(\tau_i)}. \quad (2.30)$$

ou encore avec des notations plus familières en Mécanique Quantique

$$|\rho, \alpha; \sigma, \beta\rangle \equiv |\rho\alpha\rangle|\sigma\beta\rangle = \sum_{\tau, \gamma, i} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle |\tau_i \gamma\rangle. \quad (2.31)$$

Les coefficients $C_{\rho, \alpha; \sigma, \beta | \tau_i, \gamma} = \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle$ sont les *coefficients de Clebsch-Gordan*. Au contraire des m_ρ de (2.24), ils n'ont aucune raison d'être entiers, comme on l'a vu au Chapitre 00 dans le cas du groupe des rotations, ni même réels en général. Supposons que les représentations considérées sont unitaires et que les bases ont été choisies orthonormées. Les coefficients de C.-G. qui représentent un changement de base orthonormée dans l'espace $E_1 \otimes E_2$ satisfont donc à des relations d'orthonormalité et de complétude

$$\sum_{\tau, \gamma, i} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle \langle \tau_i \gamma | \rho, \alpha'; \sigma, \beta' \rangle^* = \delta_{\alpha, \alpha'} \delta_{\beta, \beta'} \quad (2.32)$$

$$\sum_{\alpha, \beta} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle \langle \tau'_j \gamma' | \rho, \alpha; \sigma, \beta \rangle^* = \delta_{\tau, \tau'} \delta_{\gamma, \gamma'} \delta_{i, j}. \quad (2.33)$$

Ceci permet d'inverser la relation (2.31) en

$$|\tau_i \gamma\rangle = \sum_{\alpha, \beta} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle^* |\rho, \alpha; \sigma, \beta\rangle \quad (2.34)$$

et justifie la notation

$$\langle \rho, \alpha; \sigma, \beta | \tau_i \gamma \rangle = \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle^* \quad (2.35)$$

$$|\tau_i \gamma\rangle = \sum_{\alpha, \beta} \langle \rho, \alpha; \sigma, \beta | \tau_i \gamma \rangle |\rho, \alpha; \sigma, \beta\rangle. \quad (2.36)$$

Finalement en appliquant une opération du groupe aux deux membres de (2.31) et en utilisant ces relations, on décompose le produit des matrices $\mathcal{D}^{(\rho)}$ et $\mathcal{D}^{(\sigma)}$ de façon tout à fait explicite

$$\mathcal{D}_{\alpha\alpha'}^{(\rho)} \mathcal{D}_{\beta\beta'}^{(\sigma)} = \sum_{\tau, \gamma, \gamma', i} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle^* \langle \tau_i \gamma' | \rho, \alpha'; \sigma, \beta' \rangle \mathcal{D}_{\gamma\gamma'}^{(\tau_i)}. \quad (2.37)$$

On verra plus bas (§ 2.4.4) une application de ces formules au théorème de Wigner-Eckart.

2.1.8 Décomposition d'une représentation d'un groupe en représentations irréductibles d'un sous-groupe

Étant un sous-groupe H d'un groupe G , toute représentation D de G fournit par restriction à H une représentation D' de ce dernier

$$\forall h \in H \quad D'(h) = D(h) . \quad (2.38)$$

Il s'agit là d'une méthode souvent utilisée pour fabriquer des représentations de H , une fois connues celles de G . En général, si D est irréductible (sur G), D' ne l'est pas (sur H), et se pose à nouveau la question de la décomposer en représentations irréductibles. Par exemple, étant donné un sous-groupe fini de $SU(2)$, il s'agit de dresser la liste (finie, comme on verra plus bas) de ses représentations irréductibles à partir de celles de $SU(2)$. Autre exemple rencontré souvent en physique : un groupe de symétrie G est "brisé" en un sous-groupe H ; comment les représentations de G se décomposent-elles en représentations de H ? Exemples : en physique des solides, le groupe $G \subset SO(3)$ de symétrie "ponctuelle" (c'est-à-dire de rotations et réflexions) d'un cristal est brisé en H par un champ extérieur ; en physique des particules, on rencontrera aux chapitres 4 et 5 les cas de $SU(2) \subset SU(3)$; $U(1) \times SU(2) \times SU(3) \subset SU(5)$, etc.

2.2 Représentations des groupes et représentations des algèbres de Lie

2.2.1 Définition. Universalité

La notion de représentation s'applique aussi aux algèbres de Lie.

On appelle représentation d'une algèbre de Lie \mathfrak{g} dans un espace vectoriel E un homomorphisme de \mathfrak{g} dans l'algèbre de Lie des opérateurs linéaires sur l'espace E , c'est-à-dire une application $X \in \mathfrak{g} \mapsto d(X) \in \text{End } E$ qui respecte linéarité et crochet de Lie : $X, Y \in \mathfrak{g}$, $[X, Y] \mapsto d([X, Y]) = [d(X), d(Y)] \in \text{End } V$. Un corollaire de cette définition est que dans toute représentation de l'algèbre, les (représentants des) générateurs satisfont les mêmes relations de commutation. Autrement dit, dans des bases bien choisies, *les constantes de structure sont les mêmes dans toutes les représentations*. Plus précisément, si t_i est une base de \mathfrak{g} , avec $[t_i, t_j] = C_{ij}^k t_k$, et si $T_i = d(t_i)$ est son image par la représentation d

$$[T_i, T_j] = [d(t_i), d(t_j)] = d([t_i, t_j]) = C_{ij}^k d(t_k) = C_{ij}^k T_k .$$

Ainsi des calculs menés dans une représentation particulière mais faisant appel *uniquement aux règles de commutation de l'algèbre de Lie* demeurent valables dans toute représentation. On a vu au Chap. 0, § 0.2.2, une illustration de cette propriété d'universalité.

En revanche, les opérateurs de Casimir prennent des valeurs différentes dans des représentations irréductibles différentes.

En parallèle avec les définitions du § 2.1.1, on définit les notions de représentation fidèle d'une algèbre de Lie (son noyau $\ker d = \{X | d(X) = 0\}$ se réduit à l'élément nul de \mathfrak{g}), de représentation réductible ou irréductible (existence ou non d'un sous-espace invariant), etc.

2.2.2 Représentations d'un groupe de Lie et de son algèbre de Lie

Toute représentation différentiable D de G dans un espace E fournit une application d de l'algèbre de Lie \mathfrak{g} dans l'algèbre des opérateurs sur E . On l'obtient en prenant la forme infinitésimale de $D(g)$, pour $g(t) = I + tX$ (ou $g = e^{tX}$)

$$d(X) := \left. \frac{d}{dt} \right|_{t=0} D(g(t)) , \quad (2.39)$$

ou encore, pour t infinitésimal,

$$D(e^{tX}) = e^{td(X)} . \quad (2.40)$$

Montrons que cette application est bien compatible avec les crochets de Lie, et que c'est donc une représentation de l'algèbre de Lie. Pour cela nous répétons la discussion du chap. 1, § 3.4, pour faire apparaître le commutateur de façon naturelle. Soient $g(t) = e^{tX}$ et $h(u) = e^{uY}$ deux sous-groupes à un paramètre, pour t et u infiniment petits et du même ordre. On a $e^{tX} e^{uY} e^{-tX} e^{-uY} = e^Z$ avec $Z = ut[X, Y] + \dots$, et donc

$$\begin{aligned} e^{d(Z)} = D(e^Z) &= D(e^{tX} e^{uY} e^{-tX} e^{-uY}) = D(e^{tX}) D(e^{uY}) D(e^{-tX}) D(e^{-uY}) \\ &= e^{td(X)} e^{ud(Y)} e^{-td(X)} e^{-ud(Y)} \\ &= e^{ut[d(X), d(Y)] + \dots} , \end{aligned} \quad (2.41)$$

d'où en identifiant les termes dominants, $d([X, Y]) = [d(X), d(Y)]$, ce qu'il fallait démontrer.

○ Ce passage d'une représentation de G à une représentation de \mathfrak{g} s'applique en particulier à une représentation de G qui joue un rôle spécial, la représentation *adjointe* de G dans son algèbre de Lie \mathfrak{g} . Cette représentation est définie par l'action suivante

$$X \in \mathfrak{g} \quad D^{\text{adj}}(g)(X) = gXg^{-1} , \quad (2.42)$$

ce qu'on note $\text{Ad } g X$. (Il faut comprendre le membre de droite de (2.42) soit comme résultant de la dérivation en $t = 0$ de $g e^{tX} g^{-1}$, soit, selon le point de vue généralement adopté dans ces notes, au sens de la multiplication matricielle, les matrices g et X agissant dans le même espace.)

La représentation adjointe de G donne lieu à une représentation de \mathfrak{g} dans l'espace \mathfrak{g} , également appelée représentation adjointe. On l'obtient en prenant la forme infinitésimale de (2.42), formellement $g = I + tY$, ou encore en considérant le sous-groupe à un paramètre engendré par $Y \in \mathfrak{g}$, $g(t) = \exp tY$ et en calculant $\text{Ad } g(t)X = g(t)Xg^{-1}(t) = X + t[Y, X] + O(t^2)$ (cf. chap.1 (3.15)), et donc

$$\left. \frac{d}{dt} \text{Ad } g(t)X \right|_{t=0} = [Y, X] = \text{ad } Y X . \quad (2.43)$$

où on retrouve (et justifie) notre notation ad du chap. 1.

Exercice : montrer que les matrices T_i définies par $(T_i)^j_k = C_{ik}^j$ satisfont les relations de l'algèbre de Lie comme conséquence de l'identité de Jacobi, et forment donc une base de générateurs dans la représentation adjointe.

N.B. À une représentation unitaire de G correspond une représentation de \mathfrak{g} par des opérateurs (ou matrices) anti-hermitien(ne)s. Les physiciens, qui aiment bien les opérateurs hermitiens, incluent en général un “ i ” devant les générateurs infinitésimaux : cf $e^{-i\psi J}$, $[J_a, J_b] = i\epsilon_{abc}J_c$, etc.

◦ Inversement, une représentation d'une algèbre de Lie \mathfrak{g} engendre une représentation de l'unique groupe G connexe et simplement connexe ayant \mathfrak{g} comme algèbre de Lie. Autrement dit si $X \mapsto d(X)$ est une représentation de l'algèbre, $e^X \mapsto e^{d(X)}$ en est une du groupe G . En effet, la formule de BCH étant “universelle”, c'est-à-dire n'impliquant que des combinaisons linéaires de crochets dans l'algèbre de Lie et étant donc insensible à la représentation de \mathfrak{g} , on a :

$$e^X e^Y = e^Z \mapsto e^{d(X)} e^{d(Y)} = e^{d(Z)},$$

ce qui prouve que l'homomorphisme des algèbres de Lie s'intègre en un homomorphisme des groupes au voisinage de l'identité. On démontre enfin qu'un tel homomorphisme infiniment différentiable et local (au voisinage de l'identité) d'un groupe *simplement connexe* G dans un groupe G' (ici, le groupe linéaire $\text{GL}(E)$) s'étend de façon unique en un homomorphisme infiniment différentiable de tout G dans G' . En résumé, il suffit donc pour trouver les représentations (éventuellement unitaires) du groupe G de trouver les représentations par des opérateurs (éventuellement antihermitiques) de son algèbre de Lie \mathfrak{g} .

C'est ce principe fondamental qui a déjà été illustré au Chap. 0 sur les deux cas concrets de $\text{SU}(2)$ et de $\text{SL}(2, \mathbb{C})$.

2.3 Représentations des groupes de Lie compacts

Dans cette section, nous allons nous intéresser aux représentations des groupes compacts sur le corps des complexes \mathbb{C} . La plupart des résultats qu'on va obtenir reposent sur le fait qu'on peut effectuer la sommation sur le groupe avec la mesure de Haar $d\mu(g)$. On comparera à l'occasion avec la situation du cas non compact. Dans toute cette discussion, il est bon d'avoir toujours à l'esprit les deux cas de référence : le groupe compact $\text{U}(1) = \{e^{ix}\}$ avec $x \in \mathbb{R}/2\pi\mathbb{Z}$ (un angle modulo 2π), et le groupe non compact \mathbb{R} , groupe additif des réels. On mentionnera aussi rapidement, “pour mémoire”, le cas des groupes finis, très proche de celui des groupes compacts.

2.3.1 Orthogonalité et complétude

Soit G un groupe compact. Nous admettrons que ses représentations irréductibles inéquivalentes sont indexées par un indice discret, qu'on mettra en position supérieure : $D^{(\rho)}$. Ces représentations sont a priori de dimension finie ou infinie, mais on verra plus bas que la dimension de $D^{(\rho)}$ notée n_ρ est en fait finie ; dans une base finie ou dénombrable, les matrices $\mathcal{D}_{\alpha\beta}^{(\rho)}$ peuvent

être supposées unitaires d'après le résultat du paragraphe 2.1.3. (Au contraire, une représentation générique d'un groupe non compact dépend d'un indice continu. Et on va voir que ses représentations unitaires sont nécessairement de dimension infinie.)

Dans nos deux cas de référence, les représentations irréductibles de $U(1)$ (donc de dimension 1 pour ce groupe abélien) sont telles que $D^{(k)}(x)D^{(k)}(x') = D^{(k)}(x+x')$, elles sont de la forme par $D^{(k)}(x) = e^{ikx}$ avec $k \in \mathbb{Z}$, cette dernière condition garantissant que la représentation est univaluée quand on change de détermination $x \rightarrow x + 2\pi n$. Pour $G = \mathbb{R}$, on peut encore prendre $x \mapsto e^{ikx}$, mais rien ne restreint $k \in \mathbb{C}$, sauf l'unitarité qui force $k \in \mathbb{R}$.

Théorème : *Pour un groupe compact, les matrices $\mathcal{D}_{\alpha\beta}^{(\rho)}$ satisfont les propriétés d'orthogonalité suivantes*

$$\int \frac{d\mu(g)}{v(G)} \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\alpha'\beta'}^{(\rho')*}(g) = \frac{1}{n_\rho} \delta_{\rho\rho'} \delta_{\alpha\alpha'} \delta_{\beta\beta'} \quad (2.44)$$

et leurs caractères satisfont donc

$$\int \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g) \chi^{(\rho')*}(g) = \delta_{\rho\rho'} . \quad (2.45)$$

Dans ces formules, $d\mu(g)$ désigne la mesure de Haar et $v(G) = \int d\mu(g)$ est le "volume du groupe".

Preuve : Soit M une matrice quelconque de dimension $n_\rho \times n_{\rho'}$. Considérons alors la matrice

$$V = \int d\mu(g') \mathcal{D}^{(\rho)}(g') M \mathcal{D}^{(\rho')\dagger}(g') . \quad (2.46)$$

Le membre de gauche de (2.44) est (à un facteur $v(G)$ près) la dérivée par rapport à $M_{\beta\beta'}$ de $V_{\alpha\alpha'}$. Les représentations étant unitaires, $\mathcal{D}^\dagger(g) = \mathcal{D}(g^{-1})$, il est facile, en utilisant l'invariance à gauche de la mesure $d\mu(g') = d\mu(gg')$, de vérifier que V satisfait

$$V \mathcal{D}^{(\rho')}(g) = \mathcal{D}^{(\rho)}(g) V \quad (2.47)$$

pour tout $g \in G$. La matrice V est donc par le lemme de Schur nulle si les représentations ρ et ρ' sont différentes, et un multiple de l'identité si $\rho = \rho'$.

a) Dans le premier cas, en choisissant une matrice M dont le seul élément non nul est $M_{\beta\beta'} = 1$ et en identifiant l'élément de matrice $V_{\alpha\alpha'}$, on obtient la propriété d'orthogonalité (2.44).

b) Si $\rho = \rho'$, choisissons d'abord $M_{11} = 1$, les autres $M_{\beta\beta'}$ nuls. On a $V = c_1 I$, où le coefficient c_1 est obtenu en prenant la trace : $c_1 n_\rho = v(G) \mathcal{D}_{11}(I) = v(G)$, ce qui prouve que la dimension n_ρ est finie.

c) En répétant l'argument avec une matrice M arbitraire, on a à nouveau $V = c_M I$ et on calcule c_M en prenant la trace : $c_M n_\rho = v(G) \text{tr} M$, ce qui, par différentiation par rapport à $M_{\beta\beta'}$, conduit à l'orthonormalité (2.44), cqfd.

La proposition (2.45) découle simplement de la précédente en prenant la trace sur $\alpha = \beta$ et $\alpha' = \beta'$.

Soulignons deux conséquences importantes de cette discussion :

- on vient de voir que toute représentation irréductible (et unitaire) d'un groupe compact est de dimension finie ;

- la relation (2.45) implique que deux représentations $D^{(\rho)}$ et $D^{(\sigma)}$ sont équivalentes (en fait identiques, compte tenu de notre convention d'indexation) ssi leurs caractères sont égaux : $\chi^{(\rho)} = \chi^{(\sigma)}$.

Cas d'un groupe non compact

La plus grande partie du calcul précédent s'applique aussi à un groupe non compact pourvu qu'il soit doté d'une mesure invariante à gauche (ce qui est vrai dans une large classe de groupes, cf chap. 1, fin du § 1.2.4) et que la représentation soit dans un espace préhilbertien séparable, donc doté d'une base discrète, et soit de carré intégrable : $\mathcal{D}_{\alpha\beta} \in L^2(G)$. En choisissant M comme en b), on a à nouveau $\int d\mu(g) = c_1 \operatorname{tr} I$. Au membre de gauche, l'intégrale sur le groupe (le "volume du groupe" G) diverge. Au membre de droite, $\operatorname{tr} I$, la dimension de la représentation, est donc infinie.

De façon générale, on peut dire que

Toute représentation unitaire de carré intégrable d'un groupe non compact est de dimension infinie.

Bien sûr, la représentation triviale $g \mapsto 1$ (qui n'est pas $L^2(G)$) échappe à cet argument.

Testons à nouveau ces résultats sur les deux cas de $U(1)$ et \mathbb{R} . Pour la représentation unitaire e^{ikx} de $U(1)$, la relation (2.44) (ou (2.45), cela ne fait pas de différence pour ces représentations de dimension 1) exprime que

$$\int_0^{2\pi} \frac{dx}{2\pi} e^{ikx} e^{-ik'x} = \delta_{kk'} ,$$

comme on sait bien. Par contre pour \mathbb{R} elle conduirait à

$$\int_{-\infty}^{\infty} dx e^{ikx} e^{-ik'x} = 2\pi \delta(k - k')$$

avec la fonction de Dirac. Bien sûr, cette expression n'a pas de sens pour $k = k'$, la représentation n'est pas de carré intégrable.

Complétude

Revenons au cas d'un groupe compact. On peut démontrer que les matrices $\mathcal{D}_{\alpha\beta}^\rho(g)$ satisfont aussi une propriété de complétude

$$\boxed{\sum_{\rho, \alpha, \beta} n_\rho \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\alpha\beta}^{(\rho)*}(g') = v(G) \delta(g, g') ,} \quad (2.48)$$

ou encore si on préfère

$$\sum_{\rho, \alpha, \beta} n_\rho \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\beta\alpha}^{(\rho)\dagger}(g') = \sum_{\rho} n_\rho \chi^{(\rho)}(g \cdot g'^{-1}) = v(G) \delta(g, g') , \quad (2.48)'$$

où $\delta(g, g')$ est la distribution de Dirac adaptée à la mesure de Haar, c'est-à-dire telle que $\int d\mu(g') f(g') \delta(g, g') = f(g)$ pour toute fonction f sur G suffisamment régulière.

Cette propriété de complétude est importante : elle nous apprend que toute fonction sur le groupe, continue ou de carré intégrable, à valeurs dans \mathbb{C} , peut être développée sur les fonctions $\mathcal{D}_{\alpha\beta}^{(\rho)}(g)$

$$f(g) = \int d\mu(g') \delta(g, g') f(g') = \sum_{\rho, \alpha, \beta} n_{\rho} \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \int \frac{d\mu(g')}{v(G)} \mathcal{D}_{\beta\alpha}^{(\rho)\dagger}(g') f(g') =: \sum_{\rho, \alpha, \beta} n_{\rho} \mathcal{D}_{\alpha\beta}^{(\rho)}(g) f_{\alpha\beta}^{(\rho)}. \quad (2.49)$$

C'est le *théorème de Peter–Weyl*, non trivial, que nous admettrons. Un corollaire dit alors que les caractères $\chi^{(\rho)}$ d'un groupe compact forment un système complet de *fonctions de classe*, c'est-à-dire invariante par $g \sim hgh^{-1}$. Autrement dit, toute fonction de classe continue peut se décomposer sur les caractères irréductibles.

Donnons la démonstration de cette dernière assertion. Soit f une fonction de classe continue, $f(g) = f(hgh^{-1})$, appliquons lui le théorème de Peter–Weyl, et examinons l'intégrale apparaissant dans (2.49) :

$$\begin{aligned} f_{\alpha\beta}^{(\rho)} &= \int \frac{d\mu(g')}{v(G)} f(g') \mathcal{D}_{\beta\alpha}^{(\rho)\dagger}(g') = \int \frac{d\mu(g')}{v(G)} f(hg'h^{-1}) \mathcal{D}_{\beta\alpha}^{(\rho)\dagger}(hg'h^{-1}) \quad \forall h \\ &= \int \frac{d\mu(h)}{v(G)} \frac{d\mu(g')}{v(G)} f(g') \mathcal{D}_{\beta\gamma}^{(\rho)\dagger}(h) \mathcal{D}_{\gamma\delta}^{(\rho)\dagger}(g') \mathcal{D}_{\delta\alpha}^{(\rho)\dagger}(h^{-1}) \\ &= \int \frac{d\mu(g')}{v(G)} f(g') \mathcal{D}_{\gamma\delta}^{(\rho)\dagger}(g') \frac{1}{n_{\rho}} \delta_{\alpha\beta} \delta_{\gamma\delta} \quad \text{par (2.44)} \\ &= \frac{1}{n_{\rho}} \int \frac{d\mu(g')}{v(G)} f(g') \chi^{(\rho)*}(g') \delta_{\alpha\beta} \end{aligned} \quad (2.50)$$

d'où il découle que (2.49) se réduit bien à un développement sur les caractères, cqfd.

Testons à nouveau ces relations de complétude sur le cas $U(1)$. Elles expriment dans ce cas

$$\sum_{k=-\infty}^{\infty} e^{ikx} e^{-ikx'} = 2\pi \delta_P(x - x') \quad (2.51)$$

où $\delta_P(x - x') = \sum_{\ell=-\infty}^{\infty} \delta(x - x' - 2\pi\ell)$ est la distribution de Dirac périodique (alias “peigne de Dirac”). Et (2.49) signifie que toute fonction périodique de période 2π (et avec des conditions adéquates de régularité) peut être représentée par sa *série* de Fourier

$$f(x) = \sum_{k=-\infty}^{\infty} e^{ikx} f_k \quad f_k = \int_{-\pi}^{\pi} \frac{dx}{2\pi} f(x) e^{-ikx}. \quad (2.52)$$

Pour le groupe non compact \mathbb{R} , la relation de complétude (qui est encore vraie dans ce cas) équivaut à la *transformation* de Fourier

$$f(x) = \int_{-\infty}^{\infty} dk \tilde{f}(k) e^{ikx} \quad \tilde{f}(k) = \int_{-\infty}^{\infty} \frac{dx}{2\pi} f(x) e^{-ikx}. \quad (2.53)$$

Le théorème de Peter–Weyl pour un groupe quelconque est donc une généralisation des décompositions de Fourier.

Le groupe des rotations dans le plan $SO(2)$ est isomorphe au groupe $U(1)$. Noter que si on s'intéresse à des représentations irréductibles réelles, la dimension n'est plus égale à 1 (sauf pour la représentation identité!) mais à 2

$$\mathcal{D}^{(k)}(\alpha) = \begin{pmatrix} \cos k\alpha & -\sin k\alpha \\ \sin k\alpha & \cos k\alpha \end{pmatrix}, \quad k \in \mathbb{N}^*, \quad \chi^{(k)}(\alpha) = 2 \cos k\alpha \quad (2.54)$$

Que deviennent les relations d'orthogonalité et de complétude?

2.3.2 Conséquences

Pour un groupe compact,

(i) toute représentation étant complètement réductible, son caractère s'écrit

$$\chi = \sum_{\rho} m_{\rho} \chi^{(\rho)} \quad (2.55)$$

et les multiplicités peuvent se calculer par la formule

$$m_{\rho} = \int \frac{d\mu(g)}{v(G)} \chi(g) \chi^{(\rho)*}(g) . \quad (2.56)$$

On a aussi $\|\chi\|^2 := \int \frac{d\mu(g)}{v(G)} |\chi(g)|^2 = \sum_{\rho} m_{\rho}^2$, un entier supérieur ou égal à 1. Par conséquent, une représentation est irréductible ssi son caractère satisfait la condition $\|\chi\|^2 = 1$. De façon générale le calcul de $\|\chi\|^2$ nous donne des indications sur le nombre de représentations irréductibles apparaissant dans la décomposition de la représentation de caractère χ , une information souvent très utile dans les problèmes évoqués aux § 2.1.7 et 2.1.8.

Plus généralement, toute fonction de classe peut se décomposer sur les caractères irréductibles (théorème de Peter-Weyl). Comme on vient de le voir, cette décomposition des fonctions de classe sur les caractères irréductibles est une généralisation de la décomposition de Fourier.

(ii) De même on peut déterminer les multiplicités dans la décomposition de Clebsch-Gordan d'un produit direct de deux représentations en projetant le produit de leurs caractères sur les caractères irréductibles.

Illustrons ceci sur le produit de deux représentations irréductibles ρ et σ

$$\begin{aligned} D^{(\rho)} \otimes D^{(\sigma)} &= \bigoplus_{\tau} m_{\tau} D^{(\tau)} \\ \chi^{(\rho)} \chi^{(\sigma)} &= \sum_{\tau} m_{\tau} \chi^{(\tau)} \\ m_{\tau} &= \int \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g) \chi^{(\sigma)}(g) \chi^{(\tau)*}(g) , \end{aligned} \quad (2.57)$$

donc la représentation τ apparaît dans le produit $\rho \otimes \sigma$ avec la même multiplicité que σ^* dans $\rho \otimes \tau^*$. Exercice : vérifier qu'en particulier, la représentation identité apparaît dans le produit des irrep ρ et σ ssi $\sigma = \rho^*$, la représentation complexe conjuguée de ρ .

Cas de $SU(2)$

C'est un bon exercice de comprendre comment les différentes propriétés discutées dans ce paragraphe sont réalisées par les matrices de représentation de $SU(2)$. Cela sera discuté en détail en TD et dans l'Appendice E.

2.3.3 Cas des groupes finis

Nous n'évoquerons que très brièvement le cas des groupes finis. En fait les théorèmes (2.44, 2.45, 2.48) et leurs conséquences (2.55, 2.56, 2.57), qui étaient basés sur l'existence d'une mesure invariante, sont bien sûr toujours vrais. Il suffit de remplacer dans les expressions de ces

théorèmes le volume $v(G)$ par l'ordre $|G|$ (=nombre d'éléments) de G , et $\int d\mu(g)$ par $\sum_{g \in G}$:

$$\frac{1}{|G|} \sum_{g \in G} \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\alpha'\beta'}^{(\rho')*}(g) = \frac{1}{n_\rho} \delta_{\rho\rho'} \delta_{\alpha\alpha'} \delta_{\beta\beta'} \quad (2.58)$$

$$\sum_{\rho, \alpha, \beta} \frac{n_\rho}{|G|} \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\alpha\beta}^{(\rho)*}(g') = \delta_{g, g'} . \quad (2.59)$$

Mais les représentations des groupes finis jouissent de propriétés supplémentaires. Montrons ainsi que les dimensions des représentations irréductibles non équivalentes vérifient

$$\sum_{\rho} n_\rho^2 = |G| . \quad (2.60)$$

Cela découle du fait que le système d'équations (2.58-2.59) peut être vu comme exprimant que la matrice $\mathcal{U}_{\rho, \alpha\beta; g} := \left(\frac{n_\rho}{|G|}\right)^{\frac{1}{2}} \mathcal{D}_{\alpha\beta}^{(\rho)}(g)$ de dimensions $\sum_{\rho} n_\rho^2 \times |G|$ satisfait $\mathcal{U}\mathcal{U}^\dagger = I$, $\mathcal{U}^\dagger\mathcal{U} = I$, ce qui n'est possible que si c'est une matrice carrée, cqfd.

Cela implique en particulier que le nombre r de représentations irréductibles inéquivalentes est fini, et nous allons montrer que

Proposition. *Le nombre r de représentations irréductibles est fini et égal au nombre m des classes \mathcal{C}_i dans le groupe.*

Preuve : En notant $\chi_j^{(\rho)}$ la valeur du caractère $\chi^{(\rho)}$ dans la classe \mathcal{C}_i , on peut récrire les relations d'orthogonalité et de complétude des caractères selon

$$\frac{1}{|G|} \sum_{i=1}^m |\mathcal{C}_i| \chi_i^{(\rho)} \chi_i^{(\rho')*} = \delta_{\rho\rho'} \quad (2.61a)$$

$$\frac{|\mathcal{C}_i|}{|G|} \sum_{\rho=1}^r \chi_i^{(\rho)} \chi_j^{(\rho)*} = \delta_{ij} . \quad (2.61b)$$

(Exercice : déduire la deuxième relation de (2.49) et (2.50), appliquées à un groupe fini.)

Mais à nouveau, ces relations expriment que la matrice $\mathcal{K}_{\rho i} := \left(\frac{|\mathcal{C}_i|}{|G|}\right)^{\frac{1}{2}} \chi_i^{(\rho)}$ de dimensions $r \times m$ satisfait $\mathcal{K}\mathcal{K}^\dagger = I$, $\mathcal{K}^\dagger\mathcal{K} = I$, donc est carrée (et unitaire), $m = r$, cqfd.

La *table de caractères* d'un groupe fini est le tableau carré constitué par les nombres (réels ou complexes) $\chi_i^{(\rho)}$, $\rho = 1, \dots, r$, $i = 1, \dots, m$. Ses lignes et colonnes satisfont les propriétés d'orthogonalité (2.61).

Illustrons cela sur l'exemple du groupe T , sous-groupe du groupe des rotations laissant invariant un tétraèdre régulier. Ce groupe d'ordre 12 a 4 classes de conjugaison \mathcal{C}_i , celle de l'identité, celle des 3 rotations de π autour d'un axe joignant les milieux d'arêtes opposées, celle des 4 rotations de $2\pi/3$ autour d'un axe passant par un sommet, et celle des 4 rotations de $-2\pi/3$, voir Fig. 2.1.

Ce groupe a donc 4 représentations irréductibles, dont on vérifie aisément à l'aide de (2.60) que les dimensions ne peuvent être que $n_\rho = 1, 1, 1$ et 3. La table des caractères est donc un tableau 4×4 dont on connaît déjà une ligne, celle de la représentation identité D_1 , et une colonne, celles des dimensions n_ρ . La représentation de spin 1 de $\text{SO}(3)$ fournit une représentation de dimension 3 de T dont le caractère χ prend les valeurs $\chi_i = 1 + 2 \cos \theta_i = (3, -1, 0, 0)$ dans les quatre classes; selon le critère du § 3.2, $\|\chi\|^2 = \sum_i \frac{|\mathcal{C}_i|}{|G|} |\chi_i|^2 = 1$ donc ce caractère est irréductible. Cela nous fournit une deuxième ligne (celle notée D_4). La représentation de spin 2 de $\text{SO}(3)$ fournit une représentation de dimension 5 qui est elle réductible (selon le même critère) en somme de 3 irreps, mais orthogonale à D_1 . C'est la somme des trois lignes D_2 , D_3 et D_4 , dans lesquelles $j = e^{2\pi i/3}$, avec $j + j^2 = -1$.

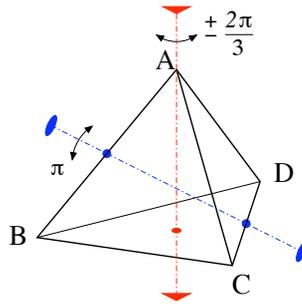


FIGURE 2.1 – Un tétraèdre, avec deux axes de rotation

↓ irreps. ρ \ Classes $C_i \rightarrow$	$\mathcal{C}(0)$	$\mathcal{C}(\pi)$	$\mathcal{C}(\frac{2\pi}{3})$	$\mathcal{C}(-\frac{2\pi}{3})$
D_1	1	1	1	1
D_2	1	1	j	j^2
D_3	1	1	j^2	j
D_4	3	-1	0	0
$ C_i $	1	3	4	4

Vérifier que les relations (2.61) sont bien satisfaites. Expliquer aussi pourquoi le groupe T n'est autre que le groupe A_4 des permutations paires de 4 objets.

2.3.4 Récapitulation

Pour un groupe *compact*, toute représentation irréductible est de dimension finie et équivalente à une représentation unitaire. Ses éléments de matrice et caractères satisfont des relations d'orthogonalité et de complétude. L'ensemble des représentations irréductibles est *discret*.

Pour un groupe *fini*, (cas qu'on n'a traité que très superficiellement dans ce cours), ces mêmes propriétés d'orthogonalité et de complétude sont satisfaites. Mais on a des propriétés supplémentaires, par exemple le nombre des représentations irréductibles inéquivalentes est *fini*, et égal au nombre de classes du groupe.

Pour un groupe *non compact*, les représentations unitaires sont généralement de dimension infinie. (Par contre il peut exister des représentations non unitaires de dimension finie, cf le cas de $SL(2, \mathbb{C})$ au Chap. 0). L'ensemble des représentations irréductibles est indexé par des paramètres discrets et *continus*.

2.4 Représentations projectives. Théorème de Wigner

2.4.1 Définition

On appelle représentation projective d'un groupe G une représentation linéaire à une phase près de ce groupe (on se restreint ici à des représentations unitaires). Pour $g_1, g_2 \in G$, on a

$$U(g_1)U(g_2) = e^{i\zeta(g_1, g_2)}U(g_1g_2) . \tag{2.62}$$

On peut toujours choisir $U(e) = I$, et on a donc $\forall g \quad \zeta(e, g) = \zeta(g, e) = 0$. On peut aussi redéfinir $U(g) \rightarrow U'(g) = e^{i\alpha(g)}U(g)$, ce qui change

$$\zeta(g_1, g_2) \rightarrow \zeta'(g_1, g_2) = \zeta(g_1, g_2) + \alpha(g_1) + \alpha(g_2) - \alpha(g_1 g_2) . \quad (2.63)$$

La fonction $\zeta(g_1, g_2)$ de $G \times G$ dans \mathbb{R} forme ce qu'on appelle une *2-cochaîne*. Elle est fermée (et on l'appelle alors *2-cocycle*) en raison de la propriété d'associativité :

$$\forall g_1, g_2, g_3 \quad (\partial\zeta)(g_1, g_2, g_3) := \zeta(g_1, g_2) + \zeta(g_1 g_2, g_3) - \zeta(g_2, g_3) - \zeta(g_1, g_2 g_3) = 0 \quad (2.64)$$

(le vérifier). En général, pour une n -cochaîne $\varphi(g_1, \dots, g_n)$, on définit l'opérateur ∂ qui fait passer des n -cochaînes aux $n + 1$ -cochaînes :

$$(\partial\varphi)(g_1, \dots, g_{n+1}) = \sum_{i=1}^n (-1)^{i+1} \varphi(g_1, g_2, \dots, (g_i g_{i+1}), \dots, g_{n+1}) - \varphi(g_2, \dots, g_{n+1}) + (-1)^n \varphi(g_1, \dots, g_n) .$$

Pour une 1-cochaîne $\alpha(g)$, $\partial\alpha(g_1, g_2) = \alpha(g_1 g_2) - \alpha(g_1) - \alpha(g_2)$, et donc (2.63) s'exprime par $\zeta' = \zeta - \partial\alpha$.

Vérifier que $\partial^2 = 0$.

La question de savoir si la représentation $U(g)$ est intrinsèquement projective ou peut être ramenée à une représentation ordinaire par un changement de phase équivaut à savoir si le cocycle ζ est *trivial*, c'est-à-dire s'il existe un $\alpha(g)$ tel que dans (2.63), $\zeta' = 0$.

Autrement dit, le 2-cocycle ζ , qui est fermé ($\partial\zeta = 0$) par (2.64), est-il *exact*, c'est-à-dire de la forme $\zeta = \partial\alpha$? C'est un problème typique de *cohomologie*. La cohomologie des groupes de Lie est un vaste sujet qui a été très étudié... mais dont nous ne dirons rien de plus dans ce cours.

On peut résumer une discussion un peu longue et complexe (esquissée plus bas au § 2.4.5) en disant que pour un groupe semi-simple G , tel $SO(n)$, l'origine de ces représentations projectives est à chercher dans le caractère non simplement connexe de G . Dans ce cas, les représentations unitaires de \tilde{G} , recouvrement universel de G , fournissent des représentations à une phase près de G . Par exemple, on retrouve que les représentations projectives de $SO(3)$ (à un signe près) sont les représentations de $SU(2)$. C'est le cas aussi du sous-groupe \mathcal{L}_+^\uparrow du groupe de Lorentz $O(1,3)$, dont le recouvrement universel est $SL(2, \mathbb{C})$.

Avant de poursuivre, il est légitime de se poser la question : pourquoi les représentations projectives intéressent-elles les physiciens? La raison est que les transformations d'un système quantique y font appel, comme on va le voir.

2.4.2 Théorème de Wigner

Soit un système quantique dont les états (purs) sont représentés par les rayons d'un espace de Hilbert \mathcal{H} ⁶ observables sont des opérateurs auto-adjoints sur \mathcal{H} . Supposons qu'il existe une transformation g du système (états *et* observables) qui laissent inchangées les quantités $|\langle \phi | A | \psi \rangle|^2$, c'est-à-dire

$$|\psi\rangle \rightarrow |^g\psi\rangle, \quad A \rightarrow {}^gA \quad \text{tel que} \quad |\langle \phi | A | \psi \rangle| = |\langle {}^g\phi | {}^gA | {}^g\psi \rangle|. \quad (2.65)$$

On démontre alors le théorème suivant

6. rayon = vecteur à un facteur scalaire près, ou à une phase près s'il est normalisé

Théorème de Wigner. *Si une bijection entre les rayons et entre les opérateurs auto-adjoints d'un espace de Hilbert \mathcal{H} préserve les modules des produits scalaires*

$$|\langle \phi|A|\psi \rangle| = |\langle {}^g\phi|{}^gA|{}^g\psi \rangle| , \quad (2.66)$$

alors cette bijection est réalisée par un opérateur $U(g)$, linéaire ou antilinéaire, unitaire sur \mathcal{H} , et unique à une phase près, c'est-à-dire

$$|{}^g\psi \rangle = U(g)|\phi \rangle , \quad {}^gA = U(g)AU^\dagger(g) ; \quad U(g)U^\dagger(g) = U(g)^\dagger U(g) = I . \quad (2.67)$$

Rappelons d'abord ce qu'on entend par opérateur antilinéaire. Un tel opérateur satisfait

$$U(\lambda|\phi \rangle + \mu|\psi \rangle) = \lambda^*U|\phi \rangle + \mu^*U|\psi \rangle \quad (2.68)$$

et son adjoint est défini par

$$\langle \phi|U^\dagger|\psi \rangle = \langle U\phi|\psi \rangle^* = \langle \psi|U\phi \rangle , \quad (2.69)$$

de façon à être compatible avec la linéarité :

$$\langle \lambda\phi|U^\dagger|\psi \rangle = \lambda^*\langle \phi|U^\dagger|\psi \rangle . \quad (2.70)$$

S'il est en outre unitaire,

$$\langle \psi|\phi \rangle^* = \langle \phi|\psi \rangle = \langle \phi|U^\dagger U|\psi \rangle = \langle U\phi|U\psi \rangle^* , \quad (2.71)$$

donc $\langle U\phi|U\psi \rangle = \langle \psi|\phi \rangle$.

La preuve du théorème est un peu laborieuse. Elle consiste à montrer que si on a une base orthonormée $|\psi_k \rangle$ de \mathcal{H} , on peut trouver des représentants $|{}^g\psi_k \rangle$ des rayons transformés tels qu'un représentant du rayon transformé de $\sum c_k|\psi_k \rangle$ soit $\sum c'_k|{}^g\psi_k \rangle$ avec ou bien **tous** les $c'_k = c_k$, ou bien **tous** les $c'_k = c_k^*$. Autrement dit, l'action $|\psi \rangle \rightarrow |{}^g\psi \rangle$ est à travers tout \mathcal{H} soit linéaire, soit antilinéaire.

Une fois connue la transformation des états par l'opérateur $U(g)$, on détermine celle des observables ${}^gA = U(g)AU^\dagger(g)$ de façon à avoir

$$\begin{aligned} \langle {}^g\phi|{}^gA|{}^g\psi \rangle &= \langle U\phi|UAU^\dagger|U\psi \rangle \\ &= \langle \phi|U^\dagger UAU^\dagger U|\psi \rangle^\# \\ &= \langle \phi|A|\psi \rangle^\# \end{aligned} \quad (2.72)$$

avec $\# =$ rien ou $*$ selon que U est linéaire ou antilinéaire.

Le cas antilinéaire n'a pas qu'un intérêt académique. Il se rencontre dans l'étude du renversement du sens du temps. En effet cette opération T laisse l'opérateur position \mathbf{x} inchangé, mais change le signe des vitesses, donc de l'impulsion \mathbf{p}

$$\mathbf{x}' = U(T)\mathbf{x}U^\dagger(T) = \mathbf{x} \quad (2.73)$$

$$\mathbf{p}' = U(T)\mathbf{p}U^\dagger(T) = -\mathbf{p} . \quad (2.74)$$

Les relations de commutation canoniques ne sont compatibles avec cette transformation que si $U(T)$ est antilinéaire

$$\begin{aligned} [x'_j, p'_k] &= -[x_j, p_k] = -i\hbar\delta_{jk} \\ &= U(T)[x_j, p_k]U^\dagger(T) = U(T)i\hbar\delta_{jk}U^\dagger(T) \end{aligned} \quad (2.75)$$

Autre argument : $U(T)$ commute avec les translations dans le temps dont le générateur est i fois l'hamiltonien : $U(T)i\hbar U^\dagger(T) = -i\hbar H$ (puisque $t \rightarrow -t$). Si U était linéaire, on conclurait que $U\hbar U^\dagger = -\hbar H$, ce qui est gênant si on veut que $\text{Spec}(H) \geq 0$!

Les transformations d'un système quantique, c'est-à-dire les bijections du théorème de Wigner, forment un groupe G : si g_1 et g_2 sont de telles bijections, leur composition g_1g_2 en est une aussi, ainsi que g_1^{-1} . Les opérateurs $U(g)$ qu'on va supposer linéaires dans la suite de ce cours forment donc une représentation à une phase près (cf l'unicité à une phase près dans le théorème), c'est-à-dire une représentation projective de G .

Une parenthèse sur un point important de terminologie

Jusqu'à ce point, nous avons mené la discussion des *transformations* d'un système quantique sans rien supposer sur son éventuelle *invariance* sous ces transformations, c'est-à-dire sur la façon dont elles affectent (ou non) sa dynamique. Ces transformations peuvent être envisagées d'un point de vue actif : on considère en parallèle le système initial et le système transformé, ou d'un point de vue passif : il s'agit du même système, examiné dans deux systèmes de coordonnées, deux référentiels, différents, obtenus l'un à partir de l'autre par la transformation considérée.

2.4.3 Invariances d'un système quantique

Supposons maintenant que sous l'action d'un certain groupe de transformations G , le système est invariant, en ce sens que sa dynamique, contrôlée par son hamiltonien H , est inchangée. On va écrire

$$H = U(g)HU^\dagger(g)$$

ou encore

$$[H, U(g)] = 0. \quad (2.76)$$

On définit donc une *invariance* (ou *symétrie*) d'un système quantique sous l'action d'un groupe G comme l'existence d'une représentation projective unitaire (linéaire ou antilinéaire) de ce groupe dans l'espace des états, qui commute avec l'hamiltonien.

- Cette situation implique l'existence de lois de conservation. En effet toute observable fonction des $U(g)$ commute avec H , donc est une quantité conservée

$$i\hbar \frac{\partial \mathcal{F}(U(g))}{\partial t} = [\mathcal{F}(U(g)), H] = 0 \quad (2.77)$$

et chacune de ses valeurs propres est un "bon nombre quantique" : si le système appartient à un sous-espace propre de \mathcal{F} au temps t , il y demeure lors de son évolution dans le temps. Si G

est un groupe de Lie, pour g une transformation infinitésimale et si T désignent les générateurs infinitésimaux dans la représentation considérée,

$$U(g) = I - i \delta\alpha^j T_j$$

(où on choisit les T auto-adjoints pour avoir U unitaire), les T_j sont des observables commutant avec H , donc des quantités conservées, mais en général, pas simultanément mesurables.

Exemples.

Groupe des translations $\longrightarrow P_\mu$ énergie-impulsion; groupe des rotations $\longrightarrow M_{\mu\nu}$ moment cinétique.

Noter encore que ces opérateurs T_i qui réalisent dans la théorie quantique les opérations infinitésimales du groupe G , forment d'un point de vue mathématique une représentation de l'algèbre de Lie \mathfrak{g} . On peut donc affirmer qu'ils satisfont les relations de commutation

$$[T_i, T_j] = iC_{ij}^k T_k \quad (2.78)$$

(avec un “ i ” parce qu'on a fait un choix d'opérateurs hermitiens). Le nombre maximal de ces opérateurs qu'on peut diagonaliser simultanément, donc de ces quantités conservées qu'on pourra fixer, dépend de la structure de \mathfrak{g} et de ces relations de commutation.

- Mais l'hypothèse d'invariance faite plus haute a une autre conséquence, d'application fréquente et importante. Si l'espace des états \mathcal{H} qui “porte une représentation” du groupe G est décomposé en représentations irréductibles, dans chaque espace $E^{(\rho)}$, supposé de multiplicité 1, l'hamiltonien est multiple de l'identité en vertu du lemme de Schur. On a donc dans ce cas une information complète sur la nature du spectre : espaces propres $E^{(\rho)}$ et multiplicités des valeurs propres \mathcal{E}_ρ de H égales à $\dim E^{(\rho)}$ ⁷. Si certains espaces de représentation $E^{(\rho)}$ apparaissent avec une multiplicité m_ρ supérieure à 1, il faut encore diagonaliser H dans la somme de ces espaces $\bigoplus_i E^{(\rho,i)}$, ce qui est tout de même plus simple que le problème de diagonalisation dans l'espace \mathcal{H} de départ. On verra plus bas que le théorème de Wigner-Eckart permet de réduire encore la complexité de ce dernier calcul. La théorie des groupes nous a donc considérablement simplifié la tâche ... mais elle ne nous fournit pas les *valeurs* des énergies propres \mathcal{E}_ρ .

Dans ce qui précède, nous avons considéré le point de vue hamiltonien. Comme on le sait, on peut mener une discussion parallèle dans le formalisme quantique – classique ou quantique –. Là, les invariances du lagrangien (ou de l'action) se traduisent par l'existence de courants de Noether de divergence nulle et de quantités conservées.

2.4.4 Transformations des observables. Théorème de Wigner–Eckart

Selon (2.67), la transformation d'un opérateur sur \mathcal{H} obéit à : $A \rightarrow U(g)AU(g)^\dagger$. Supposons qu'on a un ensemble de tels opérateurs, A_α , $\alpha = 1, 2, \dots$, qui se transforment linéairement les uns dans les autres dans ces transformations, c'est-à-dire qui forment une représentation :

$$A_\alpha \rightarrow U(g)A_\alpha U(g)^\dagger = \sum_{\alpha'} A_{\alpha'} \mathcal{D}_{\alpha'\alpha}(g) . \quad (2.79)$$

7. Il peut arriver que la multiplicité d'une valeur propre de H soit plus élevée, soit à cause de l'existence d'un groupe de symétrie plus grand que G , soit parce que certaines représentations viennent en paires complexes conjuguées, soit pour une autre raison.

Si la représentation D est irréductible, les opérateurs A_α forment ce qu'on appelle un opérateur (ou “tenseur”) irréductible. Par exemple, en physique atomique, l'opérateur moment cinétique \vec{J} et l'opérateur moment dipolaire électrique $\sum_i q_i \vec{r}_i$ se transforment comme des vecteurs sous l'effet des rotations. Utilisant les notations de la sect. 2, supposons que les A_α se transforment par la représentation irréductible $D^{(\rho)}$ et appliquons les sur des états $|\sigma\beta\rangle$ se transformant selon la représentation irréductible $D^{(\sigma)}$. L'état résultant se transforme selon

$$U(g)A_\alpha|\sigma\beta\rangle = U(g)A_\alpha U(g)^\dagger U(g)|\sigma\beta\rangle = \mathcal{D}_{\alpha'\alpha}^{(\rho)}(g)\mathcal{D}_{\beta'\beta}^{(\sigma)}(g)A_{\alpha'}|\sigma\beta'\rangle \quad (2.80)$$

c'est-à-dire selon le produit tensoriel des représentations $D^{(\rho)}$ et $D^{(\sigma)}$. Comme on l'a fait en (2.37), on peut développer sur des représentations irréductibles

$$\mathcal{D}_{\alpha'\alpha}^{(\rho)}(g)\mathcal{D}_{\beta'\beta}^{(\sigma)}(g) = \sum_{\tau,\gamma,\gamma',i} \langle \tau_i\gamma|\rho,\alpha;\sigma,\beta\rangle \langle \tau_i\gamma'|\rho,\alpha';\sigma,\beta'\rangle^* \mathcal{D}_{\gamma'\gamma}^{(\tau)}(g) . \quad (2.81)$$

Supposons maintenant que le groupe G est compact (ou fini). Les matrices des représentations satisfont les propositions d'orthogonalité (2.44). On peut alors écrire

$$\begin{aligned} \langle \tau\gamma|A_\alpha|\sigma\beta\rangle &= \langle \tau\gamma|U(g)^\dagger U(g)A_\alpha|\sigma\beta\rangle && \forall g \in G \\ &= \int \frac{d\mu(g)}{v(G)} \langle \tau\gamma|U(g)^\dagger U(g)A_\alpha|\sigma\beta\rangle \\ &= \int \frac{d\mu(g)}{v(G)} \sum_{\alpha',\beta',\gamma'} \mathcal{D}_{\gamma'\gamma}^{(\tau)*}(g) \langle \tau\gamma'|A_{\alpha'}|\sigma\beta'\rangle \mathcal{D}_{\alpha'\alpha}^{(\rho)}(g)\mathcal{D}_{\beta'\beta}^{(\sigma)}(g) \\ &= \frac{1}{n_\tau} \sum_{\alpha',\beta',\gamma',i} \langle \tau_i\gamma|\rho,\alpha;\sigma,\beta\rangle \langle \tau_i\gamma'|\rho,\alpha';\sigma,\beta'\rangle^* \langle \tau_i\gamma'|A_{\alpha'}|\sigma\beta'\rangle . \end{aligned} \quad (2.82)$$

Notons

$$\langle \tau \parallel A \parallel \sigma \rangle_i := \frac{1}{n_\tau} \sum_{\alpha',\beta',\gamma'} \langle \tau_i\gamma'|\rho,\alpha';\sigma,\beta'\rangle^* \langle \tau_i\gamma'|A_{\alpha'}|\sigma\beta'\rangle . \quad (2.83)$$

Il en découle que (théorème de Wigner–Eckart) :

$$\langle \tau\gamma|A_\alpha|\sigma\beta\rangle = \sum_{i=1}^{m_\tau} \langle \tau \parallel A \parallel \sigma \rangle_i \langle \tau_i\gamma|\rho,\alpha;\sigma,\beta\rangle \quad (2.84)$$

dans laquelle les éléments de matrice “réduits” $\langle . \parallel A \parallel . \rangle_i$ sont indépendants de α, β, γ . L'élément de matrice du membre de gauche dans (2.84) s'annule si le coefficient de Clebsch-Gordan est nul (en particulier si la représentation τ n'apparaît pas dans le produit de ρ et σ). Ce théorème a de nombreuses conséquences en physique atomique et nucléaire, où il occasionne des “règles de sélection”. Voir par exemple à l'Appendice E.3 le cas des opérateurs moments multipolaires électriques.

Ce théorème nous permet aussi de simplifier encore le problème de diagonalisation de l'hamiltonien H mentionné à la fin du § 2.4.3, quand un espace de représentation apparaît avec une multiplicité m_ρ . En repérant par un indice $i = 1, \dots, m_\rho$ les différentes copies de la représentation, on a en vertu de (2.84)

$$\langle \rho\alpha i|H|\rho\alpha' i'\rangle = \delta_{\alpha\alpha'} \langle \rho i \parallel H \parallel \rho i'\rangle \quad (2.85)$$

et le problème n'implique plus que la diagonalisation d'une matrice de taille $m_\rho \times m_\rho$.

Exercice. Pour le groupe $SO(3)$, soit K_1^m les composantes d'un opérateur irréductible vectoriel (par exemple, l'opérateur moment dipolaire de l'Appendice B.3). Montrer en utilisant le théorème de Wigner-Eckart que

$$\langle j, m_1 | K_1^m | j, m_2 \rangle = \langle j, m_1 | J^m | j, m_2 \rangle \frac{\langle \vec{J} \cdot \vec{K} \rangle_j}{j(j+1)}$$

où $\langle \vec{J} \cdot \vec{K} \rangle_j$ désigne la valeur moyenne de $\vec{J} \cdot \vec{K}$ dans l'état j . En d'autres termes, on peut remplacer \vec{K} par sa projection $\vec{J} \frac{\langle \vec{J} \cdot \vec{K} \rangle_j}{j(j+1)}$.

2.4.5 Forme infinitésimale d'une représentation projective. Extension centrale

Si G est un groupe de Lie, d'algèbre de Lie \mathfrak{g} , soit t_a une base de \mathfrak{g}

$$[t_a, t_b] = C_{ab}{}^c t_c .$$

Dans une représentation projective, examinons la composition de deux transformations infinitésimales de la forme $I + \alpha t_a$ et $I + \beta t_b$. Comme $\zeta(I, g) = \zeta(g, I) = 0$, $\zeta(I + \alpha t_a, I + \beta t_b)$ est d'ordre $\alpha\beta$

$$i\zeta(I + \alpha t_a, I + \beta t_b) = \alpha\beta z_{ab} . \tag{2.86}$$

Les t_a sont représentés par T_a , et en développant au deuxième ordre, on trouve

$$e^{-i\zeta(I+\alpha t_a, I+\beta t_b)} U(e^{\alpha t_a}) U(e^{\beta t_b}) = U(e^{\alpha t_a} e^{\beta t_b}) = U\left(e^{(\alpha t_a + \beta t_b)} e^{\frac{1}{2}\alpha\beta[t_a, t_b]}\right)$$

et donc, avec $U(e^{\alpha t_a}) = e^{\alpha T_a}$ etc,

$$\alpha\beta \left(-z_{ab} I + \frac{1}{2}[T_a, T_b] - \frac{1}{2}C_{ab}{}^c T_c \right) = 0$$

(ce qui prouve que z_{ab} doit être antisymétrique en a, b). On trouve donc que les relations de commutation des T sont modifiées par un *terme central* (c'est-à-dire commutant avec tous les autres générateurs)

$$[T_a, T_b] = C_{ab}{}^c T_c + 2z_{ab} I .$$

L'existence de représentation projective peut donc se traduire par la réalisation d'une *extension centrale* de l'algèbre de Lie. On appelle ainsi la nouvelle algèbre de Lie engendrée par les T_a et par un (ou plusieurs) nouveau(x) générateur(s) C_{ab} commutant avec tous les T_a (et entre eux)

$$[T_a, T_b] = C_{ab}{}^c T_c + C_{ab} \quad [C_{ab}, T_c] = 0 \quad [C_{ab}, C_{cd}] = 0 . \tag{2.87}$$

(Dans une représentation irréductible de l'algèbre, le lemme de Schur nous assure que $C_{ab} = c_{ab} I$.) La trivialité (ou non-trivialité) du cocycle ζ se traduit sous forme infinitésimale par la possibilité (ou l'impossibilité) d'éliminer le terme central par une redéfinition des T

$$T_a \rightarrow \tilde{T}_a = T_a + X_a \quad [\tilde{T}_a, \tilde{T}_b] = C_{ab}{}^c \tilde{T}_c , \tag{2.88}$$

en exploitant les contraintes sur les $C_{ab}{}^c$ et C_{ab} provenant de l'identité de Jacobi.

Exercice. Écrire la contrainte supplémentaire que l'identité de Jacobi met sur les constantes $C_{ab}{}^c$ et C_{ab} . Montrer que $C_{ab} = C_{ab}{}^c D_c$ en fournit une solution et qu'une redéfinition telle (2.88) est alors possible.

On démontre (Bargmann) que pour un groupe de Lie connexe G , les cocycles sont triviaux si

1. il n'existe pas d'extension centrale non triviale de \mathfrak{g} ;
2. G est simplement connexe.

En ce qui concerne le point 1, un théorème de Bargmann nous assure qu'il n'existe pas d'extension centrale non triviale pour tout algèbre semi-simple, comme celles des groupes classiques $SU(n)$, $SO(n)$, $Sp(2n)$. C'est donc plutôt le point 2 qui nous intéresse.

Si le groupe G n'est pas simplement connexe, on étudie les représentations (disons unitaires) de son recouvrement universel \tilde{G} , qui sont des représentations à une phase près de G (le groupe $\pi_1(G) = \tilde{G}/G$ est représenté sur $U(1)$). C'est le cas des groupes $SO(n)$ et de leur recouvrement universel $Spin(n)$, (par exemple du groupe $SO(3)$), ou du groupe de Lorentz $O(1,3)$, comme rappelé plus haut.

★

Bibliographie sommaire (suite)

Outre les références déjà données dans l'Introduction et au chap. 1,

Théorie générale des représentations

[Ki] A.A. Kirillov, *Elements of the theory of representations*, Springer.

[Kn] A. Knapp, *Representation Theory of semi-simple groups*, Princeton U. Pr.

[FH] W. Fulton and J. Harris, *Representation Theory*, Springer.

Pour la démonstration du théorème de Peter-Weyl, voir par exemple

[BrD] T. Bröcker and T. tom Dieck, *Representations of compact Lie groups*, Springer.

Pour la démonstration du théorème de Wigner, voir E. Wigner, [Wi], ou A. Messiah, [M] t. 2, p 774, ou S. Weinberg, [Wf] chap 2, app A.

Sur les représentations projectives, voir

[Ba] V. Bargmann, *Ann. Math.* **59** (1954) 1-46, or

S. Weinberg [Wf] Chap 2.7.

★

Appendix D. “Tenseurs, vous avez dit tenseurs ?”

Le mot de tenseur recouvre plusieurs concepts reliés mais pas tout à fait identiques. Le but de cet appendice est d'éclaircir ces choses...

D.1. Définition algébrique

Soient deux espaces vectoriels E et F , leur produit tensoriel est par définition l'espace vectoriel $E \otimes F$ engendré par les paires (x, y) , $x \in E$, $y \in F$, notées $x \otimes y$. Un élément de $E \otimes F$ peut donc s'écrire

$$z = \sum_{\alpha} x^{(\alpha)} \otimes y^{(\alpha)} \quad (\text{D.1})$$

avec une somme finie sur des vecteurs $x^{(\alpha)} \in E$, $y^{(\alpha)} \in F$ (on a absorbé dans le vecteur $x^{(\alpha)}$ un éventuel coefficient scalaire λ_{α} dans la combinaison linéaire).

Si A , resp. B , est un opérateur linéaire agissant dans E , resp. F , $A \otimes B$ est l'opérateur linéaire agissant dans $E \otimes F$ selon

$$A \otimes B(x \otimes y) = Ax \otimes By \quad (\text{D.2})$$

$$A \otimes B \sum_{\alpha} (x^{(\alpha)} \otimes y^{(\alpha)}) = \sum_{\alpha} Ax^{(\alpha)} \otimes By^{(\alpha)} \quad (\text{D.3})$$

En particulier si E et F ont deux bases e_i et f_j , $z = x \otimes y = \sum_{i,j} x^i y^j e_i f_j$, la base de $E \otimes F$ et les composantes de z sont indexées par des paires d'indices (i, j) , et $A \otimes B$ est décrit dans cette base par une matrice qu'on lit sur

$$(A \otimes B)z = \sum_{i,i',j,j'} A_{ii'} B_{jj'} x^{i'} y^{j'} e_i f_j =: (A \otimes B)_{ii';jj'} z^{i'j'} e_i \otimes f_j \quad (\text{D.4})$$

soit

$$(A \otimes B)_{ij;i'j'} = A_{ii'} B_{jj'} , \quad (\text{D.5})$$

formule qui est parfois prise comme définition du produit tensoriel de deux matrices.

D.2. Action d'un groupe

Si un groupe G admet des représentations D et D' dans deux espaces vectoriels E et F , $x \in E \mapsto D(g)x = e_i \mathcal{D}_{ij} x^j$, *ibid.* pour $y \in F$, on définit la représentation produit tensoriel $D \otimes D'$ dans $E \otimes F$ par

$$D(g) \otimes D'(g)(x \otimes y) = D(g)x \otimes D'(g)y \quad (\text{D.6})$$

en accord avec (D.2). La matrice de $D \otimes D'$ dans une base $e_i \otimes f_j$ est $\mathcal{D}_{ii'} \mathcal{D}'_{jj'}$.

Autre façon de dire les choses : si $x \ll$ se transforme par la représentation D et y par D' , sous l'action de $g \in G$, $x' = D(g)x$, $y' = D'(g)y$, $x \otimes y \mapsto x' \otimes y'$, avec

$$(x' \otimes y')^{ij} = x^i y^j = \mathcal{D}_{ii'} \mathcal{D}'_{jj'} x^{i'} y^{j'} , \quad (\text{D.7})$$

autre formule parfois prise comme définition d'un tenseur (sous l'action de G).

La construction que l'on vient de faire des tenseurs z^{ij} de rang 2 peut s'itérer pour construire des produits tensoriels $E_1 \otimes E_2 \otimes \dots \otimes E_p$ et des tenseurs $z^{i_1 \dots i_p}$ de rang p . C'est ce que nous avons fait au Chap. 0, § 0.3.3, dans la construction des représentations de SU(2) par produits tensoriels symétrisés de la représentation de spin $\frac{1}{2}$, ou au § 6.3 pour celles de SL(2,ℂ) par produits tensoriels symétrisés des deux représentations à indices pointés $(0, \frac{1}{2})$ ou non pointés $(\frac{1}{2}, 0)$.

Appendix E. Compléments sur les matrices de représentation de SU(2)

On revient ici sur les matrices \mathcal{D}^j des représentations de SU(2) définies au § 0.3.3 du Chap. 0.

E.1. Orthogonalité, complétude, caractères

Nous faisons appel à la mesure invariante sur le groupe $SU(2)$ introduite au Chap. 1 (§ 1.2.4 et Appendice C) pour énoncer des propriétés d'orthogonalité et de complétude des matrices \mathcal{D}^j . Nous avons construit au Chap. 0 *toutes* les représentations unitaires de $SU(2)$. Selon la discussion du § 2.3, les éléments des matrices \mathcal{D}^j satisfont des propriétés d'orthogonalité et de complétude

$$(2j+1) \int \frac{d\mu(U)}{2\pi^2} \mathcal{D}_{mn}^j(U) \mathcal{D}_{m'n'}^{j'*}(U) = \delta_{jj'} \delta_{mm'} \delta_{nn'} \quad (\text{E.1})$$

$$\sum_{jmn} (2j+1) \mathcal{D}_{mn}^j(U) \mathcal{D}_{mn}^{j*}(U') = 2\pi^2 \delta(U, U').$$

La fonction $\delta(U, U')$ qui apparaît dans le second membre de (E.1) est celle adaptée à la mesure $d\mu(U)$, telle que $\int d\mu(U') \delta(U, U') f(U') = f(U)$; dans les angles d'Euler α, β, γ par exemple,

$$\delta(U, U') = 8\delta(\alpha - \alpha') \delta(\cos \beta - \cos \beta') \delta(\gamma - \gamma'), \quad (\text{E.2})$$

(voir Appendice C du Chap. 1). La signification de la seconde équation (E.1) est que les fonctions $\mathcal{D}_{mn}^j(U)$ forment une base complète sur l'espace des fonctions (continues ou de carré intégrable) sur le groupe $SU(2)$. C'est le théorème de Peter-Weyl, qui généralise donc le théorème de Fourier. Les caractères des représentations de $SU(2)$ se déduisent des expressions précédentes

$$\begin{aligned} \chi_j(U) = \chi_j(\psi) &= \text{tr } \mathcal{D}^j(\mathbf{n}, \psi) = \sum_{m=-j}^j e^{im\psi} \\ &= \frac{\sin\left(\frac{2j+1}{2}\psi\right)}{\sin\frac{\psi}{2}}. \end{aligned} \quad (\text{E.3})$$

Noter que ces expressions sont des polynômes (dits de Tchebichev de 2ème espèce, Chebyshev dans la transcription anglo-saxonne) de la variable $2 \cos \frac{\psi}{2}$ (voir l'exercice D en fin de chapitre). En particulier

$$\chi_0(\psi) = 1 \quad \chi_{\frac{1}{2}}(\psi) = 2 \cos \frac{\psi}{2} \quad \chi_1(\psi) = 1 + 2 \cos \psi \quad \text{etc.} \quad (\text{E.4})$$

On est alors en mesure de vérifier toutes les propriétés attendues

unitarité et réalité	$\chi_j(U^{-1}) = \chi_j^*(U) = \chi_j(U)$	
parité et périodicité	$\chi_j(-U) = \chi_j(2\pi + \psi) = (-1)^{2j} \chi_j(U)$	(E.5)
orthogonalité	$\int_0^{2\pi} d\psi \sin^2 \frac{\psi}{2} \chi_j(\psi) \chi_{j'}(\psi) = \pi \delta_{jj'}$	
complétude	$\sum_{j=0, \frac{1}{2}, \dots} \chi_j(\psi) \chi_j(\psi') = \frac{\pi}{\sin^2 \frac{\psi}{2}} \delta(\psi - \psi') = \frac{\pi}{2 \sin \frac{\psi}{2}} \delta(\cos \frac{\psi}{2} - \cos \frac{\psi'}{2})$	

La dernière exprime que les caractères forment une base complète des fonctions de classe, c'est-à-dire des fonctions paires périodiques de $\frac{1}{2}\psi$. On retrouve là une variante du développement de Fourier.

Les formules de multiplicité (2.57) conduisent-elle bien aux formules connues (2.28)?

E.2. Fonctions spéciales. Harmoniques sphériques

On a déjà vu à plusieurs reprises qu'un générateur infinitésimal agissait dans chaque représentation comme un opérateur différentiel. Cette propriété est vraie en particulier dans le cas de SU(2) qui nous occupe ici : les générateurs J_i apparaissent comme des opérateurs différentiels par rapport aux paramètres de la rotation, cf. le cas d'un sous-groupe à un paramètre $\exp -iJ\psi$ où $J = i\partial/\partial\psi$. Cela va donner lieu à des équations différentielles satisfaites par les $\mathcal{D}_{m'm}^j$ et faire apparaître leur relation avec des "fonctions spéciales" de la Physique Mathématique.

On a déjà noté que la discussion des matrices \mathcal{D} de Wigner au Chap. 0 § 0.3.3 s'applique non seulement aux matrices de SU(2) mais aussi à des matrices quelconques $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ du groupe linéaire GL(2, \mathbb{C}). L'équation (0.70) du Chap. 0 est donc toujours vraie dans ce cas

$$P_{jm}(\xi', \eta') = \sum_{m'} P_{jm'}(\xi, \eta) \mathcal{D}_{m'm}^j(A). \quad (E.70)$$

La combinaison $(a\xi + c\eta)^{j+m}(b\xi + d\eta)^{j-m}$ satisfait évidemment

$$\left(\frac{\partial^2}{\partial a \partial d} - \frac{\partial^2}{\partial b \partial c} \right) (a\xi + c\eta)^{j+m}(b\xi + d\eta)^{j-m} = 0 \quad (E.6)$$

donc en raison de l'indépendance des $P_{jm}(\xi, \eta)$, les $\mathcal{D}_{m'm}^j(A)$ satisfont la même équation. Si maintenant on impose que $d = a^*$, $c = -b^*$, mais $\rho^2 = |a|^2 + |b|^2$ est arbitraire, les matrices A satisfont $AA^\dagger = \rho^2 I$, $\det A = \rho^2$, donc $A = \rho U$, $U \in \text{SU}(2)$, et (E.6) conduit à

$$\Delta_4 \mathcal{D}_{m'm}^j(A) = 4 \left(\frac{\partial^2}{\partial a \partial a^*} + \frac{\partial^2}{\partial b \partial b^*} \right) \mathcal{D}_{m'm}^j(A) = 0 \quad (E.7)$$

où Δ_4 est le laplacien dans l'espace \mathbb{R}^4 des variables u_0, \mathbf{u} , avec $a = u_0 + iu_3$, $b = u_1 + iu_2$. En coordonnées polaires,

$$\Delta_4 = \frac{\partial^2}{\partial \rho^2} + \frac{3}{\rho} \frac{\partial}{\partial \rho} + \frac{1}{\rho^2} \Delta_{S^3} \quad (E.8)$$

où le dernier terme, laplacien sur la sphère S^3 , n'agit que sur les "variables angulaires" $U \in \text{SU}(2)$ (voir Chap. 0, App. 0). Les fonctions \mathcal{D}^j étant homogènes de degré $2j$ en a, b, c, d donc en ρ , on a finalement

$$-\frac{1}{4} \Delta_{S^3} \mathcal{D}_{m'm}^j(U) = j(j+1) \mathcal{D}_{m'm}^j(U). \quad (E.9)$$

Par exemple, en utilisant la paramétrisation des angles d'Euler, on trouve que (voir (0.122))

$$\left\{ \frac{1}{\sin \beta} \frac{\partial}{\partial \beta} \sin \beta \frac{\partial}{\partial \beta} + \frac{1}{\sin^2 \beta} \left[\frac{\partial^2}{\partial \alpha^2} + \frac{\partial^2}{\partial \gamma^2} - 2 \cos \beta \frac{\partial^2}{\partial \alpha \partial \gamma} \right] + j(j+1) \right\} \mathcal{D}^j(\alpha, \beta, \gamma)_{m'm} = 0. \quad (E.10)$$

Pour $m = 0$ (donc j nécessairement entier), la dépendance en γ disparaît (cf. (00.3.14)). Choisissons par exemple $\gamma = 0$ et effectuons le changement de notations $(j, m') \rightarrow (l, m)$ et $(\beta, \alpha) \rightarrow (\theta, \phi)$, afin de retrouver des notations traditionnelles. L'équation se réduit à

$$\left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + l(l+1) \right] \mathcal{D}_{m0}^l(\phi, \theta, 0) = 0. \quad (E.11)$$

L'opérateur différentiel constitué des deux premiers termes est le laplacien Δ_{S^2} sur la sphère unité S^2 . L'équation (E.11) définit donc les *harmoniques sphériques* $Y_l^m(\theta, \phi)$ comme vecteurs propres du laplacien Δ_{S^2} . La normalisation correcte est que

$$\left[\frac{2l+1}{4\pi} \right]^{\frac{1}{2}} \mathcal{D}_{m0}^l(\phi, \theta, 0) = Y_l^{m*}(\theta, \phi) . \quad (\text{E.12})$$

• Introduisons encore les *polynômes et fonctions de Legendre* $P_l(u)$ et $P_l^m(u)$ définies pour l entier et $u \in [-1, 1]$ par

$$P_l(u) = \frac{1}{2^l l!} \frac{d^l}{du^l} (u^2 - 1)^l \quad (\text{E.13})$$

$$P_l^m(u) = (1 - u^2)^{\frac{1}{2}m} \frac{d^m}{du^m} P_l(u) \quad \text{pour } 0 \leq m \leq l . \quad (\text{E.14})$$

Les polynômes de Legendre $P_l(u)$ sont des polynômes orthogonaux sur l'intervalle $[-1, 1]$ avec le poids 1 : $\int_{-1}^1 du P_l(u) P_{l'}(u) = \frac{2}{2l+1} \delta_{ll'}$. Les premiers P_l sont

$$P_0 = 1 \quad P_1 = u \quad P_2 = \frac{1}{2}(3u^2 - 1) \quad P_3 = \frac{1}{2}(5u^3 - 3u) , \dots \quad (\text{E.15})$$

tandis que $P_l^0 = P_l$, $P_l^1 = (1 - u^2)^{\frac{1}{2}} P_l'$, etc. Les harmoniques sphériques sont reliées aux fonctions de Legendre $P_l^m(\cos \theta)$ (pour $m \geq 0$) par

$$Y_l^m(\theta, \phi) = (-1)^m \left[\frac{(2l+1)(l-m)}{4\pi(l+m)} \right]^{\frac{1}{2}} P_l^m(\cos \theta) e^{im\phi} \quad (\text{E.16})$$

donc

$$\mathcal{D}_{m0}^l(0, \theta, 0) = d_{m0}^l(\theta) = (-1)^m \left[\frac{(l-m)}{(l+m)} \right]^{\frac{1}{2}} P_l^m(\cos \theta) = \left(\frac{4\pi}{2l+1} \right)^{\frac{1}{2}} Y_l^{m*}(\theta, 0) . \quad (\text{E.17})$$

En particulier, $d_{00}^l(\theta) = P_l(\cos \theta)$. En général, $d_{m'm}^l(\theta)$ est relié au polynôme de Jacobi

$$P_l^{(\alpha, \beta)}(u) = \frac{(-1)^l}{2^l l!} (1-u)^{-\alpha} (1+u)^{-\beta} \frac{d^l}{du^l} [(1-u)^{\alpha+l} (1+u)^{\beta+l}] \quad (\text{E.18})$$

par

$$d_{m'm}^j(\theta) = \left[\frac{(j+m')!(j-m')!}{(j+m)!(j-m)!} \right]^{\frac{1}{2}} \left(\cos \frac{\theta}{2} \right)^{m+m'} \left(\sin \frac{\theta}{2} \right)^{m-m'} P_{j-m'}^{(m'-m, m'+m)}(\cos \theta) . \quad (\text{E.19})$$

Polynômes de Jacobi et de Legendre relèvent de la théorie générale des polynômes orthogonaux dont on montre qu'ils satisfont des relations de récurrence linéaires à trois termes. Ils satisfont en outre des équations différentielles. C'est ainsi que les polynômes de Jacobi sont orthogonaux pour la mesure

$$\int_{-1}^1 du (1-u)^\alpha (1+u)^\beta P_j^{(\alpha, \beta)}(u) P_{j'}^{(\alpha, \beta)}(u) = \delta_{jj'} \frac{2^{\alpha+\beta+1} \Gamma(l+\alpha+1) \Gamma(l+\beta+1)}{(2l+\alpha+\beta+1)! \Gamma(l+\alpha+\beta+1)} \quad (\text{E.20})$$

et satisfait la relation de récurrence

$$\begin{aligned} & 2(l+1)(l+\alpha+\beta+1)(2l+\alpha+\beta) P_{l+1}^{(\alpha, \beta)}(u) \\ &= (2l+\alpha+\beta+1)[(2l+\alpha+\beta)(2l+\alpha+\beta+2)u + \alpha^2 - \beta^2] P_l^{(\alpha, \beta)}(u) - 2(l+\alpha)(l+\beta)(2l+\alpha+\beta+2) P_{l-1}^{(\alpha, \beta)}(u) . \end{aligned} \quad (\text{E.21})$$

Le polynôme de Jacobi $P_l^{(\alpha,\beta)}(u)$ est solution de l'équation différentielle

$$\{(1-u^2)\frac{d^2}{du^2} + [\beta - \alpha - (2 + \alpha + \beta)u]\frac{d}{du} + l(l + \alpha + \beta + 1)\}P_l^{(\alpha,\beta)}(u) = 0. \quad (\text{E.22})$$

Les polynômes de Legendre correspondent au cas $\alpha = \beta = 0$. Ces relations apparaissent ici comme reliées à celles des \mathcal{D}^j . Cela est un phénomène général : de nombreuses fonctions spéciales (Bessel, etc) sont reliées à des matrices de représentations de groupes. La théorie des groupes permet donc de mettre dans une perspective géométrique des résultats de l'analyse classique.

• Revenons aux harmoniques sphériques et à leurs propriétés.

(i) Elles satisfont les équations différentielles

$$(\Delta_{S^2} + l(l+1))Y_l^m = 0 \quad (\text{E.23})$$

$$J_z Y_l^m = -i \frac{\partial}{\partial \phi} Y_l^m = m Y_l^m \quad (\text{E.24})$$

et peuvent s'écrire

$$Y_l^m(\theta, \phi) = \frac{(-1)^l}{2^l l!} \sqrt{\frac{(2l+1)(l+m)!}{4\pi(l-m)!}} e^{im\phi} \sin^{-m}\theta \left(\frac{d}{d\cos\theta}\right)^{l-m} \sin^{2l}\theta. \quad (\text{E.25})$$

(ii) Elles sont normalisées à 1 sur la sphère unité et plus généralement y satisfont des propriétés d'orthogonalité et de complétude

$$\int d\Omega Y_l^{m*} Y_{l'}^{m'} = \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin\theta Y_l^{m*} Y_{l'}^{m'} = \delta_{ll'} \delta_{mm'} \quad (\text{E.26})$$

$$\begin{aligned} \sum_{l=0}^{\infty} \sum_{m=-l}^l Y_l^{m*}(\theta, \phi) Y_l^m(\theta', \phi') &= \delta(\Omega - \Omega') = \frac{\delta(\theta - \theta') \delta(\phi - \phi')}{\sin\theta} \\ &= \delta(\cos\theta - \cos\theta') \delta(\phi - \phi') \end{aligned} \quad (\text{E.27})$$

(iii) On peut considérer $Y_l^m(\theta, \phi)$ comme fonction du vecteur unitaire \mathbf{n} d'angles directeurs θ, ϕ . Si le vecteur \mathbf{n} est transformé en \mathbf{n}' par la rotation R , on a

$$Y_l^m(\mathbf{n}') = Y_l^{m'}(\mathbf{n}) \mathcal{D}^l(R)_{m'm} \quad (\text{E.28})$$

ce qui exprime que les Y_l^m se transforment comme des vecteurs de la représentation de spin l .

(iv) On vérifie sur l'expression ci-dessus les relations de symétrie en m

$$Y_l^{m*}(\theta, \phi) = (-1)^m Y_l^{-m}(\theta, \phi) \quad (\text{E.29})$$

et de parité

$$Y_l^m(\pi - \theta, \phi + \pi) = (-1)^l Y_l^m(\theta, \phi). \quad (\text{E.30})$$

Noter qu'à $\theta = 0$, $Y_l^m(0, \phi)$ s'annule sauf pour $m = 0$, cf. (E.13, E.16).

(v) Les harmoniques sphériques satisfont aussi des relations de récurrence de deux types : celles issues de l'action de J_{\pm} , opérateurs différentiels qui agissent selon (0.116)

$$e^{\pm i\phi} \left[\pm \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right] Y_l^m = \sqrt{l(l+1) - m(m \pm 1)} Y_l^{m \pm 1} \quad (\text{E.31})$$

et celles provenant de la multiplication des représentations de spin 1 et l ,

$$\sqrt{2l+1} \cos \theta Y_l^m = \left(\frac{(l+m)(l-m)}{2l-1} \right)^{\frac{1}{2}} Y_{l-1}^m + \left(\frac{(l+m+1)(l-m+1)}{2l+3} \right)^{\frac{1}{2}} Y_{l+1}^m. \quad (\text{E.32})$$

On a plus généralement la formule de produit

$$Y_l^m(\theta, \phi) Y_{l'}^{m'}(\theta, \phi) = \sum_L \langle lm; l'm' | L, m+m' \rangle \left[\frac{(2l+1)(2l'+1)}{4\pi(2L+1)} \right]^{\frac{1}{2}} Y_L^{m+m'}(\theta, \phi). \quad (\text{E.33})$$

(vi) Finalement citons la très utile “formule d’addition”

$$\frac{2l+1}{4\pi} P_l(\cos \theta) = \sum_{m=-l}^l Y_l^m(\mathbf{n}) Y_l^{m*}(\mathbf{n}') \quad (\text{E.34})$$

où θ désigne l’angle entre les directions \mathbf{n} et \mathbf{n}' . Cette formule peut se vérifier en démontrant que le membre de droite satisfait bien les équations différentielles satisfaites par P_l (exercice 1 ci-dessous).

Exercices.

1. Démontrer que le polynôme de Legendre P_l vérifie

$$(\Delta_{S^2} + l(l+1)) P_l(\mathbf{n} \cdot \mathbf{n}') = 0$$

comme fonction de \mathbf{n} ou de \mathbf{n}' , ainsi que $(\mathbf{J} + \mathbf{J}') P_l = 0$ où \mathbf{J} et \mathbf{J}' sont les générateurs des rotations de \mathbf{n} et \mathbf{n}' respectivement. En déduire qu’il a un développement sur les harmoniques sphériques donné par le théorème d’addition de (E.34) (On rappelle que $P_l(1) = 1$).

2. Montrer qu’une fonction génératrice des polynômes de Legendre est

$$\frac{1}{\sqrt{1-2ut+t^2}} = \sum_{l=0}^{\infty} t^l P_l(u). \quad (\text{E.35})$$

On pourra vérifier que l’équation différentielle des P_l (cas particulier de (E.22) pour $\alpha = \beta = 0$) est bien satisfaite et que les coefficients P_l apparaissant dans cette formule sont bien des polynômes en u . En déduire l’identité (on suppose $r' < r$),

$$\frac{1}{|\vec{r} - \vec{r}'|} = \sum_{l=0}^{\infty} \frac{r'^l}{r^{l+1}} P_l(\cos \theta) = \sum_{l,m} \frac{4\pi}{2l+1} \frac{r'^l}{r^{l+1}} Y_l^{m*}(\mathbf{n}) Y_l^m(\mathbf{n}'). \quad (\text{E.36})$$

Les expressions des premiers Y_l^m peuvent être utiles

$$\begin{aligned} Y_0^0 &= \frac{1}{\sqrt{4\pi}} \\ Y_1^0 &= \sqrt{\frac{3}{4\pi}} \cos \theta & Y_1^{\pm 1} &= \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi} \\ Y_2^0 &= \sqrt{\frac{5}{16\pi}} (3 \cos^3 \theta - 1) & Y_2^{\pm 1} &= \mp \sqrt{\frac{15}{8\pi}} \cos \theta \sin \theta e^{\pm i\phi} & Y_2^{\pm 2} &= \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi}. \end{aligned} \quad (\text{E.37})$$

E.3. Applications physiques

E.3.1. Moments multipolaires

On considère un potentiel créé par une distribution de charge statique $\rho(\vec{r})$

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{d^3r' \rho(\vec{r}')}{|\vec{r} - \vec{r}'|}$$

et on le développe sur les harmoniques sphériques selon (E.36). Il vient

$$\phi(\vec{r}) = \frac{1}{\epsilon_0} \sum_{l,m} \frac{1}{2l+1} \frac{Y_l^{m*}(\mathbf{n})}{r^{l+1}} Q_{lm} \quad (\text{E.38})$$

où les Q_{lm} , définis par

$$Q_{lm} = \int d^3r' \rho(\vec{r}') r'^l Y_l^m(\mathbf{n}') \quad (\text{E.39})$$

sont les *moments multipolaires* de la distribution de charge ρ . Par exemple, si $\rho(\vec{r}) = \rho(r)$ est invariant par rotation, seul Q_{00} est non nul, égal à la charge totale (à $1/\sqrt{4\pi}$ près)

$$Q_{00} = \frac{Q}{\sqrt{4\pi}} = \sqrt{4\pi} \int r^2 dr \rho(r) \quad \phi(r) = \frac{Q}{4\pi\epsilon_0 r}.$$

Pour un $\rho(\vec{r})$ quelconque, les trois composantes de Q_{1m} reconstruisent le moment dipolaire $\int d^3r' \rho(\vec{r}') \vec{r}'$. Plus généralement, sous l'effet des rotations, les Q_{lm} forment les composantes d'un opérateur tensoriel se transformant selon la représentation de spin l (et cf. (E.30), de parité $(-1)^l$).

En Mécanique Quantique, les Q_{lm} deviennent des opérateurs. On peut leur appliquer le théorème de Wigner-Eckart et en conclure que

$$\langle j_1, m_1 | Q_{lm} | j_2, m_2 \rangle = \langle j_1 || Q_l || j_2 \rangle \langle j_1, m_1 | l, m; j_2, m_2 \rangle$$

avec un élément de matrice réduit indépendant des m . En particulier, si $j_1 = j_2 = j$, la valeur moyenne de Q_l n'est non nulle que pour $l \leq 2j$.

E.3.2. Etats propres de moment angulaire en Mécanique Quantique

Les harmoniques sphériques peuvent s'interpréter comme les fonctions d'onde dans les coordonnées θ, ϕ des états propres du moment angulaire $\vec{L} = \hbar \vec{J} = \hbar \vec{r} \wedge \vec{\nabla}$

$$Y_l^m(\theta, \phi) = \langle \theta, \phi | l, m \rangle$$

en analogie avec

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

(On a pris $\hbar = 1$.) En particulier, supposons que dans un processus de collision décrit par un Hamiltonien invariant par rotation, un état d'impulsion initiale \vec{p}_i selon l'axe des z , (c'est-à-dire

$\theta = \phi = 0$), interagit avec un certain centre diffuseur et ressort dans un état d'impulsion \vec{p}_f , avec $|p_i| = |p_f| = p$, selon la direction $\mathbf{n} = (\theta, \phi)$. On écrit l'amplitude

$$\begin{aligned}
 \langle p, \theta, \phi | \mathcal{T} | p, 0, 0 \rangle &= \sum_{l'm'm'} Y_l^m(\theta, \phi) \langle p, l, m | \mathcal{T} | p, l', m' \rangle Y_{l'}^{m'*}(0, 0) \\
 &= \sum_{lm} Y_l^m(\theta, \phi) \langle p, l, m | \mathcal{T} | p, l, m \rangle Y_l^{m*}(0, 0) \\
 &= \sum_l \frac{2l+1}{4\pi} \mathcal{T}_l(p) P_l(\cos \theta)
 \end{aligned} \tag{E.40}$$

selon à nouveau la formule d'addition et $\langle plm | \mathcal{T} | pl'm' \rangle = \delta_{l'l} \delta_{mm'} \mathcal{T}_l(p)$ exprimant l'invariance par rotation. C'est le *développement en ondes partielles* de l'amplitude de diffusion, très utile dans l'analyse des résultats expérimentaux et dans la modélisation.

★

Exercices et problèmes du chapitre 2

A. Représentations unitaires d'un groupe simple

Soit G un groupe simple non abélien. Soit D une représentation unitaire de G .

1. Montrer que $\det D$ est une représentation de dimension 1 du groupe, et que c'est un homomorphisme du groupe dans le groupe $U(1)$.
2. Que peut-on dire du noyau K de cet homomorphisme? Montrer que tout "commutateur" $g_1 g_2 g_1^{-1} g_2^{-1}$ appartient à K et que K ne peut donc être trivial.
3. En conclure que la représentation est unimodulaire (de déterminant 1).
4. Peut-on appliquer cet argument à $SO(3)$? à $SU(2)$?

B. Représentation adjointe

1. Montrer que si l'algèbre de Lie \mathfrak{g} d'un groupe de Lie G est simple, la représentation adjointe de G est irréductible.
2. Montrer que si \mathfrak{g} est semi-simple, sa représentation adjointe est fidèle : $\ker \text{ad} = 0$.

C. Produit tensoriel $D \otimes D^*$

Soit G un groupe compact. Soient $D^{(\rho)}$ ses représentations irréductibles. On note $D^{(1)}$ la représentation identité, $D^{(\bar{\rho})}$ la représentation conjuguée de $D^{(\rho)}$.

Quelle est la multiplicité de $D^{(1)}$ dans la décomposition en représentations irréductibles de $D^{(\rho)} \otimes D^{(\bar{\rho})}$?

D. Polynômes de Tchebichev

Soit l'expression

$$U_l = \frac{\sin(l+1)\theta}{\sin \theta},$$

où l est un entier ≥ 0 .

1. Par un calcul trigonométrique élémentaire, exprimer $U_{l-1} + U_{l+1}$ en fonction de U_l , avec un coefficient indépendant de l .
2. En déduire que U_l est un polynôme en $z = 2 \cos \theta$ de degré l , qu'on notera $U_l(z)$.
3. Quelle est l'interprétation groupiste du résultat obtenu en 1)?
4. Avec le minimum de calculs supplémentaires, que peut-on dire de

$$\frac{2}{\pi} \int_{-1}^1 dz (1-z^2)^{\frac{1}{2}} U_l(z) U_{l'}(z)$$

et

$$\frac{2}{\pi} \int_{-1}^1 dz (1-z^2)^{\frac{1}{2}} U_l(z) U_{l'}(z) U_{l''}(z) ?$$

Les $U_l(z)$ sont les polynômes de Tchebichev (Chebyshev dans la transcription anglo-saxonne) de 2ème espèce. Ils sont orthogonaux (la première des relations de la question 4) et satisfont une relation de récurrence à trois termes (question 1), qui sont deux propriétés générales des polynômes orthogonaux.

E. Harmoniques sphériques

Montrer que l'intégrale

$$\int d\Omega Y_{l_1}^{m_1}(\theta, \phi) Y_{l_2}^{m_2}(\theta, \phi) Y_{l_3}^{m_3}(\theta, \phi)$$

est proportionnelle au coefficient de Clebsch-Gordan $(-1)^{m_3} \langle l_1, m_1; l_2, m_2 | l_3, -m_3 \rangle$, avec un coefficient indépendant des m qu'on déterminera.

Problème I. Décomposition d'une amplitude

On considère deux représentations réelles et unitaires (ρ) et (σ) d'un groupe de Lie simple compact G de dimension d . On note $|\rho, \alpha\rangle$, resp. $|\sigma, \beta\rangle$, deux bases de ces représentations, et $T_{\alpha\alpha'}^{(\rho)a}$, resp. $T_{\beta\beta'}^{(\sigma)a}$, $a = 1, \dots, d$, les matrices de représentation d'une base de l'algèbre de Lie orthonormée pour la métrique de Killing. Ces matrices sont supposées **antisymétriques** réelles et satisfont donc $\text{tr } T^a T^b = -\delta_{ab}$. On va s'intéresser à la quantité

$$X_{\alpha\beta; \alpha'\beta'} := \sum_{a=1}^d T_{\alpha\alpha'}^{(\rho)a} T_{\beta\beta'}^{(\sigma)a}. \quad (2.89)$$

Pour simplifier l'écriture, on supposera que toutes les représentations apparaissant dans le produit tensoriel des représentations (ρ) et (σ) sont réelles et sans multiplicité. Soit $|\tau\gamma\rangle$ une base d'une telle représentation. On introduit alors les coefficients de Clebsch-Gordan (réels) qu'on écrit comme des matrices

$$\left(\mathcal{M}^{(\tau\gamma)}\right)_{\alpha\beta} = \langle \tau\gamma | \rho\alpha; \sigma\beta \rangle. \quad (2.90)$$

1. Rappeler pourquoi ces coefficients satisfont des propriétés d'orthogonalité et de complétude qu'on écrira.
2. En déduire qu'on peut écrire

$$X_{\alpha\beta; \alpha'\beta'} = - \sum_{\tau\gamma} \left(\mathcal{M}^{(\tau\gamma)}\right)_{\alpha\beta} \left(T^{(\rho)a} \mathcal{M}^{(\tau\gamma)} T^{(\sigma)a}\right)_{\alpha'\beta'}. \quad (2.91)$$

3. En faisant agir le générateur infinitésimal T^a sur les deux membres de la relation

$$|\rho\alpha; \sigma\beta\rangle = \sum_{\tau, \gamma} \left(\mathcal{M}^{(\tau\gamma)}\right)_{\alpha\beta} |\tau\gamma\rangle \quad (2.92)$$

montrer qu'on obtient

$$\sum_{\gamma'} T_{\gamma\gamma'}^{(\tau)a} \left(\mathcal{M}^{(\tau\gamma')}\right)_{\alpha\beta} = \sum_{\alpha'} \left(\mathcal{M}^{(\tau\gamma)}\right)_{\alpha'\beta} (T^{(\rho)a})_{\alpha'\alpha} + \sum_{\beta'} \left(\mathcal{M}^{(\tau\gamma)}\right)_{\alpha\beta'} (T^{(\sigma)a})_{\beta'\beta} \quad (2.93)$$

ou encore en termes de matrices de dimensions $\dim(\rho) \times \dim(\sigma)$

$$\sum_{\gamma'} T_{\gamma\gamma'}^{(\tau)a} \mathcal{M}^{(\tau\gamma')} = -T^{(\rho)a} \mathcal{M}^{(\tau\gamma)} + \mathcal{M}^{(\tau\gamma)} T^{(\sigma)a}. \quad (2.94)$$

4. En utilisant de façon répétée cette relation (2.94) dans (2.91), montrer qu'on a

$$X_{\alpha\beta; \alpha'\beta'} = \frac{1}{2} \sum_{\tau\gamma} (C_\rho + C_\sigma - C_\tau) \left(\mathcal{M}^{(\tau\gamma)}\right)_{\alpha\beta} \left(\mathcal{M}^{(\tau\gamma)}\right)_{\alpha'\beta'} \quad (2.95)$$

où les C sont les opérateurs de Casimir quadratiques, par exemple

$$C_\rho = - \sum_a (T^{(\rho)a})^2. \quad (2.96)$$

5. Pourquoi peut-on dire que les "grandes représentations" τ tendent à rendre le coefficient $(C_\rho + C_\sigma - C_\tau)$ de plus en plus négatif? On pourra prendre l'exemple de $SU(2)$ avec ρ et σ deux représentations de spin (entier) égal à j .
6. Pouvez-vous imaginer une théorie de champs dans laquelle le coefficient $X_{\alpha\beta; \alpha'\beta'}$ apparaîtrait dans une amplitude de diffusion à deux corps (à l'approximation en arbres)? Quelle conséquence la propriété discutée aurait-elle sur cette amplitude?

A. Représentations réelles et quaternioniques

1. Soit G un groupe compact. Si $D(g)$ est une représentation de G , montrer que $D^{-1T}(g)$ en est une autre, qu'on appelle la représentation *contragrédiente*.
2. Rappeler sommairement pourquoi on peut toujours supposer que les représentations de G sont unitaires, ce que l'on fera dans la suite.
3. Montrer que la représentation contragrédiente s'identifie alors à la représentation complexe conjuguée.
4. On suppose que la représentation unitaire D est (unitairement) équivalente à sa contragrédiente (ou sa conjuguée). Montrer qu'il existe une matrice S (unitaire) telle que

$$D = SD^{-1T}S^{-1} \quad (2.97)$$

5. Montrer que (2.97) signifie que la forme bilinéaire S est invariante.
6. Cette forme est-elle dégénérée ?
7. Montrer en manipulant (2.97) que

$$DSS^{-1T} = SS^{-1T}D. \quad (2.98)$$

8. Montrer alors que si D est irréductible, $S = \lambda S^T$, avec $\lambda^2 = 1$.
9. En conclure que la forme invariante S est soit symétrique soit antisymétrique.
Dans le premier cas (symétrique), la représentation est dite réelle, dans le second (S antisymétrique), elle est dite pseudoréelle (ou quaternionique). On peut montrer que dans le premier cas, il existe une base sur \mathbb{R} dans laquelle les matrices de la représentation sont réelles, et qu'il n'en existe pas dans le second.
10. Connaissez-vous un exemple du second cas ?

B. Indicatrice de Frobenius–Schur

1. Soit G un groupe fini ou un groupe de Lie compact. On repère ses représentations irréductibles par un indice ρ et on note $\chi^{(\rho)}(g)$ leur caractère. Soit $\chi(g)$ le caractère d'une représentation arbitraire, réductible ou non.

(a) Pour toute fonction F sur le groupe fini G , on note $\langle F \rangle$ sa moyenne

$$\langle F \rangle = \frac{1}{|G|} \sum_{g \in G} F(g). \quad (2.99)$$

Par quoi faut-il remplacer cette définition dans le cas d'un groupe de Lie compact (et d'une fonction F continue) ?

- (b) - Rappeler pourquoi $\langle \chi \rangle$ est un entier et ce qu'il vaut.
- Si $\bar{\rho}$ dénote la représentation conjuguée de la représentation irréductible ρ , rappeler pourquoi $\langle \chi^{(\rho)} \chi^{(\bar{\rho})} \rangle = 1$ et ce que cela implique sur la décomposition en représentations irréductibles de $\rho \otimes \bar{\rho}$.
- (c) Montrer qu'une représentation irréductible ρ est équivalente à $\bar{\rho}$ si et seulement si

$$\left\langle \left(\chi^{(\rho)}(g) \right)^2 \right\rangle = 1.$$

Que vaut cette expression si ρ n'est pas équivalente à $\bar{\rho}$?

2. On considère la représentation $D^{(\rho)}$ agissant dans un espace E , et son carré tensoriel $D^{(\rho) \otimes 2}$, qui agit sur les tenseurs de rang 2 de $E \otimes E$.

- (a) Écrire explicitement l'action de $D^{(\rho)\otimes 2}$ sur un tenseur $t = \{t^{ij}\}$,

$$t^{ij} \mapsto t'^{ij} = \dots$$

- (b) Montrer que tout tenseur de rang 2, $t = \{t^{ij}\}$, est la somme d'un tenseur t_S symétrique et d'un tenseur t_A antisymétrique dans leurs deux indices, se transformant selon des représentations indépendantes. Écrire explicitement les matrices de transformation de t_S et t_A en veillant bien aux propriétés de symétrie des objets considérés.

- (c) Montrer que les caractères des représentations des tenseurs symétriques et antisymétriques sont respectivement

$$\chi^{(\rho\otimes\rho)_S^A}(g) = \frac{1}{2} \left((\chi^{(\rho)}(g))^2 \pm \chi^{(\rho)}(g^2) \right). \quad (2.100)$$

- (d) Que valent ces caractères pour $g = e$, l'identité dans le groupe ? Ces résultats étaient-ils prévisibles ?

3. On définit alors l'indicatrice de Frobenius-Schur de la représentation irréductible ρ par

$$\text{ind}(\rho) = \left\langle \chi^{(\rho)}(g^2) \right\rangle. \quad (2.101)$$

- (a) Montrer en utilisant les résultats du 2. que l'on peut écrire

$$\text{ind}(\rho) = \langle \chi^{(\rho\otimes\rho)_S} \rangle - \langle \chi^{(\rho\otimes\rho)_A} \rangle.$$

- (b) Montrer en utilisant les résultats du 1. que

$$\langle (\chi^{(\rho)}(g))^2 \rangle = \langle \chi^{(\rho\otimes\rho)_S} \rangle + \langle \chi^{(\rho\otimes\rho)_A} \rangle$$

prend la valeur 0 ou 1, selon des cas que l'on précisera

- (c) - Montrer que $\langle \chi^{(\rho\otimes\rho)_S} \rangle$ et $\langle \chi^{(\rho\otimes\rho)_A} \rangle$ sont des entiers non négatifs, et qu'ils fournissent une multiplicité que l'on précisera.

- Montrer que finalement l'indicatrice de Frobenius-Schur (2.101) ne peut prendre que les trois valeurs 0 et ± 1 selon des cas que l'on précisera.

- (d) Commentez la relation entre cette discussion et celle de l'exercice A.

4. \star On se restreint au cas d'un groupe fini. Pour tout $h \in G$, on définit $Q(h) := \sum_{\rho} \text{ind}(\rho) \chi^{(\rho)}(h)$. Démontrer le

Théorème. $Q(h) = \#\{g \in G | g^2 = h\}$

Chapitre 3

Algèbres de Lie simples, classification et représentations

3.1 Sous-algèbre de Cartan. Racines. Forme canonique de l'algèbre.

On considère une algèbre de Lie \mathfrak{g} semi-simple (pas d'idéal abélien). On se propose de construire une forme canonique des relations de commutation calquée sur le cas de $SU(2)$

$$[J_z, J_{\pm}] = \pm J_{\pm} \quad [J_+, J_-] = 2J_z . \quad (3.1)$$

Il sera important de considérer l'algèbre sur \mathbb{C} , c'est-à-dire d'utiliser \mathbb{C} comme corps de nombres (au prix de la complexifier si elle était réelle). La représentation adjointe va être utilisée. Comme elle est fidèle pour une algèbre semi-simple ($\text{ad } X = 0 \Rightarrow X = 0$), cf exercice B du Chap. 2, on ne perd pas d'information.

Il peut être utile de se rappeler que l'algèbre complexe a une version réelle compacte dans laquelle les constantes de structure réelles conduisent à une forme de Killing définie négative, et, les représentations γ étant unitarisables, les éléments de l'algèbre de Lie (générateurs infinitésimaux) peuvent être considérés comme hermitiens soit comme antihermitiens, selon nos besoins.

3.1.1 Sous-algèbre de Cartan

On définit d'abord la notion de *sous-algèbre de Cartan*. On appelle ainsi une *sous-algèbre abélienne maximale* de \mathfrak{g} telle que tous ses éléments sont diagonalisables (donc simultanément diagonalisables) dans la représentation adjointe, donc dans toute représentation. Le fait qu'une telle algèbre existe est non trivial et doit être établi, mais nous l'admettrons.

Si on choisit de travailler avec la version réelle unitaire de la représentation adjointe, les éléments de \mathfrak{g} sont des matrices hermitiennes, et les matrices de \mathfrak{h} qui commutent entre elles sont simultanément diagonalisables.

Cette algèbre de Cartan n'est pas unique, mais on démontre que deux choix distincts sont reliés par un automorphisme de l'algèbre.

Ainsi si \mathfrak{g} est l'algèbre de Lie d'un groupe de Lie G et si \mathfrak{h} est une sous-algèbre de Cartan de \mathfrak{g} , tout conjugué $g\mathfrak{h}g^{-1}$ de \mathfrak{h} par un élément quelconque g de G est aussi une sous-algèbre de Cartan.

Soit \mathfrak{h} une telle sous-algèbre de Cartan. Soit ℓ sa dimension, elle est indépendante du choix de \mathfrak{h} et on l'appelle le *rang* de l'algèbre \mathfrak{g} . Pour $\mathfrak{su}(2)$, ce rang est 1, (le choix de J_z par exemple); pour $\mathfrak{su}(n)$, le rang est $n - 1$. En effet, pour $\mathfrak{su}(n)$, une algèbre de Cartan est engendrée¹ par les matrices diagonales de trace nulle. Une base en est donnée par les $n - 1$ matrices

$$H_1 = \text{diag}(1, -1, 0, \dots, 0), H_2 = \text{diag}(0, 1, -1, 0, \dots, 0), \dots, H_{n-1} = \text{diag}(0, \dots, 0, 1, -1) . \quad (3.2)$$

Une matrice quelconque de l'algèbre de Lie (dans cette représentation), (anti-)hermitienne et de trace nulle, est diagonalisable par une transformation unitaire; sa forme diagonale est de trace nulle et s'exprime donc comme combinaison linéaire des h_j ; la matrice de départ est donc conjuguée, par une transformation unitaire, d'une combinaison linéaire des h_j . Cette propriété est générale, et on démontre (Cartan, cf [Bu], chapitre 16) que

Si \mathfrak{g} est l'algèbre de Lie du groupe G , tout élément de \mathfrak{g} est conjugué par G d'un élément de \mathfrak{h} .

Application. Forme canonique des matrices antisymétriques. En utilisant ce théorème, démontrer la

Proposition *Si $A = A^* = -A^T$ est une matrice réelle antisymétrique de dimension N , alors on peut trouver une matrice orthogonale réelle O telle que $A = ODO^T$ où $D = \text{diag}\left(\begin{pmatrix} 0 & \mu_j \\ -\mu_j & 0 \end{pmatrix}_{j=1, \dots, n}\right)$ si la dimension*

$N = 2n$ et $D = \text{diag}\left(0, \begin{pmatrix} 0 & \mu_j \\ -\mu_j & 0 \end{pmatrix}_{j=1, \dots, n}\right)$ si $N = 2n + 1$, avec des μ_j réels.

Si on s'autorise à complexifier les matrices orthogonales, on peut complètement diagonaliser la matrice A sous la forme $D = \text{diag}\left(\begin{pmatrix} i\mu_j & 0 \\ 0 & -i\mu_j \end{pmatrix}_{j=1, \dots, n}\right)$ ou $D = \text{diag}\left(0, \begin{pmatrix} i\mu_j & 0 \\ 0 & -i\mu_j \end{pmatrix}_{j=1, \dots, n}\right)$. Pour une démonstration ne faisant appel qu'à la théorie des matrices, voir par exemple [M.L. Mehta, *Elements of Matrix Theory*, p 41].

3.1.2 Base canonique de l'algèbre de Lie

Soit $H_i, i = 1, \dots, \ell$ une base de \mathfrak{h} . Il est commode de choisir les $\text{ad } H_i$ hermitiens. Par définition $[H_i, H_j] = 0$ (sous-algèbre abélienne), ou plus précisément puisqu'on est dans la représentation adjointe,

$$[\text{ad } H_i, \text{ad } H_j] = 0 . \quad (3.3)$$

On peut diagonaliser simultanément ces $\text{ad } H_i$. On en connaît déjà des vecteurs propres de valeur propre nulle puisque $\forall i, j, \text{ad } H_i H_j = 0$, et on peut trouver un ensemble de vecteurs propres E_α indépendants des H_j

$$\text{ad } H_i E_\alpha = \alpha_{(i)} E_\alpha \quad (3.4)$$

c'est-à-dire un ensemble d'éléments de \mathfrak{g} tels que

$$[H_i, E_\alpha] = \alpha_{(i)} E_\alpha , \quad (3.5)$$

1. Nous nous plaçons momentanément dans la représentation "de définition" (matrices $n \times n$) et non plus dans la représentation adjointe.

avec des $\alpha_{(i)}$ pas tous nuls (sans quoi la sous-algèbre abélienne \mathfrak{h} ne serait pas maximale).

L'espace \mathfrak{h}^* . Dans ces expressions, les $\alpha_{(i)}$ sont les valeurs propres des opérateurs $\text{ad } H_i$. Puisqu'on a choisi les $\text{ad } H_i$ hermitiens, leurs valeurs propres $\alpha_{(i)}$ sont réelles. Par combinaison linéaire, pour un élément arbitraire de \mathfrak{h} écrit $H = \sum_i h^i H_i$,

$$\text{ad } H E_\alpha = \alpha(H) E_\alpha, \quad (3.6)$$

où la valeur propre de $\text{ad } H$ sur E_α est $\alpha(H) := \sum_i h^i \alpha_{(i)}$, qui est une *forme linéaire* sur \mathfrak{h} . (Rappel : d'une façon générale, les formes linéaires sur un espace vectoriel E forment un espace vectoriel E^* , appelé l'*espace dual* de l'espace E .) On peut donc considérer la *racine* α de composantes $\alpha_{(i)}$ comme un vecteur de l'espace dual de \mathfrak{h} , donc $\alpha \in \mathfrak{h}^*$. Noter que $\alpha(H_i) = \alpha_{(i)}$.

Les racines jouissent des propriétés suivantes

- (i) si α est une racine, $-\alpha$ en est une autre ;
- (ii) l'espace propre correspondant à la valeur propre α est de dimension 1 (pas de multiplicité) ;
- (iii) si α est une racine, les seules racines de la forme $\lambda\alpha$ sont $\pm\alpha$;
- (iv) les racines α engendrent tout l'espace dual \mathfrak{h}^* . (voir exercice).

Pour une preuve des points (i–iii), voir plus bas, pour (iv), voir Exercice A.

Nombre de racines. Les matrices H_j ayant été supposées diagonalisables, le nombre de leurs vecteurs propres E_α plus celui de leurs vecteurs propres H_i de valeur propre nulle doit être égal à la dimension de l'espace, ici la dimension d de la représentation adjointe (donc la dimension de l'algèbre \mathfrak{g}). Comme toute racine non nulle vient accompagnée de son opposée, le nombre de racines α non nulles est pair et égal à $d - \ell$. On note Δ l'ensemble des racines.

Dans la base $\{H_i, E_\alpha\}$ de \mathfrak{g} , la forme de Killing prend une forme simple

$$(H_i, E_\alpha) = 0 \quad (E_\alpha, E_\beta) = 0 \quad \text{sauf si } \alpha + \beta = 0. \quad (3.7)$$

En effet $(H, [H', E_\alpha]) = \alpha(H')(H, E_\alpha)$, mais aussi, en utilisant la définition de la forme de Killing et la cyclicité de la trace

$$(H, [H', E_\alpha]) = \text{tr}(\text{ad } H [\text{ad } H', \text{ad } E_\alpha]) = \text{tr}([\text{ad } H, \text{ad } H'] \text{ad } E_\alpha) = 0 \quad (3.8)$$

puisque $[\text{ad } H, \text{ad } H'] = 0$. Il en découle que $\forall H, H' \in \mathfrak{h}$, $\alpha(H')(H, E_\alpha) = 0$, donc que $(H, E_\alpha) = 0$. De même

$$([H, E_\alpha], E_\beta) = \alpha(H)(E_\alpha, E_\beta) = -(E_\alpha, [H, E_\beta]) = -\beta(H)(E_\alpha, E_\beta) \quad (3.9)$$

à nouveau par la cyclicité de la trace, et donc $(E_\alpha, E_\beta) = 0$ si $\exists H : (\alpha + \beta)(H) \neq 0$, c'est-à-dire si $\alpha + \beta \neq 0$.

Noter que le point (i) découle simplement de (3.7) : si $-\alpha$ n'était pas racine, E_α serait orthogonal à tous les éléments de la base donc à tout élément de \mathfrak{g} , et la forme serait dégénérée, contrairement à l'hypothèse de semi-simplicité. Pour une démonstration élégante des points (ii) et (iii), voir [OR, p. 29].

La restriction de cette forme à la sous-algèbre de Cartan \mathfrak{h} est non-dégénérée, sans quoi on aurait $\exists H \in \mathfrak{h}$, $\forall H' \in \mathfrak{h} : (H, H') = 0$, mais $(H, E_\alpha) = 0$, autrement dit $\forall X \in \mathfrak{g}$, on aurait $(H, X) = 0$ donc la forme serait dégénérée sur \mathfrak{g} , contrairement à l'hypothèse de semi-simplicité et à un des théorèmes de Cartan (cf. Chap. 1, § 4.4). La forme de Killing étant non-dégénérée sur \mathfrak{h} , elle induit un isomorphisme entre \mathfrak{h} et \mathfrak{h}^* : à $\alpha \in \mathfrak{h}^*$ on associe l'unique $H_\alpha \in \mathfrak{h}$ tel que

$$\forall H \in \mathfrak{h} \quad (H_\alpha, H) := \alpha(H), \quad (3.10)$$

et $\alpha_{(i)} = \alpha(H_i) = (H_\alpha, H_i)$. (Autrement dit on résout le système linéaire $g_{ij}h_\alpha^j = \alpha_{(i)}$ qui est de Cramer puisque $g_{ij} = (H_i, H_j)$ est inversible.) On a aussi une forme bilinéaire sur \mathfrak{h}^* héritée de la forme de Killing

$$\langle \alpha, \beta \rangle := (H_\alpha, H_\beta) , \quad (3.11)$$

dont nous allons faire usage au § 2 pour étudier la géométrie du système de racines.

Il reste à découvrir les relations de commutation (les crochets) des E_α entre eux. En utilisant l'identité de Jacobi, on trouve que

$$\text{ad } H_i[E_\alpha, E_\beta] = [H_i, [E_\alpha, E_\beta]] = [E_\alpha, [H_i, E_\beta]] - [E_\beta, [H_i, E_\alpha]] = (\alpha + \beta)_{(i)}[E_\alpha, E_\beta] . \quad (3.12)$$

En invoquant l'absence de multiplicité, on voit que trois cas se présentent. Si $\alpha + \beta$ est une racine, $[E_\alpha, E_\beta]$ est proportionnel à $E_{\alpha+\beta}$, avec un coefficient de proportionnalité $N_{\alpha\beta}$ dont on montrera plus bas (§ 2.1 et exercice C) qu'il est non nul. Si $\alpha + \beta \neq 0$ n'est pas une racine, $[E_\alpha, E_\beta]$ doit s'annuler. Enfin si $\alpha + \beta = 0$, $[E_\alpha, E_{-\alpha}]$ est un vecteur propre de tous les $\text{ad } H_i$ de valeur propre nulle, donc $[E_\alpha, E_{-\alpha}] = H \in \mathfrak{h}$. Pour déterminer cet H , procédons comme dans (3.9)

$$\begin{aligned} (H_i, [E_\alpha, E_{-\alpha}]) &= \text{tr}(\text{ad } H_i [\text{ad } E_\alpha, \text{ad } E_{-\alpha}]) = \text{tr}([\text{ad } H_i, \text{ad } E_\alpha] \text{ad } E_{-\alpha}) \\ &= \alpha_{(i)}(E_\alpha, E_{-\alpha}) = (H_i, H_\alpha)(E_\alpha, E_{-\alpha}) \end{aligned} \quad (3.13)$$

donc

$$[E_\alpha, E_{-\alpha}] = (E_\alpha, E_{-\alpha})H_\alpha . \quad (3.14)$$

Pour récapituler, nous avons construit une base canonique de l'algèbre \mathfrak{g}

$$\boxed{\begin{aligned} [H_i, H_j] &= 0 \\ [H_i, E_\alpha] &= \alpha_{(i)}E_\alpha \\ [E_\alpha, E_\beta] &= \begin{cases} N_{\alpha\beta}E_{\alpha+\beta} & \text{si } \alpha + \beta \text{ est une racine} \\ (E_\alpha, E_{-\alpha})H_\alpha & \text{si } \alpha + \beta = 0 \\ 0 & \text{sinon} \end{cases} \end{aligned}} \quad (3.15)$$

Jusqu'à ce point, nous n'avons pas fixé la normalisation des H_i et des E_α . Il est usuel de choisir, en accord avec (3.7)

$$\boxed{(H_i, H_j) = \delta_{ij} \quad (E_\alpha, E_\beta) = \delta_{\alpha+\beta,0} .} \quad (3.15)$$

(En effet la restriction de la forme de Killing à \mathfrak{h} , après la multiplication par i à laquelle on a procédé pour rendre les $\text{ad } H_i$ hermitiens, est définie positive.) Avec cette normalisation, H_α défini plus haut par (3.10) satisfait aussi

$$H_\alpha = \alpha.H := \alpha_{(i)}H_i . \quad (3.16)$$

operators”) sur un vecteur propre E_β donné. On a vu que si α et β sont deux racines avec $\alpha + \beta \neq 0$, il peut se faire que $\beta \pm \alpha$ soient aussi des racines. Soit $p \leq 0$ le plus petit entier tel que $(\text{ad } E_{-\alpha})^{|p|} E_\beta$ soit non nul, donc que $\beta + p\alpha$ soit racine, et soit $q \geq 0$ le plus grand entier tel que $(\text{ad } E_\alpha)^q E_\beta$ soit non nul, donc que $\beta + q\alpha$ soit racine. On appelle l’ensemble de racines $\{\beta + p\alpha, \beta + (p+1)\alpha, \dots, \beta, \dots, \beta + q\alpha\}$ la α -chaîne passant par β . Notons que les $E_{\beta'}$, quand β' parcourt cette chaîne, forment une base d’une représentation de dimension finie de l’algèbre $\text{su}(2)$ engendrée par H_α et $E_{\pm\alpha}$. D’après ce que nous savons de ces représentations de $\text{su}(2)$, les valeurs propres minimale et maximale de H_α sont opposées

$$\langle \alpha, \beta + p\alpha \rangle = -\langle \alpha, \beta + q\alpha \rangle$$

soit $2\langle \beta, \alpha \rangle = -(q+p)\langle \alpha, \alpha \rangle$, donc avec la notation (3.21)

$$m = -p - q. \quad (3.22)$$

Mais cette construction nous montre aussi que $\beta - m\alpha = \beta + (p+q)\alpha$ est dans la α -chaîne passant par β , (puisque $p \leq -m \leq q$), donc que c’est une racine.

Remarque. La discussion du § 3.1 a laissé les coefficients $N_{\alpha\beta}$ indéterminés. On démontre (cf Exercice B) en utilisant les relations de commutation des E le long d’une chaîne que les coefficients $N_{\alpha\beta}$ satisfont des relations non linéaires et qu’ils sont déterminés à des choix de signes près par la géométrie du système de racines selon

$$|N_{\alpha\beta}| = \sqrt{\frac{1}{2}(1-p)q\langle \alpha, \alpha \rangle}. \quad (3.23)$$

Noter que comme annoncé, $N_{\alpha\beta}$ s’annule si $q = 0$, c’est-à-dire si $\alpha + \beta$ n’est pas une racine.

Groupe de Weyl

Pour tout vecteur x dans l’espace des racines \mathfrak{h}^* , définissons la transformation linéaire

$$w_\alpha(x) = x - 2\frac{\langle \alpha, x \rangle}{\langle \alpha, \alpha \rangle} \alpha. \quad (3.24)$$

C’est une réflexion dans l’hyperplan orthogonal à α passant par l’origine : $(w_\alpha)^2 = I$, $w_\alpha(\alpha) = -\alpha$, et $w_\alpha(x) = x$ si x est orthogonal à α . C’est bien sûr une isométrie, préservant le produit scalaire : $\langle w_\alpha(x), w_\alpha(y) \rangle = \langle x, y \rangle$. On appelle w_α une réflexion de Weyl. Par définition le *groupe de Weyl* W est le groupe engendré par les w_α , c’est-à-dire l’ensemble de tous les produits possibles de w_α pour des racines α . D’après la remarque suivant (3.22), si α et β sont deux racines, $w_\alpha(\beta) = \beta - m\alpha$ est une racine. L’ensemble des racines est donc préservé par l’action du groupe de Weyl. Le groupe W est complètement déterminé par son action sur les racines, qui consiste donc à les permuter. C’est un sous-groupe du groupe de permutations de l’ensemble fini Δ , c’est donc un groupe fini².

2. Cette propriété n’est pas triviale : en général, quand on se donne m vecteurs dans l’espace euclidien \mathbb{R}^m , le groupe engendré par les réflexions dans les hyperplans qui leur sont orthogonaux est infini. Il faut des configurations bien particulières des vecteurs pour que le groupe soit fini. Les groupes de réflexion finis ont été classifiés par Coxeter. Les groupes de Weyl des algèbres simples forment un sous-ensemble des groupes de Coxeter.

Exemple : pour l'algèbre $\mathfrak{su}(n)$, on montre que $W = \mathcal{S}_n$, le groupe de permutations de n objets, cf plus bas § 3.2.

Signature d'un élément de W . Soit $w \in W$, écrit comme le produit de r réflexions élémentaires (3.24) : $w = w_{\alpha_r} \dots w_{\alpha_2} w_{\alpha_1}$. On définit sa *signature* comme $\text{sign}(w) := (-1)^r$. Cela généralise la notion familière dans le groupe $W = \mathcal{S}_n$, et l'on montre que cette définition est cohérente et indépendante de l'écriture de w comme produit.

Notons que si $\beta_+ = \beta + q\alpha$ est la plus haute racine dans la α -chaîne passant par β , et $\beta_- = \beta + p\alpha$ la plus basse, $w_\alpha(\beta_\pm) = \beta_\mp$ et plus généralement, les racines de la chaîne sont échangées deux par deux par l'action de w_α . La chaîne est donc invariante par w_α . Cela est une généralisation de la symétrie des "multiplets" de $\mathfrak{su}(2)$ $(-j, -j+1, \dots, j-1, j)$, et cela s'applique à toute α -chaîne passant par tout β , et donc à tout l'ensemble des racines. On en conclut que

L'ensemble des racines est invariant par W .

Racines positives, racines simples. Matrices de Cartan

Les racines ne sont pas indépendantes dans \mathfrak{h}^* . On montre qu'on peut subdiviser leur ensemble Δ en racines dites "positives" et racines "négatives", l'opposé d'une racine positive étant négative, et trouver une base $\alpha_i, i = 1, \dots, \ell$ de ℓ racines *simples*, telles que toute racine positive (resp. négative) s'exprime comme combinaison linéaire à coefficients entiers positifs ou nuls (resp. négatifs ou nuls) de ces racines simples. Une racine simple ne peut donc s'écrire comme somme de deux racines positives (pourquoi?).

Ni le choix d'un ensemble de racines positives, ni celui de la base de racines simples n'est unique. On passe d'un ensemble de racines simples à un autre par une opération du groupe de Weyl.

Si α et β sont des racines simples, $\alpha - \beta$ ne peut être une racine (pourquoi?). L'entier p dans la discussion précédente est donc nul et $m = -q \leq 0$. Il en découle que $\langle \alpha, \beta \rangle \leq 0$.

Le produit scalaire de deux racines simples est négatif ou nul. (3.25)

On définit alors la matrice de Cartan comme

$$C_{ij} = 2 \frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle}. \quad (3.26)$$

Attention, cette matrice n'est pas *a priori* symétrique³. Ses éléments diagonaux valent 2, ses éléments non diagonaux sont des entiers négatifs ou nuls.

Il faut se rappeler que le produit scalaire qui figure au numérateur de (3.26) est défini positif. Selon l'inégalité de Schwarz, $\langle \alpha, \beta \rangle^2 \leq \langle \alpha, \alpha \rangle \langle \beta, \beta \rangle$ avec égalité seulement si α et β sont colinéaires. Cette propriété, avec les propriétés d'intégrité de leurs éléments, suffit à classifier toutes les matrices de Cartan possibles, comme on va le voir.

Écrivons $\langle \alpha_i, \alpha_j \rangle = \|\alpha_i\| \|\alpha_j\| \cos \widehat{\alpha_i, \alpha_j}$. Notons d'abord qu'en multipliant ou en divisant les deux équations (3.21) pour la paire $\{\alpha_i, \alpha_j\}, i \neq j$, à savoir $C_{ij} = m_i \leq 0$ et $C_{ji} = m_j \leq 0$,

3. Attention aussi que certains auteurs appellent matrice de Cartan la transposée de (3.26)!

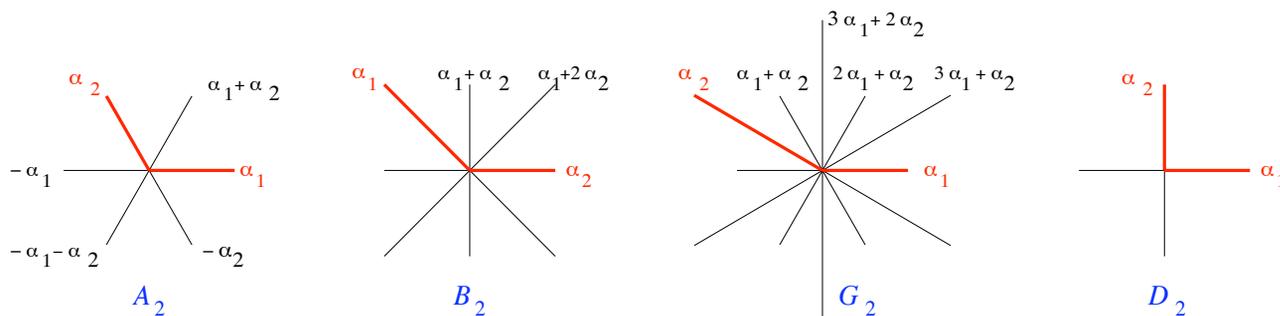


FIGURE 3.1 – Systèmes de racines de rang 2. Les deux racines simples positives sont en trait gras. Pour les algèbres B_2 , G_2 et D_2 , on n’a noté que les racines positives. The two racines simples are drawn in thick lines. For the algebras B_2 , G_2 et D_2 , only racines positives have been labelled.

et en utilisant la propriété (3.25), on obtient que si $i \neq j$,

$$\left. \begin{aligned} \cos \widehat{\alpha_i, \alpha_j} &= -\frac{1}{2} \sqrt{m_i m_j} \\ \frac{\|\alpha_i\|}{\|\alpha_j\|} &= \sqrt{\frac{m_i}{m_j}} \end{aligned} \right\} \text{ avec } m_i, m_j \in \mathbb{N}, \quad (3.27)$$

et la valeur -1 du cosinus est interdite, parce que $\alpha_i \neq -\alpha_j$ par hypothèse, si bien que les seules valeurs possibles de ce cosinus sont $0, -\frac{1}{2}, -\frac{\sqrt{2}}{2}, -\frac{\sqrt{3}}{2}$, c’est-à-dire que les seuls angles possibles entre racines simples sont $\frac{\pi}{2}, \frac{2\pi}{3}, \frac{3\pi}{4}$ ou $\frac{5\pi}{6}$, avec des rapports de longueurs des racines respectivement égaux à $1, \sqrt{2}, \sqrt{3}$.

Il n’existe bien sûr qu’une seule algèbre de rang 1, c’est l’algèbre $\mathfrak{su}(2)$ complexifiée (3.1) ou (3.17). On lui donnera désormais le nom de A_1 . Il est ensuite aisé de classifier les algèbres possibles de rang 2. Les quatre cas sont représentés sur la Fig. 3.1, avec leurs matrices de Cartan s’écrivant

$$A_2 : \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad B_2 : \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix} \quad G_2 : \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \quad D_2 : \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}. \quad (3.28)$$

La nomenclature, A_2, B_2, G_2 et D_2 , est conventionnelle, ainsi que la numérotation des racines. Le dernier cas, D_2 , qui a $\langle \alpha_1, \alpha_2 \rangle = 0$, est mentionné pour mémoire : il correspond à une algèbre semi-simple, somme directe de deux algèbres A_1 . (Rien ne force d’ailleurs ses deux racines à avoir la même longueur).

En général, si on peut séparer l’ensemble des racines en deux sous-ensembles mutuellement orthogonaux, on voit que l’algèbre de Lie se décompose en somme directe de deux algèbres, et vice versa. Se rappelant que toute algèbre semi-simple peut toujours se décomposer en somme directe de sous-algèbres simples (fin du Chapitre 1), dans la suite, on ne considérera que des algèbres simples.

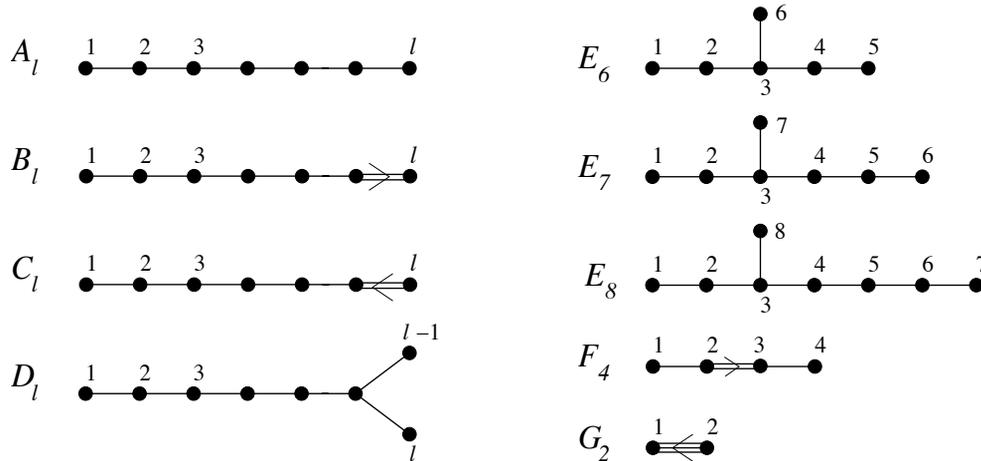


FIGURE 3.2 – Diagrammes de Dynkin

Diagramme de Dynkin

En rang plus élevé, c'est-à-dire en dimension de l'espace des racines plus élevée, il devient difficile de représenter le système des racines. On adopte une autre représentation, en codant la matrice de Cartan dans un diagramme de la façon suivante. À chaque racine simple α_i est associé un vertex i du diagramme. Deux vertex sont unis par une ligne ssi $\langle \alpha_i, \alpha_j \rangle \neq 0$; la ligne est simple si $C_{ij} = C_{ji} = -1$ (angle de $2\pi/3$, longueurs égales); elle est double (resp. triple) si $C_{ij} = -2$ (resp. -3) et $C_{ji} = -1$ (angle de $\frac{3\pi}{4}$ resp. $\frac{5\pi}{6}$, avec un rapport des longueurs de $\sqrt{2}$, resp. $\sqrt{3}$) et porte alors une flèche (ou plutôt un signe $>$) entre i et j indiquant quelle racine est la plus longue. (Attention que là encore, certains auteurs utilisent la convention opposée pour l'orientation de ces flèches!)

3.2.2 Systèmes de racines des algèbres simples. La classification de Cartan

Une analyse de tous les cas possibles a mené Cartan⁴ à une classification des algèbres de Lie simples complexes en termes de 4 familles infinies et de 5 cas exceptionnels. La notation traditionnelle est la suivante

$$A_\ell, \quad B_\ell, \quad C_\ell, \quad D_\ell, \quad E_6, \quad E_7, \quad E_8, \quad F_4, \quad G_2. \tag{3.29}$$

Dans chaque cas, l'indice indique le rang de l'algèbre. La géométrie des systèmes de racines est codée dans les diagrammes de Dynkin de la Fig. 3.2.

La preuve est un peu laborieuse et sera omise ici. Elle consiste à exploiter le fait que la matrice de Cartan est définie positive pour montrer successivement qu'au plus un seul de ses éléments non diagonaux peut être différent de 0 ou -1 (une seule arête du diagramme de Dynkin peut être multiple), que le diagramme ne contient pas de boucle, que la seule coordonnée possible d'un vertex est 0, 1 ou 2 et qu'un diagramme a au plus un vertex de coordonnée 3, etc, et finalement que la liste des diagrammes possibles se réduit à celle de la Fig. 3.2.

4. Ce travail de classification, entrepris par Killing a été complété et corrigé par É. Cartan, puis simplifié par van der Waerden, Dynkin, ...

Les quatre familles infinies sont identifiées aux algèbres de Lie (complexifiées) de groupes classiques

$$A_\ell = \mathfrak{sl}(\ell + 1, \mathbb{C}), \quad B_\ell = \mathfrak{so}(2\ell + 1, \mathbb{C}), \quad C_\ell = \mathfrak{sp}(2\ell, \mathbb{C}), \quad D_\ell = \mathfrak{so}(2\ell, \mathbb{C}) \quad (3.30)$$

ou à leur unique forme réelle compacte, respectivement $A_\ell = \mathfrak{su}(\ell + 1)$, $B_\ell = \mathfrak{so}(2\ell + 1)$, $C_\ell = \mathfrak{usp}(\ell)$, $D_\ell = \mathfrak{so}(2\ell)$.

Les “algèbres exceptionnelles” E_6, \dots, G_2 ont pour dimensions respectives 78, 133, 248, 52 et 14. Ce sont les algèbres de groupes de Lie ... exceptionnels! G_2 est le groupe d’automorphismes d’octonions, F_4 est lui-même un groupe d’automorphismes de matrices d’octonions, etc.

Parmi ces algèbres, les algèbres A, D, E , dont toutes les racines ont même longueur, sont dites *simplement lacées*. Il est curieux de remarquer que de nombreux problèmes, sous-groupes finis de $\mathfrak{su}(2)$, singularités “simples”, théories conformes “minimales”, etc, sont classifiés selon ce schéma ADE ... mais cela est une autre histoire! (Voir par exemple http://www.scholarpedia.org/article/A-D-E_Classification_of_Conformal_Field_Theories.)

Les formes réelles de ces algèbres simples complexes ont aussi été classifiées par Cartan. On trouve 12 séries infinies et 23 cas exceptionnels!

3.2.3 Base de Chevalley de l’algèbre

Il existe une autre base de l’algèbre de Lie \mathfrak{g} , dite base de Chevalley, avec des crochets qui ont l’avantage de ne dépendre que de la matrice de Cartan. Soient h_i, e_i et $f_i, i = 1, \dots, \ell$, des générateurs attachés aux racines simples α_i selon

$$e_i = \left(\frac{2}{\langle \alpha_i, \alpha_i \rangle} \right)^{\frac{1}{2}} E_{\alpha_i}, \quad f_i = \left(\frac{2}{\langle \alpha_i, \alpha_i \rangle} \right)^{\frac{1}{2}} E_{-\alpha_i}, \quad h_i = \frac{2\alpha_i \cdot H}{\langle \alpha_i, \alpha_i \rangle}. \quad (3.31)$$

Leurs relations de commutation s’écrivent

$$\begin{aligned} [h_i, h_j] &= 0 \\ [h_i, e_j] &= C_{ji} e_j \\ [h_i, f_j] &= -C_{ji} f_j \\ [e_i, f_j] &= \delta_{ij} h_j \end{aligned} \quad (3.32)$$

(le vérifier). L’algèbre est engendrée par les e_i, f_i, h_i et tous leurs commutateurs contraints par (3.32) et par les “relations de Serre”

$$\begin{aligned} \text{ad}(e_i)^{1-C_{ji}} e_j &= 0 \\ \text{ad}(f_i)^{1-C_{ji}} f_j &= 0. \end{aligned} \quad (3.33)$$

Cela prouve que toute l’algèbre est bien codée dans la donnée des racines simples et de leur géométrie (la matrice de Cartan ou le diagramme de Dynkin).

Noter aussi la propriété remarquable, pas évidente *a priori*, que dans cette base, toutes les constantes de structure (coefficients des relations de commutation) sont des *entiers*.

3.2.4 Coracines. Plus haute racine. Nombres, exposants de Coxeter

Nous donnons ici quelques compléments sur des notations et concepts qu’on rencontre dans l’étude des algèbres de Lie simples et de leurs systèmes de racines.

Comme la combinaison

$$\alpha_i^\vee := \frac{2}{\langle \alpha_i, \alpha_i \rangle} \alpha_i, \quad (3.34)$$

pour α_i une racine simple, apparaît fréquemment, on lui donne le nom de *coracine*. La matrice de Cartan se réécrit donc

$$C_{ij} = \langle \alpha_i, \alpha_j^\vee \rangle . \quad (3.35)$$

La *racine la plus haute* θ est la racine positive qui a la propriété que la somme de ses composantes dans une base de racines simples est maximale : on montre que cela suffit à la caractériser de façon unique. Ses composantes dans la base des racines et celle des coracines

$$\theta = \sum_i a_i \alpha_i \quad , \quad \frac{2}{\langle \theta, \theta \rangle} \theta = \sum_i a_i^\vee \alpha_i^\vee , \quad (3.36)$$

appelées aussi *indices de Kac* ou *indices de Kac duaux*, jouent aussi un rôle, en particulier par leurs sommes,

$$h = 1 + \sum_i a_i \quad , \quad h^\vee = 1 + \sum_i a_i^\vee . \quad (3.37)$$

Ces nombres h et h^\vee sont respectivement le *nombre de Coxeter* et le *nombre de Coxeter dual*. Quand il faut choisir une normalisation des racines, ce que nous n'avons pas fait encore, on choisit en général d'imposer que $\langle \theta, \theta \rangle = 2$.

Enfin la diagonalisation de la matrice de Cartan symétrisée

$$\widehat{C}_{ij} := 2 \frac{\langle \alpha_i, \alpha_j \rangle}{\sqrt{\langle \alpha_i, \alpha_i \rangle \langle \alpha_j, \alpha_j \rangle}} \quad (3.38)$$

donne un spectre de valeurs propres

$$\text{valeurs propres de } \widehat{C} = \left\{ 4 \sin^2 \left(\frac{\pi}{2h} m_i \right) \right\}, \quad i = 1, \dots, \ell , \quad (3.39)$$

faisant apparaître un nouvel ensemble d'entiers m_i , les *exposants de Coxeter*, compris entre 1 et $h - 1$, avec de possibles multiplicités. Ces nombres sont importants à plusieurs titres. Ils contiennent des informations utiles sur le groupe de Weyl. En leur ajoutant 1, (ce qui les rend ≥ 2), on obtient les degrés des opérateurs de Casimir algébriquement indépendants, ou encore les degrés où le groupe de Lie a une cohomologie non triviale, etc etc.

Exemples : pour A_{n-1} alias $\mathfrak{su}(n)$, racines et coracines sont identiques. La plus haute racine $\theta = \sum_i \alpha_i$, donc $h = h^\vee = n$, les exposants de Coxeter sont $1, 2, \dots, n - 1$. Pour D_n alias $\mathfrak{so}(2n)$, racines et coracines sont encore identiques, $\theta = \alpha_1 + 2\alpha_2 + \dots + 2\alpha_{n-2} + \alpha_{n-1} + \alpha_n$, $h = 2n - 2$, et les exposants sont $1, 3, \dots, 2n - 3, n - 1$, avec donc $n - 1$ double si n est pair.

Voir l'Appendice F pour des données sur les algèbres simples classiques.

3.3 Représentations des algèbres semi-simples

3.3.1 Poids. Réseau des poids

Nous étudions maintenant les représentations des algèbres semi-simples, en suivant une démarche parallèle à celle des sections précédentes. Dans tout ce qui suit, "représentation" signifie représentation de dimension finie. On supposera aussi cette représentation unitaire : c'est le cas qui nous intéresse pour la construction des représentations de groupes compacts. Les éléments de la sous-algèbre de Cartan commutant entre eux, ils commutent aussi dans toute représentation. Notant avec des "bras" et des "kets" les vecteurs de cette représentation, et écrivant simplement X (au lieu de $d(X)$) pour le représentant de l'élément $X \in \mathfrak{g}$, on peut trouver une base $|\lambda_a\rangle$ qui diagonalise simultanément les éléments de l'algèbre de Cartan

$$H|\lambda_a\rangle = \lambda(H)|\lambda_a\rangle \quad (3.40)$$

ou de façon équivalente

$$H_i|\lambda_a\rangle = \lambda_{(i)}|\lambda_a\rangle, \quad (3.41)$$

avec une valeur propre λ qui est à nouveau une forme sur l'espace \mathfrak{h} , donc un élément de \mathfrak{h}^* , l'espace des racines. On donne le nom de *poids* à un tel vecteur $\lambda = (\lambda_{(i)})$ de \mathfrak{h}^* . Noter que pour une représentation unitaire, les H sont hermitiens, donc λ est à valeurs réelles : les poids sont des vecteurs réels de \mathfrak{h}^* . La valeur propre λ pouvant apparaître avec une multiplicité, nous avons doté les états propres d'un indice a de multiplicité. L'ensemble des poids d'une représentation forme dans l'espace \mathfrak{h}^* le *diagramme des poids* de la représentation, voir Fig. 3.5 ci-dessous pour des exemples dans le cas de $\mathfrak{su}(3)$.

La représentation adjointe est une représentation particulière de l'algèbre dont les poids *non nuls* sont les racines. Les racines étudiées aux paragraphes précédents appartiennent donc à l'ensemble des poids dans \mathfrak{h}^* .

Les vecteurs $|\lambda_a\rangle$ formant une base de la représentation, leur nombre total, multiplicité incluse, égale la dimension de l'espace E de représentation. Cet espace E contient des sous-espaces de représentation de chacune des algèbres $\mathfrak{su}(2)$ que nous avons identifiées au § 3.2, engendrées par $\{H_\alpha, E_\alpha, E_{-\alpha}\}$. Par le même argument qu'au § 3.2, on va montrer que tout poids λ satisfait

$$\forall \alpha, \quad 2 \frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} = m' \in \mathbb{Z}, \quad (3.42)$$

et réciproquement, on montre que tout $\lambda \in \mathfrak{h}^*$ satisfaisant (3.42) est le poids d'une représentation de dimension finie. On peut donc utiliser (3.42) comme autre définition d'un poids. Pour se convaincre que le poids de toute représentation satisfait (3.42), on peut, comme au § 3.2, définir la chaîne maximale des poids passant par λ

$$\lambda + p'\alpha, \dots, \lambda, \dots, \lambda + q'\alpha \quad p' \leq 0, \quad q' \geq 0,$$

qui forment une représentation de la sous-algèbre $\mathfrak{su}(2)$, et montrer alors que $m' = -p' - q'$. Soient p' le plus petit entier ≤ 0 tel que $(E_{-\alpha})^{|p'|}|\lambda_a\rangle \neq 0$, et q' le plus grand entier ≥ 0 tel que $(E_\alpha)^{q'}|\lambda_a\rangle \neq 0$, H_α a pour valeurs propres sur ces deux vecteurs $\langle \lambda, \alpha \rangle + p'\langle \alpha, \alpha \rangle$, resp $\langle \lambda, \alpha \rangle + q'\langle \alpha, \alpha \rangle$. En exprimant que les valeurs propres de $2H_\alpha/\langle \alpha, \alpha \rangle$ sont entières et opposées, on a

$$2q' + 2 \frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} = 2j \quad 2p' + 2 \frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} = -2j. \quad (3.43)$$

En soustrayant membre à membre, on a $q' - p' = 2j$, la longueur de la chaîne est bien $2j + 1$ (dimension de la représentation de spin j de $\mathfrak{su}(2)$), tandis qu'en ajoutant membre à membre pour éliminer $2j$, on a

$$2 \frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} = -(q' + p') =: m', \quad \text{comme annoncé en (3.42).}$$

Cette chaîne est invariante sous l'action de la réflexion de Weyl w_α . (Cela est une généralisation de la symétrie des "multiplets" de $\mathfrak{su}(2)$ $(-j, -j + 1, \dots, j - 1, j)$.) Plus généralement, l'ensemble des poids est invariant par le groupe de Weyl : si λ est un poids d'une représentation, $w_\alpha(\lambda)$ en est un aussi, et on montre qu'il est de même multiplicité. Le diagramme des poids d'une représentation est donc invariant sous l'action de W .

L'ensemble des poids est découpé par le groupe de Weyl W en “chambres”, en nombre égal à l'ordre de W . La chambre associée à l'élément w de W est le cône

$$\mathcal{C}_w = \{ \lambda \mid \langle w\lambda, \alpha_i \rangle \geq 0, \forall i = 1, \dots, \ell \}, \quad (3.44)$$

où les α_i sont les racines simples. (Ce n'est pas tout à fait une partition, car il existe des poids sur les “murs” de séparation.) La *chambre fondamentale* est \mathcal{C}_1 , correspondant à l'identité dans W . Les poids appartenant à cette chambre fondamentale sont appelés *poids dominants*. Tout poids peut être amené dans \mathcal{C}_1 par une opération de W : il est sur l'“orbite” (pour le groupe de Weyl) d'un unique poids dominant. Parmi les poids de toute représentation irréductible, il en existe au moins un qui est dans \mathcal{C}_1 .

Par ailleurs, de $[H_i, E_\alpha] = \alpha_{(i)} E_\alpha$, il découle que

$$H_i E_\alpha |\lambda_a\rangle = ([H_i, E_\alpha] + E_\alpha H_i) |\lambda_a\rangle = (\alpha_{(i)} + \lambda_{(i)}) E_\alpha |\lambda_a\rangle$$

donc que $E_\alpha |\lambda_a\rangle$, s'il n'est pas nul, est un vecteur de poids $\lambda + \alpha$. Or dans une représentation irréductible, tous les vecteurs s'obtiennent à partir les uns des autres par de telles actions de E_α , d'où il découle que

▷ *Deux poids de la même représentation irréductible diffèrent par une combinaison à coefficients entiers de racines,*

(mais cette combinaison n'est en général pas une racine).

On introduit alors un ordre partiel sur les poids d'une même représentation : on dit que $\lambda' > \lambda$ si $\lambda' - \lambda = \sum_i n_i \alpha_i$, avec des coefficients n_i entiers non négatifs. Il existe (parmi les poids de cette représentation) un unique plus haut poids Λ , dont on montre qu'il est de multiplicité 1. Le vecteur de plus haut poids sera noté $|\Lambda\rangle$ (sans indice a). Il est tel que pour toute racine positive $E_\alpha |\Lambda\rangle = 0$, (sans quoi il ne serait pas le plus haut), donc $q' = 0$ dans l'équation (3.43) ci-dessus, et $\langle \Lambda, \alpha \rangle = \frac{1}{2} \langle \alpha, \alpha \rangle j > 0$, Λ est un poids dominant.

▷ *Le plus haut poids d'une représentation est un poids dominant, $\Lambda \in \mathcal{C}_1$.*

Ce vecteur de plus haut poids caractérise la représentation irréductible. (Dans le cas de $\mathfrak{su}(2)$, ce serait un vecteur $|j, m = j\rangle$.) Autrement dit, deux représentations sont équivalentes ssi elles ont le même plus haut poids.

On introduit ensuite les *indices de Dynkin* du poids λ par

$$\lambda_i = 2 \frac{\langle \lambda, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} \in \mathbb{Z} \quad (3.45)$$

avec α_i les racines simples. Pour un poids dominant, donc pour tout poids le plus haut d'une représentation irréductible, ces indices sont non négatifs, c'est-à-dire dans \mathbb{N} .

Les *poids fondamentaux* Λ_i satisfont par définition

$$2 \frac{\langle \Lambda_j, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = \delta_{ij}. \quad (3.46)$$

Leur nombre égale le rang ℓ de l'algèbre, ils constituent une base de \mathfrak{h}^* . Chacun d'eux est le plus haut poids d'une représentation irréductible dite *fondamentale*; il y a donc ℓ représentations

fondamentales. Nous avons donc obtenu

▷ Toute représentation irréductible est caractérisée par son plus haut poids,

et par abus de notation, on désignera par (Λ) la représentation irréductible de poids le plus haut Λ .

▷ Tout plus haut poids se décompose sur les poids fondamentaux, avec pour composantes ses indices de Dynkin (3.45),

$$\Lambda = \sum_{j=1}^{\ell} \lambda_j \Lambda_j \quad , \quad \lambda_i \in \mathbb{N} . \quad (3.47)$$

et tout Λ de la forme (3.47) est le plus haut poids d'une représentation irréductible.

Autrement dit, la connaissance des poids fondamentaux suffit à construire toutes les représentations irréductibles de l'algèbre.

Montrer en utilisant les propriétés énoncées ci-dessus que le plus haut poids de la représentation adjointe ne peut être que θ , défini en eq. (3.36).

Réseaux des poids et des racines

En général, étant donnée une base de vecteurs e_1, \dots, e_p dans un espace de dimension p , on appelle réseau engendré par ces vecteurs l'ensemble des vecteurs $\sum_{i=1}^p z^i e_i$ avec des coefficients $z^i \in \mathbb{Z}$. On note encore ce réseau $\mathbb{Z}e_1 + \dots + \mathbb{Z}e_p$.

Le réseau P des poids est le réseau engendré par les ℓ poids fondamentaux Λ_j . Le réseau Q des racines est celui engendré par les ℓ racines simples α_i . C'est un sous-réseau de P . Tout poids d'une représentation irréductible se trouve sur P .

On peut considérer les différentes classes du groupe additif P par rapport à son sous-groupe Q . Ces *classes de congruence*, dont le nombre fini $|P/Q|$ se trouve être égal au déterminant de la matrice de Cartan, rassemblent tous les poids qui diffèrent par un vecteur du réseau des racines. Dans le cas de $\mathfrak{su}(n)$, il y a n de ces classes, nous y reviendrons. (Exercice : prouver la propriété que $|P/Q| = \det C$. Indication : comparer les deux déterminants des Λ_i et des α_i dans la base des coracines.)

On peut aussi introduire le réseau Q^\vee engendré par les ℓ coracines α_i^\vee (cf § 2.4). C'est le "dual" de P puisque $\langle \alpha_i^\vee, \Lambda_j \rangle \in \mathbb{Z}$.

On montre enfin que les sous-groupes du groupe fini P/Q sont aussi isomorphes aux groupes d'homotopie des groupes G ayant \mathfrak{g} comme algèbre de Lie ! Par exemple pour $\mathfrak{su}(n)$, on verra plus bas que $P/Q = \mathbb{Z}_n$, et ces sous-groupes sont caractérisés par un diviseur d de n . Pour chacun d'eux, $SU(n)/\mathbb{Z}_d$ a pour algèbre de Lie $\mathfrak{su}(n)$. Le cas $n = 2$, avec $SU(2)$ et $SO(3)$, nous est familier.

Dimension et opérateur de Casimir

Il est utile de connaître la dimension d'une représentation de plus haut poids donné et la valeur qu'y prend l'opérateur de Casimir quadratique. Elles sont données en termes du *vecteur de Weyl* ρ , défini par deux formules (non trivialement !) équivalentes

$$\begin{aligned} \rho &= \frac{1}{2} \sum_{\alpha > 0} \alpha \\ &= \sum_j \Lambda_j . \end{aligned} \quad (3.48)$$

Une formule remarquable, due à Weyl, exprime la dimension de la représentation de plus haut poids Λ comme un produit sur les racines positives

$$\dim(\Lambda) = \prod_{\alpha > 0} \frac{\langle \Lambda + \rho, \alpha \rangle}{\langle \rho, \alpha \rangle} \quad (3.49)$$

tandis que la valeur propre du Casimir quadratique est

$$C_2(\Lambda) = \frac{1}{2} \langle \Lambda, \Lambda + 2\rho \rangle . \quad (3.50)$$

Une question liée est celle de la trace des générateurs de \mathfrak{g} dans la représentation (Λ) . Soit t_a une base de \mathfrak{g} telle que $\text{tr } t_a t_b = T_A \delta_{ab}$, avec un coefficient T_A dont le signe dépend des conventions (t hermitiques ou antihermitiques, cf Chapitre 1). Dans la représentation de plus haut poids Λ , on doit avoir (voir ci-dessous Exercice B du Chap. 5)

$$\text{tr } d_\Lambda(t_a) d_\Lambda(t_b) = T_A \delta_{ab} . \quad (3.51)$$

Mais dans cette base, l'opérateur de Casimir quadratique s'écrit $C_2 = \sum_a (d_\Lambda(t_a))^2$ donc en prenant la trace,

$$\begin{aligned} \text{tr } C_2 &= \sum_a \text{tr } (d_\Lambda(t_a))^2 = T_A \sum_a 1 = T_A \dim \mathfrak{g} \\ &= C_2(\Lambda) \text{tr } I_\Lambda = C_2(\Lambda) \dim(\Lambda) \end{aligned} \quad (3.52)$$

d'où

$$T_\Lambda = C_2(\Lambda) \frac{\dim(\Lambda)}{\dim \mathfrak{g}} , \quad (3.53)$$

une formule souvent utile dans les calculs (théories de jauge...). Dans la représentation adjointe, $\dim(\Lambda_A) = \dim \mathfrak{g}$, donc $T_{\Lambda_A} = C_2(\Lambda_A)$.

Il existe une flopée de formules variées, souvent intrigantes, reliant différents aspects de la théorie des algèbres de Lie et de leurs représentations. Ainsi par exemple la "formule étrange" de Freudenthal–de Vries qui relie les normes des vecteurs ρ et θ à la dimension de l'algèbre et au nombre de Coxeter : $\langle \rho, \rho \rangle = \frac{\hbar}{24} \langle \theta, \theta \rangle \dim \mathfrak{g}$.

Il existe encore une formule (Freudenthal) qui décrit la multiplicité d'un poids λ au sein d'une représentation de plus haut poids Λ . Et aussi, question liée, une formule de Weyl qui donne l'expression du caractère $\chi_\Lambda(e^H)$ de cette représentation évalué sur un élément du "tore de Cartan", qui résulte de l'exponentiation de l'algèbre de Cartan \mathfrak{h} .

Représentation conjuguée

Étant donnée une représentation de plus haut poids Λ , sa représentation conjuguée est en général non équivalente. On sait caractériser quel est son plus haut poids $\bar{\Lambda}$ grâce au groupe de Weyl. La non équivalence des représentations (Λ) et $(\bar{\Lambda})$ a à voir avec les symétries du diagramme de Dynkin. Pour les algèbres de type B, C, E_7, E_8, F_4, G_2 pour lesquelles il n'existe pas de telle symétrie non triviale, les représentations sont auto-conjuguées. C'est aussi le cas de D_{2r} . Pour les autres, la conjugaison correspond à la symétrie suivante sur les indices de Dynkin

$$\begin{aligned} A_\ell = \text{su}(\ell + 1) & \quad \lambda_i \leftrightarrow \lambda_{\ell+1-i} & \ell > 1 \\ D_{2r+1} = \text{so}(4r + 2) & \quad \lambda_\ell \leftrightarrow \lambda_{\ell-1} , & \ell = 2r + 1 \\ E_6 & \quad \lambda_i \leftrightarrow \lambda_{6-i}, & i = 1, 2 . \end{aligned} \quad (3.54)$$

où l'indexation des poids fondamentaux, donc celle des indices de Dynkin, suit celle des racines simples, cf Fig. 3.2.

3.3.2 Racines et poids de $\mathfrak{su}(n)$

Examinons maintenant ce que sont concrètement les poids et donc les représentations irréductibles de $\mathfrak{su}(n)$.

On choisit d'abord une paramétrisation commode de l'espace \mathfrak{h}^* , qui est, rappelons-le, de dimension $n - 1$. Soient e_i , $i = 1, \dots, n$, n vecteurs de $\mathfrak{h}^* = \mathbb{R}^{n-1}$ (donc nécessairement liés), satisfaisant $\sum_1^n e_i = 0$. On les obtient à partir d'une base orthonormée \hat{e}_i de \mathbb{R}^n en projetant les \hat{e}_i dans un hyperplan orthogonal à $\hat{\rho} := \sum_{i=1}^n \hat{e}_i$, donc $e_i = \hat{e}_i - \frac{1}{n}\hat{\rho}$. Il est commode de se placer dans l'hyperplan $\sum_i^n x^i = 1$ de l'espace \mathbb{R}^n . Ces vecteurs ont pour produits scalaires

$$\langle e_i, e_j \rangle = \delta_{ij} - \frac{1}{n}. \quad (3.55)$$

En termes de ces vecteurs, les racines positives de $\mathfrak{su}(n) = A_{n-1}$, en nombre égal à $|\Delta_+| = n(n-1)/2$, sont

$$\alpha_{ij} = e_i - e_j, \quad 1 \leq i < j \leq n, \quad (3.56)$$

et les $\ell = n - 1$ racines simples sont

$$\alpha_i = \alpha_{i i+1} = e_i - e_{i+1}, \quad 1 \leq i \leq n - 1. \quad (3.57)$$

Ces racines ont été normalisées à $\langle \alpha, \alpha \rangle = 2$. La somme des racines positives se calcule aisément

$$\begin{aligned} 2\rho &= (n-1)e_1 + (n-3)e_2 + \dots + (n-2i+1)e_i + \dots - (n-1)e_n \\ &= (n-1)\alpha_1 + 2(n-2)\alpha_2 + \dots + i(n-i)\alpha_i + \dots + (n-1)\alpha_{n-1}. \end{aligned} \quad (3.58)$$

On vérifie que la matrice de Cartan est bien

$$C_{ij} = \langle \alpha_i, \alpha_j \rangle = \begin{cases} 2 & \text{si } i = j \\ -1 & \text{si } i = j \pm 1 \end{cases}$$

en accord avec le diagramme de Dynkin de type A_{n-1} . Les poids fondamentaux Λ_i $i = 1, \dots, n-1$ s'écrivent alors aisément

$$\Lambda_i = \sum_{j=1}^i e_j, \quad (3.59)$$

$$e_1 = \Lambda_1, \quad e_i = \Lambda_i - \Lambda_{i-1} \text{ for } i = 2, \dots, n-1, \quad e_n = -\Lambda_{n-1} \quad (3.60)$$

avec des produits scalaires

$$\langle \Lambda_i, \Lambda_j \rangle = \frac{i(n-j)}{n}, \quad i \leq j. \quad (3.61)$$

Le groupe de Weyl $W \equiv \mathcal{S}_N$ agit sur les racines et les poids en permutant les e_i : $w \in W \leftrightarrow \bar{w} \in \mathcal{S}_N : w(e_i) = e_{\bar{w}(i)}$.

Dimension de la représentation de poids Λ

En réunissant les formules (3.49) et (3.56), démontrer l'expression suivante

$$\dim(\Lambda) = \prod_{1 \leq i < j \leq n} \frac{f_i - f_j + j - i}{j - i} \quad \text{où} \quad f_i := \sum_{k=i}^{n-1} \lambda_k, \quad f_n = 0. \quad (3.62)$$

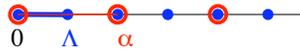


FIGURE 3.3 – Poids de $\mathfrak{su}(2)$. Les parties positives des réseaux des poids (petits points) et des racines (gros points)

Représentations conjuguées

Si $\Lambda = (\lambda_1, \dots, \lambda_{n-1})$ est le plus haut poids d'une représentation irréductible de $\mathfrak{su}(n)$, $\bar{\Lambda} = (\lambda_{n-1}, \dots, \lambda_1)$ est celui de la représentation conjuguée, en général pas équivalente. On note que ni la dimension, ni la valeur de l'opérateur de Casimir ne distinguent les représentations Λ et $\bar{\Lambda}$.

“ n -alité”

Les classes de congruence de P par rapport à Q sont en nombre n . On les distingue par la valeur de

$$\nu(\lambda) := \lambda_1 + 2\lambda_2 + \dots + (n-1)\lambda_{n-1} \pmod{n}, \quad (3.63)$$

à laquelle on peut donner le vilain nom de “ n -alité”, par extension de la “trialité” de $\mathfrak{su}(3)$. Les éléments du réseau des racines ont donc $\nu(\lambda) = 0$.

Exemples de $\mathfrak{su}(2)$ et $\mathfrak{su}(3)$

Dans le cas de $\mathfrak{su}(2)$, il n'y a qu'un poids fondamental $\Lambda = \Lambda_1$ et une racine positive α , normalisée à $\langle \alpha, \alpha \rangle = 2$, donc $\langle \Lambda, \alpha \rangle = 1$, $\langle \Lambda, \Lambda \rangle = \frac{1}{2}$. Donc $\alpha = 2\Lambda$, Λ correspond à la représentation de spin $\frac{1}{2}$, α à celle de spin 1. Le réseau des poids et celui des racines sont simples à dessiner, voir Fig. 3.3. L'indice de Dynkin λ_1 s'identifie à l'entier $2j$, les deux classes de congruence de P par rapport à Q correspondent aux représentations de spin entier ou demi-entier, la dimension $\dim(\Lambda) = \lambda_1 + 1 = 2j + 1$ et l'opérateur de Casimir $C_2(\Lambda) = \frac{1}{4}\lambda_1(\lambda_1 + 2) = j(j + 1)$, en accord avec les expressions bien connues.

Pour $\mathfrak{su}(3)$, le réseau des poids est triangulaire, voir Fig. (3.4) sur laquelle on a indiqué la trialité $\tau(\lambda) := \lambda_1 + 2\lambda_2 \pmod{3}$ et porté les poids fondamentaux et les poids les plus hauts des premières représentations. Selon l'usage, les représentations sont repérées par leur dimension ⁵

$$\dim(\Lambda) = \frac{1}{2}(\lambda_1 + 1)(\lambda_2 + 1)(\lambda_1 + \lambda_2 + 2), \quad (3.64)$$

complétée par une barre pour distinguer la représentation de sa conjuguée, chaque fois que nécessaire. La conjuguée de la représentation de poids $\Lambda = (\lambda_1, \lambda_2)$ est $\bar{\Lambda} = (\lambda_2, \lambda_1)$. Seules les représentations sur la bissectrice de la chambre de Weyl sont donc autoconjuguées.

Exercice. Calculer la valeur propre de l'opérateur de Casimir quadratique en termes des indices de Dynkin λ_1, λ_2 grâce aux formules (3.50) et (3.58).

L'ensemble des poids des représentations les plus basses est représenté sur la Fig. 3.5, après rotation des axes des figures précédentes. L'axe horizontal, porté par α_1 , et l'axe vertical, porté

⁵. ce qui est parfois ambigu ; par exemple, identifier sur la Fig. (3.4) le poids d'une autre représentation de dimension 15.

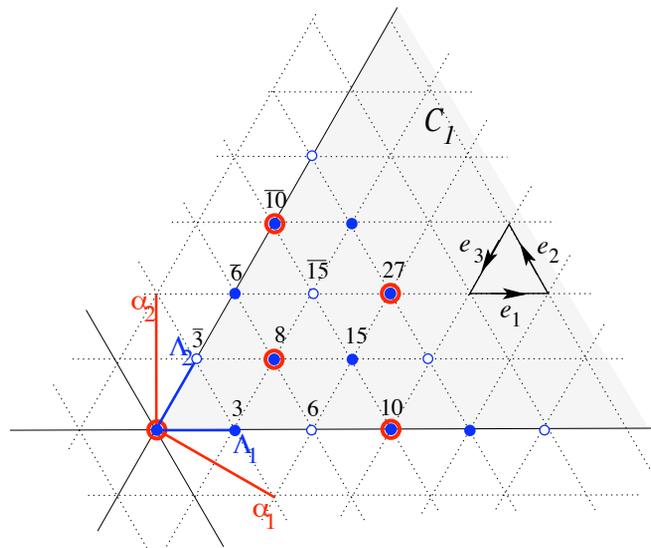


FIGURE 3.4 – Poids de $su(3)$. Seule la première chambre de Weyl C_1 a été détaillée, avec quelques plus hauts poids. Les poids de trialité 0 (réseau des racines) sont représentés par un gros disque, ceux de trialité 1, resp. 2, par un disque plein, resp. vide.

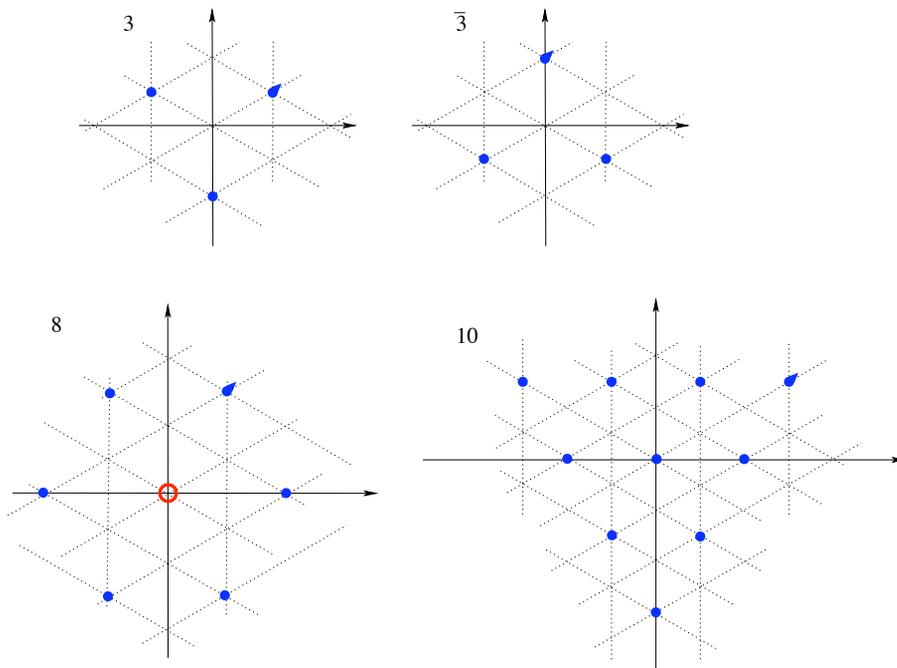


FIGURE 3.5 – Les diagrammes de poids des représentations les plus basses de $su(3)$, désignées par leur dimension. Noter qu’une rotation de 30° du réseau des poids a été effectuée par rapport aux figures précédentes. Dans chaque représentation, le poids le plus haut est indiqué par une petite indentation. Les points ont une multiplicité 1, le disque ouvert a la multiplicité 2.

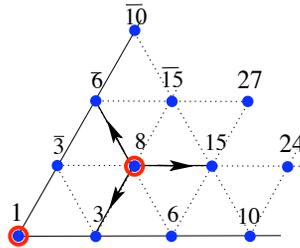


FIGURE 3.6 – Opération de produit tensoriel de la représentation 8 par la représentation 3 figurée sur le diagramme des poids de $\mathfrak{su}(3)$.

par Λ_2 , vont en effet acquérir une signification physique : celle d’axes des coordonnées d’isospin et d’“hypercharge”, voir chapitre suivant.

N.B. On a détaillé le cas de $\mathfrak{su}(n)$. Des formules analogues pour les racines, poids fondamentaux, etc des autres algèbres simples sont bien sûr connues explicitement et tabulées dans la littérature. Voir par exemple l’appendice A pour la fiche signalétique des algèbres “classiques” de type A, B, C, D , et Bourbaki, chap.6 pour plus de détails sur les autres algèbres.

3.4 Produit tensoriel des représentations de $\mathfrak{su}(n)$

3.4.1 Règles de Littlewood–Richardson et de Racah–Speiser

Étant données deux représentations irréductibles de $\mathfrak{su}(n)$ (ou de toute autre algèbre de Lie), c’est un problème souvent rencontré de décomposer leur produit tensoriel en somme de représentations irréductibles. Si on ne désire obtenir que les multiplicités et qu’on dispose de tables de caractères du *groupe* correspondant, on peut utiliser les formules démontrées au Chap. 2, § 2.3.2.

Il existe aussi des règles assez complexes donnant la décomposition en représentations irréductibles d’un produit de deux représentations irréductibles (Λ) et (Λ') de $\mathfrak{su}(n)$. Ce sont les *règles de Littlewood–Richardson*, qui font appel à la représentation en tableaux d’Young (cf § suivant). Mais il est souvent plus simple de procéder de proche en proche, en notant que la représentation irréductible (Λ') se trouve dans un produit adéquat de représentations fondamentales, et en examinant les produits successifs de la représentation Λ par ces représentations fondamentales. Par l’associativité du produit tensoriel, on ramène le problème de départ à celui du produit tensoriel de (Λ) par les diverses représentations fondamentales.

Cette dernière opération est aisée à décrire sur le réseau des poids. Étant donné le plus haut poids Λ dans la première chambre de Weyl \mathcal{C}_1 , le produit tensoriel de (Λ) par la représentation fondamentale de plus haut poids Λ_i se décompose en représentations irréductibles de la manière suivante : on ajoute de toutes les façons possibles les $\dim(\Lambda_i)$ poids de la fondamentale au vecteur Λ et on ne garde comme plus hauts poids de la décomposition que les poids résultant de cette addition qui appartiennent à \mathcal{C}_1 .

Illustrons ceci sur le cas de $\mathfrak{su}(3)$. Supposons que l’on désire calculer la décomposition de $8 \otimes 8$.

On sait que la représentation 8 (adjointe) se trouve dans le produit des deux fondamentales 3 et $\bar{3}$. Les poids de la représentation fondamentale “3” de plus haut poids $\Lambda_1 = e_1$ sont e_1, e_2, e_3 . Ceux de la représentation fondamentale “ $\bar{3}$ ” sont leurs opposés. Avec la règle ci-dessus, on obtient donc

$$\begin{aligned} 3 \otimes 3 &= \bar{3} \oplus 6 & 3 \otimes \bar{3} &= 1 \oplus 8 \\ 3 \otimes 6 &= 8 \oplus 10 & 3 \otimes \bar{6} &= \bar{3} \oplus \bar{15} \\ 3 \otimes 8 &= 3 \oplus \bar{6} \oplus 15 \\ 3 \otimes 15 &= 6 \oplus \bar{15} \oplus 24 \end{aligned} \tag{3.65}$$

etc, et leurs conjuguées, cf Fig. 3.6. En général on ajoute les trois vecteurs $e_1 = (1, 0)$, $e_2 = (-1, 1)$ et $e_3 = (0, -1)$ (dans la base Λ_1, Λ_2) à $\Lambda = (\lambda_1, \lambda_2)$: les poids les plus hauts de la décomposition sont donc $(\lambda_1 + 1, \lambda_2)$, $(\lambda_1 - 1, \lambda_2 + 1)$ et $(\lambda_1, \lambda_2 - 1)$, dont on élimine ceux qui ont un indice de Dynkin négatif. Noter la cohérence avec la trialité : toutes les représentations apparaissant au membre de droite ont la même trialité, somme (modulo 3) des trialités de celles du membre de gauche. Par exemple, $\tau(3) = 1$, $\tau(15) = 1$, $\tau(6) = 2$, $\tau(\bar{15}) = 2$, etc.

En itérant cette opération, on peut alors calculer

$$8 \otimes (1 \oplus 8) = 8 \otimes 3 \otimes \bar{3} = (3 \oplus \bar{6} \oplus 15) \otimes \bar{3} = 1 \oplus 8 \oplus 8 \oplus 8 \oplus 10 \oplus \bar{10} \oplus 27$$

d'où on tire

$$8 \otimes 8 = 1 \oplus 8_s \oplus 8_a \oplus 10 \oplus \bar{10} \oplus 27 . \tag{3.66}$$

Dans cette dernière expression, on a ajouté un indice s ou a pour distinguer les deux copies de la représentation 8 : l'une est symétrique, l'autre antisymétrique dans l'échange des deux représentations 8 du membre de gauche. Cette relation nous sera très utile dans le chapitre suivant, dans l'étude de la symétrie SU(3) de saveur.

Exercice. Vérifier selon la même méthode que $3 \otimes 3 \otimes 3 = 1 \oplus 8 \oplus 8 \oplus 10$.

Bien que fastidieux, ce processus est simple et systématique. Il existe une règle un peu plus élaborée pour le produit tensoriel de deux représentations de plus hauts poids quelconques (Λ) et (Λ') d'une algèbre quelconque, voir ci-dessous. Il existe aussi des codes qui calculent ces décompositions, tel l'étonnant LiE, voir <http://www.mathlabo.univ-poitiers.fr/~maavl/LiE/form.html>

Une généralisation des règles précédentes, valide pour toute algèbre simple \mathfrak{g} , est donnée par l'*algorithme de Racah-Speiser*, qui fournit les multiplicités $N_{\lambda\mu}^\nu$, pour des poids les plus hauts λ et μ de \mathfrak{g} (les notations précédentes ont été changées $\Lambda \rightarrow \lambda$, $\Lambda' \rightarrow \mu$)

$$(\lambda) \otimes (\mu) = \oplus N_{\lambda\mu}^\nu (\nu) . \tag{3.67}$$

Considérons l'ensemble des poids $\sigma = \lambda' + \mu + \rho$ où λ' parcourt le diagramme de poids $[\lambda]$ de la représentation (λ) et ρ est le vecteur de Weyl. Trois cas peuvent advenir :

- i) si tous les indices de Dynkin de σ sont strictement positifs, $\lambda' + \mu$ contribue à la somme sur ν dans (3.67) avec une multiplicité égale à la multiplicité de σ (c'est-à-dire de λ') ;
- ii) si σ ou l'une de ses images par le groupe de Weyl a un indice de Dynkin nul, c'est-à-dire si σ est sur le mur d'une chambre de Weyl, $\lambda' + \mu$ ne contribue pas à la somme sur ν ;

- iii) si σ a un (ou des) indice(s) de Dynkin négatif(s) (mais aucun nul), et n'est pas du cas discuté en (ii), il peut être appliqué dans la chambre de Weyl fondamentale par un unique élément w du groupe de Weyl. Le poids $w[\sigma] - \rho$ contribue alors par $\text{sign}(w)$ fois la multiplicité de λ' à la somme sur ν , où $\text{sign}(w)$ est la signature de l'élément w de W , définie plus haut au § 3.2.1.

Ceci est résumé dans la formule

$$N_{\lambda\mu}^\nu = \sum_{\lambda' \in [\lambda]} \sum_{\substack{w \in W \\ w[\lambda' + \mu + \rho] - \rho \in \mathcal{P}_+}} \text{sign}(w) \delta_{\nu, w[\lambda' + \mu + \rho] - \rho} \quad (3.68)$$

dans laquelle \mathcal{P}_+ est la chambre de Weyl fondamentale (murs compris) : $\nu \in \mathcal{P}_+ \Leftrightarrow \nu_i \geq 0 \forall i = 1, \dots, n$.

3.4.2 Construction tensorielle explicite des représentations de $\text{SU}(2)$ et $\text{SU}(3)$

Considérons un vecteur $V \in \mathbb{C}^n$ de la représentation de définition de $\text{SU}(n)$. Sous l'action de $U \in \text{SU}(n)$, $V \mapsto V' = UV$, ou en composantes $v^i \mapsto v'^i = U^i_j v^j$, avec des indices $i, j = 1, \dots, n$. Soit W un vecteur qui se transforme par la représentation complexe conjuguée, (tel $W = V^*$), soit $W \mapsto W' = U^*W$. Il est naturel de noter les composantes de W avec des indices inférieurs, puisque $U^* = (U^\dagger)^T$, donc $w'_i = w_j (U^\dagger)^j_i$. On note alors que $V \cdot W := v^i w_i$ est invariant, en vertu de $U^\dagger U = I$. Autrement dit le tenseur mixte δ^i_j est invariant

$$\delta'^i_j = U^i_{i'} U^\dagger{}^{j'}_j \delta^{i'}_{j'} = (U \cdot U^\dagger)^i_j = \delta^i_j .$$

Formons maintenant des tenseurs de rang (p, m) , à p indices supérieurs et m indices inférieurs, se transformant comme $V^{\otimes p} \otimes W^{\otimes m}$, donc selon

$$t'^{i_1 \dots i_p}_{k_1 \dots k_m} = U^{i_1}_{j_1} \dots U^{i_p}_{j_p} U^\dagger{}^{l_1}_{k_1} \dots U^\dagger{}^{l_m}_{k_m} t^{j_1 \dots j_p}_{l_1 \dots l_m} . \quad (3.69)$$

- Dans le cas de $\text{SU}(2)$, on sait que les représentations U et U^* sont équivalentes. Cela résulte de l'existence d'une matrice $C = i\sigma_2$, telle que $CUC^{-1} = U^*$, donc CV^* se transforme comme V . Ou encore, puisque $C_{ij} = \epsilon_{ij}$ et $\epsilon_{i'j'} U^{i'}_i U^{j'}_j = \epsilon_{ij} \det U = \epsilon_{ij}$, le tenseur antisymétrique ϵ , invariant et inversible ($\epsilon^{ij} = -\epsilon_{ij}$, $\epsilon_{ij} \epsilon^{jk} = \delta^k_i$), peut être utilisé pour monter et descendre les indices, ($v_i := \epsilon_{ij} v^j$, soit $v_1 = v^2$, $v_2 = -v^1$); et il suffit de ne considérer que les tenseurs de rang p à indices supérieurs. Pour toute paire d'indices, par exemple i_1 et i_2 , un tel tenseur peut être écrit comme somme de composantes symétrique et antisymétrique dans ces indices

$$t^{i_1 i_2 \dots i_p} = t^{[i_1, i_2] \dots i_p} + t^{\{i_1, i_2\} \dots i_p}$$

avec $t^{[i_1, i_2] \dots i_p} := \frac{1}{2}(t^{i_1 i_2 \dots i_p} - t^{i_2 i_1 \dots i_p})$ et $t^{\{i_1, i_2\} \dots i_p} := \frac{1}{2}(t^{i_1 i_2 \dots i_p} + t^{i_2 i_1 \dots i_p})$. Pour la composante antisymétrique, on peut écrire $t^{[i_1, i_2] \dots i_p} = \epsilon^{i_1 i_2} \tilde{t}^{i_3 \dots i_p}$, avec $\tilde{t}^{i_3 \dots i_p} = -\frac{1}{2} \epsilon_{ab} t^{abi_3 \dots i_p}$, et donc réduire son rang⁶. Donc seuls les tenseurs complètement symétriques dans tous leurs p indices supérieurs fournissent des représentations irréductibles, et on retrouve ainsi une nouvelle fois la construction de toutes les représentations irréductibles de $\text{SU}(2)$ par produits tensoriels symétrisés de la représentation de dimension 2, cf Chap. 0, et le rang p s'identifie à $2j$. On vérifie en particulier

6. Il est bon de se rappeler les identités $\epsilon_{ab} \epsilon_{cd} = \delta_{ac} \delta_{bd} - \delta_{ad} \delta_{bc}$ et donc $\epsilon_{ab} \epsilon_{bc} = -\delta_{ac}$.

que le nombre de composantes indépendantes d'un tenseur de rang p complètement symétrique dans l'espace \mathbb{C}^2 est $p + 1$, puisque ces composantes possèdent $0, 1, \dots, p$ indices égaux à 1, les autres étant égaux à 2. La dimension de cette représentation est bien $2j + 1$.

Un tenseur de rang p complètement symétrique sera représenté par un "diagramme d'Young" avec p boîtes $\underbrace{\square\square\square}_p$. (La définition précise d'un diagramme d'Young sera donnée au paragraphe suivant.) Prenons $p = 3$ pour fixer les idées. Le produit tensoriel d'un tel tenseur de rang 3 par un tenseur de rang 1 sera figuré par

$$\square\square\square \otimes \square = \square\square\square\square \oplus \square\square$$

ce qui signifie, en termes de composantes,

$$4t^{ijk}u^l = (t^{ijk}u^l + t^{jkl}u^i + t^{ikl}u^j + t^{ijl}u^k) + (t^{ijk}u^l - t^{jkl}u^i) + (t^{ijk}u^l - t^{ikl}u^j) + (t^{ijk}u^l - t^{ijl}u^k)$$

où le premier terme est complètement symétrique dans ses $p + 1 = 4$ indices, et les suivants sont antisymétriques en (i, l) , (j, l) ou (k, l) . Selon l'argument précédent, ces derniers peuvent être réduits à des tenseurs de rang 2,

$$(t^{ijk}u^l - t^{jkl}u^i) = \epsilon^{il}\tilde{t}^{jk} \quad , \quad \tilde{t}^{jk} = -\epsilon_{ab}t^{ajk}u^b$$

ce que l'on représente en effaçant les colonnes à deux boîtes, soit

$$\square\square\square \otimes \square = \square\square\square\square \oplus \square\square$$

où on reconnaît la règle familière $j \otimes \frac{1}{2} = (j + \frac{1}{2}) \oplus (j - \frac{1}{2})$.

Exercice : reproduire avec cette méthode la règle de décomposition de $j \otimes j'$.

- Dans le cas de $SU(n)$, $n > 2$, il faut considérer les tenseurs avec les deux types d'indices, supérieur et inférieur, et les réduire. Mais ce n'est que dans le cas de $SU(3)$ que cette construction va nous fournir toutes les représentations irréductibles. Pour $n > 3$ il faudrait introduire d'autres tenseurs se transformant selon les autres représentations fondamentales de $SU(n)$ que la représentation de définition (de dimension n) et sa conjuguée.

Restreignons nous dans la fin de ce paragraphe au cas de $SU(3)$. Les tenseurs sont donc du type $t_{j_1 \dots j_m}^{i_1 \dots i_p}$, ($i, j = 1, 2, 3$), se transformant selon la représentation $3^{\otimes p} \otimes \bar{3}^{\otimes m}$. On a encore un tenseur ϵ invariant, mais cette fois de rang 3

$$\epsilon_{i'j'k'}U^{i'}_i U^{j'}_j U^{k'}_k = \epsilon_{ijk} \det U = \epsilon_{ijk} \quad ,$$

qui nous permet de troquer toute paire d'indices antisymétriques supérieurs pour un indice inférieur et vice versa, et de réduire ainsi le rang. Mais on peut aussi contracter toute paire d'indices supérieur et inférieur en un invariant, selon une remarque au début du paragraphe. On se contente donc de considérer les tenseurs complètement symétriques et de trace nulle de rang (p, m) . On peut démontrer que ces tenseurs forment une représentation irréductible, qui n'est autre que celle de poids le plus haut $p\Lambda_1 + m\Lambda_2$ dans les notations du § 3.2. Nous nous contenterons de vérifier que les dimensions de ces représentations sont bien en accord avec celles

données en (3.64), see Exercise E. A cette représentation on associe à nouveau un diagramme d'Young à deux lignes, la première à $p + m$, la seconde à m boîtes.

Les règles de produit tensoriel, en particulier celles par les représentations fondamentales 3 et $\bar{3}$ (cf § 3.4.1), se retrouvent aussi dans ce langage : il faut ajouter la boîte nouvelle de toutes les façons possibles au diagramme, et effacer toute colonne de hauteur 3 (ce qui reflète la propriété que $\det U = 1$). Exercice : étudier la réduction de $\begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & & \\ \hline \end{array} \otimes \square$ et retrouver la règle graphique du § 3.4.1 dans ce langage.

Un cas particulier que nous allons beaucoup utiliser au chapitre suivant est le suivant : la représentation adjointe est celle des tenseurs de rang (1,1) et de trace nulle. Cela n'est pas étonnant : la représentation adjointe est celle engendrée par l'algèbre de Lie $\mathfrak{su}(3)$, donc par les matrices (anti)hermitiennes 3×3 de trace nulle. Un tenseur de cette représentation se transforme par $t_j^i \mapsto t_j^i = U^i_{i'} U^{*j'}_{j} t_{j'}^{i'}$, soit sous forme matricielle

$$t' = UtU^\dagger, \tag{3.70}$$

ce qui à nouveau n'est pas pour nous surprendre, cf la définition de la représentation adjointe au chapitre 2. Quel diagramme d'Young associe-t-on à la représentation adjointe ?

3.5 Tableaux d'Young et représentations de $GL(n)$ et $SU(n)$

La construction précédente s'étend à $\mathfrak{su}(n)$, en fait au groupe $GL(n)$, et fait intervenir des opérations de symétrisation ou d'antisymétrisation liées au groupe symétrique des permutations S_m . On va se borner à quelques brèves indications.

Soit $E = \mathbb{C}^n$ l'espace vectoriel de dimension n . Le groupe $GL(n, \mathbb{C})$, ou $GL(n)$ en abrégé, est naturellement représenté dans E

$$g \in GL(n), \quad x \in E \mapsto x' = g.x. \tag{3.71}$$

Formons la puissance tensorielle m -ième de E : $F = E^{\otimes m} = E \otimes \dots \otimes E$. Dans F , le groupe $GL(n)$ agit par une représentation, la puissance tensorielle m -ième de (3.71)

$$g \in GL(n), \quad D(g)x^{(1)} \otimes \dots \otimes x^{(m)} = g.x^{(1)} \otimes \dots \otimes g.x^{(m)} \tag{3.72}$$

qui est en général réductible. Mais dans F , agit aussi le groupe symétrique S_m selon

$$\sigma \in S_m, \quad \widehat{D}(\sigma)x^{(1)} \otimes \dots \otimes x^{(m)} = x^{(\sigma^{-1}1)} \otimes \dots \otimes x^{(\sigma^{-1}m)}. \tag{3.73}$$

Faisons le choix d'une base e_i dans E , et notons g_{ij} les éléments de matrice de $g \in GL(n)$ dans cette base. La représentation de $GL(n)$ dans F a pour matrice

$$\mathcal{D}(g)_{\{i_1 \dots i_m\} \{j_1 \dots j_m\}} = \prod_{k=1}^m g_{i_k j_k} \tag{3.74}$$

et celle de S_m

$$\widehat{D}(\sigma)_{\{i_1 \dots i_m\} \{j_1 \dots j_m\}} = \prod_{k=1}^m \delta_{i_{\sigma k} j_k}. \tag{3.75}$$

Un tenseur t , élément de F , a dans cette base des composantes t^i . et se transforme sous l'action de $g \in GL(n)$, resp. de $\sigma \in S_m$, en un tenseur t' de composantes $t'^i = \mathcal{D}_{i,j} t^j$, resp. $\widehat{D}_{i,j} t^j$. Ces deux ensembles de matrices

commutent

$$\begin{aligned} \sum_{\{j\}} \mathcal{D}(g)_{\{i\},\{j\}} \widehat{\mathcal{D}}(\sigma)_{\{j\},\{k\}} &= \prod_l g_{i_l j_l} \delta_{j_l, k_{\sigma^{-1}l}} = \prod_l g_{i_{\sigma l} k_l} \\ &= \sum_{\{j\}} \widehat{\mathcal{D}}(\sigma)_{\{i\},\{j\}} \mathcal{D}(g)_{\{j\},\{k\}} . \end{aligned} \tag{3.76}$$

Définissons alors un *diagramme d'Young*. Un diagramme d'Young est formé de m cases disposées en k lignes de longueur non croissante : $f_1 \geq f_2 \geq \dots \geq f_k$, $\sum f_i = m$. Voici un exemple pour $m = 8$, ayant $f_1 = 4, f_2 = 2, f_3 = 2$



On décore ensuite les m boîtes d'un diagramme d'Young, le transformant ainsi en un *tableau*, en y distribuant les entiers de 1 à m . Un *tableau standard* est un tableau dans lequel les entiers sont croissants dans chaque ligne de gauche à droite, et dans chaque colonne de haut en bas.

Le nombre n_Y de tableaux standards obtenus à partir d'un diagramme Y est calculé comme suit. On définit les nombres $\ell_i = f_i + k - i, i = 1, \dots, k$. Ils forment une suite strictement décroissante : $\ell_1 > \ell_2 > \dots > \ell_k$. On démontre alors que

$$n_Y = \frac{n!}{\prod_i \ell_i!} \prod_{i < j} (\ell_i - \ell_j) \tag{3.77}$$

où le produit du numérateur vaut 1 s'il n'y a qu'une seule ligne.

La théorie des représentations du groupe symétrique S_m nous apprend alors qu'il y a une bijection entre les diagrammes d'Young à m boîtes Y et les représentations irréductibles $\widehat{\mathcal{D}}_Y$. La dimension de la représentation $\widehat{\mathcal{D}}_Y$ est donnée par le nombre (3.77) de tableaux standards de type Y .

On appelle tenseur de type (de symétrie) Y un tenseur se transformant par S_m selon cette représentation $\widehat{\mathcal{D}}_Y$. La commutation des matrices $\mathcal{D}(g)$ et $\widehat{\mathcal{D}}(\sigma)$ garantit alors que les tenseurs de type Y forment un sous-espace invariant pour l'action de $GL(n)$.

Exemple. Considérons les cas de $m = 2$ et $m = 3$. Dans le premier cas, les tenseurs de rang 2 peuvent se décomposer en leurs parties symétrique et antisymétrique qui se transforment indépendamment sous l'action de $GL(n)$

$$t^{i_1 i_2} = \frac{1}{2} (t^{i_1 i_2} + t^{i_2 i_1}) + \frac{1}{2} (t^{i_1 i_2} - t^{i_2 i_1}) .$$

Cette décomposition correspond aux deux tableaux d'Young à 2 boîtes, disposées horizontalement ou verticalement. Pour le rang 3, on écrit de même les tenseurs associés aux 4 tableaux d'Young standards

$$\boxed{1|2|3} \quad A = t^{i_1 i_2 i_3} + t^{i_2 i_3 i_1} + t^{i_3 i_1 i_2} + t^{i_2 i_1 i_3} + t^{i_3 i_2 i_1} + t^{i_1 i_3 i_2} \tag{3.78}$$

$$\begin{array}{|c|} \hline 1 \\ \hline 2 \\ \hline 3 \\ \hline \end{array} \quad B = t^{i_1 i_2 i_3} + t^{i_2 i_3 i_1} + t^{i_3 i_1 i_2} - t^{i_2 i_1 i_3} - t^{i_3 i_2 i_1} - t^{i_1 i_3 i_2} \tag{3.79}$$

$$\begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & \\ \hline \end{array} \quad C_1 = t^{i_1 i_2 i_3} - t^{i_2 i_3 i_1} + t^{i_2 i_1 i_3} - t^{i_3 i_2 i_1} \tag{3.80}$$

$$\begin{array}{|c|c|} \hline 1 & 3 \\ \hline 2 & \\ \hline \end{array} \quad D_1 = t^{i_1 i_2 i_3} - t^{i_2 i_1 i_3} + t^{i_3 i_2 i_1} - t^{i_3 i_1 i_2} \tag{3.81}$$

où pour alléger les notations on n'a pas fait figurer les indices i_1, i_2, i_3 sur A, \dots, D_1 . Tout tenseur se décompose sur cette base :

$$6t^{i_1 i_2 i_3} = A + B + 2(C_1 + D_1) .$$

Les indices 1 sur C et D rappellent que sous l'action du groupe S_3 , ces objets se mélangent avec une autre combinaison $C_2 = t^{i_1 i_3 i_2} - t^{i_3 i_1 i_2} + t^{i_2 i_3 i_1} - t^{i_1 i_2 i_3}$ (resp. $D_2 = t^{i_2 i_1 i_3} + t^{i_2 i_3 i_1} - t^{i_1 i_2 i_3} - t^{i_1 i_3 i_2}$) des t^{ijk} en des représentations de dimension 2. Au contraire l'action du groupe $GL(n)$ mélange entre elles les différentes composantes du tenseur A , celles du tenseur B , etc. Les tenseurs C et D se transforment selon des représentations équivalentes.

Tous les tableaux d'Young, cependant, ne contribuent pas aux représentations de $GL(n)$ pour n donné. Il est clair qu'un tableau à $k > n$ lignes implique une antisymétrisation de k indices prenant leurs valeurs dans

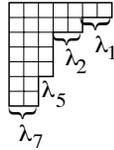


FIGURE 3.7 – Correspondance entre diagramme d'Young et poids le plus haut (ou ses indices de Dynkin). Ici $Y \leftrightarrow \Lambda = (2, 2, 0, 0, 1, 0, 2)$

$\{1, \dots, n\}$ et donne donc un résultat nul. En revanche il est aisé de voir que tout tableau à $k \leq n$ lignes donne lieu à une représentation. On démontre, et nous admettrons, que cette représentation de $GL(n)$ est irréductible et que sa dimension est

$$\dim_Y^{(n)} = \frac{\Delta(f_1 + n - 1, f_2 + n - 2, \dots, f_n)}{\Delta(n - 1, n - 2, \dots, 0)} \tag{3.82}$$

où $\Delta(a_1, a_2, \dots, a_n) = \prod_{i < j} (a_i - a_j)$ est le *déterminant de Vandermonde* des a et les f_i désignent les longueurs des lignes du tableau Y . C'est un polynôme de degré $m = \sum f_i$ en n . Comparer à (3.62).

Dans le cas d'un tableau à une ligne, la formule résulte d'un argument combinatoire simple. La dimension est égale au nombre de composantes du tenseur complètement symétrique t_{i_1, \dots, i_m} où on peut supposer $1 \leq i_1 \leq i_2 \leq \dots \leq i_n \leq n$. Il s'agit de disposer de toutes les façons possibles $n - 1$ signes $<$ parmi les m indices i_1, \dots, i_m pour marquer les blocs successifs de $1, 2, \dots, n$. La dimension cherchée est donc le coefficient binomial C_{n+m-1}^n , en accord dans ce cas particulier avec (3.82).

Dans l'exemple précédent avec $m = 3$, les deux derniers tenseurs C_1 et D_1 se transforment selon des représentations équivalentes. On dit donc que $E^{\otimes 3}$ se décompose en

$$\square\square\square + \begin{matrix} \square \\ \square \\ \square \end{matrix} + 2 \begin{matrix} \square & \square \\ & \square \end{matrix}$$

où la troisième apparaît avec la multiplicité deux. En général, la multiplicité dans $E^{\otimes m}$ d'une représentation de $GL(n)$ indexée par un tableau d'Young est égale à la dimension de la représentation correspondante de S_m .

Cette remarquable relation entre les représentations de S_m et de $GL(n)$ est due à Frobenius et Weyl et appelée *dualité de Frobenius–Weyl*.

On peut étendre ces considérations à d'autres groupes de transformations linéaires, $SL(n)$, $O(n)$, $U(n)$, ... En raison des conditions additionnelles sur les matrices g dans ces groupes, une réduction supplémentaire des représentations peut être possible. Par exemple, on a vu à la sect. 2.2 du chap. 2 que la puissance $E^{\otimes 2}$ de l'espace euclidien à 3 dimensions se réduisait sous l'action de $SO(3)$ en trois sous-espaces, correspondant à des tenseurs de symétrie définie *et de trace nulle* et à un scalaire invariant.

Relations entre les diagrammes d'Young et les poids de $su(n)$

Donnons pour finir la relation entre les deux descriptions des représentations irréductibles obtenues pour $SU(n)$ ou son algèbre de Lie $su(n)$. Dans ce cas, il suffit de limiter le nombre de lignes du diagramme d'Young Y à $k \leq n - 1$ pour obtenir toutes les représentations irréductibles. Le i -ème poids fondamental est représenté par un diagramme d'Young fait d'une colonne de hauteur i , par exemple $\Lambda_3 = \begin{matrix} \square \\ \square \\ \square \end{matrix}$. Et la correspondance entre le poids le plus haut Λ avec ses indices de Dynkin λ_i et le tableau Y avec ses lignes de longueur f_i est comme suit

$$\Lambda = (\lambda_1, \dots, \lambda_{n-1}) \leftrightarrow Y = (f_i = \sum_{j=i}^{n-1} \lambda_j) . \tag{3.83}$$

Autrement dit, λ_k est le nombre de colonnes de Y de hauteur k , voir Fig. 3.7.

★

Références additionnelles

La description de la construction des poids et racines peut se trouver dans de nombreuses références données plus haut (Bump ; Bröcker et Dieck ; Gilmore ...) mais encore dans J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Graduate Texts in Mathematics 9, Springer.

Le chapitre 13 du “Gros livre jaune” de P. Di Francesco, P. Mathieu et D. Sénéchal, [DFMS], *Conformal Field Theory*, Springer, est une mine d’informations sur les algèbres de Lie simples, leurs représentations, les produits tensoriels d’ycelles ...

Pour le calcul des expressions explicites des constantes $N_{\alpha\beta}$, voir [Gi], ou Wybourne, *Classical groups for physicists*, John Wiley.

Sur les octonions et les groupes exceptionnels, voir l’article très complet (et disponible en ligne) de John C. Baez, *Bull. Amer. Math. Soc.* 39 (2002), 145-205.

Pour la classification des formes réelles, voir S. Helgason, *Differential Geometry, Lie groups and Symmetric spaces*, Academic Press, 1978, ou Kirillov, *op. cit.* au chap. 2 .

Appendix F. Les algèbres classiques de type A, B, C, D

F.1 $\mathfrak{sl}(N) = A_{N-1}$

Rang = $l = N - 1$, dimension $N^2 - 1$, nombre de Coxeter $h = N$, nombre de Coxeter dual $h^\vee = N$.

$e_i, i = 1, \dots, N$ un ensemble de vecteurs de \mathbb{R}^N tels que $\sum_1^N e_i = 0, \langle e_i, e_j \rangle = \delta_{ij} - \frac{1}{N}$.

Racines $\alpha_{ij} = e_i - e_j, i \neq j = 1, \dots, N$; racines positives $\alpha_{ij} = e_i - e_j, i < j$; leur nombre $|\Delta_+| = N(N-1)/2$; racines simples $\alpha_i := \alpha_{i, i+1} = e_i - e_{i+1} i = 1, \dots, N-1$.

Racine la plus haute $\theta = \alpha_1 + \dots + \alpha_{N-1} = 2e_1 + e_2 + \dots + e_{N-1} = \Lambda_1 + \Lambda_{N-1} = (1, 0, \dots, 0, 1)$.

Somme des racines positives

$$\begin{aligned} 2\rho &= (N-1)e_1 + (N-3)e_2 + \dots + (N-2i+1)e_i + \dots - (N-1)e_N \\ &= (N-1)\alpha_1 + 2(N-2)\alpha_2 + \dots + i(N-i)\alpha_i + \dots + (N-1)\alpha_{N-1}. \end{aligned} \quad (\text{F.1})$$

$$\text{Matrice de Cartan } \langle \alpha_i, \alpha_j \rangle = \begin{cases} 2 & \text{si } i = j \\ -1 & \text{si } i = j \pm 1 \\ 0 & \text{sinon} \end{cases}$$

Poids fondamentaux $\Lambda_i i = 1, \dots, N-1, \Lambda_i = \sum_{j=1}^i e_j, e_1 = \Lambda_1, e_i = \Lambda_i - \Lambda_{i-1}$ pour $i = 2, \dots, N-1, e_N = -\Lambda_{N-1}$.

$\langle \Lambda_i, \Lambda_j \rangle = \frac{i(N-j)}{N}$ pour $i \leq j$.

Groupe de Weyl : $W \equiv \mathcal{S}_N$ agit sur les poids en permutant les $e_i : w \in W \leftrightarrow \bar{w} \in \mathcal{S}_N : w(e_i) = e_{\bar{w}(i)}$

Exposants de Coxeter $\{1, 2, \dots, N-1\}$.

F.2 $\mathfrak{so}(2l+1) = B_l, l \geq 2$

Rang = l , dimension $l(2l+1)$, nombre de Coxeter $h = 2l$, nombre de Coxeter dual $h^\vee = 2l-1$

$e_i, i = 1, \dots, l, \langle e_i, e_j \rangle = \delta_{ij}$ une base de \mathbb{R}^l .

Racines $\pm e_i, 1 \leq i \leq l$ et $\pm e_i \pm e_j, 1 \leq i < j \leq l$. Base de racines simples $\alpha_i = e_i - e_{i+1}, i = 1, \dots, l-1$, et $\alpha_l = e_l$.

Racines positives

$$\begin{aligned} e_i &= \sum_{i \leq k \leq l} \alpha_k, \quad 1 \leq i \leq l, \\ e_i - e_j &= \sum_{i \leq k < j} \alpha_k, \quad 1 \leq i < j \leq l, \\ e_i + e_j &= \sum_{i \leq k < j} \alpha_k + 2 \sum_{j \leq k \leq l} \alpha_k, \quad 1 \leq i < j \leq l, \end{aligned} \quad (\text{F.2})$$

leur nombre est $|\Delta_+| = l^2$.

Racine la plus haute $\theta = e_1 + e_2 = \alpha_1 + 2\alpha_2 + \dots + 2\alpha_l$.

Somme des racines positives

$$\begin{aligned} 2\rho &= (2l-1)e_1 + (2l-3)e_2 + \dots + (2l-2i+1)e_i + \dots + 3e_{l-1} + e_l \\ &= (2l-1)\alpha_1 + 2(2l-2)\alpha_2 + \dots + i(2l-i)\alpha_i + \dots + l^2\alpha_l. \end{aligned} \quad (\text{F.3})$$

$$\text{Matrice de Cartan } \langle \alpha_i, \alpha_j^\vee \rangle = \begin{cases} 2 & \text{si } 1 \leq i = j \leq l \\ -1 & \text{si } 1 \leq i = (j \pm 1) \leq l-1 \\ -2 & \text{si } i = l-1, j = l \\ -1 & \text{si } i = l, j = l-1 \\ 0 & \text{sinon} \end{cases}$$

Poids fondamentaux $\Lambda_i = \sum_{j=1}^i e_j$, $i = 1, \dots, l-1$, $\Lambda_l = \frac{1}{2} \sum_{j=1}^l e_j$; donc $e_1 = \Lambda_1 = (1, 0, \dots, 0)$, $e_i = \Lambda_i - \Lambda_{i-1} = (0, \dots, -1, 1, 0, \dots)$, $i = 2, \dots, l-1$, $e_l = 2\Lambda_l - \Lambda_{l-1} = (0, \dots, 0, -1, 2)$.

Indices de Dynkin des racines

$\alpha_1 = (2, -1, 0, \dots)$, $\alpha_i = (0, \dots, -1, 2, -1, 0, \dots)$, $i = 2, \dots, l-2$; $\alpha_{l-1} = (0, \dots, 0, -1, 2, -2)$; $\alpha_l = (0, \dots, 0, -1, 2)$ et $\theta = (0, 1, 0, \dots, 0)$

Groupe de Weyl : $W \equiv \mathcal{S}_l \times (\mathbb{Z}_2)^l$, d'ordre $2^l \cdot l!$, agit sur les poids en permutant les e_i et $e_i \mapsto (\pm 1)_i e_i$.

Exposants de Coxeter $\{1, 3, 5, \dots, 2l-1\}$.

F.3. $\mathfrak{sp}(2l) = C_l$, $l \geq 2$

Rang = l , dimension $l(2l+1)$, nombre de Coxeter $h = 2l$, nombre de Coxeter dual $h^\vee = l+1$

e_i , $i = 1, \dots, l$, $\langle e_i, e_j \rangle = \frac{1}{2} \delta_{ij}$ une base de \mathbb{R}^l (**Attention!** au facteur 2 qui assure la normalisation $\theta^2 = 2$).

Base de racines simples $\alpha_i = e_i - e_{i+1}$, $i = 1, \dots, l-1$, et $\alpha_l = 2e_l$.

Racines $\pm 2e_i$, $1 \leq i \leq l$ et $\pm e_i \pm e_j$, $1 \leq i < j \leq l$.

Racines positives

$$\begin{aligned} e_i - e_j &= \sum_{i \leq k < j} \alpha_k, \quad 1 \leq i < j \leq l, \\ e_i + e_j &= \sum_{i \leq k < j} \alpha_k + 2 \sum_{j \leq k < l} \alpha_k + \alpha_l, \quad 1 \leq i < j \leq l, \\ 2e_i &= 2 \sum_{i \leq k < l} \alpha_k + \alpha_l, \quad 1 \leq i \leq l, \end{aligned} \quad (\text{F.4})$$

leur nombre est $|\Delta_+| = l^2$.

Racine la plus haute $\theta = 2e_1 = 2\alpha_1 + 2\alpha_2 + \dots + 2\alpha_{l-1} + \alpha_l$.

Somme des racines positives

$$\begin{aligned} 2\rho &= 2le_1 + (2l-2)e_2 + \dots + (2l-2i+2)e_i + \dots + 4e_{l-1} + 2e_l \\ &= 2l\alpha_1 + 2(2l-1)\alpha_2 + \dots + i(2l-i+1)\alpha_i + \dots + (l-1)(l-2)\alpha_{l-1} + \frac{1}{2}l(l+1)\alpha_l. \end{aligned} \quad (\text{F.5})$$

$$\text{Matrice de Cartan } \langle \alpha_i, \alpha_j^\vee \rangle = \begin{cases} 2 & \text{si } 1 \leq i = j \leq l \\ -1 & \text{si } 1 \leq i = (j \pm 1) \leq l-1 \\ -1 & \text{si } i = l-1, j = l \\ -2 & \text{si } i = l, j = l-1 \\ 0 & \text{sinon} \end{cases}$$

Poids fondamentaux $\Lambda_i = \sum_{j=1}^i e_j$, $i = 1, \dots, l$, donc $e_1 = \Lambda_1 = (1, 0, \dots, 0)$, $e_i = \Lambda_i - \Lambda_{i-1} = (0, \dots, -1, 1, 0, \dots)$, $i = 2, \dots, l$.

Indices de Dynkin des racines

$\alpha_1 = (2, -1, 0, \dots)$, $\alpha_i = (0, \dots, -1, 2, -1, 0, \dots)$, $i = 2, \dots, l-1$; $\alpha_l = (0, \dots, 0, -2, 2)$ and $\theta = (2, 0, \dots, 0)$

Groupe de Weyl : $W \equiv \mathcal{S}_l \times (\mathbb{Z}_2)^l$, d'ordre $2^l \cdot l!$, agit sur les poids en permutant les e_i et $e_i \mapsto (\pm 1)_i e_i$.

F.4. $\mathfrak{so}(2l) = D_l$, $l \geq 3$

Rang = l , dimension $l(2l-1)$, nombre de Coxeter = nombre de Coxeter dual $h = 2l-2 = h^\vee$

e_i , $i = 1, \dots, l$, $\langle e_i, e_j \rangle = \delta_{ij}$ une base de \mathbb{R}^l .

Base de racines simples $\alpha_i = e_i - e_{i+1}$, $i = 1, \dots, l-1$, et $\alpha_l = e_{l-1} + e_l$.

Racines positives

$$\begin{aligned}
 e_i - e_j &= \sum_{i \leq k < j} \alpha_k, \quad 1 \leq i < j \leq l, \\
 e_i + e_j &= \sum_{i \leq k < j} \alpha_k + 2 \sum_{j \leq k < l-1} \alpha_k + \alpha_{l-1} + \alpha_l, \quad 1 \leq i < j \leq l-1, \\
 e_i + e_l &= \sum_{i \leq k \leq l-2} \alpha_k + \alpha_l, \quad 1 \leq i \leq l-1,
 \end{aligned} \tag{F.6}$$

leur nombre est $|\Delta_+| = l(l-1)$.

Racine la plus haute $\theta = e_1 + e_2 = \alpha_1 + 2\alpha_2 + \dots + 2\alpha_{l-2} + \alpha_{l-1} + \alpha_l$.

Somme des racines positives

$$\begin{aligned}
 2\rho &= 2(l-1)e_1 + 2(l-2)e_2 + \dots + 2e_{l-1} \\
 &= 2(l-1)\alpha_1 + 2(2l-3)\alpha_2 + \dots + i(2l-i-1)\alpha_i + \dots + \frac{l(l-1)}{2}(\alpha_{l-1} + \alpha_l).
 \end{aligned} \tag{F.7}$$

Groupe de Weyl : $W \equiv \mathcal{S}_l \times (\mathbb{Z}_2)^{l-1}$, of order $2^{l-1} \cdot l!$, agit sur les poids en permutant les e_i et $e_i \mapsto (\pm 1)_i e_i$, avec $\prod_i (\pm 1)_i = 1$.

Exposants de Coxeter $\{1, 3, 5, \dots, 2l-3, l-1\}$, avec $l-1$ apparaissant donc deux fois si l est pair.

$$\text{Matrice de Cartan } \langle \alpha_i, \alpha_j \rangle = \begin{cases} 2 & \text{si } 1 \leq i = j \leq l \\ -1 & \text{si } 1 \leq i = (j \pm 1) \leq l-2 \\ -1 & \text{si } (i, j) = (l-2, l) \text{ or } (l, l-2) \\ 0 & \text{sinon} \end{cases}$$

Poids fondamentaux $\Lambda_i = \sum_{j=1}^i e_j = \alpha_1 + 2\alpha_2 + \dots + (i-1)\alpha_{i-1} + i(\alpha_i + \dots + \alpha_{l-2}) + \frac{i}{2}(\alpha_{l-1} + \alpha_l)$ pour $i = 1, \dots, l-2$; $\Lambda_{l-1} = \frac{1}{2}(e_1 + \dots + e_{l-1} - e_l) = \frac{1}{2}(\alpha_1 + 2\alpha_2 + \dots + (l-2)\alpha_{l-2}) + \frac{l}{2}\alpha_{l-1} + \frac{l-2}{2}\alpha_l$; $\Lambda_l = \frac{1}{2}(e_1 + \dots + e_{l-1} + e_l) = \frac{1}{2}(\alpha_1 + 2\alpha_2 + \dots + (l-2)\alpha_{l-2}) + \frac{l-2}{2}\alpha_{l-1} + \frac{l}{2}\alpha_l$.

Pour les algèbres exceptionnelles de types E, F, G , voir Bourbaki.

Exercices et Problème du chapitre 3

A. Algèbre de Cartan et racines

1. Montrer que tout élément X de \mathfrak{g} peut s'écrire $X = \sum x^i H_i + \sum_{\alpha \in \Delta} x^\alpha E_\alpha$ avec les notations du § 3.1.2.

Pour H quelconque dans l'algèbre de Cartan, calculer l'action de $\text{ad } H$ sur un tel vecteur X ; en déduire que $\text{ad } H \text{ad } H' X = \sum_{\alpha \in \Delta} x^\alpha \alpha(H) \alpha(H') E_\alpha$ et que (compte tenu que les espaces propres pour chaque racine α ont dimension 1, cf (ii) du § 3.1.2), la forme de Killing s'écrit

$$(H, H') = \text{tr}(\text{ad } H \text{ad } H') = \sum_{\alpha \in \Delta} \alpha(H) \alpha(H'). \quad (3.82)$$

2. On veut montrer que les racines α définies par (3.5) ou (3.6) engendrent bien tout l'espace \mathfrak{h}^* dual de la sous-algèbre de Cartan \mathfrak{h} . Montrer que s'il n'en était pas ainsi, il existerait un élément H de \mathfrak{h} tel que

$$\forall \alpha \in \Delta \quad \alpha(H) = 0. \quad (3.83)$$

Montrer en utilisant (3.82) qu'alors, $\forall H' \in \mathfrak{h}$, $(H, H') = 0$. Pourquoi cela est-il impossible dans une algèbre semi-simple? (cf la discussion précédant l'équation (3.10)).

3. Variante de l'argument précédent : sous l'hypothèse du 2. et donc de (3.83), montrer que H commuterait avec tous les H_i et tous les E_α , donc appartiendrait au centre de \mathfrak{g} . Montrer que le centre d'une algèbre est un idéal abélien. Que peut-on dire du centre d'une algèbre semi-simple? Conclure.

B. Calcul des $N_{\alpha\beta}$

1. Montrer que les constantes réelles $N_{\alpha\beta}$ satisfont $N_{\alpha\beta} = -N_{\beta\alpha}$ et, en conjuguant $[E_\alpha, E_\beta] = N_{\alpha\beta} E_{\alpha+\beta}$ que

$$N_{\alpha\beta} = -N_{-\alpha, -\beta}. \quad (3.84)$$

2. Soient trois racines satisfaisant $\alpha + \beta + \gamma = 0$. En écrivant l'identité de Jacobi pour le triplet $E_\alpha, E_\beta, E_\gamma$, montrer que $\alpha_{(i)} N_{\beta\gamma} + \text{cycl.} = 0$. En déduire la relation

$$N_{\alpha\beta} = N_{\beta, -\alpha-\beta} = N_{-\alpha-\beta, \alpha}. \quad (3.85)$$

3. Considérant la α -chaîne passant par β et les deux entiers p et q définis au § 3.2.1, écrire l'identité de Jacobi pour $E_\alpha, E_{-\alpha}$ et $E_{\beta+k\alpha}$, avec $p \leq k \leq q$, et montrer qu'elle implique

$$\langle \alpha, \beta + k\alpha \rangle = N_{-\alpha, \beta+k\alpha} N_{\alpha, \beta+(k-1)\alpha} + N_{\beta+k\alpha, \alpha} N_{-\alpha, \beta+(k+1)\alpha}.$$

Soit $f(k) := N_{\alpha, \beta+k\alpha} N_{-\alpha, -\beta-k\alpha}$. Montrer en utilisant les relations (3.85) que l'équation précédente se réécrit

$$\langle \alpha, \beta + k\alpha \rangle = f(k) - f(k-1). \quad (3.86)$$

4. Que valent $f(q)$ et $f(q-1)$? Montrer que la récurrence (3.86) se résout en

$$f(k) = -(N_{\alpha, \beta+k\alpha})^2 = (k-q) \langle \alpha, \beta + \frac{1}{2}(k+q+1)\alpha \rangle. \quad (3.87)$$

Que vaut $f(p-1)$? Montrer que l'on retrouve l'expression (3.21). Montrer que la formule (3.87) est en accord avec (3.23). Il reste à déterminer le signe de la racine carrée... cf [Gi].

C. Étude des algèbres $B_l = \mathfrak{so}(2l+1)$ et G_2

1. $\mathfrak{so}(2l+1) = B_l$, $l \geq 2$

a. Quelle est la dimension du groupe $\text{SO}(2l+1)$ ou de son algèbre de Lie $\mathfrak{so}(2l+1)$?

b. Quel est le rang de l'algèbre? (on cherchera à diagonaliser une matrice de $\mathfrak{so}(2l+1)$ sur \mathbb{C} , ou à l'écrire comme une diagonale de blocs 2×2 , cf § 3.1)

c. Combien l'algèbre a-t-elle de racines? Combien de racines positives? Combien de simples?

d. Soit e_i , $i = 1, \dots, l$, une base orthonormée de \mathfrak{R}^l , $\langle e_i, e_j \rangle = \delta_{ij}$. On considère l'ensemble de vecteurs

$$\Delta = \{\pm e_i, 1 \leq i \leq l\} \cup \{\pm e_i \pm e_j, 1 \leq i < j \leq l\}$$

Quel est le cardinal de Δ ? Δ décrit l'ensemble des racines de l'algèbre $\mathfrak{so}(2l+1)$.

e. Une base de racines simples est donnée par $\alpha_i = e_i - e_{i+1}$, $i = 1, \dots, l-1$, et $\alpha_l = e_l$. Expliquer pourquoi les racines

$$\begin{aligned} e_i &= \sum_{i \leq k \leq l} \alpha_k, \quad 1 \leq i \leq l, \\ e_i - e_j &= \sum_{i \leq k < j} \alpha_k, \quad 1 \leq i < j \leq l, \\ e_i + e_j &= \sum_{i \leq k < j} \alpha_k + 2 \sum_{j \leq k \leq l} \alpha_k, \quad 1 \leq i < j \leq l, \end{aligned} \quad (3.88)$$

se qualifient comme racines positives.

Vérifier cette assertion sur le cas de $B_2 = \mathfrak{so}(5)$.

f. Calculer la matrice de Cartan et vérifier qu'elle est en accord avec le diagramme de Dynkin donné en cours.

g. Calculer la somme ρ des racines positives.

h. Le groupe de Weyl est le produit ("semi-direct") $W \equiv \mathcal{S}_l \times (\mathbb{Z}_2)^l$, qui agit sur les e_i (et donc sur les poids et racines) par permutation et par changements de signe indépendants $e_i \mapsto (\pm 1)_i e_i$. Quel est son ordre? Dans le cas de B_2 , dessiner la première chambre de Weyl.

i. Montrer que les vecteurs $\Lambda_i = \sum_{j=1}^i e_j$, $i = 1, \dots, l-1$, $\Lambda_l = \frac{1}{2} \sum_{j=1}^l e_j$ sont les poids fondamentaux.

j. Calculer en utilisant la formule de Weyl : $\dim(\Lambda) = \prod_{\alpha > 0} \frac{\langle \Lambda + \rho, \alpha \rangle}{\langle \rho, \alpha \rangle}$ la dimension des 2 représentations fondamentales de B_2 et celle de poids le plus haut $2\Lambda_2$. Au vu de ces dimensions, à quoi correspondent ces représentations de $\mathrm{SO}(5)$?

k. Dessiner sur un même dessin les racines et les premiers poids de $\mathfrak{so}(5)$.

2. G_2

Dans l'espace \mathfrak{R}^2 , on considère trois vecteurs e_1, e_2, e_3 de somme nulle, $\langle e_i, e_j \rangle = \delta_{ij} - \frac{1}{3}$, puis on construit les 12 vecteurs

$$\pm(e_1 - e_2), \pm(e_1 - e_3), \pm(e_2 - e_3), \pm(2e_1 - e_2 - e_3), \pm(2e_2 - e_1 - e_3), \pm(2e_3 - e_1 - e_2)$$

Ils constituent le système de racines de G_2 , comme on va le vérifier.

a. Que peut-on dire alors de la dimension de l'algèbre G_2 ?

b. Montrer que $\alpha_1 = e_1 - e_2$ et $\alpha_2 = -2e_1 + e_2 + e_3$ sont deux racines simples, en accord avec le diagramme de Dynkin de G_2 donné dans le cours. Calculer la matrice de Cartan.

c. Que sont les racines positives? Calculer le vecteur ρ , demi-somme des racines positives.

d. Quel est le groupe d'invariance du diagramme des racines? Montrer qu'il est d'ordre 12 et que c'est le groupe de Weyl de G_2 . Dessiner la première chambre de Weyl.

e. Vérifier que les poids fondamentaux sont

$$\Lambda_1 = 2\alpha_1 + \alpha_2 \quad \Lambda_2 = 3\alpha_1 + 2\alpha_2$$

f. Que sont les dimensions de représentations fondamentales?

g. Dans les deux cas de B_2 et G_2 , on constate que le poids le plus haut de la représentation adjointe est donné par la racine la plus haute. Expliquer pourquoi cela est vrai en général.

3. Un peu de physique (à voir après étude du Chap. 4)

Pourquoi les groupes $\mathrm{SO}(5)$ ou G_2 ne convenaient-ils pas comme groupe de symétrie étendant le $\mathrm{SU}(2)$ d'isospin, sachant qu'on avait observé plusieurs "octets" de particules?

D. Systèmes de racines. Repléments de diagrammes de Dynkin

On considère les racines simples α_i de l'algèbre $\mathfrak{su}(2n)$, numérotées comme dans le cours. (Attention! nous disons bien $2n$!)

1. Quel est le rang de cette algèbre? Que valent les $\langle \alpha_i, \alpha_j \rangle$? Dessiner le diagramme de Dynkin correspondant. Quelle est la symétrie de ce diagramme?
2. On définit alors $\beta_i = (\alpha_i + \alpha_{2n-i})/\sqrt{2}$, pour $i = 1, \dots, n-1$ et $\beta_n = \alpha_n/\sqrt{2}$. Calculer les $\langle \beta_i, \beta_j \rangle$.
3. Montrer que les β forment un système de racines que l'on précisera.
4. Plus généralement, tout système de racines simples de même longueur peut être "replié" selon une symétrie éventuelle de son diagramme de Dynkin et donne alors un autre diagramme de Dynkin. *Sans aucun calcul*, quel diagramme devrait-on obtenir de cette manière à partir du diagramme E_6 ?

E. Dimensions des représentations de $SU(3)$

Admettant que la construction du § 3.4.2, celle des tenseurs de rang (p, m) complètement symétriques et de trace nulle de \mathbb{C}^3 , fournit bien les représentations irréductibles de poids le plus haut (p, m) de $SU(3)$, on se propose de calculer la dimension $d(p, m)$ de l'espace de ces tenseurs.

1. Montrer, en étudiant le produit de deux tenseurs de rang $(p, 0)$ et $(0, m)$ et en séparant les termes de trace (contenant un δ_j^i entre des indices supérieur et inférieur) que $(p, 0) \otimes (0, m) = ((p-1, 0) \otimes (0, m-1)) \oplus (p, m)$ et donc que

$$d(p, m) = d(p, 0)d(0, m) - d(p-1)d(0, m-1).$$

2. Montrer par un calcul semblable à celui de $SU(2)$ que

$$d(p, 0) = d(0, p) = \frac{1}{2}(p+1)(p+2).$$

3. En déduire l'expression de $d(p, m)$ et comparer à (3.64).

Problème : algèbre de Lie à identifier

1. Rappel. Étant données deux matrices A et B , carrées de dimension 2×2 , on définit la matrice $A \otimes B$

$$(A \otimes B)_{ij;kl} := A_{ik}B_{jl}$$

et on convient d'ordonner les paires (ij) ou (kl) selon l'ordre lexicographique 11, 12, 21, 22.

- (a) Montrer que le produit de deux telles matrices satisfait

$$(A \otimes B) \cdot (C \otimes D) = (A \cdot C) \otimes (B \cdot D).$$

- (b) En déduire une expression du commutateur $[A \otimes B, C \otimes D]$ en termes des commutateurs $[A, C]$ et $[B, D]$ (avec des coefficients qui peuvent encore impliquer les matrices A, \dots, D).
2. On considère alors les 3 matrices de Pauli σ_a , $a = 1, 2, 3$ et la matrice identité I en dimension 2. On forme les 10 matrices

$$A_a = \sigma_a \otimes I, B_a = \sigma_a \otimes \sigma_1, C_a = \sigma_a \otimes \sigma_3, D = I \otimes \sigma_2.$$

(Dans la mesure du possible, on s'abstiendra d'écrire l'expression explicite de ces matrices.)

- (a) Avec le minimum de calculs, calculer les commutateurs de ces matrices et montrer qu'elles forment une algèbre de Lie \mathfrak{g} . On admettra que cette algèbre est simple.
- (b) Soient $H_1 = A_3$ et $H_2 = C_3$. Pourquoi peut-on dire qu'elles appartiennent à une sous-algèbre de Cartan? Que signifie l'assertion: "elles engendrent une sous-algèbre de Cartan"? On admettra que c'est bien le cas. Quel est le rang de l'algèbre \mathfrak{g} ?
- (c) Montrer que l'on peut trouver 4 combinaisons linéaires $X(\epsilon_1, \epsilon_2) = (A_1 + \epsilon_1 C_1) + \epsilon_2 i(A_2 + \epsilon_1 C_2)$ avec $\epsilon_1, \epsilon_2 = \pm 1$ telles que $[H_i, X(\epsilon_1, \epsilon_2)] = \gamma_i(\epsilon_1, \epsilon_2)X(\epsilon_1, \epsilon_2)$, et déterminer les $\gamma_i(\epsilon_1, \epsilon_2)$. (Indication: ces $\gamma_i(\epsilon_1, \epsilon_2)$ prennent les valeurs ± 2 .)

- (d) De même montrer que $B_1 \pm iB_2$ et $B_3 \pm iD$ ont également des propriétés de commutation simples avec H_1 et H_2 . (Indication : les “valeurs propres” sont cette fois $0, \pm 2$.)
 - (e) Que peut-on dire des racines de l’algèbre \mathfrak{g} ? On en donnera les composantes dans une base de l’algèbre de Cartan (duale).
3. On va maintenant identifier plus précisément \mathfrak{g} .
- (a) Donner un système de racines positives, puis un système de racines simples.
 - (b) Calculer la matrice de Cartan. Identifier de quelle algèbre de Lie il s’agit dans la classification de Cartan.
 - (c) Dans le plan des racines, dessiner les racines simples, l’ensemble des racines. Indiquer quelle est la chambre de Weyl fondamentale.
 - (d) Calculer les composantes des poids fondamentaux et les indiquer sur la figure précédente.
 - (e) Quel est le vecteur de Weyl? Calculer la dimension des deux représentations fondamentales. A quoi correspondent-elles d’un point de vue géométrique?

Chapitre 4

Symétries globales en physique des particules

La physique des particules va nous offrir un terrain de choix pour illustrer les différentes manifestations de symétries en physique. Nous ne nous occuperons dans ce chapitre et le suivant que de “symétries internes”, excluant les symétries d’espace-temps.

Nous allons examiner tour à tour différents types de symétries et leurs réalisations, comme symétrie exacte ou brisée explicitement, spontanément ou par des anomalies quantiques. Nous consacrerons pas mal d’attention au groupe de saveur $SU(3)$.

4.1 Symétries globales exactes ou brisées. Brisure spontanée

4.1.1 Panorama. Symétries exactes ou brisées

Les transformations dont on s’occupera dans ce chapitre sont des *transformations globales*. Un groupe G agit sur les degrés de liberté de chaque champ $\phi(x)$ de la même façon en tous les points x de l’espace-temps. Par exemple, G agit sur ϕ par l’intermédiaire d’une représentation, et à chaque élément g du groupe correspond une matrice ou opérateur $D(g)$, indépendamment du point x

$$\phi(x) \mapsto D(g)\phi(x) . \quad (4.1)$$

Dans une théorie quantique, selon le théorème de Wigner, on suppose que cette transformation est aussi réalisée sur les états de l’espace de Hilbert par un opérateur unitaire $U(g)$; en tant qu’opérateur, $\phi(x) \mapsto U(g)\phi(x)U^\dagger(g)$.

Cette transformation peut être une *symétrie* de la dynamique, auquel cas $U(g)$ commute avec l’hamiltonien du système, ou dans le langage lagrangien, elle laisse le lagrangien invariant et donne donc naissance à des courants de Noether j_μ^a de divergence nulle et à des charges

conservées $Q^a = \int d\mathbf{x} j_0^a(\mathbf{x}, t)$, $a = 1, \dots, \dim G$. Ces charges agissent sur les champs comme générateurs infinitésimaux, classiquement au sens du crochet de Poisson, $\{Q^a, \phi(x)\} \delta\alpha_a = \delta\phi(x)$, et si tout se passe bien dans la théorie quantique, comme opérateurs dans l'espace de Hilbert dotés de relations de commutation avec les champs $[Q^a, \phi(x)] \delta\alpha_a = -i\hbar\delta\phi(x)$ et entre eux $[Q^a, Q^b] = iC_{ab}^c Q^c$. Une question importante va en effet être de savoir si une symétrie apparente au niveau classique, disons sur le lagrangien, est bien réalisée dans la théorie quantique.

- Un exemple de *symétrie exacte* est fourni par l'invariance de groupe $U(1)$, associé à la conservation de la charge électrique. Un champ portant une charge électrique q (fois $|e|$) est un champ complexe, il se transforme sous l'action du groupe $U(1)$ selon la représentation irréductible indexée par l'entier q

$$\phi(x) \mapsto e^{iq\alpha} \phi(x) ; \quad \phi^\dagger(x) \mapsto e^{-iq\alpha} \phi^\dagger(x).$$

S'il y a invariance (du lagrangien) quand tous les champs se transforment ainsi, on a alors un courant de Noether $j^\mu(x)$, somme des contributions des différents champs chargés, de divergence nulle, $\partial_\mu j^\mu(x) = 0$, et la charge associée Q est conservée. La théorie quantique est l'électrodynamique quantique, et on y démontre que la symétrie classique par le groupe $U(1)$ ainsi que la conservation du courant (et l'invariance de jauge) sont bien préservées par la quantification et en particulier par la renormalisation, par exemple que toutes les charges électriques se renormalisent de la même façon, cf cours de Théorie Quantique des Champs.

D'autres invariances et lois de conservation de nature similaire sont celles associées aux charges baryoniques ou leptoniques, conservées (jusqu'à plus ample informé ...).

- Une symétrie peut aussi être *brisée* explicitement. Par exemple le lagrangien contient des termes non invariants sous l'action de G . Dans ce cas, les courants de Noether ne sont pas conservés, mais leur divergence s'écrit

$$\partial_\mu j_i^\mu(x) = \frac{\partial \mathcal{L}(x)}{\partial \alpha^i}. \quad (4.2)$$

Nous verrons plus bas avec $SU(3)$ un exemple de symétrie brisée (ou approchée).

Certains types de brisures, dites "douces" (soft), sont telles que la symétrie est restaurée à courte distance ou haute énergie. C'est par exemple le cas de l'invariance d'échelle (par dilatations d'espace-temps), brisée par la présence de toute échelle de masse dans la théorie, mais restaurée –de façon un peu subtile– à courte distance, cf. l'étude du groupe de renormalisation dans les cours de théorie des champs.

- Un mécanisme plus subtil de brisure de symétrie est celui de *brisure spontanée de symétrie*. On appelle ainsi les situations où l'état fondamental du système ne possède pas une symétrie apparente sur le lagrangien. L'exemple le plus simple présentant ce phénomène est celui d'un système classique à un degré de liberté décrit par le potentiel "à double puits" de la Fig. 1(a). Bien que le potentiel exhibe une symétrie \mathbb{Z}_2 manifeste par $x \rightarrow -x$, le système choisit un état fondamental dans l'un des deux minima du potentiel, ce qui brise la symétrie. Ce mécanisme joue un rôle fondamental en physique, avec des manifestations dans des situations très variées, de la matière condensée –ferromagnétisme, superfluidité, supraconductivité ...– à la physique des particules –symétrie chirale, phénomène de Brout–Englert–Higgs– et à la cosmologie.

▷ *Exemple. Brisure spontanée du modèle $O(n)$*

Le lagrangien du "modèle $O(n)$ " bosonique (minkovskien, ici), pour un champ $\phi = \{\phi^i\}$ réel à

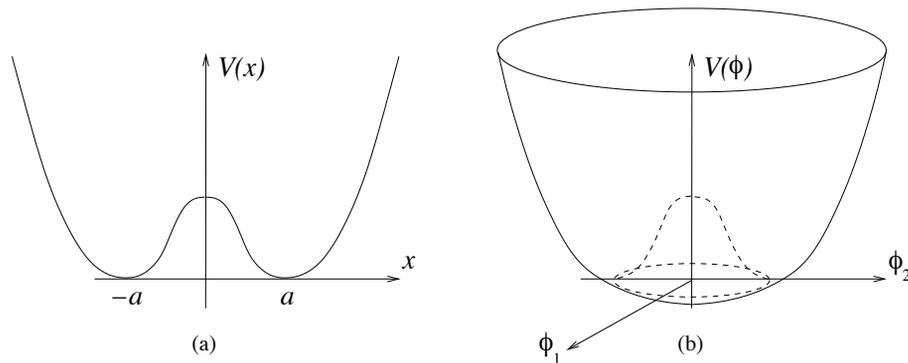


FIGURE 4.1 – Potentiels (a) à “double puits” ; (b) en “chapeau mexicain”

n composantes,

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4}(\phi^2)^2 \quad (4.3)$$

est invariant sous l’effet des rotations de $O(n)$. Le courant de Noether $j_\mu^a = \partial_\mu\phi^i(T^a)_{ij}\phi^j$ (avec T^a antisymétrique réelle) a une divergence nulle, ce qui implique la conservation d’une “charge” etc. Le minimum du potentiel correspond à l’état fondamental, alias le vide, de la théorie. Si le paramètre m^2 est choisi négatif, le minimum du potentiel $V = \frac{1}{2}m^2\phi^2 + \frac{\lambda}{4}(\phi^2)^2$ n’est plus en $\phi^2 = 0$ mais en une certaine valeur v^2 de ϕ^2 telle que $-m^2 = \lambda v^2$, cf Fig. 1(b). Le champ ϕ “choisit” spontanément une direction \hat{n} ($\hat{n}^2 = 1$) dans l’espace interne, dans laquelle sa valeur moyenne dans le vide (“vev” dans le jargon franglais) est non nulle

$$\langle 0|\phi|0\rangle = v\hat{n} . \quad (4.4)$$

Cette “vev” brise le groupe d’invariance $G = O(n)$ de départ en son sous-groupe H qui laisse invariant le vecteur $\langle 0|\phi|0\rangle = v\hat{n}$, soit un groupe isomorphe à $O(n-1)$. Que cette valeur moyenne dans le vide d’un champ non invariant par le groupe soit non nulle, $\langle 0|\phi|0\rangle \neq 0$, est le signal que le vide n’est pas invariant : on est bien dans un cas de *symétrie brisée spontanément*. C’est le mécanisme à l’œuvre dans un ferromagnétique à basse température, par exemple, où l’aimantation non nulle signale la brisure spontanée de la symétrie d’isotropie spatiale.

Exercice (cf cours de F. David) : Posant $\phi = (v+\sigma)\hat{n} + \pi$, où π désigne les $n-1$ composantes du champ ϕ orthogonales à $\langle \phi \rangle = v\hat{n}$, calculer les termes de $V(\sigma, \pi)$ linéaires et quadratiques dans les champs σ et π ; vérifier que le terme linéaire en σ s’annule (minimum du potentiel), que σ a un terme de masse non nul, mais que les π sont de masse nulle, ce sont les *bosons de Nambu-Goldstone* de la symétrie brisée spontanément. Il s’agit là d’un phénomène général : toute symétrie continue brisée spontanément s’accompagne de l’apparition d’excitations de masse nulle en nombre égal à celui des générateurs de la symétrie brisée (*théorème de Goldstone*). Plus précisément quand un groupe G se brise spontanément en un sous-groupe H (le groupe de symétrie résiduelle, groupe d’invariance du fondamental), il apparaît un nombre $d(G) - d(H)$ de bosons de Goldstone de masse nulle. Dans l’exemple précédent, $G = O(n)$, $H = O(n-1)$, $d(G) - d(H) = n - 1$.

Donnons une démonstration simple de ce théorème dans le cas d’une théorie lagrangienne des champs. On écrit $\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - V(\phi)$ avec des notations très génériques, ϕ désigne un ensemble de champs $\{\phi_i\}$ sur lequel

agit un groupe de transformations G . Le potentiel V est supposé invariant sous l'action de transformations infinitésimales $\delta^a \phi_i$, $a = 1, \dots, \dim G$. Par exemple pour des transformations linéaires : $\delta^a \phi_i = T_{ij}^a \phi_j$. On a donc

$$\frac{\partial V(\phi(x))}{\partial \phi_i(x)} \delta^a \phi_i(x) = 0 .$$

Dérivons cette équation par rapport à $\phi_j(x)$ (en omettant l'argument x partout)

$$\frac{\partial V}{\partial \phi_i} \frac{\partial \delta^a \phi_i}{\partial \phi_j} + \frac{\partial^2 V}{\partial \phi_i \partial \phi_j} \delta^a \phi_i = 0$$

et évaluons la en $\phi(x) = v$, un minimum (constant, indépendant de x) du potentiel : le premier terme s'annule, le second nous dit que

$$\left. \frac{\partial^2 V}{\partial \phi_i \partial \phi_j} \right|_{\phi=v} \delta^a v_i = 0 , \quad (4.5)$$

en notant (un peu abusivement) $\delta^a v_i = \delta^a \phi_i|_{\phi=v}$. Par ailleurs on quantifie la théorie au voisinage de ce minimum v ("vide" de la théorie) en écrivant $\phi(x) = v + \varphi(x)$ et en développant

$$V(\phi) = V(v) + \frac{1}{2} \left. \frac{\partial^2 V}{\partial \phi_i \partial \phi_j} \right|_{\phi=v} \varphi_i \varphi_j + \dots$$

et les masses des champs φ se lisent alors sur la forme quadratique. Or (4.5) nous apprend que la "matrice de masse" $\left. \frac{\partial^2 V}{\partial \phi_i \partial \phi_j} \right|_{\phi=v}$ a autant de "modes zéros" (vecteurs propres de valeur propre nulle) qu'il y a de variations indépendantes $\delta^a v_i \neq 0$. Si H est le groupe d'invariance de v , $\delta^a v_i \neq 0$ pour les générateurs de G qui ne sont pas générateurs de H , et il y a donc bien $\dim G - \dim H$ modes de masse nulle, cqfd.

4.1.2 Brisure de la symétrie chirale

Considérons un lagrangien impliquant des fermions de masse nulle,

$$\mathcal{L} = \bar{\psi} i \not{\partial} \psi + g(\bar{\psi} \gamma^\mu \psi)(\bar{\psi} \gamma_\mu \psi) , \quad (4.6)$$

où $\psi = \{\psi_i\}_{i=1, \dots, N}$ est un vecteur à N composantes qui sont des champs de 4-spinors. Noter l'absence de terme de masse $\bar{\psi} \psi$. Ce lagrangien est invariant sous l'action des deux types de transformations infinitésimales

$$\begin{aligned} \delta_A \psi(x) &= \delta A \psi(x) \\ \delta_B \psi(x) &= \delta B \gamma_5 \psi(x) , \end{aligned} \quad (4.7)$$

où les matrices A et B sont antihermitiennes infinitésimales, de taille $N \times N$; elles agissent sur les indices de "saveur" i mais pas sur les indices spinoriels, et commutent donc avec les matrices γ . On rappelle que γ_5 est hermitienne et anticommute avec les γ_μ . Vérifier que $\delta_A \bar{\psi} = -\bar{\psi} \delta A$, $\delta_B \bar{\psi} = \bar{\psi} \delta B \gamma_5$. Les courants de Noether conservés sont respectivement

$$J_\mu^a = \bar{\psi} T^a \gamma_\mu \psi \quad J_\mu^{a(5)} = \bar{\psi} T^a \gamma_5 \gamma_\mu \psi , \quad (4.8)$$

avec T^a des générateurs infinitésimaux du groupe unitaire $U(N)$. Les transformations de la première ligne sont dites "vectorielles", celle de la seconde ligne, qui impliquent γ_5 , sont dites "axiales". On peut aussi redire cela en termes de transformations indépendantes de $\psi_L :=$

$\frac{1}{2}(I - \gamma_5)\psi$ et de $\psi_R := \frac{1}{2}(I + \gamma_5)\psi$; on se rappelle que $(\gamma_5)^2 = I$, et que $\frac{1}{2}(I \pm \gamma_5)$ sont donc des projecteurs; on a donc $\bar{\psi}_L = \frac{1}{2}\bar{\psi}(I + \gamma_5)$, etc, et

$$\mathcal{L} = \bar{\psi}_L i \not{\partial} \psi_L + \bar{\psi}_R i \not{\partial} \psi_R + (\bar{\psi}_L \gamma_\mu \psi_L + \bar{\psi}_R \gamma_\mu \psi_R)(\bar{\psi}_L \gamma^\mu \psi_L + \bar{\psi}_R \gamma^\mu \psi_R)$$

qui est clairement invariant par les transformations unitaires finies $\psi_L \rightarrow U_1 \psi_L$, $\psi_R \rightarrow U_2 \psi_R$, avec $U_1, U_2 \in U(N)$. Le groupe de *symétrie chirale* est donc $U(N) \times U(N)$.

Si maintenant nous introduisons un terme de masse $\delta\mathcal{L} = -m\bar{\psi}\psi$ (qui “couple” les composantes ψ_L et ψ_R : $\delta\mathcal{L} = -m(\bar{\psi}_R \psi_L + \bar{\psi}_L \psi_R)$), la symétrie “vectorielle” est préservée, mais l’axiale ne l’est plus et donne lieu à une divergence

$$\partial^\mu J_\mu^{a(5)}(x) \propto m\bar{\psi}T^a\gamma_5\psi . \quad (4.9)$$

Le groupe de symétrie résiduelle est $U(N)$, sous-groupe “diagonal” de $U(N) \times U(N)$ (diagonal en ce sens que l’on prend $U_1 = U_2$ dans les transformations de $\psi_{L,R}$.)

La symétrie axiale peut aussi être brisée spontanément. Partons d’un lagrangien somme de termes du type (4.6) avec $N = 2$ et (4.3) pour $n = 4$, avec un terme couplant les fermions à quatre champs de bosons, nommés traditionnellement σ et $\boldsymbol{\pi}$

$$\mathcal{L} = \bar{\psi}(i\not{\partial} + g(\sigma + i\boldsymbol{\pi}\cdot\boldsymbol{\tau}\gamma_5))\psi + \frac{1}{2}((\partial\boldsymbol{\pi})^2 + (\partial\sigma)^2) - \frac{1}{2}m^2(\sigma^2 + \boldsymbol{\pi}^2) - \frac{\lambda}{4}(\sigma^2 + \boldsymbol{\pi}^2)^2 , \quad (4.10)$$

dans laquelle les matrices de Pauli ont été désignées exceptionnellement par $\boldsymbol{\tau}$ pour ne pas les confondre avec le champ σ . Le groupe de symétrie est $U(2) \times U(2)$, avec les champs ψ_L , ψ_R et $\sigma + i\boldsymbol{\pi}\cdot\boldsymbol{\tau}$ se transformant respectivement par les représentations $(\frac{1}{2}, 0)$, $(0, \frac{1}{2})$ et $(\frac{1}{2}, \frac{1}{2})$ de $SU(2) \times SU(2)$ (cf exercice A). Si $m^2 < 0$, le champ $\phi = (\sigma, \boldsymbol{\pi})$ acquiert une vev, qu’on peut orienter dans la direction σ en ayant introduit au préalable un petit terme de brisure explicite $\delta\mathcal{L} = c\sigma$, l’analogie d’un petit champ magnétique, qu’on fait tendre vers zéro par la suite. La vev est donnée comme plus haut par $v^2 = -m^2/\lambda$, et en réécrivant le champ $\sigma(x) = \sigma'(x) + v$, où le champ σ' a maintenant une valeur moyenne nulle dans le vide, on voit que les fermions ont acquis un terme de masse $m_\psi = -gv$, tandis que les $\boldsymbol{\pi}$ sont de masse nulle. Ce lagrangien, le modèle σ de Gell-Mann–Lévy, a été proposé comme modèle expliquant la brisure de l’invariance chirale et la faible masse des mésons π , quasi-bosons de Nambu–Goldstone (“quasi” parce que la symétrie chirale n’est qu’approchée avant d’être brisée spontanément). Nous en retrouverons des éléments dans le modèle standard.

4.1.3 Brisures quantiques. Anomalies

Un autre mode de brisure de symétrie, de nature purement quantique, se manifeste dans les *anomalies* des théories quantiques de champs. Une symétrie, apparente au niveau classique du lagrangien, est brisée par l’effet des “corrections quantiques”. C’est par exemple ce qui se produit avec certaines symétries chirales, du type qu’on vient d’étudier : un courant axial de divergence classiquement nulle peut acquérir par un “effet à une boucle” une divergence $\partial_\mu J_5^\mu \neq 0$. Dans le cas où le courant “anormal” est le courant de Noether d’une symétrie classique interne, cette symétrie est brisée par l’anomalie quantique, ce qui peut donner lieu à des effets physiques intéressants (cf. discussion de la désintégration $\pi^0 \rightarrow \gamma\gamma$, par exemple dans [IZ] chap 11). Mais dans une théorie comme une théorie de jauge où la conservation du courant axial est cruciale pour assurer la

cohérence –renormalisabilité, unitarité–, l’anomalie constitue un danger potentiel qu’il faut contrôler. C’est ce qui se produit dans le Modèle Standard, et nous y reviendrons au chap. 5. Un autre exemple est fourni par l’invariance par dilatation d’une théorie de masse nulle, cf l’étude du groupe de renormalisation dans le cours de F. David.

4.2 La symétrie de saveur SU(3) et le modèle des quarks

Une symétrie approchée de grande importance est celle de SU(3) de saveur, à laquelle nous consacrons le reste de ce chapitre.

4.2.1 Pourquoi le groupe SU(3) ?

On a vu (Chap. 0) que, selon l’observation remarquable de Heisenberg (1932), si on néglige leurs interactions faibles et électromagnétiques, les *hadrons*, c’est-à-dire les particules soumises aux interactions fortes telles le proton et le neutron, les mésons π etc, sont classifiés en “multiplets” d’un groupe SU(2) d’isospin. Ou dit autrement, l’hamiltonien (ou le lagrangien) des interactions fortes est invariant sous l’action de ce groupe SU(2) et le groupe SU(2) est représenté dans l’espace des états hadroniques par des représentations unitaires. Proton et neutron appartiennent à une représentation de dimension 2, d’isospin $\frac{1}{2}$, les trois pions π^\pm, π^0 forment une représentation de dimension 3, d’isospin 1, etc. La charge électrique Q de chacune des particules que nous venons de citer est reliée à la valeur propre de la troisième composante I_z de l’isospin par la relation

$$Q = \frac{1}{2}\mathcal{B} + I_z \quad [\text{for SU}(2)] \quad (4.11)$$

où apparaît un nouveau nombre quantique \mathcal{B} , la *charge baryonique*, supposée conservée (additivement) dans *toutes* les interactions (jusqu’à nouvel ordre). \mathcal{B} vaut 0 pour les mésons π , et 1 pour les “baryons” que sont le proton et le neutron, -1 pour leurs antiparticules, 4 pour une particule α (noyau d’helium), etc.

Cette relation entre Q et I_z doit être amendée pour de nouvelles familles de mésons K^\pm, K^0, \bar{K}^0 ou de baryons $\Lambda^0, \Sigma, \Xi, \dots$ découverts dans les années 50. On leur attribue un nouveau nombre quantique, l’*étrangereté* S . Cette étrangereté est également supposée conservée (additivement) dans les interactions fortes. Ainsi, si S vaut -1 pour le Λ^0 et $+1$ pour le K^+ et le K^0 , le processus $p + \pi^- \rightarrow \Lambda^0 + K^0$ conserve l’étrangereté, tandis que la désintégration observée $\Lambda^0 \rightarrow p + \pi^-$ viole cette loi de conservation, car elle procède par les interactions faibles. La relation (4.11) doit être modifiée en la *relation de Gell-Mann–Nishima*

$$Q = \frac{1}{2}\mathcal{B} + \frac{1}{2}S + I_z = \frac{1}{2}Y + I_z, \quad (4.12)$$

où on a introduit l’*hypercharge* Y , qui, à ce stade, vaut $Y = \mathcal{B} + S$.

Ces lois de conservation et différentes propriétés des mésons et baryons découverts alors, en particulier leur organisation en “octets”, ont conduit au début des années 60 Gell-Mann et Ne’eman à postuler l’existence d’un groupe SU(3) de symétrie approchée des interactions fortes. Les nombres quantiques conservés et simultanément mesurables I_z et Y sont interprétés

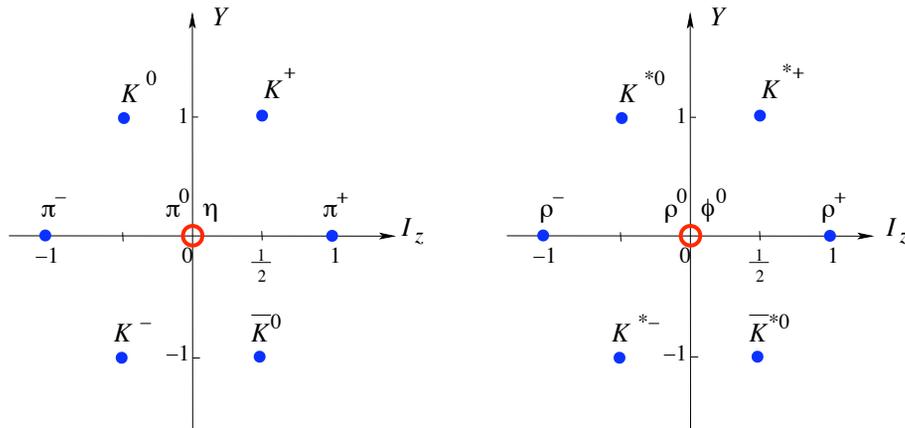


FIGURE 4.2 – Les octets de mésons. pseudoscalaires ($J^P = 0^-$) et vecteurs ($J^P = 1^-$)

comme les valeurs propres de deux charges commutantes, c'est-à-dire de deux éléments d'une algèbre de Cartan de rang 2, et c'est l'algèbre de SU(3) qui est le candidat naturel, puisque possédant une représentation de dimension 8 (cf exercice C du chap. 3).

Dans la représentation 3 de définition de SU(3), on construit une base de l'algèbre de Lie su(3), faite de 8 matrices hermitiennes λ_a qui jouent le rôle des matrices de Pauli σ_i pour su(2). Ces matrices sont normalisées par

$$\text{tr } \lambda_a \lambda_b = 2\delta_{ab} . \tag{4.13}$$

λ_1 et λ_2 , λ_4 et λ_5 , λ_6 et λ_7 ont les mêmes éléments de matrice que σ_1 et σ_2 en position

$\begin{pmatrix} \cdot & * & \cdot \\ * & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}$, $\begin{pmatrix} \cdot & \cdot & * \\ \cdot & \cdot & \cdot \\ * & \cdot & \cdot \end{pmatrix}$ et $\begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & * \\ \cdot & * & \cdot \end{pmatrix}$ respectivement, où les points signifient des zéros. Les deux générateurs de l'algèbre de Cartan sont

$$\lambda_3 = \begin{pmatrix} 1 & \cdot & \cdot \\ \cdot & -1 & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & \cdot & \cdot \\ \cdot & 1 & \cdot \\ \cdot & \cdot & -2 \end{pmatrix} . \tag{4.14}$$

Les charges I_z et Y sont alors les représentants dans la représentation considérée de $\frac{1}{2}\lambda_3$ et de $\frac{1}{\sqrt{3}}\lambda_8$.

Voir l'exercice B pour le changement de coordonnées de (λ_1, λ_2) (indices de Dynkin d'une représentation, à ne pas confondre avec les matrices précédentes!) en (I_z, Y) .

Les matrices λ_a satisfont des relations de commutation

$$[\lambda_a, \lambda_b] = 2if_{abc}\lambda_c \tag{4.15}$$

avec les constantes de structure (réelles, complètement antisymétriques) f_{abc} de l'algèbre su(3). Il est utile de considérer aussi les anticommutateurs

$$\{\lambda_a, \lambda_b\} = \frac{4}{3}\delta_{ab}I + 2d_{abc}\lambda_c . \tag{4.16}$$

Grâce à (4.13), (4.15) et (4.16) peuvent se récrire comme $\text{tr}([\lambda_a, \lambda_b]\lambda_c) = 4if_{abc}$, $\text{tr}(\{\lambda_a, \lambda_b\}\lambda_c) = 4d_{abc}$. On trouve ces nombres f et d tabulés dans la littérature ... mais on les recalcule aisément! Attention, au contraire

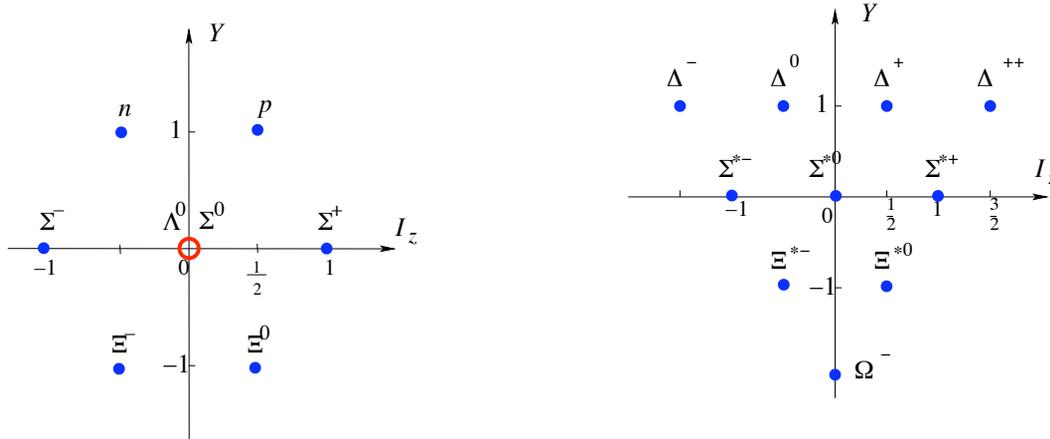


FIGURE 4.3 – L’octet ($J^P = \frac{1}{2}^+$) et le décuplet ($J^P = \frac{3}{2}^+$) de baryons.

de (4.15), la relation (4.16) et les constantes d_{abc} (réelles, complètement symétriques) sont propres à la représentation de dimension 3.

Les hadrons s’organisent en représentations de $SU(3)$. Chaque multiplet regroupe des particules de même spin J et parité P . C’est ainsi que deux octets de mésons de J^P égal à 0^- ou 1^- et qu’un octet et un “décuplet” de baryons de charge baryonique $\mathcal{B} = 1$ sont aisément identifiés. Contrairement à la symétrie d’isospin, la symétrie $SU(3)$ ¹ n’est pas une symétrie exacte des interactions fortes. Les règles de conservation ou de sélection auxquelles elle donne lieu ne sont qu’approchées.

À ce point, on peut s’interroger sur l’absence d’autres représentations de tripléte nulle, telle la représentation 27, ou de celles de tripléte non nulle, comme la 3 et la $\bar{3}$. On y reviendra au § 4.2.5.

4.2.2 Conséquences de la symétrie $SU(3)$

Les octets de champs

Concentrons nous sur les deux octets de baryons $\mathcal{N} = (N, \Sigma, \Xi, \Lambda)$ et de mésons pseudoscalaires $\mathcal{P} = (\pi, K, \eta)$. Au vu de ce que l’on a dit au Chap. 3, § 3.4.2, à savoir que la représentation adjointe est faite de tenseurs de rang (1, 1) et de trace nulle, il est naturel de regrouper les 8 champs associés à ces particules sous forme d’une matrice de trace nulle.

$$\Phi = \begin{pmatrix} \frac{1}{\sqrt{2}}\pi^0 - \frac{1}{\sqrt{6}}\eta & \pi^+ & K^+ \\ \pi^- & -\frac{1}{\sqrt{2}}\pi^0 - \frac{1}{\sqrt{6}}\eta & K^0 \\ K^- & \bar{K}^0 & \sqrt{\frac{2}{3}}\eta \end{pmatrix}, \quad (4.17)$$

1. dite “de saveur”, selon la terminologie moderne, mais appelée “symétrie unitaire” ou “voie octuple” à l’époque de Gell-Mann et Ne’eman...

et

$$\Psi = \begin{pmatrix} \frac{1}{\sqrt{2}}\Sigma^0 - \frac{1}{\sqrt{6}}\Lambda & \Sigma^+ & p \\ \Sigma^- & -\frac{1}{\sqrt{2}}\Sigma^0 - \frac{1}{\sqrt{6}}\Lambda & n \\ \Xi^- & \Xi^0 & \sqrt{\frac{2}{3}}\Lambda \end{pmatrix}. \quad (4.18)$$

Pour s'assurer que les assignements de champs/particules aux différents éléments de matrice sont corrects, il suffit de vérifier leurs nombres quantiques de charge et d'hypercharge. Les générateurs de charge Q et d'hypercharge Y

$$Q = I_z + \frac{1}{2}Y = \frac{1}{3} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad Y = \frac{1}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad (4.19)$$

agissent dans la représentation adjointe par commutation et on a bien

$$[Q, \Phi] = \begin{pmatrix} 0 & \pi_+ & K_+ \\ -\pi_- & 0 & 0 \\ -K_- & 0 & 0 \end{pmatrix} \quad [Y, \Phi] = \begin{pmatrix} 0 & 0 & K_+ \\ 0 & 0 & K_0 \\ -K_- & -\bar{K}_0 & 0 \end{pmatrix}.$$

Exercice : (i) sans aucun calcul, que doit valoir $[I_z, \Phi]$? Vérifier.

(ii) Calculer $\text{tr } \Phi^2$; en quoi le résultat justifie-t-il le choix de normalisation dans (4.17)? Voir aussi le Problème 2.c.

Produits tensoriels dans SU(3) et couplages invariants

On rappelle que pour SU(3), avec les notations du chapitre 3,

$$8 \otimes 8 = 1 \oplus 8 \oplus 8 \oplus 10 \oplus \bar{10} \oplus 27. \quad (4.20)$$

(Notons au passage que la multiplicité 2 de la représentation 8 reflète l'existence des deux tenseurs invariants indépendants f_{abc} et d_{abc} (cf. 4.15) et (4.16).) Montrons que cela a des implications immédiates sur le nombre de couplages invariants entre champs.

• On se propose d'écrire un lagrangien invariant par SU(3) impliquant les champs d'octets Φ et Ψ précédents. Quel est le nombre de "couplages de Yukawa", c'est-à-dire de la forme $\bar{\Psi}\Phi\Psi$, invariants par SU(3)? Autrement dit, quel est le nombre d'invariants dans $8 \otimes 8 \otimes 8$? Selon un raisonnement fait au chap 2, (cf § 2.3.2), ce nombre est égal au nombre de fois où la représentation 8 apparaît dans $8 \otimes 8$, soit selon (4.20), 2. Il y a donc deux couplages de Yukawa invariants indépendants. Si on écrit les deux octets de champs Ψ et Φ sous forme de matrices 3×3 de trace nulle comme au sous-paragraphe précédent, $\Psi = \{\psi_j^i\}$ et $\Phi = \{\phi_k^i\}$, ces deux couplages s'écrivent

$$\text{tr } \bar{\Psi}\Psi\Phi = \bar{\psi}_j^i \psi_i^k \phi_k^j \quad \text{and} \quad \text{tr } \bar{\Psi}\Phi\Psi = \bar{\psi}_j^i \phi_i^k \psi_k^j \quad (4.21)$$

(cette écriture omet les indices des spineurs de Dirac, l'éventuelle matrice γ_5 etc). On préfère souvent réécrire ces deux termes en termes de leurs somme et différence, donc de $\text{tr } \bar{\Psi}[\Phi, \Psi]$ et $\text{tr } \bar{\Psi}\{\Phi, \Psi\}$, appelés terme f et terme d, par référence à (4.15) et (4.16).

• Autre question de même nature : quel est *a priori* le nombre d'amplitudes invariantes par SU(3) dans la diffusion de deux particules des octets \mathcal{N} et \mathcal{P} : $\mathcal{N}_i + \mathcal{P}_i \rightarrow \mathcal{N}_f + \mathcal{P}_f$? (On ne prend en compte que l'invariance par SU(3), en ne considérant pas d'éventuelles symétries discrètes.) Il s'agit donc de chercher le nombre d'invariants dans la 4ème puissance tensorielle de la représentation 8. Ou encore de façon équivalente, le nombre de fois où l'on trouve la même représentation dans les deux produits $8 \otimes 8$ et $8 \otimes 8$. Si m_i sont les multiplicités apparaissant dans $8 \otimes 8$, soit $m_1 = 1, m_8 = 2$, etc, cf (4.20), ce nombre est $\sum_i m_i^2 = 8$. Il y a donc huit amplitudes invariantes. Autrement dit on peut écrire *a priori* l'amplitude de diffusion sous la forme

$$\langle \mathcal{N}_f \mathcal{P}_f | \mathcal{T} | \mathcal{N}_i \mathcal{P}_i \rangle = \sum_{r=1}^8 A_r(s, t) \langle (I, I_z, Y)_{(\mathcal{N}_f)}, (I, I_z, Y)_{(\mathcal{P}_f)} | r, (I, I_z, Y)_{(r)} \rangle \langle r, (I, I_z, Y)_{(r)} | (I, I_z, Y)_{(\mathcal{N}_i)}, (I, I_z, Y)_{(\mathcal{P}_i)} \rangle \quad (4.22)$$

(avec s et t les invariants relativistes usuels $s = (p_1 + p_2)^2$, $t = (p_1 - p_3)^2$), toute la dépendance dans la nature des particules membres des octets, repérées par les valeurs de leur isospin et hypercharge, étant contenue dans des coefficients de Clebsch-Gordan de SU(3).

• Soient Φ_i , $i = 1, 2, 3, 4$ quatre champs d'octets distincts. Combien de couplages de degré 4 invariants par SU(3) peut-on former avec ces quatre champs ? D'une part, le raisonnement précédent nous donne huit couplages ; de l'autre, il est clair que pour toute permutation P de $\{1, 2, 3, 4\}$, les termes $\text{tr}(\Phi_{P_1} \Phi_{P_2} \Phi_{P_3} \Phi_{P_4})$ et $\text{tr}(\Phi_{P_1} \Phi_{P_2}) \text{tr}(\Phi_{P_3} \Phi_{P_4})$ sont invariants par SU(3). Un décompte rapide donne 9 termes différents, en contradiction avec l'argument précédent. Quelle est l'origine de cette contradiction ? Pour en savoir plus, rendez-vous au Problème 1 en fin de chapitre...

4.2.3 Brisures électromagnétiques de la symétrie SU(3)

La symétrie SU(3) est brisée, on l'a dit, par les interactions fortes. Bien sûr, tout comme la symétrie SU(2) d'isospin, elle l'est aussi par les interactions électromagnétiques et les interactions faibles. Nous n'examinerons pas ici l'effet de ces dernières mais décrirons deux conséquences des brisures fortes et électromagnétiques.

Le lagrangien d'interaction d'une particule de charge q avec le champ électromagnétique A s'écrit

$$\mathcal{L}_{em} = -qj^\mu A_\mu \quad (4.23)$$

où j est le courant électrique. Le champ A est invariant par les transformations de SU(3), mais comment j se transforme-t-il ? On connaît la transformation de sa charge $Q = \int d^3x j_0(\mathbf{x}, t)$, puisque selon (4.12), Q est une combinaison linéaire de deux générateurs Y et I_z . Q se transforme donc selon la représentation adjointe (8, alias (1, 1) en termes d'indices de Dynkin). Et il est naturel de supposer que le courant j se transforme de la même manière. C'est d'ailleurs ce que l'on trouve quand on construit le courant j^μ comme courant de Noether de la symétrie U(1) (exercice, le vérifier).

Moments magnétiques

Les facteurs de forme électromagnétiques de l'octet de baryons sont définis par

$$\langle B|j_\mu(x)|B'\rangle = e^{ikx}\bar{u}(F_e^{BB'}(k^2)\gamma_\mu + F_m^{BB'}(k^2)\sigma_{\mu\nu}k^\nu)u' \quad (4.24)$$

où \bar{u} et u' sont des spineurs de Dirac décrivant respectivement les baryons B et B' ; k est la quadri-impulsion transférée de B' à B . F_e est le facteur de forme électrique, si $B = B'$, $F_e(0) = q_B$, charge électrique de B , tandis que F_m est le facteur de forme magnétique et $F_m^{BB}(0)$ donne le moment magnétique du baryon B . On veut calculer ces facteurs au premier ordre électromagnétique et à l'ordre zéro dans les autres termes brisant éventuellement la symétrie.

D'un point de vue groupiste, l'élément de matrice $\langle B|j_\mu(x)|B'\rangle$ relève du théorème de Wigner-Eckart : il y a deux façons de projeter 8×8 sur 8 (cf l'équ. (4.2) du chap 3), (ou encore, il y a deux façons de construire un invariant avec $8 \otimes 8 \otimes 8$). Il y a donc deux "éléments de matrice réduits", donc deux amplitudes indépendantes pour chacun des deux facteurs de forme, complétées par des coefficients de Clebsch-Gordan de SU(3). Par un argument similaire à (4.21), on vérifie que l'on peut écrire

$$F_{e,m}^{BB'}(k^2) = F_{e,m}^{(1)}(k^2) \text{tr } \bar{B}QB' + F_{e,m}^{(2)}(k^2) \text{tr } \bar{B}B'Q$$

où Q est la matrice de (4.19)

$$Q = \begin{pmatrix} \frac{2}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & -\frac{1}{3} \end{pmatrix},$$

et $\text{tr } \bar{B}QB'$ signifie le coefficient de $\bar{B}B'$ dans la trace matricielle $\text{tr } \bar{\Psi}Q\Psi$, et de même pour $\text{tr } \bar{B}B'Q$. Par exemple, le moment magnétique du neutron $\mu(n)$ est proportionnel au terme magnétique en $\bar{n}n$, soit $-\frac{1}{3}(F_m^{(1)} + F_m^{(2)})$. Les quatre fonctions $F_{e,m}^{(1,2)}$ sont inconnues (leur calcul ferait appel à la théorie des interactions fortes) mais on peut les éliminer et trouver des relations

$$\begin{aligned} \mu(n) &= \mu(\Xi^0) = 2\mu(\Lambda) = -2\mu(\Sigma^0) & \mu(\Sigma^+) &= \mu(p) \\ \mu(\Xi^-) &= \mu(\Sigma^-) = -(\mu(p) + \mu(n)) & \mu(\Sigma^0 \rightarrow \Lambda) &= \frac{\sqrt{3}}{2}\mu(n), \end{aligned} \quad (4.25)$$

où la dernière quantité est le *moment magnétique de transition* $\Sigma^0 \rightarrow \Lambda$. Ces relations sont en accord qualitatif avec les valeurs expérimentales.

Les moments magnétiques des "hypérons" (baryons de masse plus élevée que les nucléons) sont mesurés par leur précession de spin dans un champ magnétique ou dans des transitions dans des atomes "exotiques" dans le noyau desquels un nucléon a été remplacé par un hypéron. Le moment magnétique de transition $\Sigma^0 \rightarrow \Lambda$ est déterminé à partir de la section efficace $\Lambda \rightarrow \Sigma^0$ dans le champ coulombien d'un noyau lourd. On lit dans les tables

$$\begin{aligned} \mu(p) &= 2.792847351 \pm 0.000000028 \mu_N & \mu(n) &= -1.9130427 \pm 0.00000005 \mu_N \\ \mu(\Lambda) &= -0.613 \pm 0.004 \mu_N & |\mu(\Sigma^0 \rightarrow \Lambda)| &= 1.61 \pm 0.08 \mu_N \\ \mu(\Sigma^+) &= 2.458 \pm 0.010 \mu_N & \mu(\Sigma^-) &= -1.160 \pm 0.025 \mu_N \\ \mu(\Xi^0) &= -1.250 \pm 0.014 \mu_N & \mu(\Xi^-) &= -0.6507 \pm 0.0025 \mu_N \end{aligned} \quad (4.26)$$

où μ_N est le magnéton nucléaire, $\mu_N = \frac{e\hbar}{2m_p} = 3.152 \cdot 10^{-14} \text{ MeV T}^{-1}$.

Écarts de masses électromagnétiques

Des hypothèses et méthodes analogues permettent de trouver des relations entre les écarts de masses électromagnétiques, entre particules de même hypercharge et isospin I , mais de charge différente, voir Problème 3.

4.2.4 Écarts de masses “forts”. Formule de masse de Gell-Mann–Okubo

Au vu des disparités entre masses au sein d’un multiplet, le terme de masse dans le lagrangien (ou l’hamiltonien) ne peut être un invariant de $SU(3)$. Gell-Mann et Okubo ont fait l’hypothèse que le terme non invariant ΔM se transforme selon la représentation 8, plus précisément, puisqu’il doit avoir un isospin et une hypercharge nuls, qu’il se transforme comme la composante η ou Λ des octets. On est donc conduit à considérer les éléments de matrice $\langle H|\Delta M|H\rangle$ pour les hadrons H d’un multiplet, et à faire appel une nouvelle fois au théorème de Wigner–Eckart. Selon les règles de décomposition de produit tensoriel données au chap. 3, la représentation 8 apparaît au plus 2 fois dans le produit d’une représentation irréductible de $SU(3)$ par sa conjuguée, (le vérifier, en se rappelant que $8 = 3 \otimes \bar{3} \ominus 1$); il y a au plus deux amplitudes indépendantes qui décrivent les écarts de masse au sein du multiplet, ce qui conduit à des relations entre ces écarts de masse.

Un argument élégant permet d’éviter le calcul des coefficients de Clebsch–Gordan et de trouver ces deux amplitudes dans toute représentation. Les huit générateurs infinitésimaux se transformant eux-mêmes selon la représentation 8 (représentation adjointe), on les dispose selon une matrice 3×3 comme précédemment

$$G = \begin{pmatrix} \frac{1}{2}Y + I_z & \sqrt{2}I_+ & * \\ \sqrt{2}I_- & \frac{1}{2}Y - I_z & * \\ * & * & -Y \end{pmatrix}$$

où les * désignent des générateurs changeant l’étrangeté qui ne nous concernent pas. (Noter que $G_{11} = I_z + \frac{1}{2}Y =$

Q , la charge électrique, est invariante par l’action (par commutation avec G) des générateurs $X = \begin{pmatrix} 0 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{pmatrix}$

qui préservent la charge électrique.) On cherche deux combinaisons des générateurs I_z et Y se transformant comme l’élément (3,3) de cette matrice. L’une est bien sûr Y lui-même, l’autre est fournie par l’élément (3,3) du cofacteur de G , $\text{cof}G_{33} = \frac{1}{4}Y^2 - I_z^2 - 2I_+I_- = \frac{1}{4}Y^2 - \vec{I}^2$.

On obtient ainsi, pour toute représentation (tout multiplet), une formule de masse

$$M = m_1 + m_2Y + m_3(I(I+1) - \frac{1}{4}Y^2) \quad (4.27)$$

ce qui laisse trois constantes indéterminées (dépendant du multiplet). Par exemple pour l’octet de baryons, on a quatre particules N , Σ , Λ et Ξ pour lesquelles $(Y, I, I(I+1) - \frac{1}{4}Y^2) = (1, \frac{1}{2}, \frac{1}{2}), (0, 1, 2), (0, 0, 0), (-1, \frac{1}{2}, \frac{1}{2})$ respectivement, et satisfaisant donc

$$M_N = m_1 + m_2 + \frac{1}{2}m_3 \quad M_\Sigma = m_1 + 2m_3 \quad (4.28)$$

$$M_\Lambda = m_1 \quad M_\Xi = m_1 - m_2 + \frac{1}{2}m_3. \quad (4.29)$$

En éliminant les trois paramètres m_1, m_2, m_3 entre ces quatre relations, on est amené à la règle de somme

$$\frac{M_{\Xi} + M_N}{2} = \frac{3M_{\Lambda} + M_{\Sigma}}{4} \quad (4.30)$$

bien vérifiée expérimentalement : on trouve $1128,5 \text{ MeV}/c^2$ au membre de gauche, $1136 \text{ MeV}/c^2$ à celui de droite². Pour le décuplet, vérifier que cette même formule donne des écarts de masse égaux entre les quatre particules Δ , Σ^* , Ξ^* et Ω^- . Cela a permis de prédire avec justesse l'existence et la masse de cette dernière particule, ce qui a été considéré comme un des grands succès de SU(3). Pour l'octet de mésons pseudoscalaires, la formule de masse implique (empiriquement) les carrés de masses

$$m_K^2 = \frac{3m_\eta^2 + m_\pi^2}{4}.$$

4.2.5 Les quarks

Les représentations 3 et $\bar{3}$ sont à ce stade les grandes absentes de la scène : parmi les particules observées, aucun “triplet” ne semble se manifester. Le modèle de Gell-Mann–Zweig fait l'hypothèse qu'un triplet (représentation 3) de quarks (u, d, s) (“up”, “down” et “strange”) et sa représentation conjuguée $\bar{3}$ d'antiquarks ($\bar{u}, \bar{d}, \bar{s}$) rassemblent les constituents élémentaires de tous les hadrons (connus à l'époque). Leurs charges et hypercharges sont respectivement

Quarks	u	d	s	\bar{u}	\bar{d}	\bar{s}
Isospin I_z	$\frac{1}{2}$	$-\frac{1}{2}$	0	$-\frac{1}{2}$	$\frac{1}{2}$	0
Baryonic charge \mathcal{B}	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$
Strangeness S	0	0	-1	0	0	1
Hypercharge Y	$\frac{1}{3}$	$\frac{1}{3}$	$-\frac{2}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$
Electric charge Q	$\frac{2}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

Table 1. Nombres quantiques des quarks u, d, s

On se rappelle (chapitre 3 § 3.4) que toute représentation irréductible de SU(3) s'obtient dans la décomposition de produits itérés des représentations 3 et $\bar{3}$; en particulier, $3 \otimes \bar{3} = 1 \oplus 8$ et $3 \otimes 3 \otimes 3 = 1 \oplus 8 \oplus 8 \oplus 10$. Les mésons et baryons observés dans la nature et répertoriés comme ci-dessus selon des représentations 8 et 10 de SU(3) sont des états liés de paires $q\bar{q}$ ou qqq , respectivement. Plus généralement, on suppose que seules les représentations de triality nulle peuvent donner lieu à des particules observables. Ainsi,

$$\begin{aligned} p = uud, \quad n = udd, \quad \Omega^- = sss, \quad \Delta^{++} = uuu, \quad \dots, \quad \Delta^- = ddd, \\ \pi^+ = u\bar{d}, \quad \pi^0 = \frac{(u\bar{u} - d\bar{d})}{\sqrt{2}}, \quad \pi^- = d\bar{u}, \quad \eta_8 = \frac{(u\bar{u} + d\bar{d} - 2s\bar{s})}{\sqrt{6}}, \quad K^+ = u\bar{s}, \quad K^0 = d\bar{s} \quad \text{etc.} \end{aligned} \quad (4.31)$$

2. Les masses observées de ces hadrons sont $M_N \approx 939 \text{ MeV}/c^2$, $M_\Lambda = 1116 \text{ MeV}/c^2$, $M_\Sigma \approx 1195 \text{ MeV}/c^2$, $M_\Xi \approx 1318 \text{ MeV}/c^2$; celles des mésons pseudoscalaires $m_\pi \approx 137 \text{ MeV}/c^2$, $m_K \approx 496 \text{ MeV}/c^2$ et $m_\eta = 548 \text{ MeV}/c^2$. Pour le décuplet, $M_\Delta \approx 1232 \text{ MeV}/c^2$, $M_{\Sigma^*} \approx 1385 \text{ MeV}/c^2$, $M_{\Xi^*} \approx 1530 \text{ MeV}/c^2$, $M_\Omega \approx 1672 \text{ MeV}/c^2$.

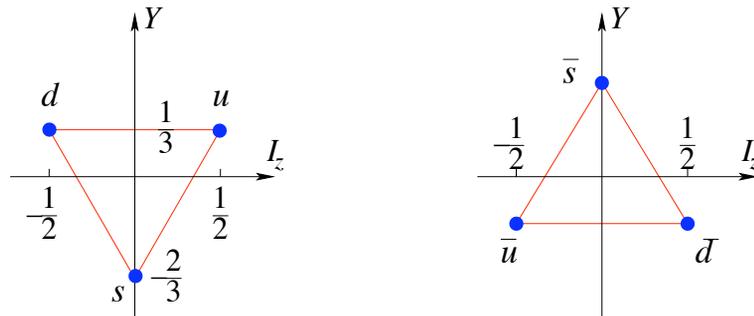


FIGURE 4.4 – Les triplets de quarks et antiquarks.

Le modèle des quarks interprète le singulet qui apparaît dans le produit $3 \times \bar{3}$ comme un état lié $\eta_1 = \frac{(u\bar{u}+d\bar{d}+s\bar{s})}{\sqrt{3}}$. Les particules physiquement observées η (masse 548 MeV) et η' (958 MeV) résultent d'un “mélange” (c'est-à-dire une combinaison linéaire) de ces η_1 et η_8 dû aux interactions brisant SU(3). Exercice : compléter sur la Fig. 4.3 les interprétations des baryons comme états liés des quarks en s'aidant des charges et nombres quantiques.

4.2.6 Courants hadroniques et interactions faibles

Les interactions faibles sont phénoménologiquement bien décrites par un lagrangien effectif de la forme “courant-courant” (Fermi)

$$\mathcal{L}_{\text{Fermi}} = -\frac{G}{\sqrt{2}} J^\rho(x) J_\rho^\dagger(x) \quad (4.32)$$

où G est la constante de Fermi qui vaut (dans les unités où $\hbar = c = 1$)

$$G = (1,026 \pm 0,001) \times 10^{-5} M_p^{-2} . \quad (4.33)$$

(Ce lagrangien d'interaction a le défaut majeur de ne pas être renormalisable, un défaut que vient corriger la théorie de jauge du Modèle Standard. À basse énergie, $\mathcal{L}_{\text{Fermi}}$ fournit toutefois une bonne description de la physique, d'où le qualificatif d'“effectif”.) Le courant J_ρ est la somme d'une contribution leptonique et d'une hadronique

$$J_\rho(x) = l_\rho(x) + h_\rho(x) \quad (4.34)$$

Le courant leptonique

$$l_\rho(x) = \bar{\psi}_e(x) \gamma_\rho (1 - \gamma_5) \psi_{\nu_e} + \bar{\psi}_\mu(x) \gamma_\rho (1 - \gamma_5) \psi_{\nu_\mu} \quad [+ \bar{\psi}_\tau(x) \gamma_\rho (1 - \gamma_5) \psi_{\nu_\tau}]$$

est la somme des contributions des familles de leptons, e , μ (et τ que nous omettrons en première analyse). Le courant hadronique si on se borne aux deux premières familles s'écrit

$$h_\rho = \cos \theta_C h_\rho^{(\Delta S=0)} + \sin \theta_C h_\rho^{(\Delta S=1)} \quad (4.35)$$

comme combinaison de courants conservant ou changeant l'étrangeté, pondérée par l'angle de Cabibbo $\theta_C \approx 0,25$. (Ce “mélange” s'étend à l'introduction de la troisième famille, cf chapitre

suisant.) Enfin chacun des courants $h_\rho^{(\Delta S=0)}, h_\rho^{(\Delta S=1)}$ est de la forme “ $V - A$ ”, selon l’idée de Feynman et Gell-Mann, c’est-à-dire est une combinaison de courants vectoriel et axial,

$$h_\rho^{(\Delta S=0)} = (V_\rho^1 - iV_\rho^2) - (A_\rho^1 - iA_\rho^2) \quad (4.36)$$

$$h_\rho^{(\Delta S=1)} = (V_\rho^4 - iV_\rho^5) - (A_\rho^4 - iA_\rho^5) . \quad (4.37)$$

Les courants vectoriels $V_\rho^{1,2,3}$ sont les courants de Noether d’isospin, les autres composantes de V_ρ ceux de la symétrie SU(3). On montre que leur conservation (exacte pour l’isospin, approchée pour les autres) implique que dans l’élément de matrice $G\langle p|h_\rho^{(\Delta S=0)}|n\rangle = \bar{u}_p\gamma_\rho(G_V(q^2) - G_A(q^2)\gamma_5)u_n$ mesuré dans la désintégration beta à transfert d’impulsion quasi-nul, le facteur de forme vectoriel $G_V(0) = G$. Au contraire, les courants axiaux ne sont pas conservés et $G_A(0)$ est “renormalisé” (c’est-à-dire habillé) par les interactions fortes, $G_A/G_V \approx 1.22$. Le courant électromagnétique n’est autre que la combinaison $j_\rho = V_\rho^3 + \frac{1}{\sqrt{3}}V_\rho^8$. Dans le modèle des quarks, ces courants hadroniques sont de la forme

$$V_\rho^a(x) = \bar{q}(x)\frac{\lambda^a}{2}\gamma_\rho q(x) \quad A_\rho^a(x) = \bar{q}(x)\frac{\lambda^a}{2}\gamma_\rho\gamma_5 q(x) . \quad (4.38)$$

Nous les retrouverons dans le Modèle Standard.

4.3 De SU(3) à SU(4) et aux six saveurs

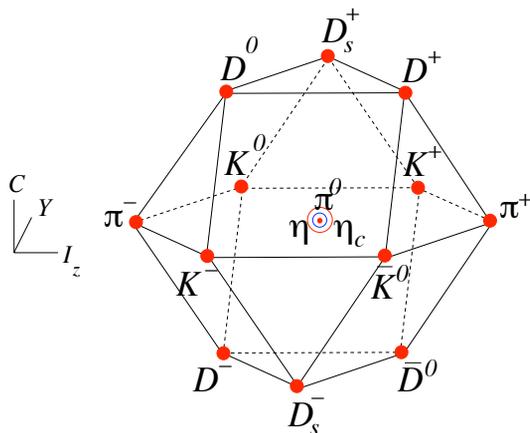
4.3.1 Nouvelles saveurs

La découverte au milieu des années 70 de particules d’un nouveau type a relancé le jeu : ces particules portent un autre nombre quantique, le “charme” (postulé antérieurement par Glashow, Iliopoulos et Maiani et par Kobayashi et Maskawa pour des raisons différentes). Cela amène à ajouter une 3ème direction à l’espace des symétries internes, en sus de l’isospin et de l’étrangeté (ou de l’hypercharge). C’est un groupe SU(4), encore plus sévèrement brisé que SU(3), qui est à l’œuvre. Les particules s’organisent en représentations de ce SU(4), etc. Une quatrième saveur, le charme, est donc ajoutée, et un quatrième quark c charmé constitue avec u, d, s la représentation 4 de SU(4), tout aussi inobservable que la 3 de SU(3), selon le même principe.

Au jour d’aujourd’hui, on pense qu’il existe en tout six saveurs, les deux dernières étant la “beauté” (*beauty* ou *bottomness*) et la vérité (...?) (*truth* ou *topness*), donc deux quarks supplémentaires b et t . Des mésons B , états liés $u\bar{b}$, $d\bar{b}$ etc, sont observés quotidiennement par exemple dans l’expérience LHC_{*b*} au LHC, tandis que les preuves expérimentales de l’existence du quark t sont plus indirectes. Le groupe hypothétique de saveur SU(6) est très fortement brisé, comme l’attestent les masses des 6 quarks ³

$$\begin{aligned} m_u &\approx 1.5 - 4 \text{ MeV}, & m_d &\approx 4 - 8 \text{ MeV}, & m_s &\approx 80 - 130 \text{ MeV} \\ m_c &\approx 1.15 - 1.35 \text{ GeV}, & m_b &\approx 4 - 5 \text{ GeV}, & m_t &\approx 175 \text{ GeV} \end{aligned} \quad (4.39)$$

3. Il faudrait bien sûr préciser le sens de la masse d’une particule qui reste invisible, ce qu’on sait faire indirectement et avec plusieurs définitions possibles, d’où les plages de valeurs données.

FIGURE 4.5 – Les mésons de spin $J^P = 0^-$ de la représentation 15 de $SU(4)$.

ce qui réduit son utilité. On peut toutefois récrire (4.12) sous la forme

$$Q = \frac{1}{2}Y + I_z \quad Y = \mathcal{B} + S + C + B + T$$

avec les différents nombres quantiques contribuant additivement à l'hypercharge. La convention est que la saveur d'un quark est nulle ou du même signe que sa charge électrique Q , voir Table 1. Ainsi $C(c) = 1$, $B(b) = -1$ etc. La Table 1 est donc à compléter comme suit

Quarks	u	d	s	c	b	t
Isospin I_z	$\frac{1}{2}$	$-\frac{1}{2}$	0	0	0	0
Charge baryonique \mathcal{B}	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
Étrangeté S	0	0	-1	0	0	0
Charme C	0	0	0	1	0	0
Beauté B	0	0	0	0	-1	0
Vérité T	0	0	0	0	0	1
Hypercharge Y	$\frac{1}{3}$	$\frac{1}{3}$	$-\frac{2}{3}$	$\frac{4}{3}$	$-\frac{2}{3}$	$\frac{4}{3}$
Charge électrique Q	$\frac{2}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$

Table 2. Nombres quantiques des quarks u , d , s , c , b , t

4.3.2 Introduction de la couleur

Des problèmes variés avec le modèle des quarks originel ont conduit à l'hypothèse (Han-Nambu) que chaque saveur est dotée d'une multiplicité 3, qui reflète l'existence d'un groupe $SU(3)$ **différent** du précédent, le *groupe de couleur* $SU(3)_c$.

Les considérations menant à cette hypothèse de triplement sont d'une part l'étude de la particule Δ^{++} , de spin $3/2$, composée de 3 quarks u . Ce système de 3 quarks a un spin $3/2$ et un moment angulaire orbital $L = 0$, qui lui donnent une fonction d'onde symétrique, en contradiction avec le caractère fermionique des quarks. Le degré de liberté supplémentaire de couleur permet une antisymétrisation supplémentaire (conduisant à un état singulet de couleur), et lève donc ce problème. D'autre part, le processus de désintégration $\pi^0 \rightarrow 2\gamma$ est proportionnel à la somme $\sum Q^2 I_z$ pour l'ensemble des constituents fermioniques élémentaires. Le proton, avec

sa charge $Q = 1$ et $I_z = \frac{1}{2}$, conduit à une valeur en accord avec l'expérience. Les quarks (u, d, s) avec les valeurs $Q = (\frac{2}{3}, \frac{1}{3}, -\frac{1}{3})$ et $I_z = (\frac{1}{2}, -\frac{1}{2}, 0)$ conduisant à un résultat trois fois trop petit, que la multiplicité de couleur vient corriger.

Selon l'hypothèse de confinement des quarks, seuls les états de la représentation 1 de $SU(3)_c$ sont observables. Les autres états, dits "colorés", sont liés de façon permanente au sein des hadrons. Cela s'applique aux quarks, mais aussi aux *gluons*, des particules vectorielles (spin 1) se transformant selon la représentation 8 de $SU(3)_c$, dont l'existence est requise par la construction de la théorie de jauge des interactions fortes, la chromodynamique quantique, voir Chap. 5.

Pour être plus précis, l'hypothèse de confinement s'applique à température nulle ou faible, la libération des quarks et gluons pouvant se produire dans la matière hadronique sous haute température ou pression (au sein du "plasma de quarks et gluons")

Le modèle des quarks avec son groupe de couleur $SU(3)_c$ est maintenant considéré comme partie intégrante de la chromodynamique quantique. Les six saveurs de quarks sont regroupées en trois "générations", (u, d) , (c, s) , (t, b) , qui sont en correspondance avec trois générations de leptons, (e^-, ν_e) , (μ^-, ν_μ) , (τ^-, ν_τ) . Cette correspondance est importante pour la cohérence du modèle standard (compensation des anomalies), voir chap. suivant.

★

Références additionnelles pour le Chapitre 4

Sur $SU(3)$ de saveur, la référence incontournable, contenant tous les articles historiques, est M. Gell-Mann et Y. Ne'eman, *The Eightfold Way*, Benjamin 1964. On y trouvera en particulier les tables de coefficients de Clebsch-Gordan pour $SU(3)$ par J.J. de Swart.

Dans la discussion des brisures de $SU(3)$, on a suivi S. Coleman, *Aspects of Symmetry*, Cambridge Univ. Press 1985.

Pour un exposé plus récent sur la physique des saveurs, voir K. Huang, *Quarks, Leptons and Gauge Fields*, World Scientific 1992.

Toutes les propriétés des particules citées se trouvent dans les tables du Particle Data Group, en ligne sur le site http://pdg.lbl.gov/2013/listings/contents_listings.html

★

Exercices et Problèmes du chapitre 4

A. Modèle sigma et brisure de la symétrie chirale

On considère le lagrangien (4.10) et on définit $W = \sigma + i\boldsymbol{\pi}\boldsymbol{\tau}$.

1. Que vaut $\det W$? Montrer que l'on peut récrire \mathcal{L} en termes de $\psi_{L,R}$ et W selon

$$\mathcal{L} = \bar{\psi}_R i \not{\partial} \psi_R + \bar{\psi}_L i \not{\partial} \psi_L + g(\bar{\psi}_L W \psi_R + \bar{\psi}_R W^\dagger \psi_L) + \mathcal{L}_K - \frac{1}{2} m^2 \det W - \frac{\lambda}{4} (\det W)^2$$

où \mathcal{L}_K est le terme cinétique des champs $(\sigma, \boldsymbol{\pi})$. On peut donner à ce terme la forme $\mathcal{L}_K = \frac{1}{2}(\det \partial_0 W - \sum_{i=1}^3 \det \partial_i W)$ (d'allure un peu étrange, mais bel et bien invariant de Lorentz!).

2. Montrer que \mathcal{L} est invariant par les transformations de $SU(2) \times SU(2)$ avec $\psi_L \rightarrow U\psi_L$, $\psi_R \rightarrow V\psi_R$, à condition que W se transforme d'une façon qu'on précisera. Justifier l'assertion faite au § 4.1.2 : ψ_L , ψ_R et W se transforment respectivement par les représentations $(\frac{1}{2}, 0)$, $(0, \frac{1}{2})$ et $(\frac{1}{2}, \frac{1}{2})$.

3. Si le champ W acquiert une vev v , par exemple selon la direction de σ , $\langle \sigma \rangle = v$, montrer que le champ ψ acquiert une masse $M = -gv$

B. Changement de base dans $SU(3)$

Dans $SU(3)$, identifier le changement de base qui fait passer des poids Λ_1, Λ_2 du chapitre 3 aux axes utilisés dans les Fig. 4.2, 4.3 et 4.4. En déduire la transformation des coordonnées (λ_1, λ_2) (indices de Dynkin) aux coordonnées physiques (I_z, Y) . Quelle est la dimension de la représentation de $SU(3)$ exprimée en termes de l'isospin et de l'hypercharge de son plus haut poids?

C. Formule de Gell-Mann-Okubo

Compléter et justifier tous les arguments esquissés aux § 4.2.2, 4.2.3 et 4.2.4. En particulier vérifier que la formule (4.27) conduit bien pour l'octet de baryon à la règle (4.30), et pour le décuplet, à des écarts de masse constants.

D. Comptage d'amplitudes

Combien d'amplitudes indépendantes sont nécessaires pour décrire la diffusion $\mathcal{BD} \rightarrow \mathcal{BD}$, où \mathcal{B} et \mathcal{D} décrivent l'octet et le décuplet de baryons?

Problèmes

1. Couplages à quatre champs invariants par $SU(3)$

On considère une matrice A , hermitienne, 3×3 et de trace nulle.

a. Montrer que l'équation caractéristique

$$A^3 - (\text{tr } A)A^2 + \frac{1}{2}((\text{tr } A)^2 - \text{tr } A^2)A - \det A = 0$$

implique une relation entre $\text{tr } A^4$ et $(\text{tr } A^2)^2$.

b. Si le groupe $SU(3)$ agit sur A par $A \rightarrow UAU^\dagger$, montrer que toute somme de produits de traces de puissances de A est invariante. On appelle une telle somme "polynôme invariant en A ". Combien y-a-t-il de tels polynômes invariants de degré 4 en A et linéairement indépendants?

c. On "polarise" alors l'identité trouvée en a., c'est-à-dire qu'on écrit $A = \sum_{i=1}^4 x_i A_i$ avec 4 matrices A_i du type précédent et 4 coefficients x_i arbitraires, et que l'on identifie le coefficient de $x_1 x_2 x_3 x_4$. Montrer que l'on obtient une identité de la forme (identité de Burgoyne)

$$\sum_P \text{tr}(A_{P_1} A_{P_2} A_{P_3} A_{P_4}) = a \sum_P \text{tr}(A_{P_1} A_{P_2}) \text{tr}(A_{P_3} A_{P_4}) \quad (4.40)$$

avec des sommes sur les permutations P de 4 éléments et un coefficient a qu'on déterminera. Combien de termes distincts apparaissent dans chacun des membres de cette identité?

d. Combien de polynômes de degré 4 quadrilatéraux en A_1, \dots, A_4 , invariants par l'action de $SU(3)$ $A_i \rightarrow UA_i U^\dagger$ et linéairement indépendants peut-on écrire? Pourquoi l'identité (4.40) est-elle utile?

2. Invariance cachée du lagrangien bosonique

On cherche à écrire un lagrangien pour le champ Φ de l'octet des mésons *pseudoscalaires*, cf (4.17).

a. Pourquoi est-il naturel d'imposer que ce lagrangien soit pair dans le champ Φ ?

b. En utilisant les résultats du Problème 1., écrire la forme la plus générale du lagrangien de degré inférieur ou égal à 4 et pair en Φ , invariant par $SU(3)$.

c. On écrit alors chaque champ complexe en distinguant sa partie réelle et sa partie imaginaire, par exemple $K^+ = \frac{1}{\sqrt{2}}(K_1 - iK_2)$, $K^- = \frac{1}{\sqrt{2}}(K_1 + iK_2)$, et de même avec K^0, \bar{K}^0 et avec π^\pm . Calculer $\text{tr } \Phi^2$ avec cette paramétrisation et montrer qu'on obtient une forme quadratique simple dans les 8 composantes réelles. Quel est le groupe d'invariance G de cette forme quadratique? Ce groupe est-il un sous-groupe de $\text{SU}(3)$?

d. En déduire que tout lagrangien de degré 4 en Φ invariant par $\text{SU}(3)$ est en fait invariant par ce groupe G .

3. Écarts de masses électromagnétiques dans un octet de $\text{SU}(3)$

Question préliminaire.

Étant donné un espace vectoriel E de dimension d , on note $E \otimes E$ l'espace des tenseurs de rang 2 et $(E \otimes E)_S$, resp. $(E \otimes E)_A$, l'espace des tenseurs de rang 2 symétriques, resp. antisymétriques, appelé encore produit tensoriel (anti)symétrisé. Quelle est la dimension des espaces $E \otimes E$, $(E \otimes E)_S$, $(E \otimes E)_A$?

On fait l'hypothèse que $\text{SU}(3)$ est une symétrie *exacte* des interactions fortes, et on se propose d'étudier les différences de masses dues aux effets électromagnétiques.

a. Combien y a-t-il de différences de masses indépendantes entre baryons de mêmes nombres quantiques I et Y mais de charges Q (ou de composantes I_z) différentes, dans l'octet de baryons $J^P = \frac{1}{2}^{+}$?

On admettra que ces effets électromagnétiques résultent de perturbations *du second ordre* dans le lagrangien $\mathcal{L}_{em}(x) = -qj^\mu(x)A_\mu(x)$. Si $|B\rangle$ est un état de baryon, il faudrait donc calculer

$$\delta M_B = \langle B | (\int d^4x \mathcal{L}_{em})^2 | B \rangle .$$

Faute de savoir calculer cet élément de matrice, on veut calculer le nombre d'amplitudes indépendantes y contribuant.

b. Expliquer pourquoi ce calcul amène à compter les invariants apparaissant dans le produit tensoriel de *quatre* représentations 8. Au vu des calculs effectués en cours, que devrait être ce nombre?

c. Mais attention ! le produit des deux lagrangiens est symétrique. En ce qui concerne le produit $\int \mathcal{L}_{em} \int \mathcal{L}_{em}$, il faut donc décomposer en représentations irréductibles le produit tensoriel symétrisé $(8 \otimes 8)_S$. Utiliser le résultat de la Question préliminaire pour calculer le nombre de tenseurs indépendants symétriques de rang 2 dans la représentation 8. Montrer que ce nombre est compatible avec la décomposition qu'on admettra

$$(8 \otimes 8)_S = 1 \oplus 8 \oplus 27 .$$

d. i) Quel est alors le nombre d'amplitudes invariantes contribuant à δM_B ?

d. ii) Quel est le nombre d'amplitudes invariantes contribuant à $\delta M_B - \delta M_{B'}$ pour deux hadrons B et B' de mêmes nombres quantiques, comme discuté au a.?

d. iii) Dans l'esprit de ce qui a été fait en cours sur les amplitudes contribuant aux moments magnétiques, pouvez-vous écrire une base d'invariants en termes des matrices Ψ , $\bar{\Psi}$ et Q ?

e. i) Montrer *a priori* que le nombre d'amplitudes calculé à la question d. ii) implique une relation entre les écarts de masse électromagnétiques.

e. ii) Calculer alors $\Delta_{em} M = \alpha \text{tr } \bar{B} Q^2 B + \beta \text{tr } \bar{B} B Q^2 + \gamma \text{tr } \bar{B} Q B Q$, (l'usage de Maple ou de Mathematica peut aider...), identifier dans cette expression les coefficients $\Delta_{em} M_p$ de $\bar{p}p$, $\Delta_{em} M_n$ de $\bar{n}n$, etc, et vérifier la relation

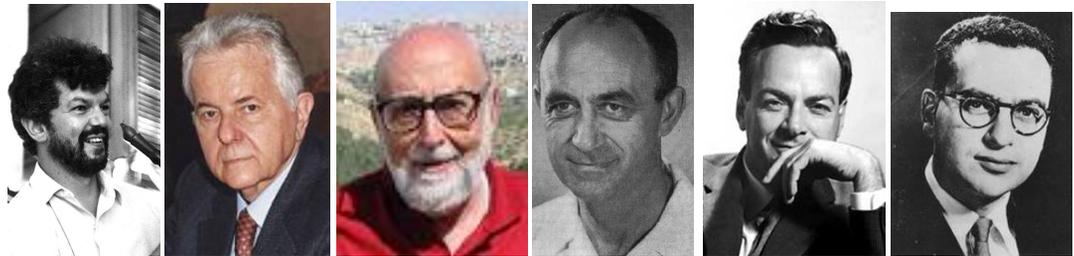
$$M_{\Xi^-} - M_{\Xi^0} = M_{\Sigma^-} - M_{\Sigma^+} + M_p - M_n . \tag{R}$$

Les valeurs expérimentales sont $M_n = 939,56 \text{ MeV}/c^2$, $M_p = 938,27 \text{ MeV}/c^2$, $M_{\Xi^-} = 1321,71 \text{ MeV}/c^2$, $M_{\Xi^0} = 1314,86 \text{ MeV}/c^2$, $M_{\Sigma^-} = 1197,45 \text{ MeV}/c^2$, $M_{\Sigma^0} = 1192,64 \text{ MeV}/c^2$, $M_{\Sigma^+} = 1189,37 \text{ MeV}/c^2$. Calculer les valeurs des deux membres de la relation (R). Commenter.

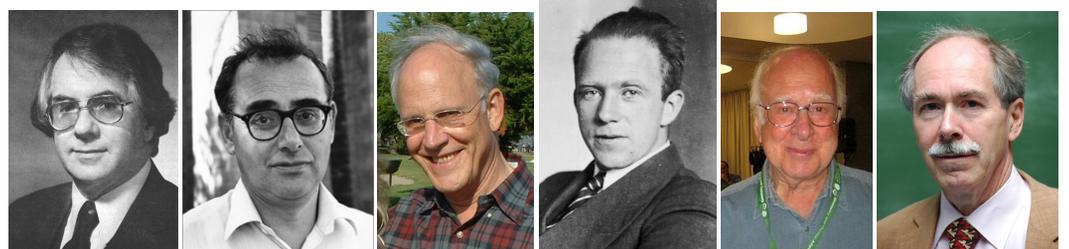
f. Octet des mésons pseudoscalaires. Pourrait-on raisonner de façon analogue avec les mésons pseudoscalaires?

g. Quid des écarts électromagnétiques au sein du décuplet $(\frac{3}{2})^+$?

*



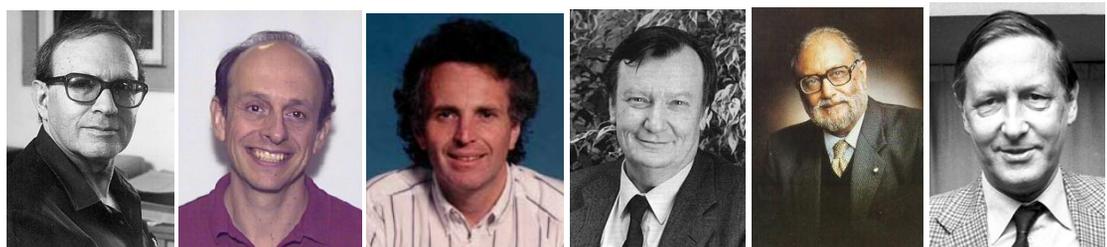
Robert Brout 1928-2011 Nicola Cabbibo 1935-2010 François Englert 1932- Enrico Fermi 1901-1954 Richard Feynman 1918-1988 Murray Gell-Mann 1929-



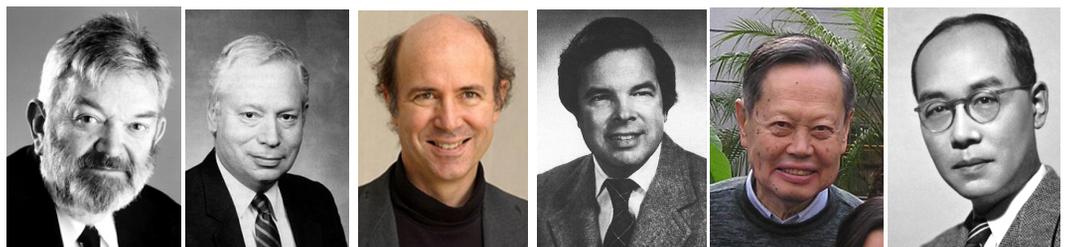
Sheldon Glashow 1932- Jeffrey Goldstone 1933- David Gross 1941- Werner Heisenberg 1901-1976 Peter Higgs 1929- Gerard 't Hooft 1946-



Jean Iliopoulos 1940- Maurice Lévy 1922- Makoto Kobayashi 1944- Luciano Maiani 1941- Toshihide Maskawa 1940- Yoichiro Nambu 1921-



Yuval Ne'eman 1925-2006 H. David Politzer 1949- Alexander Polyakov 1945- Carlo Rubbia 1934- Abdus Salam 1926-1996 Simon van der Meer 1925-2011



Martin Veltman 1931- Steve Weinberg 1933- Frank Wilczek 1951- Kenneth Wilson 1936-2013 Chen Ning Yang 1922- Hideki Yukawa 1907-1981

Quelques-uns des physiciens mentionnés dans la deuxième partie de ces notes

Chapitre 5

Théories de jauge. Modèle standard

En dehors d'un bref commentaire sur l'invariance de jauge de l'électrodynamique classique, les transformations qu'on a rencontrées jusqu'à maintenant dans ce cours étaient des transformations globales, indépendantes du point d'espace-temps où elles s'appliquent. Un autre type de symétrie, beaucoup plus contraignant sur la dynamique du système, consiste à supposer que la transformation est *locale*. En chaque point d'espace-temps, agit une copie distincte du groupe de transformations. Une telle symétrie, appelée *symétrie de jauge*, est familière en électrodynamique. Son extension à des groupes de transformations non abéliens par Yang et Mills s'est avérée être une des idées théoriques les plus fécondes de la seconde moitié du XXème siècle. Un cours entier devrait lui être consacré. Plus modestement, le présent chapitre en donnera une introduction élémentaire.

5.1 Invariance de jauge. Couplage minimal. Lagrangien de Yang–Mills

5.1.1 Invariance de jauge de l'électrodynamique

L'étude de l'électrodynamique a familiarisé avec la notion d'invariance locale. Le lagrangien de l'électrodynamique

$$\mathcal{L} = \bar{\psi}(i\cancel{\partial} - e\cancel{A} - m)\psi - \frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) \quad (5.1)$$

est invariant sous l'effet de transformations de jauge infinitésimale

$$\begin{aligned} \delta A_\mu(x) &= -\partial_\mu \delta\alpha(x) \\ \delta\psi(x) &= ie\delta\alpha(x)\psi(x), \end{aligned} \quad (5.2)$$

puisque le tenseur de champ

$$F_{\mu\nu} = (\partial_\mu A_\nu - \partial_\nu A_\mu)$$

est invariant, et que la combinaison

$$i\mathcal{D}\psi(x) := (i\partial - eA)\psi(x)$$

se transforme comme ψ . La forme finie de ces transformations est aussi aisée à écrire

$$\begin{aligned} A_\mu(x) &\mapsto A_\mu(x) - \partial_\mu\alpha(x) \\ \psi(x) &\mapsto e^{ie\alpha(x)}\psi(x), \end{aligned} \quad (5.3)$$

ce qui montre bien que les transformations sont celles d'une version locale (dépendant de x) du groupe $U(1)$ ou \mathbb{R} (voir ci-dessous). Les transformations globales correspondantes sont celles qui conduisent à un courant de Noether conservé, lié à la conservation de la charge électrique. Le lagrangien met aussi en évidence le "couplage minimal" du champ ψ au champ électromagnétique¹. Tout autre champ chargé de charge q se couple au champ électromagnétique par un terme impliquant la "dérivée covariante" $i\partial_\mu - qA_\mu(x)$.

C'est par exemple le cas d'un champ ϕ de boson chargé, donc complexe, dont la contribution au lagrangien s'écrit

$$\delta\mathcal{L} = [(\partial_\mu - iqA_\mu)\phi^*][(\partial^\mu + iqA_\mu)\phi] - V(\phi^*\phi) \quad (5.4)$$

qui est bien invariant sous $\phi(x) \mapsto e^{iq\alpha(x)}\phi(x)$, $A_\mu(x) \mapsto A_\mu(x) - \partial_\mu\alpha(x)$.

Noter que si le champ A est couplé à plusieurs champs de charges q_1, q_2, \dots , demander que le groupe de jauge est $U(1)$ (plutôt que \mathbb{R}), c'est-à-dire identifier $\alpha(x)$ et $\alpha(x) + 2\pi x$ (x un réel donné), impose que $xq_1, xq_2, \dots \in \mathbb{Z}$ et donc que les charges q_1, q_2, \dots sont *commensurables*. Cela peut donc expliquer la quantification de la charge électrique observée dans la nature.

5.1.2 Extension non abélienne de Yang–Mills

Selon l'observation brillante de Yang et Mills (1954), cette construction se transpose au cas d'un groupe de Lie non abélien G , avec toutefois quelques intéressantes petites modifications... Soit ψ un champ (que nous notons comme un champ fermionique, mais la chose est sans importance) se transformant par G selon une certaine représentation \mathcal{D} . Soient T_a les représentants des générateurs infinitésimaux dans cette représentation, nous les supposons antihermitiens : $[T_a, T_b] = C_{ab}{}^c T_c$; la transformation infinitésimale de ψ s'écrit donc

$$\delta\psi(x) = T_a\delta\alpha^a\psi(x). \quad (5.5)$$

(Dans ce paragraphe, on notera t_a les matrices correspondantes dans la représentation adjointe.) Pour étendre l'idée de transformation locale, nous avons besoin d'un champ de jauge A_μ , permettant de construire une dérivée covariante $D_\mu\psi$. Il est naturel de considérer que A_μ vit dans l'algèbre de Lie de G (puisqu'il est associé à des transformations infinitésimales du groupe), et qu'il porte donc un indice de la représentation adjointe

$$A_\mu(x) = \{A_\mu^a(x)\} \quad (5.6)$$

1. Un terme supplémentaire dans le lagrangien de la forme $\bar{\psi}[\gamma_\mu, \gamma_\nu]\psi F^{\mu\nu}$ serait invariant de jauge invariant mais non minimal.

ou encore A_μ est représenté dans toute représentation par la matrice antihermitique²

$$A_\mu(x) = T_a A_\mu^a(x) . \quad (5.7)$$

La dérivée covariante s'écrit

$$D_\mu \psi(x) := (\partial_\mu - A_\mu(x))\psi(x) , \quad (5.8)$$

ou encore, en composantes

$$D_\mu \psi_A(x) := (\partial_\mu \delta_{AB} - A_\mu^a(x) (T_a)_A^B) \psi_B(x) . \quad (5.9)$$

Cette dérivée se transforme bien comme ψ , à l'instar du cas abélien, à condition qu'on impose à A_μ de se transformer selon

$$\begin{aligned} \delta A_\mu^a(x) &= \partial_\mu \delta \alpha^a(x) + C_{bc}^a \delta \alpha^b(x) A_\mu^c(x) \\ &= (\partial_\mu \delta \alpha_b^a - A_\mu^c(x) (t_c)^a_b) \delta \alpha^b(x) = (D_\mu \delta \alpha)^a(x) . \end{aligned} \quad (5.10)$$

Au terme $\partial_\mu \delta \alpha^a(x)$ près, on voit que $\{A_\mu^a\}$ se transforme bien selon la représentation adjointe (dont les matrices sont $(t_c)^a_b = -C_{bc}^a$). Enfin un tenseur de champ se transformant de façon covariante (c'est-à-dire sans terme inhomogène en $\partial \delta \alpha^a(x)$) peut être construit

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - [A_\mu, A_\nu] \quad (5.11)$$

ou en composantes

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - C_{bc}^a A_\mu^b A_\nu^c . \quad (5.12)$$

On démontre, au prix d'un peu d'algèbre, et en utilisant l'identité de Jacobi, que

$$\delta F_{\mu\nu}^a(x) = C_{bc}^a \delta \alpha^b(x) F_{\mu\nu}^c(x) , \quad (5.13)$$

qui est bien une transformation infinitésimale dans la représentation adjointe.

Il est en fait profitable, et peut-être plus éclairant, de regarder l'effet d'une transformation *finie* locale $g(x)$ du groupe G ,

$$\begin{aligned} \psi(x) &\mapsto \mathcal{D}(g(x))\psi(x) \\ A_\mu = A_\mu^a T_a &\mapsto \mathcal{D}(g(x))(-\partial_\mu + A_\mu(x))\mathcal{D}(g^{-1}(x)) , \end{aligned} \quad (5.14)$$

où \mathcal{D} est la représentation portée par ψ , et pour la dérivée covariante agissant sur ψ on a

$$D_\mu \psi(x) \mapsto \mathcal{D}(g(x))D_\mu \psi(x) \quad (5.15)$$

ou encore³

$$D_\mu \mapsto \mathcal{D}(g(x))D_\mu \mathcal{D}(g^{-1}(x)) . \quad (5.16)$$

2. Attention, cette convention implique que certaines expressions diffèrent par un facteur i du cas abélien.

3. Attention aux notations : dans cette équation, qui porte sur un opérateur différentiel, la dérivée ∂_μ contenue dans D_μ agit sur tout ce qui est à sa droite, tandis que dans la deuxième équation (5.14), elle n'agit que sur $\mathcal{D}(g^{-1}(x))$.

Or on vérifie aisément que dans une représentation donnée

$$[D_\mu, D_\nu] = -F_{\mu\nu} := -F_{\mu\nu}^a T_a \quad (5.17)$$

d'où il découle que $F_{\mu\nu}(x) \mapsto \mathcal{D}(g(x))F_{\mu\nu}\mathcal{D}(g^{-1}(x))$, et qu'en particulier, dans la représentation adjointe, la transformation finie de $F_{\mu\nu} = F_{\mu\nu}^a t_a$ est

$$F_{\mu\nu}(x) \mapsto g(x)F_{\mu\nu}(x)g^{-1}(x) , \quad (5.18)$$

dont (5.13) est la version infinitésimale.

“Pure jauge”

Si le tenseur $F_{\mu\nu}$ s'annule dans le voisinage d'un point x_0 , on peut écrire localement (c'est-à-dire dans ce voisinage) $A_\mu(x)$ comme une “pure jauge”, c'est-à-dire

$$F_{\mu\nu} = 0 \iff A_\mu(x) = (\partial_\mu g(x))g^{-1}(x) . \quad (5.19)$$

L'appellation “pure jauge” se justifie par le fait qu'un tel $A_\mu(x) = (\partial_\mu g(x))g^{-1}(x)$ est le transformé de jauge d'un champ de jauge ... nul! Le \Leftarrow se démontre au prix d'une ligne de calcul, pour le \Rightarrow , voir quelques lignes plus bas... Insistons sur le caractère local de cette propriété.

Transport parallèle le long d'une courbe

Un autre objet intéressant est l'élément du groupe attaché à une courbe C allant de x_0 à x

$$\gamma(C) := P \exp \left(\int_C dx^\mu A_\mu(x) \right) \quad (5.20)$$

où $A = A^a t_a$ est pris dans la représentation adjointe et où le symbole P signifie qu'une paramétrisation $x(s)$ de la courbe ayant été choisie, les termes dans le développement de l'exponentielle sont ordonnés de droite à gauche selon les s croissants (cf le T -produit en théorie quantique des champs). On montre que sous l'effet de la transformation de jauge (5.14)

$$\gamma(C) \mapsto g(x)\gamma(C)g^{-1}(x_0) . \quad (5.21)$$

Plus généralement, pour toute représentation \mathcal{D} et avec $A = A^a T_a$, (5.20) définit un $\gamma_{\mathcal{D}}(C)$ dans la représentation \mathcal{D} se transformant selon $\gamma_{\mathcal{D}}(C) \mapsto \mathcal{D}(g(x))\gamma_{\mathcal{D}}(C)\mathcal{D}(g^{-1}(x_0))$.

Exercice. Démontrer cette assertion en considérant d'abord une trajectoire infinitésimale de x à $x + dx$, c'est-à-dire $\gamma(C) \approx 1 + A_\mu(x)dx^\mu$, et en effectuant une transformation de jauge finie $A_\mu(x) \rightarrow g(x)(-\partial_\mu + A_\mu(x))g^{-1}(x)$, montrer que $\gamma(C) \rightarrow g(x + dx)\gamma(C)g^{-1}(x)$. Le résultat pour une courbe finie s'ensuit en combinant ces éléments de courbe infinitésimaux.

Étant donné un objet, tel le champ ψ , se transformant selon une représentation \mathcal{D} , le rôle de $\gamma_{\mathcal{D}}(C)$ est de “transporter” $\psi(x_0)$ en un objet noté ${}^t\psi(x)$ se transformant comme $\psi(x)$. Montrer que pour un trajet infinitésimal $(x, x + dx)$ la différence ${}^t\psi(x + dx) - \psi(x + dx)$ fait apparaître de façon naturelle la dérivée covariante.

Considérons alors le cas où $x = x_0$. La boucle C est fermée et $\gamma(C)$ se transforme de façon covariante, $\gamma(C) \mapsto g(x_0)\gamma(C)g^{-1}(x_0)$. Examinons à nouveau le cas d’une boucle infinitésimale. On montre alors que

$$\gamma(C) \approx \exp \frac{1}{2} \int_{\mathcal{S}} dx^\mu \wedge dx^\nu F_{\mu\nu} , \quad (5.22)$$

où l’intégrale s’effectue sur une surface infinitésimale \mathcal{S} s’appuyant sur C .

Exercice : Démontrer cette assertion en considérant un circuit carré élémentaire s’étendant à partir de x le long des axes de coordonnées μ et ν : $(x \rightarrow x+dx^\mu \rightarrow x+dx^\mu+dx^\nu \rightarrow x+dx^\nu \rightarrow x)$, et développer au second ordre en dx pour obtenir $\gamma(C) \approx 1 + dx^\mu dx^\nu F_{\mu\nu}$ (sans sommation sur μ, ν). Indication : la formule du commutateur du Chap. 1, (1.20), simplifie beaucoup le calcul !

Cela a une conséquence immédiate : si $F = 0$, tout $\gamma(C)$ de la forme (5.20) n’est pas sensible à de petites variations du contour C à extrémités x_0 et x fixées et ne dépend donc que de x_0 et x . Le $g(x, x_0) := \gamma(C)$ qui en résulte satisfait $(\partial_\mu - A_\mu)g(x, x_0) = 0$, (vérifier !), ce qui achève la démonstration de (5.19).

Boucle de Wilson

Revenons au cas d’une boucle fermée C dans (5.20). Comme on vient de le noter, $\gamma(C)$ se transforme de façon covariante, $\gamma(C) \mapsto g(x_0)\gamma(C)g^{-1}(x_0)$. Sa trace

$$W(C) = \text{tr } \gamma(C) = \text{tr } P \exp \oint dx^\mu A_\mu(x) \quad (5.23)$$

est donc invariante. Toute quantité physique doit être “invariante de jauge”, c’est-à-dire invariante par une transformation de jauge. C’est le cas de $\text{tr } F_{\mu\nu} F^{\mu\nu}$, $\bar{\psi}(i\cancel{\partial} - \cancel{A})\psi$ etc. L’intérêt de $W(C)$ est d’être une quantité invariante *non locale*, dépendant du contour C . Noter qu’elle dépend de la représentation dans laquelle on évalue $A = A^a T_a$. Cette *boucle de Wilson* a été proposée par Wilson et Polyakov comme permettant la mesure du potentiel d’interaction entre les particules se propageant le long de C , et comme fournissant donc un bon indicateur du confinement. Voir plus bas au § 5.3.1, et voir le Problème I à la fin de ce chapitre pour une version discrétisée de cette quantité.

5.1.3 Géométrie des champs de jauge

Les considérations qui précèdent montrent que la théorie des champs de jauge a une forte coloration géométrique. Le langage approprié pour discuter ces choses est celui de la *théorie des fibrés*, fibré principal pour le groupe de jauge, fibré vectoriel pour chacun des champs de matière comme ψ , au dessus de l’espace de base qui est l’espace-temps. Le champ de jauge est une connexion sur le fibré, qui permet de définir un transport parallèle de point à point. Le tenseur $F_{\mu\nu}$ en est la courbure, ce qu’exprime (5.17) ou (5.22). Toutes ces notions sont définies localement, dans un système de coordonnées locales (une carte), et les changements de carte implique des transformations de la forme (5.14). Ce langage devient particulièrement utile quand on s’intéresse aux propriétés topologiques (instantons etc) ou globales (problème de Gribov) des théories de jauge. Pour la simple introduction aux propriétés de symétrie locale et à la construction *perturbative* du modèle standard, nous n’en aurons pas besoin.

5.1.4 Lagrangien de Yang–Mills

Le lagrangien décrivant le champ de jauge couplé à un champ de matière tel ψ via le couplage minimal s'écrit donc

$$\mathcal{L} = \frac{1}{2g^2} \text{tr} (F_{\mu\nu} F^{\mu\nu}) + \bar{\psi} (i(\not{\partial} - \not{A}) - m) \psi, \quad (5.24)$$

avec un paramètre, le couplage g . La valeur de ce couplage est évidemment liée à la normalisation des matrices T_a intervenant dans $F_{\mu\nu} = F_{\mu\nu}^a T_a$. On montre (cf Exercice B en fin de chapitre) que pour toute algèbre de Lie *simple* on peut choisir une base telle que dans toute représentation R , $\text{tr} T_a T_b = -T_R \delta_{ab}$, avec T_R un coefficient réel positif dépendant du groupe et de la représentation. On choisira pour écrire $F_{\mu\nu}$ la représentation fondamentale de dimension la plus basse (la représentation de définition de dimension N dans le cas de $SU(N)$) avec la normalisation $T_f = \frac{1}{2}$, donc $\text{tr} T_a T_b = -\frac{1}{2} \delta_{ab}$. Au lagrangien \mathcal{L} , on peut ajouter la contribution de champs de bosons, etc. Noter que la représentation portée par les fermions ou les autres champs de matière, qui apparaît dans la dérivée covariante $D_\mu = \partial_\mu - A_\mu^a T_a$, peut différer de cette représentation fondamentale.

Tel quel, le lagrangien \mathcal{L} de (5.24) ressemble beaucoup au lagrangien du cas abélien (5.1), après qu'on a effectué le changement $A \rightarrow gA$.

Retenons les éléments les plus marquants de cette construction :

- comme dans le cas abélien, le principe d'invariance de jauge implique un couplage *minimal* de type universel, via la dérivée covariante ; (bien entendu, l'addition de termes invariants de jauge “non minimaux”, tel $\bar{\psi} \sigma_{\mu\nu} F^{\mu\nu} \psi$, est possible mais limité par la contrainte de renormalisabilité) ;
- contrairement au cas abélien où chaque charge est indépendante et non quantifiée (tout au moins si le groupe de jauge est \mathbb{R} et non pas $U(1)$), la constante de couplage g de tous les champs au champ de jauge est la même à l'intérieur de chaque composante simple du groupe de jauge ;
- comme dans le cas abélien, le champ de jauge vient naturellement sans terme de masse : un terme de masse $\frac{1}{2} M^2 A_\mu A^\mu$ brise en effet l'invariance de jauge. Cela est très embarrassant pour les applications physiques, les champs vectoriels (de spin 1) de masse nulle étant exceptionnels dans la nature (le champ électromagnétique et ses excitations photoniques étant le contre-exemple de base), et nous contraindra soit à introduire des mécanismes “doux” de brisure de l'invariance de jauge (brisure spontanée) pour y remédier, soit à invoquer le mécanisme du confinement pour cacher les gluons de masse nulle ;
- contrairement au cas abélien, le champ de jauge “porte une charge du groupe” : on a vu que pour les transformations globales (c'est-à-dire indépendantes de x) du groupe G , A_μ se transforme selon la représentation adjointe. La propriété du champ d'être chargé a des implications importantes dans de nombreux phénomènes, les effets infra-rouges entre autres, mais aussi ultraviolets (signe de la fonction β), comme on verra plus bas.

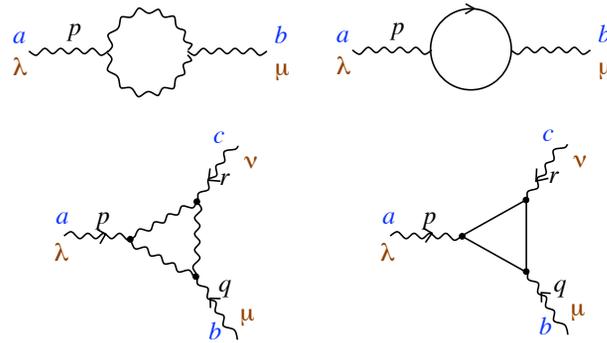


FIGURE 5.1 – Quelques diagrammes à une boucle dans une théorie de jauge

5.1.5 Quantification. Règles de Feynman.

La quantification de la théorie de Yang–Mills nécessite de surmonter de sérieuses difficultés que nous nous contenterons d’évoquer. Comme dans l’électrodynamique, la forme quadratique du champ de jauge dans le lagrangien,

$$(\partial_\mu A_\nu - \partial_\nu A_\mu)^2 \quad \text{ou dans l’espace de Fourier} \quad A_\mu(-k)(k^\mu k^\nu - k^2 g^{\mu\nu})A_\nu(k)$$

est dégénérée, donc non inversible, ce qui est un reflet de l’invariance de jauge ; en conséquence, le propagateur du champ A_μ n’est *a priori* pas défini. Il faut donc “fixer la jauge”, en imposant une condition non-invariante de jauge, et la procédure de Faddeev et Popov, justifiée par leur étude générale de la quantification des *systemes contraints*, conduit via l’introduction de champs supplémentaires à des règles de Feynman explicites, (voir par exemple [IZ, chap. 12] et les cours du second semestre).

On démontre, et cela a été une étape décisive dans la construction du Modèle Standard⁴, que la théorie de jauge ainsi quantifiée est renormalisable : toutes les divergences ultraviolettes apparaissant dans les calculs de diagrammes de Feynman, peuvent, à tout ordre fini de la série des perturbations, être absorbées dans une redéfinition des paramètres –couplages, normalisation des champs, masses– du lagrangien. Cette procédure de renormalisation préserve l’invariance de jauge.

Ainsi à l’ordre à une boucle, les diagrammes de la Fig. 5.1 ont des divergences qui peuvent être absorbées dans un changement de la normalisation du champ A_μ (“renormalisation de fonction d’onde”) et une renormalisation de la constante de couplage g

$$g \mapsto g_0 = \left(1 - \frac{g^2}{(4\pi)^2} \left(\frac{11}{3} C_2 - \frac{4}{3} T_f \right) \log \frac{\Lambda}{\mu} \right) g, \quad (5.25)$$

où Λ est une échelle de coupure ultraviolette (“cutoff”) et μ une échelle de masse indispensable à la définition de la procédure de renormalisation. T_f a été défini plus haut, juste après (5.24), et C_2 est la valeur de l’opérateur de Casimir quadratique dans la représentation adjointe, $C_{acd}C_{bcd} = C_2 \delta_{ab}$, donc $C_2 = c_2(\text{adj})$ avec les notations de l’exercice A.1, et $C_2 = N$ pour $SU(N)$, cf exercice A.2.

4. G. ’t Hooft et M. Veltman, prix Nobel 1999

5.2 Champs de jauge massifs

5.2.1 Interactions faibles et bosons intermédiaires

On a vu au chapitre 4 (équ. (4.32)) que le lagrangien de Fermi

$$\mathcal{L}_{\text{Fermi}} = -\frac{G}{\sqrt{2}} J^\rho(x) J_\rho^\dagger(x) \quad (5.26)$$

fournissait une bonne description de la physique de basse énergie des interactions faibles : processus leptoniques tel $\bar{\nu}_e e^- \rightarrow \bar{\nu}_e e^-$ ou $\bar{\nu}_\mu \mu^-$, semi-leptoniques comme $\pi^+ \rightarrow \mu^+ \nu_\mu$ ou la désintégration β du neutron $n \rightarrow p e^- \bar{\nu}_e$, ou non leptoniques : $\Lambda \rightarrow p \pi^-$, $K^0 \rightarrow \pi \pi$, etc. Mais qu'il n'était théoriquement pas satisfaisant, puisqu'il conduit à une théorie non renormalisable, rendant impossible tout calcul au delà du "terme de Born", le premier ordre de la série des perturbations, lequel viole l'unitarité.

La violation de l'unitarité apparaît dans le calcul de la section efficace totale σ de tout processus, au premier ordre de la série des perturbations. Un simple argument dimensionnel donne à haute énergie

$$\sigma \sim \text{const. } G^2 s$$

où s est le carré de l'énergie dans le centre de masse. Mais ce comportement est en conflit avec des résultats généraux basés sur l'unitarité qui prévoient que σ doit décroître dans chaque onde partielle comme $1/s$. Une violation de l'unitarité du terme de Born est donc attendue à une énergie de l'ordre de $\sqrt{s} \sim G^{-\frac{1}{2}} \sim 300$ GeV. Et la non-renormalisabilité de la théorie empêche d'améliorer ce terme de Born par le calcul de "corrections radiatives", c'est-à-dire de termes plus élevés de la série des perturbations.

L'idée est alors de regarder $\mathcal{L}_{\text{Fermi}}$ comme l'approximation d'une théorie où le courant chargé J^ρ est couplé à un champ vectoriel chargé W de masse M , dans la limite de grande masse M^5 . Considérons le nouveau lagrangien

$$\mathcal{L}_{\text{int.boson}} = g J^\rho(x) W_\rho^\dagger(x) + \text{h.c.} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + M^2 W_\rho^\dagger W^\rho. \quad (5.27)$$

Dans la limite de grande masse M , on peut négliger le terme cinétique $-\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$ par rapport au terme de masse, le champ W devient un simple champ auxiliaire sans dynamique sur lequel on peut intégrer en "complétant le carré", et on retrouve $\mathcal{L}_{\text{Fermi}}$ à condition que

$$\frac{G}{\sqrt{2}} = \frac{g^2}{M^2} \quad (5.28)$$

reliant le nouveau couplage g au couplage de Fermi. La théorie (5.27) avec son "boson intermédiaire" W , vecteur des interactions faibles, est-elle la bonne théorie des interactions faibles ? En fait le propagateur du champ massif W est

$$-i \frac{g_{\mu\nu} - \frac{k_\mu k_\nu}{M^2}}{k^2 - M^2} \quad (5.29)$$

qui se comporte mal pour $k \gg M$ et rend à nouveau la théorie non-renormalisable : on n'a fait que déplacer le problème. La solution vient d'une manière douce et subtile (!) d'introduire la masse du champ W , via un mécanisme de brisure spontanée de symétrie.

5. L'inverse de la masse M représente la portée des interactions faibles qu'on sait courte, et la masse M est donc élevée (de l'ordre de 100 GeV, comme nous verrons).

5.2.2 Théorie de jauge à symétrie brisée spontanément. Mécanisme de Brout–Englert–Higgs

Revenons au cas abélien décrit par (5.1), (5.4) et supposons maintenant que le potentiel V est tel que son minimum est localisé à une valeur non nulle de $\phi^* \phi$. En conséquence, le champ ϕ acquiert une vev $\langle \phi \rangle = v/\sqrt{2} \neq 0$. Reparamétrisant le champ ϕ selon

$$\phi(x) = e^{iq\theta(x)/v} \frac{v + \varphi(x)}{\sqrt{2}} \quad (5.30)$$

avec v réel et φ hermitien, et accompagnant cela d'une transformation de jauge

$$\begin{aligned} \phi(x) &\mapsto \phi'(x) = e^{-iq\theta(x)/v} \phi(x) = \frac{v + \varphi(x)}{\sqrt{2}} \\ A_\mu(x) &\mapsto A'_\mu(x) = A_\mu(x) + \frac{1}{v} \partial_\mu \theta(x) \end{aligned} \quad (5.31)$$

et de la transformation correspondante pour les éventuels autres champs chargés ($\psi \dots$), on voit que le lagrangien $\delta\mathcal{L}$ de (5.4) s'écrit

$$\begin{aligned} \delta\mathcal{L} &= (\partial_\mu - iqA'_\mu)\phi'(\partial^\mu + iqA'^\mu)\phi' - V(\phi'^2) \\ &= \frac{1}{2}|(\partial_\mu - iqA'_\mu)\varphi|^2 + \frac{1}{2}q^2v^2A'_\mu A'^\mu - V\left(\frac{1}{2}(v + \varphi)^2\right). \end{aligned} \quad (5.32)$$

Au final, on voit que la brisure spontanée de la symétrie $U(1)$ par le champ bosonique ϕ conduit à l'apparition d'un terme de masse du champ de jauge A'_μ ! On note aussi que le champ θ qui en l'absence du champ de jauge, aurait été le champ de Goldstone, a purement et simplement disparu, “avalé” par le nouveau mode massif (“longitudinal”) du vecteur A'_μ ; le nombre total de degrés de liberté de ces champs n'est donc pas modifié. C'est le mécanisme de Brout–Englert–Higgs⁶, dans sa version abélienne. Si le boson ϕ est couplé à un champ de fermions ψ par un terme du type $\bar{\psi}\phi\psi$, l'apparition de sa “vev” donne lieu à un terme de masse $\frac{qv}{\sqrt{2}}\bar{\psi}\psi$ pour le ψ .

Remarque importante. Bien comprendre que dans ce mécanisme, la symétrie globale est brisée spontanément, mais que l'invariance de jauge est bien toujours là. À preuve, la *direction* dans laquelle “pointe” le champ scalaire n'est pas observable, (invariance de jauge), on sait seulement que son *module* v est non nul (brisure spontanée)

Ce mécanisme de Brout–Englert–Higgs (B-E-H) s'étend à un groupe de jauge non abélien. Les détails dépendent du schéma de brisure et du choix de représentation pour le champ bosonique. En général, si le groupe G est brisé en un sous-groupe H , les $r = \dim G - \dim H$ bosons de Goldstone, qui sont en correspondance avec les générateurs du quotient (“coset”) G/H , se muent en modes longitudinaux de r vecteurs. Il reste $\dim H$ champs vectoriels de masse nulle. Exemple : le modèle standard électrofaible du § 5.3.2 : $G = SU(2) \times U(1)$, $H = U(1)$ (pas celui qu'on croit !), trois champs de jauge deviennent massifs, un demeure de masse nulle.

Une étape cruciale dans la construction du modèle standard a été de comprendre que ce mécanisme de brisure spontanée de symétrie dans une théorie de jauge, décrit ici au niveau

6. F. Englert et P. Higgs, prix Nobel 2013

classique, est compatible avec la quantification de la théorie. La renormalisabilité à 4 dimensions de la théorie de jauge n'est pas affectée par cette brisure, et la théorie obtenue est bien unitaire. Seuls les états physiques (champs de jauge massifs ou de masse nulle, bosons ayant subsisté à la brisure etc) participent à la somme sur les états intermédiaires dans la relation d'unitarité.

5.3 Le modèle standard

Ce qu'on appelle actuellement le modèle standard de la physique des particules est une théorie de jauge basée sur un groupe de jauge non simple : $SU(3) \times SU(2) \times U(1)$, dans lequel les différents facteurs jouent des rôles bien distincts. Comme le groupe a trois facteurs, la théorie dépend *a priori* de trois constantes de couplage indépendantes et possède des champs de jauge pour chacun, qui sont couplés aux champs de matière, quarks et leptons, ainsi qu'à des champs de bosons qui jouent un rôle auxiliaire mais crucial !

5.3.1 Le secteur fort

Le groupe $SU(3)$ est celui de la couleur (cf chap. 4, § 4.3.2). Les champs de jauge A_μ portent des indices de la représentation adjointe (de dimension 8). Les particules associées, ou *gluons*, sont des particules de spin 1 et de masse nulle, jamais observées directement jusqu'à présent. Les champs de gluons sont couplés aux degrés de liberté de couleur des champs fermioniques de quarks, ψ_{Ai} , qui portent un indice A de la représentation 3 (ou $\bar{3}$ pour les $\bar{\psi}$) (et aussi un indice de saveur $i = u, d, s, c, b, t$, sur lequel $SU(3)_c$ n'agit pas). La théorie ainsi définie est la Chromodynamique Quantique (QCD dans l'acronyme anglais). Elle décrit la physique de toutes les interactions fortes. Son lagrangien est du type (5.24), avec des masses fermioniques dépendant de la saveur, engendrées par le secteur faible.

Liberté asymptotique

Connaissant les règles de Feynman et la renormalisabilité de la théorie, cf § 5.1.5, on peut calculer la renormalisation de la constante de couplage g , (5.25), et la fonction beta correspondante. On trouve⁷

$$\beta(g^2) = -\Lambda \frac{\partial}{\partial \Lambda} g^2(\Lambda)|_{g_0} = -2 \frac{g^4}{(4\pi)^2} \left(\frac{11}{3} C_2 - \frac{4}{3} T_f \right) + O(g^6) \quad (5.33)$$

Il apparaît donc que cette fonction beta est négative au voisinage de $g = 0$, tant que le coefficient $\frac{11}{3} C_2 - \frac{4}{3} T_f > 0$ (pas trop de champs de matière!), autrement dit que $g = 0$ est un point fixe attractif ultraviolet du groupe de renormalisation : $\beta(g^2(\lambda)) = \frac{dg^2(\lambda)}{d \log \lambda} < 0 \Rightarrow g^2(\lambda) \sim (b \log \lambda)^{-1} \rightarrow 0$ quand $\lambda \rightarrow \infty$, avec $b =$ coefficient du terme $-g^4$ dans (5.33).

C'est la *liberté asymptotique*, une propriété fondamentale des interactions fortes.

Exercice : combien de triplets de quarks sont compatibles avec la liberté asymptotique de la QCD ?

Cette théorie de jauge non abélienne est la seule théorie des champs locale et renormalisable à 4 dimensions

7. David J. Gross, H. David Politzer, Frank Wilczek, prix Nobel 2004

à posséder cette propriété de liberté asymptotique. Comme telle, elle est la seule compatible avec les résultats des expériences de diffusion profondément inélastique de leptons sur des hadrons, qui révèlent une structure interne de ces derniers faite de constituants ponctuels quasi-libres à très courte distance (cf les cours du second semestre sur la chromodynamique quantique).

Ce groupe de jauge $SU(3)$ est non brisé, ni explicitement, ni spontanément. Ceci est essentiel pour la cohérence du scénario imaginé pour expliquer le confinement des quarks et gluons (cf chap. 4, § 4.3.2.) : les particules non singulets du groupe de jauge sont réputées inobservables, car soumises à des interactions d'intensité croissant avec la distance quand on cherche à les séparer.

Cette propriété d'“esclavage infra-rouge” (c'est-à-dire à grande distance) est le pendant de celle de “liberté asymptotique”, à courte distance. Elle montre que le phénomène de confinement est un phénomène de couplage fort, par essence non-perturbatif, c'est-à-dire inaccessible aux calculs de la théorie des perturbations.

Une approche non-perturbative qui a fourni de nombreux résultats qualitatifs et quantitatifs est la discrétisation de la chromodynamique en une *théorie de jauge sur un réseau*. Cela a ouvert la voie à l'utilisation de méthodes empruntées à la Mécanique Statistique des modèles sur réseau, analytiques (calculs de couplage fort ou de haute température, méthode de champ moyen, ...) ou numériques (Monte-Carlo). Le scénario de confinement semble confirmé dans cette approche par l'étude de la valeur moyenne de la boucle de Wilson définie plus haut (§ 5.1.2). Selon l'idée de Wilson et Polyakov, pour une boucle rectangulaire C de dimensions $T \times R$, $T \gg R$, et portant la représentation σ du groupe de jauge, $W^{(\sigma)}(C)$ décrit la propagation pendant le temps T d'une paire de particules statiques (de masse très grande), figées à une distance relative R . On cherche à calculer le potentiel entre ces charges statiques

$$V_{\sigma}(R) = - \lim_{T \rightarrow \infty} \frac{1}{T} \log W^{(\sigma)}(C) .$$

Si la boucle de Wilson a une “loi d'aire”, $\log W(C) \sim -\kappa RT$, le potentiel entre les charges statiques croît linéairement, $V \sim \kappa R$, ce qui est en accord avec l'idée de confinement. C'est ce qui se passe en général dans une théorie de jauge sur réseau en couplage fort, voir le Problème I en fin de chapitre. Les calculs de Monte-Carlo confirment que ce comportement persiste aux couplages faibles pertinents pour la théorie continue (le couplage de la théorie sur réseau est le couplage effectif à l'échelle de la maille du réseau a , donc selon la liberté asymptotique, $g_0^2 = g^2(\Lambda = 1/a) \rightarrow 0$), et permettent même de déterminer numériquement le coefficient κ , ou *tension de corde*.

La QCD est toujours un sujet d'étude très actif. Les interactions fortes sont en effet omniprésentes et l'observation de toute autre interaction, de tout autre effet, présuppose une connaissance aussi précise que possible de la contribution forte. Dans l'analyse des données de LHC les calculs des contributions de QCD gardent une importance fondamentale : la “nouvelle physique” ne pourra être identifiée que si le fonds du Modèle Standard est parfaitement connu. De plus, l'étude de l'hadronisation des quarks et gluons, de la diffusion profondément inélastique et d'autres phénomènes hadroniques demeure un sujet très “chaud” et un point-clé où la théorie se confronte à l'expérience.

5.3.2 Le secteur électro-faible, une esquisse.

La théorie de jauge de groupe $SU(2) \times U(1)$ décrit les interactions électro-faibles (modèle de Glashow–Salam–Weinberg⁸). On parle parfois d'isopin faible et d'hypercharge faible pour désigner les générateurs de ces groupes $SU(2)$ et $U(1)$. Nous nous contenterons de présenter les

8. S. Glashow, A. Salam, S. Weinberg, prix Nobel 1979

grandes lignes de la construction, sans bien expliquer les raisons qui ont conduit aux choix de groupes, de représentations etc.

Appelons A_μ^a , W_μ^i et B_μ les champs de jauge de SU(3), SU(2) et U(1) respectivement. Les quarks et leptons *gauches*, $\psi_L := \frac{1}{2}(1-\gamma_5)\psi$, et *droits*, $\psi_R := \frac{1}{2}(1+\gamma_5)\psi$, sont couplés aux champs W_μ et B_μ de façon différente. On peut écrire la dérivée covariante de l'un de ces champs selon

$$D_\mu\psi = (\partial_\mu - g_3 A_\mu^a T_a - g_2 W_\mu^j t_j - i\frac{g_1}{2} y B_\mu)\psi \quad (5.34)$$

où T_a , resp. t_j denotent des générateurs infinitésimaux antihermitiens de SU(3) et SU(2) dans la représentation de ψ ; les représentations assignées à chaque champ, lepton ou quark, gauche ou droit, sont la représentation triplet de SU(3)_c pour les quarks et la triviale pour les leptons, bien sûr, et pour la partie électro-faible, données dans la Table ci-dessous.

Une conséquence remarquable de l'utilisation de SU(2) comme groupe de symétrie des interactions faibles est qu'en plus des deux courants chargés $J_\mu^{1,2}$ (ou J_μ^\pm) de la théorie de Fermi apparaît une troisième composante J_μ^3 . Ce courant neutre, qui *n'est pas* le courant électromagnétique et qui est couplé au champ de jauge W_μ^3 , est nécessairement présent et contribue par exemple à la diffusion $e^-\nu_\mu \rightarrow e^-\nu_\mu$ interdite dans la théorie de Fermi. La découverte expérimentale de ces courants neutres (1973)⁹ a été la première confirmation de la validité du Modèle Standard.

Quarks & Leptons	(ν_L^e, e_L)	ν_R^e	e_R	(u_L, d_L)	u_R	d_R
Isospin faible t_z	$(\frac{1}{2}, -\frac{1}{2})$	0	0	$(\frac{1}{2}, -\frac{1}{2})$	0	0
Hypercharge faible y	$(-1, -1)$	0	-2	$(\frac{1}{3}, \frac{1}{3})$	$\frac{4}{3}$	$-\frac{2}{3}$
Charge électrique $Q = \frac{1}{2}y + t_z$	$(0, -1)$	0	-1	$(\frac{2}{3}, -\frac{1}{3})$	$\frac{2}{3}$	$-\frac{1}{3}$

Table 1. Nombres quantiques faibles des leptons ν^e et e et des quarks u, d .

Les choses se répètent à l'identique pour les autres générations.

Le groupe U(1)_{em} de l'électromagnétisme va maintenant être identifié par les charges des champs. C'est un "mélange" du facteur U(1) initial et d'un sous-groupe U(1) de SU(2). Ce mélange est caractérisé par un angle θ_W , dit *angle de Weinberg* : si on note B_μ et W_μ les champs de jauge des groupes U(1) et SU(2) respectivement, le champ électromagnétique est $A_\mu^{em} = \cos\theta_W B_\mu + \sin\theta_W W_\mu^3$, la combinaison orthogonale correspondant à un autre champ vectoriel nommé Z^0 .

Examinons les termes de "courant neutre" couplant par exemple l'électron et son neutrino aux bosons neutres W^3 et B . On les lit sur les dérivées covariantes (5.34) avec les nombres quantiques de la Table 1

$$\frac{1}{2}i[\bar{e}_L(-g_2 W_\mu^3 - g_1 B_\mu)\gamma^\mu e_L + \bar{e}_R(-2g_1 B_\mu)\gamma^\mu e_R + \bar{\nu}_e(g_2 W_\mu^3 - g_1 B_\mu)\gamma^\mu \nu_e]$$

La rotation $W^3 = \cos\theta_W Z^0 + \sin\theta_W A$, $B = -\sin\theta_W Z^0 + \cos\theta_W A$ doit être telle que la charge électrique e (couplage au A) est la même pour e_L et e_R et nulle pour ν_e . Il vient

$$2e = g_2 \sin\theta_W + g_1 \cos\theta_W = 2g_1 \cos\theta_W \quad \text{et} \quad g_2 \sin\theta_W - g_1 \cos\theta_W = 0$$

9. En lire l'histoire dans <http://cerncourier.com/cws/article/cern/29168>

qui sont bien compatibles et fournissent

$$\tan \theta_W = \frac{g_1}{g_2} \quad e = g_1 \cos \theta_W = g_2 \sin \theta_W . \quad (5.35)$$

Le résultat de ce calcul ne dépend bien sûr pas de la représentation à laquelle on l'applique. A ce stade nous n'avons fait qu'un changement de paramètres, $(g_1, g_2) \mapsto (e, \theta_W)$ mais ces derniers sont physiquement observables.

Le lagrangien contient aussi un couplage à un champ bosonique, supposé être un doublet complexe de $SU(2)$ $\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$ (isospin faible $\frac{1}{2}$, hypercharge faible $y = +1$), et donc $D_\mu \Phi = (\partial_\mu - ig_2 W_\mu^i \frac{\tau_i}{2} - \frac{i}{2} g_1 B_\mu) \Phi$. Le champ Φ est doté d'un potentiel $V(\Phi)$ en "chapeau mexicain", responsable de la brisure spontanée de $SU(2) \times U(1)$ en $U(1)_{em}$, et donc de l'apparition des masses des champs vectoriels selon le mécanisme décrit au § 5.2.2, et même de celle des fermions. Ce champ (2 composantes complexes, donc 4 hermitiennes) voit trois de ces composantes disparaître au profit des modes longitudinaux des champs de jauge devenus massifs. Une seule des quatre composantes demeure, et c'est le *boson de Higgs* que cette composante φ crée qui a été découverte en 2012 dans les expériences ATLAS et CMS du LHC. En parallèle, trois des quatre champs de jauge, les W^\pm et le Z^0 , deviennent massifs, le quatrième, le champ électromagnétique A demeurant de masse nulle.

La brisure de la symétrie $SU(2) \times U(1)$ par le champ Φ s'effectue dans une direction qui préserve $U(1)_{em}$. (Ou plus exactement c'est la direction de cette brisure qui détermine ce qu'on appelle $U(1)_{em}$.) On écrit, en généralisant (5.30) au groupe $SU(2)$ de générateurs $i\frac{\tau^j}{2}$ ($\tau^j =$ matrices de Pauli)

$$\Phi(x) = e^{i\xi_j(x)\frac{\tau^j}{2v}} \begin{pmatrix} 0 \\ \frac{v+\varphi(x)}{\sqrt{2}} \end{pmatrix} ,$$

que l'on accompagne d'une transformation de jauge, ce qui fait disparaître les champs ξ_j et donne pour les champs W et B la forme quadratique (termes de masse)

$$\mathcal{L}_{(2)} = \frac{1}{8} v^2 [(g_1 B - g_2 W^3)^2 + g_2^2 ((W^1)^2 + (W^2)^2)]$$

C'est bien comme attendu la composante $Z^0 = (g_1 B - g_2 W^3)/\sqrt{g_1^2 + g_2^2}$ qui devient massive, ainsi que $W^{1,2}$, tandis que la combinaison orthogonale $A = (g_2 B + g_1 W^3)/\sqrt{g_1^2 + g_2^2}$ demeure de masse nulle. On trouve

$$M_{W^\pm} = \frac{1}{2} v g_2 \quad M_{Z^0} = \frac{1}{2} v \sqrt{g_1^2 + g_2^2} \quad (5.36)$$

et en utilisant (5.35), la relation $\frac{G}{\sqrt{2}} = \frac{g_2^2}{8M_W^2}$ qu'on lit sur le lagrangien et la valeur expérimentale de e et de $G = 10^{-5} m_p^2$

$$M_{W^\pm} \approx \frac{38}{\sin \theta_W} \text{ GeV} \quad M_{Z^0} = \frac{M_W}{\cos \theta_W} \approx \frac{38}{\sin \theta_W \cos \theta_W} \text{ GeV} .$$

Ces expressions subissent ensuite de petites corrections perturbatives. Enfin la masse du fameux boson de Higgs φ n'est pas prédite par la théorie. Des expériences successives ont progressivement exclu des domaines de masses de plus en plus étendus, en réduisant la "fenêtre" possible de 100–200 GeV à 120–130 GeV. Les résultats de l'été 2012 ont finalement identifié une particule de masse 125.9 ± 0.4 GeV, et les déterminations de son spin, de ses modes de désintégration, etc, ont confirmé qu'il s'agit bien de la particule de Higgs attendue. Voir les cours de P. Binétruy et P. Fayet au second semestre pour plus de détails.

Les "bosons intermédiaires" associés aux champs vectoriels massifs W^\pm et Z^0 ont été découverts expérimentalement dès la fin des années 70¹⁰; ils ont des masses $M_{W^\pm} = 80.4$ GeV

10. Carlo Rubbia et Simon van der Meer, prix Nobel 1984

et $M_{Z^0} = 91.2 \text{ GeV}$ compatibles avec une valeur de l'angle de Weinberg donnée par

$$\sin^2 \theta_W \approx 0.23 , \quad (5.37)$$

également compatible avec tous les autres résultats expérimentaux.

Au total, le lagrangien décrivant toutes les interactions en dehors de la gravitation a la forme remarquablement simple et compacte

$$\mathcal{L} = -\frac{1}{4} \mathbf{F}_{\mu\nu} \mathbf{F}^{\mu\nu} + \sum_{\substack{\text{left et right} \\ \text{quarks \& leptons}}} \bar{\psi} \gamma^\mu D_\mu \psi + |D\Phi|^2 - V(\Phi) + \text{Higgs} - \text{fermions couplings} , \quad (5.38)$$

où $\mathbf{F}_{\mu\nu}$ désigne les trois tenseurs de champs de jauge \mathbf{A} , \mathbf{W} et A . Noter que l'invariance $SU(2) \times U(1)$ interdit les couplages entre fermions gauches et droits (qui se transforment sous des représentations différentes), et donc interdit des termes de masse. La seule échelle de masse se trouve dans $V(\Phi)$, et ce sont le mécanisme de Higgs et le couplage du champ Φ aux fermions –leptons et quarks– qui donnent lieu à l'apparition des masses de fermions et de (certains) bosons-vecteurs. Ce couplage, dit de Yukawa, est de la forme générale, (écrite ici pour les quarks),

$$\mathcal{L}_Y = -Y_{ij}^d \bar{\psi}_{Li} \cdot \Phi d_{Rj} - Y_{ij}^u \bar{\psi}_{Li} \cdot \tilde{\Phi}^\dagger u_{Rj} + \text{h.c.} , \quad (5.39)$$

avec des matrices *a priori* arbitraires Y_{ij}^d , Y_{ij}^u : $i, j = 1, 2, 3$ sont des indices de génération, le point dénote le produit scalaire des doublets d'isospin Φ et $\tilde{\Phi}^\dagger = \begin{pmatrix} \phi^{0\dagger} \\ -\phi^{+\dagger} \end{pmatrix}$ avec les doublets des quarks

$$\psi_{Li} = \begin{pmatrix} u_i \\ d_i \end{pmatrix}_L = \left(\begin{pmatrix} u \\ d \end{pmatrix}_L , \begin{pmatrix} c \\ s \end{pmatrix}_L , \begin{pmatrix} t \\ b \end{pmatrix}_L \right) .$$

Des couplages du même type apparaissent entre leptons et champs scalaires.

La vev $v/\sqrt{2}$ du champ ϕ^0 donne alors naissance à une “matrice de masse”. Une complication de la théorie décrite par (5.38) est que la diagonalisation de cette matrice de masse des quarks fait intervenir une rotation par une matrice unitaire de (u_L, c_L, t_L) et de (d_L, s_L, b_L) par rapport à la base couplée aux champs de jauge dans (5.38) : si (u_L, c_L, t_L) et (d_L, s_L, b_L) désignent maintenant les états propres de masse, le courant hadronique chargé couplé au champ W^+ est

$$J_\mu = (\bar{u} \bar{c} \bar{t})_L \gamma_\mu M \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L \quad (5.40)$$

avec M la *matrice unitaire de Cabibbo-Kobayashi-Maskawa*¹¹. Ce mécanisme généralise à 3 générations le mélange par l'angle de Cabibbo rencontré au chapitre 4, (équ. (2.20)) dans le cas de 2 générations. On écrit la matrice M sous la forme

$$M = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}$$

11. M. Kobayashi, T. Maskawa, prix Nobel 2008, avec Y. Nambu

avec 4 angles δ et θ_{ij} , ($c_{ij} = \cos \theta_{ij}$ et $s_{ij} = \sin \theta_{ij}$), et $\theta_{12} = \theta_C =$ angle de Cabibbo. Expérimentalement $0 \ll \theta_{13} \ll \theta_{23} \ll \theta_{12} \ll \pi/2$. La mesure précise des éléments de matrice de M est actuellement l'objet d'une activité intense, en relation avec l'étude de la violation de la symétrie CP (due en grande partie à la phase $e^{i\delta}$) et des "oscillations de saveurs".

C'est tout un cours qui serait nécessaire pour rendre compte des détails et des succès du modèle standard, cf les cours du 2ème semestre...

5.4 Compléments

5.4.1 Modèle standard et au delà.

Le modèle standard est à la fois remarquablement vérifié et peu satisfaisant. En dehors de la présence de neutrinos massifs, dont on est maintenant convaincu et qui nécessite de petits amendements au lagrangien (5.38), on n'a à ce jour observé aucun désaccord significatif entre les résultats expérimentaux et les prédictions du modèle. Les aspects non satisfaisants du modèle standard sont pourtant nombreux : le nombre jugé excessif (une vingtaine) de paramètres libres dans le modèle, le manque de "naturel" de la façon dont certains termes doivent être ajustés de façon extrêmement fine ; la question du mécanisme B-E-H qui semble être confirmé par la découverte du boson de Higgs, mais que certains physiciens considèrent comme une construction *ad hoc* ; etc.

Il faut mentionner les tentatives d'améliorer le modèle standard en fusionnant les 3 groupes de jauge au sein d'un plus grand groupe d'une théorie "grand-unifiée" (GUT en anglais ;-). On y consacre le paragraphe suivant.

Les extensions les plus en vogue du modèle standard sont en définitive celles basées sur la supersymétrie. Le "MSSM", ("*Maximally Supersymmetric (extension of the) Standard Model*"), ou le "NMSSM" ("*Next-to-...*"), résolvent le problème de hiérarchie, prédisent une convergence des couplages électro-faibles et fort à haute énergie (voir ci-dessous), et prédisent aussi l'existence de partenaires supersymétriques pour toutes les particules connues. Les résultats à venir prochainement du LHC pourraient valider ou contredire tel ou tel modèle...

5.4.2 Théories grand-unifiées ou GUTs

L'observation que les trois constantes de couplage g_1, g_2, g_3 semblent à partir de leurs valeurs mesurées aux énergies actuelles converger sous l'effet du groupe de renormalisation vers une valeur commune à une énergie d'environ 10^{15} ou 16 GeV a été une forte incitation dans le sens d'une grande unification, voir Fig. 5.2. La théorie grand-unifiée qui en résulte doit non seulement être une théorie de jauge dotée d'un seul couplage si le groupe d'unification G est simple, mais aussi être capable de prédire le contenu en champs et particules de matière selon les représentations de $SU(3) \times SU(2) \times U(1)$ à partir de représentations du groupe G . Pour des raisons variées, le groupe $SU(5)$ est le meilleur candidat. Cette GUT possède $\dim SU(5) = 24$ champs de jauge.

La raison principale du choix de $SU(5)$ vient du nombre de fermions chiraux par génération. Chaque génération du Modèle Standard contient deux saveurs de quarks venant chacune en 3 couleurs, plus un lepton, et chacun de ces 6+1 champs peut avoir deux chiralités, plus un neutrino supposé de masse nul et chiral. Au total il y a 15 fermions chiraux par génération. (Se rappeler que l'antiparticule d'un fermion droit est gauche :

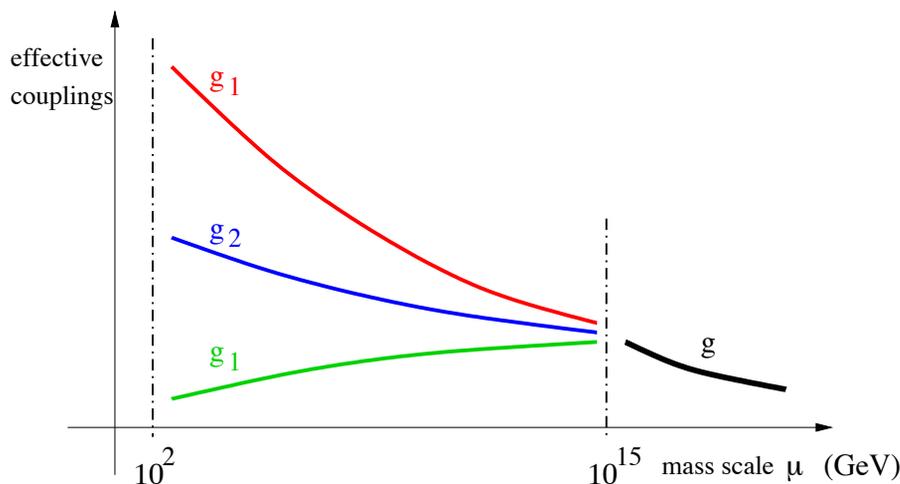


FIGURE 5.2 – Évolutions schématisées des 3 couplages effectifs du modèle standard et de celui de la théorie grand-unifiée.

on peut se contenter de raisonner sur des fermions gauches.) On cherche donc un groupe G simple possédant une représentation (réductible ou irréductible) de dimension 15 pouvant regrouper tous les fermions gauches de chaque génération. Le seul candidat est en définitive le groupe $SU(5)$ qui possède des représentations de dimension 15 : la représentation tensorielle symétrique, et des représentations somme de 5 (ou $\bar{5}$) et 10 (ou $\bar{10}$).

Le groupe $SU(5)$ des matrices unitaires 5×5 contient un sous-groupe $SU(3)$ (sous-matrices 3×3 du coin supérieur gauche), un sous-groupe $SU(2)$ (blocs 2×2 du coin inférieur droit), ce qui donne les générateurs correspondants de $SU(3) \times SU(2)$; le sous-groupe $U(1)$ est engendré par la matrice diagonale et de trace nulle $\text{diag}(-\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}, \frac{1}{2}, \frac{1}{2})$. Il est clair que ces trois groupes commutent entre eux.

Il faut alors décomposer tous les champs (les représentations 5, 10, 15 et 24) en représentations de $SU(3) \times SU(2)$. Cet exercice montre que la représentation 15 est à écarter et que la représentation réductible $\bar{5} \oplus 10$ est la représentation appropriée pour les champs de fermions : la $\bar{5}$ se décompose en représentations $(\bar{3}, 1) \oplus (1, 2)$ et contient les antiquarks \bar{d}_L et les leptons gauches e_L^- et ν_e ; la 10 se décompose en $(1, 1) \oplus (3, 2) \oplus (\bar{3}, 1)$ contenant le lepton gauche e_L^+ singlet de $SU(2)$ et de $SU(3)$, les deux quarks gauches u_L, d_L qui forment un doublet de $SU(2)$ et les antiquarks \bar{u}_L .

De même, les 24 champs de jauge incorporent les 8 champs de gluons, les 3+1 vecteurs du secteur électrofaible, plus 12 champs supplémentaires, qui acquièrent une masse très grande lors de la brisure attendue de $SU(5) \rightarrow SU(3) \times SU(2) \times U(1)$.

La brisure $SU(5) \rightarrow SU(3) \times SU(2) \times U(1)$ devrait intervenir à une énergie de grand-unification de l'ordre de 10^{15} ou 10^{16} GeV, énergie à laquelle les couplages g_3, g_2, g_1 de $SU(3), SU(2)$ et $U(1)$ semblent converger (Fig. 5.2). Les générateurs infinitésimaux étant maintenant rigidement liés au sein du groupe simple $SU(5)$, on peut relier la charge électrique et le couplage au champ de jauge de $SU(2)$ et prédire l'angle de Weinberg : on trouve que $\sin^2 \theta = \frac{3}{8}, \dots$ mais ce calcul s'applique à l'énergie d'unification ! L'angle est renormalisé entre cette énergie et les énergies de la physique actuelle.

Une conséquence frappante de l'unification quarks-leptons au sein de multiplets est la violation des conservations séparées des nombres leptonique et baryonique. En particulier, l'existence de termes d'interaction, par exemple $X^\rho (\bar{d}\gamma_\rho e^+ + \bar{u}^c \gamma_\rho u)$, avec un des champs de jauge supplémentaires (les matrices des générateurs sont omises), permet la désintégration du proton $p = d u u \rightarrow d \bar{d} e^+ = \pi^0 e^+$, et par d'autres canaux encore. Il faut donc calculer soigneusement si le taux de désintégration est compatible avec les données expérimentales sur la durée de vie du proton (borne actuelle $10^{32 \pm 1}$ ans), ... ce qui n'est pas le cas !

Il faudrait encore montrer dans quelle représentation se placent les champs bosoniques de Higgs pour

permettre la brisure en deux étapes $SU(5) \rightarrow SU(3) \times SU(2) \times U(1) \rightarrow SU(3) \times U(1)$ à deux échelles extrêmement différentes.

Au final, la GUT $SU(5)$

- incorpore par construction la structure des générations de fermions ;
- place dans une même représentation leptons et quarks et explique donc la commensurabilité de leurs charges électriques et la compensation des anomalies (voir § suivant) ;
- réduit le nombre de paramètres du modèle standard et prédit la valeur de l'angle de Weinberg (à l'échelle d'unification) ;
mais à l'inverse
- elle n'explique pas le pourquoi des trois générations observées ;
- elle n'éclaire pas la question du "naturel" (que nous avons juste évoquée) ni celle reliée de la "hiérarchie" (pourquoi le rapport M_{GUT}/M_W est-il aussi grand ?) ;
- enfin, défaut fatal, elle prédit des effets tels la désintégration du proton à des taux qui semblent incompatibles avec les observations.

C'est ce dernier point qui a conduit à abandonner ce schéma d'unification et à lui préférer des voies supersymétriques.

5.4.3 Anomalies

On a mentionné au chapitre 4 l'existence des anomalies chirales, affectant le courant axial $J_\mu^{(5)}$ de la symétrie classique $U(1)$. Dans la théorie de jauge du Modèle Standard, les champs de jauge électro-faibles sont couplés de façon différente aux fermions gauches et droits, autrement dit, ils sont couplés aux courants axiaux, cf le lagrangien

$$\mathcal{L} = i\bar{\psi}(\not{\partial} - A)\frac{(1 - \gamma_5)}{2}\psi$$

qui contient un terme $A_a^\mu J_{\mu a}$ avec $J_{\mu a} = \bar{\psi}T_a\frac{(1 - \gamma_5)}{2}\psi$. Classiquement ce courant $J_{\mu a}$ devrait avoir une dérivée covariante (dans la représentation adjointe) nulle si la masse des fermions s'annule. On peut à nouveau effectuer le calcul de la divergence (covariante) de ce courant à l'ordre à une boucle, et on trouve que

$$D_\mu J^\mu = \frac{i}{24\pi^2}\partial_\mu\epsilon^{\mu\nu\rho\sigma}\text{tr}T_a(A_\nu\partial_\rho A_\sigma + \frac{1}{2}A_\nu A_\rho A_\sigma) .$$

Curieusement le membre de droite n'est pas invariant de jauge (mais sa forme ne doit rien au hasard et est dictée par des considérations géométriques que nous ne discuterons pas). L'anomalie de ce courant "non-singlet" (c'est-à-dire portant une représentation non triviale du groupe de jauge) brise donc l'invariance de jauge. Ce faisant elle met en danger toute la cohérence, renormalisabilité et unitarité, de la construction de la théorie. On conçoit que le contrôle de cette anomalie soit crucial pour la construction d'une théorie physique.

Or on constate que le coefficient "groupiste" de l'anomalie est proportionnel à

$$d_{abc} = \text{tr}(T_a\{T_b, T_c\})$$

où $\{T_b, T_c\}$ est l'anticommutateur des générateurs infinitésimaux, cf Exercice B.3.

En pratique on s'assure de l'annulation de l'anomalie dans deux cas :

- (a) Supposons que les fermions appartiennent tous à des représentations réelles ou pseudoréelles. On rappelle (cf chap 2) que l'on désigne ainsi les situations où la représentation est équivalente à sa représentation conjuguée, $T_a^* = CT_a C^{-1}$. On s'est placé dans des représentations unitaires où les T_a sont antihermitiques, $T_a = -T_a^\dagger = -T_a^{T*}$. On vérifie alors (cf Exercice B.3) que le coefficient groupiste $d_{abc} = -d_{abc} = 0$ s'annule et l'anomalie avec lui. Ainsi les théories (quadridimensionnelles) de groupe $SU(2)$ (dont les représentations sont réelles ou pseudoréelles) n'ont pas d'anomalie.
- (b) Une autre situation est celle où il y a compensation des anomalies venant des différentes représentations portées par les fermions. C'est ce qui se passe dans le modèle standard. Selon l'argument du a),

il n'y a pas d'anomalie associée aux seuls courants d'isospin faible, couplés au champ de jauge SU(2). Mais il peut a priori y en avoir avec les courants d'hypercharge faible (groupe U(1)), ainsi que des anomalies mixtes, par exemple 1 courant U(1) et deux SU(2) etc. Il faut donc vérifier que pour tous les choix de trois générateurs indexés par a, b, c , la constante d_{abc} s'annule quand on somme sur toutes les représentations de fermions. En définitive, on montre que tout se réduit à l'annulation de $\text{tr}(t_3^2 Q)$ pour chaque génération, qui est bien satisfaite. C'est encore ce qui se passe pour la théorie SU(5) discutée au § précédent : on montre que pour chaque génération, les contributions des représentations $\bar{5}$ et 10 se compensent.

★

Références additionnelles pour le Chapitre 5

Sur les aspects géométriques des théories de jauge et une introduction à la théorie des fibrés, voir par exemple M. Daniel et C. Viallet, *The geometric setting of gauge theories of the Yang-Mills type*, *Rev. Mod. Phys.* **52** (1980) 175-197.

Sur les théories de jauge, Yang-Mills, le modèle standard, etc, on peut consulter tout livre de théorie quantique des champs postérieur à 1975, par exemple [IZ], [PS], [Wf], [Z-J].

Sur les aspects “groupistes” des théories de jauge, voir L. O’Raifeartaigh, *op. cit.*.

Une très bonne revue de la grand-unification est donnée dans *Introduction to unified theories of weak, electromagnetic and strong interactions - SU(5)*, A. Billoire et A. Morel, rapport Saclay DPh-T/80/068, disponible sur le site du Master ICFP.

Pour une revue détaillée du Modèle Standard et une compilation de toutes les propriétés connues des particules élémentaires, voir *The Review of Particle Physics*, sur <http://pdg.lbl.gov/> déjà cité au Chap. 4.

★

Exercices et Problèmes pour le chapitre 5

A. Champ de jauge non abélien

1. Compléter les démonstrations de (5.21) et de (5.22).
2. Pour un champ de jauge non abélien A , soit F son tenseur de champ. Montrer que la dérivée covariante de F est telle que

$$D_{\mu a}^b F_{\nu\rho} t^a = [D_{\mu}, F_{\nu\rho}] = \partial_{\mu} F_{\nu\rho} - [A_{\mu}, F_{\nu\rho}] .$$

Démontrer alors l'identité

$$[D_{\mu}, F_{\nu\rho}] + [D_{\nu}, F_{\rho\mu}] + [D_{\rho}, F_{\mu\nu}] = 0 .$$

Rappeler quelle est la version abélienne de cette identité et son interprétation.

3. Soit l'opérateur $\mathcal{D} = \not{\partial} - \not{A}$ agissant sur des fermions de Dirac dans la représentation R . On veut calculer \mathcal{D}^2 . En écrivant $D_{\mu} D_{\nu} \gamma^{\mu} \gamma^{\nu} = \frac{1}{2} D_{\mu} D_{\nu} \{\gamma^{\mu}, \gamma^{\nu}\} + \frac{1}{2} [D_{\mu}, D_{\nu}] \gamma^{\mu} \gamma^{\nu}$, montrer qu'on peut écrire \mathcal{D}^2 comme somme de $D^2 = D_{\mu} D^{\mu}$ et d'un terme de la forme $a F_{\mu\nu} \sigma^{\mu\nu}$, où $\sigma^{\mu\nu} = \frac{i}{2} [\gamma^{\mu}, \gamma^{\nu}]$. Calculer a .

B. Facteurs groupistes...

1. Opérateurs de Casimir

Soient G un groupe de Lie simple et compact de dimension d , R une de ses représentations, que l'on suppose irréductible et unitaire. Soient t_a une base de l'algèbre de Lie \mathfrak{g} de G , T_a ses représentants dans la représentation R . Les t_a et T_a sont choisis antihermitiens. On considère alors la forme bilinéaire sur l'algèbre de Lie définie par

$$(X, Y)^{(R)} = \text{tr} (T_a T_b) x^a y^b$$

si $X = x^a t_a$ et $Y = y^b t_b \in \mathfrak{g}$ (avec sommation sur les indices répétés).

- (a) Démontrer que cette forme est invariante en ce sens que

$$\forall Z \in \mathfrak{g} \quad ([X, Z], Y)^{(R)} + (X, [Y, Z])^{(R)} = 0 .$$

On rappelle que toute forme bilinéaire invariante sur une algèbre de Lie simple est proportionnelle à la forme de Killing.

- (b) Démontrer que l'on peut choisir une base des t_a et donc des T_a telle que

$$\text{tr} (T_a T_b) = -T_R \delta_{ab}$$

avec T_R un coefficient dépendant de la représentation.

- (c) Quel est le signe de T_R ?
- (d) On considère alors l'opérateur de Casimir quadratique

$$C_2^{(R)} = - \sum_{\alpha} (T_{\alpha})^2 .$$

Sur combien de valeurs de a somme-t-on dans cette expression ?

- (e) Rappeler pourquoi $C_2^{(R)}$ est un multiple de l'identité dans l'espace de représentation de R

$$C_2^{(R)} = c_2(R) \mathcal{I} .$$

- (f) En quoi les hypothèses de simplicité de G et d'irréductibilité de R sont-elles importantes pour ce résultat ?
- (g) Quel est le signe de $c_2(R)$? Justifier.

- (h) Montrer que T_R est relié à la valeur de l'opérateur de Casimir quadratique $c_2(R)$. Pour cela, on pourra calculer de deux façons différentes la quantité

$$\text{tr} \sum_a (T_a)^2 .$$

- (i) À quoi se réduit cette relation pour la représentation adjointe de G ?
 (j) On normalise les générateurs (antihermitiens) de $SU(N)$ à être tels que dans la représentation de définition $\text{tr} T_a T_b = -\frac{1}{2} \delta_{ab}$, soit $T_f = \frac{1}{2}$. Cela est-il bien vérifié par les générateurs infinitésimaux $i \frac{\sigma_a}{2}$ de $SU(2)$? Quelle est alors la valeur de c_2 dans cette représentation de définition ?

2. Calculs de traces et de Casimir dans les représentations de $SU(N)$

- (a) Montrer que l'expression (3.50) du chap. 3, $c_2(\Lambda) = \frac{1}{2} \langle \Lambda, \Lambda + 2\rho \rangle$, se réécrit $c_2(\Lambda) = \frac{1}{2} (\langle \Lambda + \rho, \Lambda + \rho \rangle - \langle \rho, \rho \rangle)$, soit pour $SU(N)$, en utilisant les expressions (3.48) et (3.61) du Chap. 3

$$c_2(\Lambda) = \frac{1}{2N} \sum_{i=1}^{N-1} \left([(\lambda_i + 1)^2 - 1] i(N - i) + 2 \sum_{j=i+1}^{N-1} [(\lambda_i + 1)(\lambda_j + 1) - 1] i(N - j) \right) .$$

- (b) Calculer cette expression pour la représentation de définition. La valeur obtenue est-elle en accord avec celle trouvée à la question 1.(j) ci-dessus ?
 (c) Rappeler pourquoi le poids le plus haut de la représentation adjointe est la plus haute racine (notée θ dans l'appendice F du chap 3). Pourquoi l'expression $\theta = \Lambda_1 + \Lambda_{N-1}$ est-elle en accord avec ce qu'on sait de la représentation adjointe ?
 (d) Calculer la valeur de $c_2(\Lambda)$ pour la représentation adjointe.
 (e) Vérifier cette valeur pour $SU(2)$ par le calcul direct de $c_2(\text{adj})$.
 (f) Quelle est la valeur de T_{adj} qu'on en déduit, au vu de la question 1.(i) ?

3. Coefficients des anomalies

Avec les mêmes notations et conventions que précédemment,

- (a) Dans le calcul de certains diagrammes de Feynman dans une théorie de jauge sur le groupe G , on rencontre le coefficient

$$d_{\alpha\beta\gamma} = \text{tr} (T_\alpha (T_\beta T_\gamma + T_\gamma T_\beta)) .$$

Montrer que $d_{\alpha\beta\gamma}$ est complètement symétrique dans ses trois indices.

- (b) On rappelle que la représentation est dite réelle ou *pseudoréelle* si elle est (unitairement) équivalente à sa conjuguée, donc si dans la base où les T_α sont antihermitiens, on peut trouver une matrice unitaire U telle que le complexe conjugué de chaque T_α vérifie

$$(T_\alpha)^* = U T_\alpha U^{-1} .$$

Montrer que si cette condition est satisfaite, $d_{\alpha\beta\gamma}$ est identiquement nul. Cette condition est importante pour assurer la cohérence de la théorie de jauge, c'est la condition de compensation des anomalies.

- (c) La représentation de spin $\frac{1}{2}$ de $SU(2)$ est-elle pseudoréelle ? Celle de spin j ? Justifier votre réponse.
 (d) Donner deux exemples de représentations (pas nécessairement irréductibles) non triviales de $SU(3)$ qui sont pseudoréelles, et deux qui ne le sont pas.
 (e) Que vaut le coefficient d pour le groupe $U(1)$ et une représentation de charge q ?

C. Brisure spontanée de SU(2)

On considère une théorie de jauge de groupe SU(2) couplée à un champ de bosons $\vec{\Phi}$ de spin 1, considéré comme un vecteur de dimension 3. On note $V(\vec{\Phi}^2)$ le potentiel de ce champ.

1. Écrire le lagrangien et les transformations de jauge des champs \vec{A}_μ et $\vec{\Phi}$.
2. On suppose qu'il y a brisure spontanée de symétrie : le champ Φ acquiert une vev v selon une direction, par exemple 3 : $\langle \vec{\Phi} \rangle = \begin{pmatrix} 0 \\ 0 \\ v \end{pmatrix}$. Quel est le groupe de symétrie résiduel ? Quel sera l'effet sur le champ A_μ ? Décrire les champs et particules physiques après brisure de symétrie.

Problème I. Théories de jauge sur réseau

Dans tout ce problème, G désigne un groupe de Lie compact, $\chi^{(\rho)}$ le caractère de sa représentation irréductible unitaire ρ .

1. Montrer que les relations d'orthogonalité des $\mathcal{D}^{(\rho)}$ impliquent les formules suivantes :

$$\int_G \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g \cdot g_1 \cdot g^{-1} \cdot g_2) = \frac{1}{n_\rho} \chi^{(\rho)}(g_1) \chi^{(\rho)}(g_2), \quad (5.41)$$

et

$$\int_G \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g \cdot g_1) \chi^{(\sigma)}(g^{-1} \cdot g_2) = \frac{\delta_{\rho, \sigma}}{n_\rho} \chi^{(\rho)}(g_1 \cdot g_2). \quad (5.42)$$

Rappeler pourquoi une représentation de G peut toujours être considérée comme unitaire et montrer qu'alors

$$\chi^{(\rho)}(g^{-1}) = \chi^{(\bar{\rho})}(g) = (\chi^{(\rho)}(g))^*, \quad (5.43)$$

où $\bar{\rho}$ est la représentation conjuguée de ρ . On fera un usage fréquent de ces trois relations dans la suite.

2. Soit χ le caractère d'une représentation réelle r (pas nécessairement irréductible) de G , β un paramètre réel.
 - a) Montrer que l'on peut développer $\exp \beta \chi(g)$ sur les caractères des représentations irréductibles de G selon

$$e^{\beta \chi(g)} = \sum_{\rho} n_{\rho} b_{\rho} \chi^{(\rho)}(g),$$

avec des fonctions $b_{\rho}(\beta)$. Exprimer la fonction $b_{\rho}(\beta)$ à l'aide d'une intégrale sur le groupe. En utilisant (5.43), montrer que les fonctions $b_{\rho}(\beta)$ sont réelles, $b_{\rho}(\beta) = (b_{\bar{\rho}}(\beta))^* = b_{\bar{\rho}}(\beta)$.

b) Montrer que b_{ρ} est non nul pourvu que la représentation ρ apparaisse dans une puissance tensorielle $r^{\otimes n}$.

c) Pour $G = \text{SU}(2)$ et $r = (j = \frac{1}{2})$, la représentation de spin $\frac{1}{2}$, la condition du **b)** est-elle satisfaite pour tout ρ ? Pourquoi ?

Si $r = (j = 1)$, quelles sont les représentations pour lesquelles b_{ρ} est a priori nul ?

d) Pour $G = \text{SU}(3)$ et $\chi = \chi^{(3)} + \chi^{(\bar{3})}$, montrer que b_{ρ} est non nul pour tout ρ .

Pour $\beta \rightarrow 0$, quel est le comportement dominant de $b_a(\beta)$ si a désigne la représentation adjointe de SU(3) ? Plus généralement quel est le comportement dominant de $b_{\rho}(\beta)$ où ρ est la représentation de plus haut poids $\Lambda = (\lambda_1, \lambda_2)$?

3. On définit un modèle de mécanique statistique à d dimensions de la façon suivante. Sur un réseau hypercubique de dimension d et de maille a , les degrés de liberté sont attachés aux liens entre sites voisins et prennent leur valeur dans le groupe compact G . A chaque lien orienté $\ell = \vec{ij}$ on associe l'élément de

G noté $g_\ell = g_{ij}$, à $-\ell = \vec{j}i$, on associe $g_{ji} = g_\ell^{-1}$. A chaque carré élémentaire (ou “plaquette”) $p = ijkl$, on associe le produit des éléments des liens :

$$g_p = g_{ij} \cdot g_{jk} \cdot g_{kl} \cdot g_{li}$$

et l’“énergie” d’une configuration de ces variables est donnée par

$$E = - \sum_{\text{plaquettes } p} \chi(g_p) \quad (5.44)$$

où χ est, comme à la question 2, le caractère d’une certaine représentation réelle du groupe. Le poids de Boltzmann est donc

$$e^{-\beta E} = \prod_p e^{\beta \chi(g_p)} \quad , \quad \beta = \frac{1}{kT}$$

et la fonction de partition s’écrit

$$Z = \prod_{\text{links } \ell} \int_G \frac{d\mu(g_\ell)}{v(G)} \prod_{\text{plaquettes}} e^{\beta \chi(g_p)} . \quad (5.45)$$

a) Montrer que l’énergie E est invariante par redéfinition des g_{ij} selon $g_{ij} \mapsto g_i \cdot g_{ij} \cdot g_j^{-1}$, où $g_i \in G$, (c’est une invariance *locale*, l’analogie dans ce formalisme discret de l’invariance de jauge étudiée dans ce chapitre), et que E ne dépend pas de l’orientation des plaquettes.

b) On cherche à mieux comprendre la relation avec le formalisme du § 5.1. Les degrés de liberté g_{ij} représentent les variables de chemin définies en (5.20), $g_{ij} = g(j, i)$ le long de l’arête du site i au site j

$$g_{ij} \equiv P \exp \int_{l=i\vec{j}} A_\mu dx^\mu$$

— Pour une maille a du réseau petite, montrer en utilisant par exemple la formule de BCH et en développant au premier ordre non nul que

$$g_p = \exp (a^2 F_{\mu\nu} + o(a^2))$$

où μ et ν désignent les directions du bord de la plaquette p . (On s’intéresse ici à une version euclidienne de la théorie de jauge, et la position des indices μ, ν n’importe pas.) Montrer alors que l’énergie E_p (5.44) s’écrit

$$E_p \sim \text{const. } a^4 (F_{\mu\nu})^2 + \text{const.}'$$

où on déterminera la première constante en fonction de la représentation choisie pour χ .

— Expliquer pourquoi le paramètre β s’identifie (à un facteur près) à l’inverse du couplage g^2 de la théorie de jauge continue. En fait il s’agit plutôt de la constante de couplage “nue” (ou non renormalisée), pourquoi ?

On se restreint d’abord pour simplicité à $d = 2$ dimensions. Pour un réseau fini de \mathcal{N} plaquettes, par exemple un rectangle de taille $L_1 \times L_2$ (voir Fig. 5.3), on désire calculer Z . On choisit des “conditions aux bords libres”, autrement dit les variables g_ℓ des bords du rectangle sont indépendantes. On s’intéresse aussi à la valeur moyenne $W^{(\sigma)}(C)$ de $\chi^{(\sigma)}(g_C)$ où g_C désigne le produit ordonné des g_ℓ le long d’une courbe fermée orientée C pour une certaine représentation irréductible σ de G

$$W^{(\sigma)}(C) := \langle \chi^{(\sigma)}(g_C) \rangle = \frac{1}{Z} \prod_{\text{liens } \ell} \int_G \frac{d\mu(g_\ell)}{v(G)} \chi^{(\sigma)} \left(\prod_{\ell \in C} g_\ell \right) \prod_p e^{\beta \chi(g_p)} . \quad (5.46)$$

c) En utilisant les résultats de la question 2, montrer que l’on peut développer chaque $\exp \beta \chi(g_p)$ sur les caractères de représentations irréductibles de G selon

$$e^{\beta \chi(g_p)} = \sum_\rho n_\rho b_\rho \chi^{(\rho)}(g_p) . \quad (5.47)$$

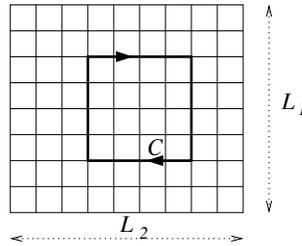


FIGURE 5.3 – Réseau carré à 2 dimensions.

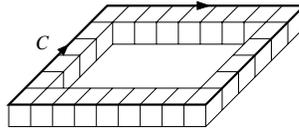


FIGURE 5.4 – Une configuration tubulaire contribuant à la boucle de Wilson.

d) On insère dans (5.45) ou (5.46) le développement (5.47) pour chaque plaquette. Montrer que si deux plaquettes ont en commun un lien ℓ , les formules de la question 1 permettent d'intégrer sur la variable g_ℓ de ce lien et que les deux représentations portées par les plaquettes s'identifient alors. En utilisant alors de façon répétée ces formules de la question 1, montrer que l'on peut intégrer sur toutes les variables g_ℓ et que

$$Z = b_1^N \quad W^{(\sigma)}(C) = n_\sigma \left(\frac{b_\sigma}{b_1} \right)^A \tag{5.48}$$

où A désigne l'aire de la courbe C , c'est-à-dire le nombre de plaquettes qu'elle encadre, et l'indice 1 se rapporte à la représentation identité.

e) On se place maintenant en dimension $d = 3$. Les variables g_ℓ sont attachées aux liens d'un réseau cubique. L'énergie est toujours donnée par (5.44), où la somme court sur toutes les plaquettes de ce réseau tridimensionnel. Comme précédemment, $W^{(\sigma)}(C) = \langle \chi^{(\sigma)}(g_C) \rangle$ reçoit des contributions de configurations de plaquettes formant une surface de bord C .

On va voir que peuvent contribuer aussi à $W^{(\sigma)}(C)$ des configurations de plaquettes formant un tube qui s'appuie sur le contour C (Fig. 5.4).

– Montrer en effet que pour une telle configuration l'application répétée des formules (5.41) et (5.42) sur toutes les variables g_ℓ conduit à l'expression suivante

$$W^{(\sigma)}(C)|_{\text{tube}} = \sum_{\rho} \left(\frac{b_\rho}{b_1} \right)^P \int_G \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g) \chi^{(\rho)}(g^{-1}) \chi^{(\sigma)}(g) \tag{5.49}$$

où P est le nombre de plaquettes constituant le tube.

– À quelle condition \mathcal{C} sur la représentation σ de la boucle C la contribution de la représentation ρ au membre de droite de (5.49) est-elle non nulle ?

– Donner un exemple pour $G = \text{SU}(2)$ de représentations σ pour laquelle cette condition \mathcal{C} n'est *jamais* satisfaite quelle que soit ρ , et donc ces configurations tubulaires absentes.

– À l'inverse donner un exemple (toujours pour $\text{SU}(2)$) d'un choix possible de σ qui la satisfait.

On admettra qu'à haute température (β petit), la contribution dominante à $W^{(\sigma)}(C)$ est du type (5.49) si la condition \mathcal{C} peut être satisfaite, et du type (5.48) dans le cas contraire.

4. L'évaluation de la valeur moyenne de la "boucle de Wilson" $W^{(\sigma)}(C)$ dans la limite d'une grande boucle C ayant la forme d'un rectangle $R \times T$ permet de calculer le potentiel $V_\sigma(R)$ entre deux particules chargées statiques séparées par la distance R , l'une portant la représentation σ du groupe et l'autre

étant son antiparticule. Plus précisément on admettra que

$$V_\sigma(R) = - \lim_{T \rightarrow \infty} \frac{1}{T} \log W^{(\sigma)}(C) .$$

Évaluer la dépendance de $V_\sigma(R)$ en R qui découle soit de (5.48), soit de la contribution à (5.49) due à la représentation ρ . Qu'en concluez-vous sur l'interaction entre les deux particules dans ces deux situations ?

Physiquement, ce type de considérations fournit un modèle discrétisé (sur réseau) et très simplifié ici (deux ou trois dimensions, pas de quarks) de la chromodynamique quantique. On peut répéter ce calcul en dimension plus élevée, où les résultats ci-dessus apparaissent comme le terme dominant dans un développement à β petit ("haute température"). Le fait que la valeur moyenne ci-dessus décroisse comme x^A ($x = b_\sigma/b_1 < 1$ pour β assez petit) pour de grandes aires est un signal du "confinement des quarks" dans cette théorie, c'est-à-dire de l'impossibilité de séparer une paire quark-antiquark à grande distance ...

Problème II. Mécanisme de BEH

I. Modèle de Georgi–Glashow

Dans un article de 1972, H. Georgi et S. Glashow proposaient un modèle d'interactions électrofaibles basé sur le groupe de jauge $SO(3)$ avec un champ de Higgs se transformant comme un triplet sous ce groupe.

1. Combien de champs de jauge possède un tel modèle ?
2. Le triplet de Higgs $\Phi = (\phi^+, \phi^0, \phi^-)$ est supposé acquérir une "vev"

$$\langle \Phi \rangle = v(0, 1, 0) .$$

Quel est le groupe H de symétrie résiduel ?

3. Que peut-on dire du spectre de masse de la théorie, après la brisure de symétrie $SO(3) \rightarrow H$? Quelle est son interprétation physique ?
4. Quelle est la différence majeure entre ce modèle et ce qui est devenu le modèle standard, en ce qui concerne les interactions faibles ? Pouvez-vous citer une découverte expérimentale qui a permis d'écarter rapidement ce modèle ?

II. Groupe de jauge $SU(n)$

On considère maintenant une théorie de jauge basée sur le groupe $SU(n)$, avec des champs de jauge couplés à un champ scalaire Φ .

1. Que peut-on dire du groupe de symétrie résiduel et des masses des champs vectoriels quand
 - (a) le champ scalaire se transforme selon une représentation fondamentale de dimension n et $\langle \Phi \rangle = v(0, 0, \dots, 0, 1)$?
 - (b) le champ scalaire se transforme selon la représentation adjointe et $\langle \Phi \rangle = v \text{diag}(1, 1, \dots, 1, -n+1)$?
2. On introduit un champ Ψ de fermions se transformant aussi comme la représentation fondamentale de dimension n (ou sa conjuguée). Quels termes de masse invariants sont possibles pour les fermions ?
3. Supposant que le champ scalaire se transforme selon la représentation adjointe,
 - (a) combien y a-t-il de couplages de type Yukawa (c'est-à-dire en $\bar{\Psi}\Psi\Phi$) invariants indépendants ?
 - (b) Écrire les couplages possibles entre ce multiplet de fermions et le champ scalaire.
 - (c) Quels termes de masse additionnels pour les fermions résultent de la brisure de symétrie envisagée au 1. ?

★

Index

- action d'un groupe dans un ensemble, 61
- adjointe, application, 45
- adjointe, représentation, 76, 99, 105, 116, 118, 123, 127, 135
- algèbre, 42
- algèbre de Lie, 42
- algèbres de Lie de dimension 3, 61
- α -chaîne, 110
- alterné, groupe, 34
- anomalies, 143, 178, 180
- axial, courant, 142
- axiale, transformation, 142

- Baker-Campbell-Hausdorff, formule de, 46
- beauté, 153
- boost, 21
- boson intermédiaire, 168
- Bratteli, diagramme de, 101
- brisure spontanée de symétrie, 140, 169
- Brout-Englert-Higgs, mécanisme de, 169
- Burgoyne, identité de, 156

- Cabibbo, angle de, 153
- Cabibbo-Kobayashi-Maskawa, matrice de, 174
- caractère, 67
- caractère de $SU(2)$, 92
- Cartan, critères de, 52
- Cartan, matrice de, 111
- Cartan, sous-algèbre de, 105
- Cartan, tore de, 119
- Casimir, opérateur de, 8, 26, 53, 61, 100, 115, 118, 179
- centralisateur, 34
- centre d'un groupe, 33
- chaîne de poids, 116
- chaîne de racines, 110

- charge baryonique, 20, 144, 176
- charme, 153
- Chevalley, base de, 114
- chirale, symétrie, 143
- classe d'homotopie, 37
- classe de conjugaison, 33
- classe par rapport à un sous-groupe, 34
- Clebsch-Gordan, coefficients de, 16, 74
- Clebsch-Gordan, décomposition de, 15, 72
- cocycle, 84
- coefficients de Clebsch-Gordan, 16
- cohomologie d'un groupe, 84
- commutant, 34
- commutateur dans un groupe, 45
- compact, 21, 40, 57
- compact, localement, 57
- compacte, algèbre de Lie, 49
- compensation des anomalies, 178, 180
- complètement réductible, représentation, 67
- complexe, représentation, 101
- complexifiée, algèbre de Lie, 49
- confinement de la couleur, 155, 171, 184
- conforme, transformation, 62
- conjuguée, représentation, 68, 119
- connexe, groupe, 37
- connexe, simplement, 37
- conservation, loi de, 86
- constantes de structure, 50
- contragrédiente, représentation, 69, 102
- coracine, 115
- "coset", 34
- couleur, 154, 170
- couplage minimal, 162
- covariante, dérivée, 162
- Coxeter, exposants de, 115

- Coxeter, groupe de, 110
 Coxeter, nombre de, 115
 critères de Cartan, 52
 crochet de Lie, 42
 cyclique, groupe , 31
- décomposition de Clebsch-Gordan, 15, 72
 dérivée covariante, 162, 163
 diagramme de Dynkin, 113
 diagramme de poids d'une représentation, 116
 difféomorphismes du cercle, 48
 dimension d'un groupe, 42
 direct, produit – de groupes, 36
 dual d'un quaternion, 55
 dual, espace, 107
 Dynkin, diagramme de, 113
 Dynkin, indices de, 117
- écart de masse électromagnétique, 150, 157
 électrodynamique, 161
 entrelacement, opérateur d', 66, 70
 enveloppante, algèbre, 53
 $\epsilon^{\mu\nu\rho\sigma}$ tenseur, 25
 équivalente, représentation, 66
 esclavage infrarouge, 171
 espace des racines, 107
 espace tangent, 43, 57
 étrangeté, 144
 Euler, angles d', 2, 13
 exponentielle, application, 44, 58
 extension centrale, 48, 89
- facteur de forme électromagnétique, 149
 Fermi, constante de, 152
 Fermi, lagrangien de, 152, 168
 fibré, 165
 fidèle, représentation, 66
 fonction de classe, 67, 80
 fondamental, groupe, 38
 formule d'Olinde Rodrigues, 1
 Freudenthal, formules de , 119
 Freudenthal–de Vries, formule étrange de , 119
- Frobenius–Schur, indicatrice de, 102
 Frobenius–Weyl, dualité de, 129
- Galilée, groupe de, 32, 40, 42
 Gell-Mann–Nishima, relation de, 144
 Gell-Mann–Okubo, formule de masse, 150, 156
 générateurs infinitésimaux, 4, 6
 générations, 155
 Glashow–Salam–Weinberg, modèle de, 172
 Goldstone, théorème de, 141
 grand-unifiée, théorie, 175
 groupe compact, 40
 groupe connexe, 37
 groupe cyclique, 31
 groupe d'homotopie, 38
 groupe de Lie, 41
 groupe du tétraèdre, 82
 groupe fondamental, 38
 groupe non compact, représentations, 79
 groupe quotient, 35
 groupe semi-simple, 36
 groupe simple, 36
 groupe symplectique, 32, 56
 groupe topologique, 36
 GUT, 175
- Haar, mesure de, 41, 59
 hadron, 19
 harmoniques sphériques, 94, 95, 99
 Higgs, boson de, 173
 homéomorphisme, 38
 homogène, espace, 61
 homomorphisme, 34
 homotopie, 37
 hypercharge, 144
 hypercharge faible, 172
- idéal, 48
 identité de Jacobi, 6, 43, 51
 indécomposable, représentation, 67
 indice d'un sous-groupe dans un groupe, 35
 invariance de jauge, 161

- invariances d'un système quantique, 86
- invariant de jauge, 165
- invariant, sous-groupe, 35
- irréductible, représentation, 13, 67
- $ISL(2, \mathbb{C})$, 23
- isospin, 19, 144
- isospin faible, 172
- isotropie, groupe d', 61

- Jacobi, identité de, 6, 43, 51
- Jacobi, polynômes de, 94
- jauge, invariance de, 161

- Kac, indices de, 115
- Killing, forme de, 51
- Kobayashi-Maskawa, matrice de, 174

- lacet, 37
- “ladder operators”, 110
- Lagrange, théorème de, 35
- laplacien, 27
- Legendre, polynômes et fonctions de, 94
- lemme de réarrangement, 40
- lemme de Schur, 70
- liberté asymptotique, 170
- Lie, algèbre de, 42
- Lie, groupe de, 41
- Littlewood-Richardson, règles de, 123
- loi de conservation, 86
- Lorentz, groupe de, 21
- Lorentz, transformation spéciale, 21

- mésons π , 19
- Maschke, théorème de, 69
- masses, matrice de, 174
- matrice de Cartan, 111
- matrice de masses, 174
- mesure de Haar, 41, 59
- mesure invariante, 40, 59
- modèle $O(n)$, 47
- modèle standard, 170
- moment magnétique, 149
- moment multipolaire, 97

- Montgomery et Zippen, théorème de, 42
- multipolaire, développement, 97

- Nambu–Goldstone, bosons de, 141
- Noether, courant de, 47, 87, 139, 140, 142, 153
- normal, sous-groupe, 35
- normalisateur, 34
- noyau, 34
- nucléon, 19

- ondes partielles, 98
- orbite d'un groupe, 61
- ordre d'un groupe fini, 31
- orthogonalité et complétude des caractères, 78
- orthogonalité et complétude des matrices \mathcal{D} ,
77

- Pauli, matrices de, 3
- Pauli-Lubanski, tenseur de, 26
- Peter–Weyl, théorème de, 80
- petit groupe, 61
- pions, 19
- poids d'une représentation, 116
- poids dominant, 117
- poids fondamental, 117
- poids le plus haut, 117
- Poincaré, algèbre de, 22
- Poincaré, groupe de, 21
- produit tensoriel de représentations, 15, 71
- produits tensoriels dans $SU(2)$, 101
- projective, représentation, 12, 83
- pseudoréelle, représentation, 68, 101, 177, 180
- pure jauge, 164

- QCD, 155
- quantité conservée, 86
- quark bottom, 153
- quark top, 153
- quarks, modèle des, 144
- quaternion dual, 55
- quaternionique, représentation, 101
- quaternions, 55
- quotient, groupe, 35

- Racah–Speiser, algorithme de, 124
 racine, 107
 racine la plus haute, 115
 racine positive, 111
 racine simple, 111
 rang d’une algèbre de Lie, 106
 rayon, 84
 recouvrement universel, 38
 réductible, représentation, 67
 réelle, représentation, 68, 101, 177, 180
 renormalisation de constant de couplage, 167
 représentation, 65
 représentation complexe, 101
 représentation d’un groupe de Lie compact, 77
 représentation d’un groupe non compact, 79
 représentation d’une algèbre de Lie, 75
 représentation projective, 12, 83
 représentation pseudoréelle, 68, 101, 177, 180
 représentation quaternionique, 101
 représentation réelle, 68, 101, 177, 180
 représentation spinorielle, 25
 représentations de $SO(1,3)$ et $SL(2,\mathbb{C})$, 21
 représentations de $SO(3)$, 12
 représentations de $su(2)$ et $SU(2)$, 9, 91
 représentations des groupes finis, 81
 réseau des poids, 118
 réseau des racines, 118
 résonance Δ , 20
 riemannienne, variété, 27, 57
 Rodrigues, formule d’Olinde -, 1
 rotations dans \mathbb{R}^3 , 1

 saveur, 153, 155
 Schur, lemme de, 70
 semi-direct, produit, 23
 semi-direct, produit – de groupes, 36
 semi-simple, algèbre de Lie, 48
 semi-simple, groupe, 36
 Serre, relations de, 114
 signature dans le groupe de Weyl, 111
 simple connexité, 37

 simple, algèbre de Lie, 48
 simple, groupe, 36
 simple, groupe de Lie, 48
 $SL(2,\mathbb{C})$, 23
 $SL(2,\mathbb{R})$, 32, 44, 56, 61
 $so(n)$, algèbre de Lie de $SO(n)$, 47
 $SO(1,3)$ et $SL(2,\mathbb{C})$, 21
 $SO(2,1)$, 61
 $SO(3)$, 1
 $so(4)$, algèbre de Lie, 49
 somme directe de représentation, 67
 sous-groupe invariant, 35
 sous-groupe normal, 35
 sphériques, coordonnées, 28
 sphériques, harmoniques, 94, 95, 99
 spineur, 25
 stabilisateur, 61
 standard, modèle, 170
 $SU(1,1)$, 61
 $SU(2)$, 3
 $su(2)$, algèbre de Lie, 7
 $SU(3)_c$ groupe de couleur, 154, 170
 $SU(3)$, groupe de saveur, 144
 $SU(4)$, 153
 $SU(5)$, grand-unification par, 175
 symétrie chirale, 143
 symétrie de jauge, 161
 symétrique, groupe, 31, 33, 34
 symboles 3- j , 17
 symboles 6- j , 18
 symplectique, groupe, 32, 56

 tableau d’Young, 128
 tableau d’Young standard, 128
 Tchebichev, polynômes de, 92, 99
 tenseur énergie–impulsion, 62
 tenseurs, 90
 tension de corde, 171
 théorème de Wigner, 84
 théorie de jauge sur réseau, 171, 181
 théorie grand-unifiée, 175

- transformation de Lorentz spéciale, 21
- “truth”, 153

- $U(n)$, 59
- unitaire, représentation, 69
- unitarisable, représentation, 69

- V-A, 153
- Vandermonde, déterminant de, 60, 129
- variété, 57
- variété riemannienne, 27, 57
- vectorielle, transformation, 142
- Virasoro, algèbre de, 48
- voisinage, 56
- vortex, 40

- Weinberg, angle de, 172, 177
- Weyl, groupe de, 110
- Weyl, vecteur de, 118
- Wigner, matrices de, 12, 14
- Wigner, théorème de, 84
- Wigner–Eckart, théorème de, 20, 87, 149, 150
- Wilson, boucle de, 165, 171, 183
- Witt, algèbre de, 48

- Yang–Mills, 162
- Yang–Mills, lagrangien de, 166
- Young, diagramme d’, 126, 128
- Young, tableau standard, 128
- Young, tableau d’, 128
- Yukawa, couplage de, 147, 174