

Invariances en physique et théorie des groupes

Jean-Bernard Zuber

Laboratoire de Physique Théorique et Hautes Énergies
Université Pierre et Marie Curie Paris 6
Tour 24, 5ème étage, Boîte 126, 4 place Jussieu, F-75252 Paris Cedex 05 France

Cours donné au M2/CFP/Parcours de Physique Théorique. Automne 2010.

Avertissement

On trouvera ci-dessous dans les chapitres intitulés 0 à 000 des notes couvrant essentiellement le contenu des cours de “préentré” donnés par Paul Windey et moi-même.

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

Les références les plus fréquemment mentionnées sont rassemblées ici.

- [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.
- [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.
- [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

Chapitre 0. Invariance relativiste et champs classiques	1
1. Principes de la Relativité restreinte. Conséquences physiques	1
1.1. Référentiels, synchronisation des horloges et simultanéité.	1
1.2. Postulats de la Relativité	2
1.3. Transformations de Lorentz.	2
1.4. Conséquences des transformations de Lorentz. Cinématique relativiste	5
2. Espace de Minkowski. Groupe de Lorentz $O(3,1)$.	9
2.1. Métrique de l'espace de Minkowski. Groupes de Lorentz et de Poincaré	9
2.2. Vecteurs et tenseurs contravariants et covariants.	11
2.3. Géométrie de l'espace de Minkowski, temps propre, lignes d'univers.	12
3. Formalisme lagrangien et relativité. Cinématique relativiste	14
3.1. Rappels : Principe de moindre action. Équations d'Euler-Lagrange. Formalisme hamiltonien, etc.	14
3.2. Particule libre relativiste	15
3.3. Cinématique relativiste : collisions et désintégration de particules	18
4. Théorie classique des champs. Équations de Maxwell et relativité	19
4.1. Théorie des champs lagrangienne	19
4.2. Symétries, lois de conservation. Courant de Noether.	21
4.3. Invariance de Lorentz des équations de Maxwell	28
4.4. Lagrangien de l'électrodynamique	30
5. Rotation du minkowskien vers l'euclidien	31
Bibliographie sommaire	32
Chapitre 00. Quelques éléments de base sur les groupes $SO(3)$, $SU(2)$ et $SL(2, \mathbb{C})$	33
1. Rotations de \mathbb{R}^3 , les groupes $SO(3)$ et $SU(2)$	33
1.1. Le groupe $SO(3)$, groupe à trois paramètres	33
1.2. Du groupe $SO(3)$ au groupe $SU(2)$	35
2. Générateurs infinitésimaux. L'algèbre de Lie $su(2)$	37
2.1. Générateurs infinitésimaux de $SO(3)$	37
2.2. Générateurs infinitésimaux dans $SU(2)$	39
2.3. Algèbre de Lie $su(2)$	40
3. Représentations de $SU(2)$	42
3.1. Représentations des groupes $SO(3)$ et $SU(2)$	42
3.2. Représentations de l'algèbre $su(2)$	43
3.3. Construction explicite	47
4. Produit direct de représentations de $SU(2)$	48
4.1. Produit direct de représentations et l'"addition de moments angulaires"	48
4.2. Coefficients de Clebsch-Gordan, symboles $3-j$ et $6-j$	50
5. Une application physique : l'isospin	53
6. Représentations de $SO(3,1)$ et $SL(2, \mathbb{C})$	55
6.1. Algèbre de Lie des groupes de Lorentz et Poincaré	55
6.2. Groupes de recouvrement de \mathcal{L}_+^\uparrow et \mathcal{P}_+^\uparrow	57
6.3. Représentations irréductibles de dimension finie de $SL(2, \mathbb{C})$	58

6.4. Représentations irréductibles unitaires du groupe de Poincaré. États à une particule.	60
Bibliographie	62
Problème	63
Appendice A. Calcul du laplacien sur les sphères S^2 et S^3	63
Chapitre 000. Équations de Klein-Gordon et Dirac	65
1. Rappels de Mécanique Quantique	65
1.1. Postulats	65
1.2. Principe de correspondance	66
2. Équation de Klein-Gordon	67
3. Équation de Dirac	69
3.1. Les matrices γ de Dirac	69
3.2. Une représentation explicite des matrices γ	71
3.3. L'algèbre de Clifford des matrices γ . La “ γ -gymnastique”.	71
3.4. Covariance relativiste. Représentation spinorielle. Parité	73
3.5. Combinaisons bilinéaires de ψ et $\bar{\psi}$	77
3.6. Représentations de Weyl et de Majorana.	79
3.7. Matrices de Dirac en métrique euclidienne	80
4. Contenu physique de l'équation de Dirac	81
4.1. Ondes planes et projecteurs	81
4.2. Couplage minimal au champ électromagnétique	84
4.3. Limite non relativiste	85
4.4. Spectre relativiste de l'atome d'hydrogène	87
5. Particules et trous. La mer de Dirac.	87
5.1. La mer de Dirac. Particules et trous, antiparticules.	87
5.2. Conjugaison de charge	87
6. Particules de masse nulle	89
6.1. Chiralité et hélicité	89
6.2. Spineurs de Weyl	90
6.3. Petite incursion dans le monde des neutrinos	90
Problème	91
Bibliographie sommaire	92
Chapitre 1. Groupes. Groupes et algèbres de Lie	
1. Généralités sur les groupes	93
1.1. Définitions de base et premiers exemples	93
1.2. Classes d'un groupe	95
1.3. Sous-groupes	95
1.4. Homomorphisme d'un groupe G dans un groupe G'	96
1.5. Classes par rapport à un sous-groupe	96
1.6. Sous-groupe invariant.	97
1.7. Groupe simple, groupe semi-simple	98
2. Groupes continus. Propriétés topologiques. Groupes de Lie.	98
2.1. Connexité	98
2.2. Simple connexité. Groupe d'homotopie. Recouvrement universel	99

2.3. Groupes compacts et non compacts	102
2.4. <i>Mesure invariante</i>	102
2.5. Groupes de Lie	103
3. Étude locale d'un groupe de Lie. Algèbre de Lie.	104
3.1. Algèbres et algèbres de Lie. Définitions	104
3.2. Espace tangent d'un groupe de Lie G	105
3.3. Relations entre l'espace tangent \mathfrak{g} et le groupe G	106
3.4. L'espace tangent comme algèbre de Lie	107
3.5. Un exemple explicite : l'algèbre de Lie de $SO(n)$	109
3.6. Un exemple de dimension infinie : l'algèbre de Virasoro	110
4. Des propriétés de l'algèbre de Lie à celles du groupe	110
4.1. Simplicité, semi-simplicité	110
4.2. Compacité. Complexifiée	112
4.3. Connexité, simple-connexité	112
4.4. Constantes de structure. Forme de Killing. Critères de Cartan	113
4.5. Opérateur(s) de Casimir	116
Bibliographie sommaire	117
Appendix A. Corps des quaternions et groupes symplectiques	118
A.1. Quaternions	118
A.2. Matrices de quaternions	119
A.3. Groupes symplectiques $Sp(2n, \mathbb{R})$ et $USp(n)$, algèbres de Lie $\mathfrak{sp}(2n)$ et $\mathfrak{usp}(n)$	119
Appendix B. Rappels et compléments de topologie et de géométrie différentielle	120
B.1. Petit lexique de quelques concepts de topologie utilisés dans ces notes	120
B.2. Notion de variété	120
B.3. Espace tangent	121
B.4. Groupe de Lie. Application exponentielle	121
Appendix C. Mesures invariantes sur $SU(2)$ et sur $U(n)$	122
Exercices pour le chapitre 1	124
Problème : Transformations conformes	125
Chapitre 2. Représentations linéaires des groupes	129
1. Définition. Représentations équivalentes. Représentations unitaires	129
1.1. Définitions de base	129
1.2. Représentations équivalentes. Caractères	130
1.3. Représentations réductibles et irréductibles	131
1.4. Lemme de Schur	135
1.5. Produit tensoriel de représentations ; décomposition de Clebsch-Gordan	136
1.6. Décomposition en représentations irréductibles d'un sous-groupe d'une représentation d'un groupe	140
2. Représentations des groupes et représentations des algèbres de Lie	140
2.1. Définition. Universalité	140
2.2. Représentations d'un groupe de Lie et de son algèbre de Lie	141
3. Représentations des groupes de Lie compacts	143
3.1. Orthogonalité et complétude	143
3.2. Conséquences	147

3.3. Cas des groupes finis	148
3.4. Récapitulation	149
4. Représentations projectives. Théorème de Wigner. Cocycles et termes centraux. Théorème de Bargmann	150
4.1. Définition	150
4.2. Théorème de Wigner.	151
4.3. Invariance d'un système quantique	153
4.4. Transformations des observables. Théorème de Wigner–Eckart	154
4.5. Forme infinitésimale d'une représentation projective. Extension centrale	156
Bibliographie sommaire	157
Appendix A. Tenseurs, vous avez dit tenseurs ?	158
Appendix B. Compléments sur les matrices de représentation de $SU(2)$	159
B.1. Orthogonalité, complétude, caractères	159
B.2. Fonctions spéciales. Harmoniques sphériques	160
B.3. Applications physiques	165
Exercices pour le chapitre 2	167
Problème 1 : Décomposition d'une amplitude	168
Problème 2 : Produit tensoriel dans $SU(2)$	169
Chapitre 3. Algèbres de Lie simples, classification et représentations.	
Racines et poids	171
1. Sous-algèbre de Cartan. Racines. Forme canonique de l'algèbre.	171
1.1. Sous-algèbre de Cartan	171
1.2. Base canonique de l'algèbre de Lie	172
2. Géométrie des systèmes de racines	175
2.1. Produits scalaires de racines. La matrice de Cartan	175
2.2. Systèmes de racines des algèbres simples. La classification de Cartan	180
2.3. Base de Chevalley de l'algèbre	181
2.4. Coracines. Plus haute racine. Nombres et exposants de Coxeter.	181
3. Représentations des algèbres semi-simples	182
3.1. Poids. Réseau des poids.	182
3.2. Racines et poids de $su(n)$	187
4. Produit tensoriel des représentations de $su(n)$	191
4.1. Règles de Littlewood-Richardson	191
4.2. Construction tensorielle explicite des représentations de $SU(2)$ et $SU(3)$	193
5. Tableaux d'Young et représentations de $GL(n)$ et $SU(n)$	195
Références additionnelles	198
Appendix A. The classical algebras of type A, B, C, D	199
Exercices	201
Chapitre 4. Symétries globales en physique des particules	205
1. Symétries globales exactes ou brisées. Brisure spontanée	205
1.1. Panorama. Symétries exactes ou brisées	205
1.2. Brisure de la symétrie chirale	208
1.3. Brisures quantiques. Anomalies	210
2. La symétrie de saveur $SU(3)$ et le modèle des quarks.	210

2.1. Pourquoi le groupe $SU(3)$?	210
2.2. Conséquences de la symétrie $SU(3)$	213
2.3. Brisures électromagnétiques de la symétrie $SU(3)$	215
2.4. Écart de masses “forts”. Formule de masse de Gell-Mann–Okubo.	217
2.5. Les quarks	218
2.6. Courants hadroniques et interactions faibles	219
3. De $SU(3)$ à $SU(4)$ et aux six saveurs	220
3.1. Nouvelles saveurs	220
3.2. Introduction de la couleur	222
Références	223
Exercices pour le chapitre 4	223
Problème : Couplages invariants par $SU(3)$ et invariance cachée	224
Chapitre 5. Théories de jauge. Modèle standard	227
1. Invariance de jauge. Couplage minimal. Lagrangien de Yang–Mills	227
1.1. Invariance de jauge de l'électrodynamique	227
1.2. Extension non abélienne de Yang–Mills	228
1.3. Géométrie des champs de jauge	231
1.4. Lagrangien de Yang–Mills	231
1.5. Quantification. Règles de Feynman.	232
1.6. Renormalisabilité	234
2. Champs de jauge massifs	234
2.1. Interactions faibles et bosons intermédiaires	234
2.2. Brisure spontanée de symétrie de jauge. Mécanisme de Brout–Englert–Higgs	235
3. Le modèle standard	237
3.1. Le secteur fort	237
3.2. Le secteur électro-faible, une esquisse.	238
4. Compléments	242
4.1. Modèle standard et au delà.	242
4.2. Théories grand-unifiées ou GUTs	243
4.3. Anomalies	244
Références	246
Exercices	246
Problème : Théories de jauge sur réseau	249

Chapitre 0

Invariance relativiste et champs classiques

1. Principes de la Relativité restreinte. Conséquences physiques

Au début du XX^{ème} siècle, diverses expériences paraissant en contradiction avec les principes de base de la physique conduisent à une refonte de ces principes. Ce sont en particulier les expériences négatives essayant de mettre en évidence la vitesse de la Terre par rapport à l’“éther” (Michelson-Morley). Cette réflexion théorique, après les travaux de Lorentz et Poincaré, culmine dans l’œuvre d’Einstein, dont nous rappelons les grandes lignes.

1.1. Référentiels, synchronisation des horloges et simultanéité.

Notions de référentiel : système de coordonnées spatiales et temporelle. Les coordonnées spatiales d’un événement sont mesurées dans un repère (disons orthogonal) par des coordonnées x, y, z (ou x^1, x^2, x^3) à l’aide de “règles” ; sa coordonnée temporelle t est mesurée par une “horloge” au repos dans le référentiel. Règles et horloge sont bien entendu idéalisées, supposées parfaites, ...

Notion de référentiel inertiel, dans lequel une particule matérielle libre (c’est-à-dire soumise à aucune force) garde une vitesse constante (a un mouvement uniforme). On va voir que dès qu’on connaît un tel référentiel¹, tout autre s’en déduisant par une translation uniforme, à vitesse constante, l’est aussi. Les lois de transformation des coordonnées d’un référentiel à l’autre vont devoir être précisées.

Synchronisation des horloges, simultanéité de deux événements. On admet qu’on peut attacher des horloges à chaque point d’un référentiel et les *synchroniser*. Selon Einstein, deux horloges en A et B sont synchronisées si pour un rayon lumineux émis de A vers B , et immédiatement réfléchi vers A , les temps t_A, t_B et t'_A où le rayon atteint A ou B satisfont

$$t_B - t_A = t'_A - t_B .$$

Noter le rôle opérationnel que la lumière joue à ce point dans la définition de la théorie. Dans la conception de la physique classique, la notion de synchronisation sous-entend l’existence de signaux de vitesse de propagation infinie : le temps t est le même en tout point du référentiel.

Cette définition de la synchronisation permet alors de parler d’événements simultanés **dans un référentiel donné** : deux tels événements ont même valeur de t .

¹ comme le fameux système de coordonnées “lié à des étoiles fixes”

1.2. Postulats de la Relativité

1. **Principe de relativité.** Les lois de la physique sont les mêmes dans tout référentiel inertiel.

Il y a déjà un principe de relativité en physique classique. Les lois de la *mécanique* y sont les mêmes dans deux référentiels inertiels. c'est-à-dire obtenu l'un par rapport à l'autre par une transformation du groupe de Galilée $\mathbf{x} \rightarrow \mathbf{x}' = \mathcal{R}\mathbf{x} + \mathbf{a} - \mathbf{v}t$, $t' = t + \Delta t$. Par exemple la loi fondamentale de la dynamique, $\mathbf{F} = m\ddot{\mathbf{x}}$ est la même dans tout référentiel inertiel, à condition que m soit constante et que \mathbf{F} se transforme comme un vecteur. Ce n'est pas le cas des lois de l'électrodynamique . . . Le postulat d'Einstein étend donc ce principe de relativité à toutes les lois de la Physique.

Ce principe de relativité va être réexprimé par la suite comme une invariance des théories physiques sous l'action de transformations de coordonnées. On parle alors d'invariance relativiste ou d'invariance de Lorentz.

2. **Vitesse de la lumière.** Tout rayon lumineux se propage dans tout référentiel inertiel R avec la même vitesse c , qu'il ait été émis par une source au repos ou en mouvement dans R .

Cette constance de la vitesse de la lumière est une idée neuve d'Einstein, en rupture avec ses contemporains qui s'ingéniaient (en vain) à mettre en évidence l'influence d'un mouvement d'une source par rapport à l'éther sur la vitesse de propagation de la lumière. La valeur de cette constante est

$$c = 299\,792\,458 \text{ ms}^{-1} .$$

Cette valeur est *exacte*. En effet depuis le début des années 60, le mètre est *défini* comme étant la longueur du trajet parcouru dans le vide par la lumière de certains lasers pendant la fraction $1/c$ de seconde. (Actuellement la seconde est la durée de 9 192 631 770 périodes de la radiation correspondant à la transition entre les deux niveaux hyperfins de l'état fondamental de l'atome de césium 133.) Le choix optimal des radiations à utiliser continue de faire l'objet de travaux en métrologie.

Un ordre de grandeur à retenir : 1 année lumière (1 ly) $\approx 10^{16}$ m (plus précisément $0.9461 \cdot 10^{16}$ m).

1.3. Transformations de Lorentz.

Soient deux référentiels R et R' , le second étant doté d'une vitesse \mathbf{v} par rapport au premier. Quelle transformation supposée *linéaire* des coordonnées $(x, y, z, t) \rightarrow (x', y', z', t')$ est compatible avec les deux principes précédents ?

On suppose qu'au temps $t = t' = 0$, l'origine $x' = y' = z' = 0$ du référentiel R' passe par celle $x = y = z = 0$ de R avec la vitesse \mathbf{v} . Supposons qu'une onde lumineuse sphérique est émise à l'origine au temps $t = t' = 0$. Selon le principe de relativité, les deux

observateurs doivent observer le même phénomène : aux instants ultérieurs, cette onde est localisée sur la sphère d'équation

$$x^2 + y^2 + z^2 = c^2 t^2 \quad (1.1)$$

dans R , et sur $x'^2 + y'^2 + z'^2 = c^2 t'^2$ dans R' , avec la même vitesse c (deuxième postulat). La transformation cherchée doit donc préserver le cône futur \mathcal{C}_+ d'équation

$$C(x, y, z, t) := x^2 + y^2 + z^2 - c^2 t^2 = 0, \quad t \geq 0$$

Proposition Les transformations linéaires cherchées préservent non seulement le cône futur \mathcal{C}_+ mais aussi la forme quadratique $x^2 + y^2 + z^2 - c^2 t^2$.

Preuve ². Soit $Q(x, y, z, t)$ la forme quadratique obtenue en substituant x', y', z', t' fonctions linéaires de x, y, z, t dans la forme $x'^2 + y'^2 + z'^2 - c^2 t'^2$. Quelles sont les formes quadratiques Q telles que $Q = 0$ contienne le cône $x^2 + y^2 + z^2 - c^2 t^2 = 0$?

Lemme Une forme quadratique $Q(x, y, z, t)$ telle que $Q = 0$ contient le cône futur $x^2 + y^2 + z^2 - c^2 t^2 = 0, t \geq 0$ est proportionnelle à la forme $C(x, y, z, t)$

$$Q(x, y, z, t) = K(x^2 + y^2 + z^2 - c^2 t^2),$$

avec K une constante.

Preuve du lemme. Donnons en la preuve en dimension d'espace d : $\mathbf{x} \in \mathbb{R}^d$. On écrit $Q(\mathbf{x}, t) = A_{ij} x^i x^j + 2t B_i x^i + C t^2$ en explicitant les différents termes ³. La forme $Q(\mathbf{x}, t)$ s'annule si $(\mathbf{x}, t) \in \mathcal{C}_+$, c'est-à-dire $t = \frac{1}{c} \sqrt{\mathbf{x}^2}$,

$$\forall \mathbf{x} \in \mathbb{R}^d \quad Q\left(\mathbf{x}, \frac{1}{c} \sqrt{\mathbf{x}^2}\right) = 0$$

et on trouve donc $A_{ij} x^i x^j + \frac{2}{c} B_i x^i \sqrt{\mathbf{x}^2} + \left(\frac{1}{c}\right)^2 C |\mathbf{x}|^2 = 0$ pour tout \mathbf{x} , ce qui n'est possible que si le terme 'croisé' est nul, $B_i = 0$, tandis que $A_{ij} x^i x^j = -\left(\frac{1}{c}\right)^2 C |\mathbf{x}|^2$. La forme Q s'écrit donc

$$Q(\mathbf{x}, t) = -\frac{C}{c^2} (\mathbf{x}^2 - c^2 t^2)$$

ce qui établit le lemme.

² Je dois cette preuve à Michel Bauer.

³ Ici et dans toute la suite, convention de sommation sur les indices répétés.

Revenant à la proposition, cela implique que les transformations linéaires cherchées préservent la forme quadratique $x^2 + y^2 + z^2 - c^2t^2$, à un facteur près indépendant des coordonnées, ne pouvant donc dépendre que de la vitesse \mathbf{v}

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = \varphi(\mathbf{v})(x^2 + y^2 + z^2 - c^2t^2) .$$

Finalement, le facteur $\varphi(\mathbf{v})$ qui satisfait $\varphi(\mathbf{v})\varphi(-\mathbf{v}) = 1$ (par composition de la transformation avec son inverse), $\varphi(\mathbf{v}) = \varphi(-\mathbf{v})$ (par symétrie de retournement des axes $\mathbf{x} \rightarrow -\mathbf{x}$), $\varphi(\mathbf{0}) = 1$ et est une fonction continue de \mathbf{v} , vaut 1.

Les transformations des coordonnées cherchées sont donc les transformations linéaires laissant invariantes la forme quadratique $x^2 + y^2 + z^2 - c^2t^2$. Elles forment un groupe, le groupe de Lorentz. Nous avons donc le

Théorème *Les transformations linéaires de coordonnées entre deux référentiels inertiels sont les transformations de Lorentz, qui préservent la forme quadratique $x^2 + y^2 + z^2 - c^2t^2$.*

Remarques et exercices.

1. Plutôt que l'argument géométrique précédent, on peut préférer l'argument physique d'Einstein (voir aussi Panofsky et Phillips) qui *construisent* les transformations de Lorentz par une suite d'expériences de pensée.
2. Peut-on relâcher l'hypothèse de linéarité ? Autrement dit, y-a-t-il des transformations non linéaires $x \rightarrow x'$ (non singulières, c'est-à-dire inversibles en tout point), qui préservent le cône ? y-en-a-t-il qui préservent la forme quadratique ? (éléments de réponse dans [Wg, p27].)

Donnons l'expression explicite de ces transformations⁴. Elles comprennent bien sûr les rotations d'espace usuelles : $\mathbf{x} \rightarrow \mathcal{R}_{\mathbf{n}}(\theta)\mathbf{x}$, $t \rightarrow t$ est la rotation d'angle θ autour d'un axe \mathbf{n} . Ainsi la rotation d'axe \hat{z}

$$\mathcal{R}_{\hat{z}}(\theta) \left\{ \begin{array}{l} x' = x \cos \theta + y \sin \theta \\ y' = -x \sin \theta + y \cos \theta \\ z' = z \quad t' = t \end{array} \right. . \quad (1.2)$$

Les “transformations de Lorentz spéciales” (ou *boosts*), qui décrivent un changement de coordonnées entre deux référentiels animés d'une vitesse relative $\mathbf{v} = v\mathbf{n}$, disons de R' par rapport à R , en sont l'analogie en géométrie hyperbolique. Ainsi pour une vitesse le long de l'axe des x

$$L_{\hat{x}}(\beta) \left\{ \begin{array}{l} x' = x \cosh \beta - ct \sinh \beta \\ ct' = -x \sinh \beta + ct \cosh \beta \\ y' = y, \quad z' = z \end{array} \right. \quad (1.3)$$

⁴ ou tout au moins de celles qui sont “connectées à l'identité”, voir la discussion plus bas.

où l'angle hyperbolique β , appelé aussi *rapidité*, est défini par

$$\cosh \beta = \gamma := \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \tanh \beta = \frac{v}{c}. \quad (1.4)$$

Ces relations prennent donc aussi la forme

$$\begin{aligned} x' &= \gamma(x - vt) & y' &= y & z' &= z & t' &= \gamma\left(t - \frac{v}{c^2}x\right) \\ \text{ou inversement} & & & & & & & \\ x &= \gamma(x' + vt') & y &= y' & z &= z' & t &= \gamma\left(t' + \frac{v}{c^2}x'\right). \end{aligned} \quad (1.5)$$

En général pour une vitesse \mathbf{v} de R' par rapport à R quelconque

$$\begin{aligned} \mathbf{x}'_{\parallel} &= \gamma(\mathbf{x}_{\parallel} - \mathbf{v}t) & \mathbf{x}'_{\perp} &= \mathbf{x}_{\perp} & t' &= \gamma\left(t - \frac{1}{c^2}\mathbf{v}\cdot\mathbf{x}\right) \\ \text{avec } \mathbf{x}_{\parallel} &:= \mathbf{v} \frac{\mathbf{v}\cdot\mathbf{x}}{v^2} & \mathbf{x}_{\perp} &:= \mathbf{x} - \mathbf{x}_{\parallel}, \end{aligned} \quad (1.6)$$

où par définition \mathbf{x}_{\parallel} est colinéaire à \mathbf{v} tandis que \mathbf{x}_{\perp} lui est orthogonal, et donc satisfait $\mathbf{x}_{\parallel}\cdot\mathbf{x}_{\perp} = 0$.)

Nous reviendrons plus bas sur ces transformations de Lorentz (§2.1).

1.4. Conséquences des transformations de Lorentz. Cinématique relativiste

Première conséquence des transformations de Lorentz : *le temps perd son caractère absolu*.

En particulier, la simultanéité de deux événements est une notion qui est liée à un référentiel donné : pour deux événements, on peut avoir $\Delta t = 0$ dans R mais $\Delta t' \neq 0$ dans R' .

• Dilatation des temps.

Une horloge placée (disons à l'origine) dans un référentiel inertiel R' en mouvement paraît retarder par un facteur γ quand elle est observée dans R : selon les formules (1.5), les battements successifs en $t' = 0, 1, 2, \dots$ secondes paraissent dans R avoir lieu en $t = 0, \gamma, 2\gamma, \dots$. Bien sûr, il y a symétrie et une horloge au repos dans R paraît retarder vue de R' .

Cette dilatation du temps s'observe couramment en physique des particules. Les muons μ^{\pm} , (des leptons chargés, c'est-à-dire des analogues environ 207 fois plus massifs de l'électron et du positron e^{\pm}), sont produits dans la haute atmosphère par les rayons cosmiques (dont les protons créent par collisions avec les noyaux de l'air des mésons π , qui se désintègrent très rapidement en μ^{\pm}). Au repos, ces μ^{\pm} se désintègrent rapidement, avec

une demi-vie de $2.19 \cdot 10^{-6}$ s. Pendant un temps aussi court, une particule ultrarelativiste, de vitesse $\approx c \approx 3 \cdot 10^8$ m parcourt environ 600 m. Or ces muons constituent une bonne part du rayonnement cosmique reçu sur terre, au niveau de la mer. Cela est rendu possible par la dilatation de leur temps de vie par un facteur γ . Pour $v = (1 - \epsilon)c$, $\epsilon \ll 1$, $\gamma \approx 1/\sqrt{2\epsilon}$; exemple $\epsilon = 5 \cdot 10^{-3}$, $\gamma = 10$.

D'autres vérifications expérimentales très précises sont obtenues au laboratoire avec des mesures d'effet Doppler sur des ions . . . (cf revue de C.M. Will citée dans les références).

• *Notion de temps propre.*

La notion de dilatation du temps peut aussi s'appliquer à un mouvement non uniforme. Soit un point M animé d'un mouvement donné dans un référentiel R "au repos" ; sa trajectoire est donnée par $\mathbf{x}(t)$ dans les coordonnées de R . On suppose M doté d'une horloge qui lui est liée. Pendant un temps dt infinitésimal, la position de M dans R change de $d\mathbf{x}$, et on peut assimiler le mouvement de M à une translation de vitesse $\mathbf{v}(t) = d\mathbf{x}(t)/dt$. On peut donc attacher de façon instantanée un référentiel inertiel $R'(t)$ à M , doté par rapport à R de la vitesse $\mathbf{v}(t)$: on pourrait parler du référentiel inertiel tangent à la trajectoire de M . Pendant ce temps infinitésimal, l'horloge liée à M avance de dt' (et sa coordonnée ne change pas, $d\mathbf{x}' = 0$). Calculant l'élément de longueur invariante dans R et dans R' , on trouve

$$ds^2 = c^2 dt'^2 = c^2 dt^2 - d\mathbf{x}^2 = (c^2 - \mathbf{v}^2) dt^2$$

donc

$$d\tau := \frac{1}{c} ds = dt' = dt \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} < dt . \quad (1.7)$$

C'est une nouvelle manifestation du ralentissement de l'horloge en mouvement par rapport à une horloge au repos. Le temps t' de l'horloge liée au mouvement, qu'on notera désormais τ , est appelé le *temps propre* du mobile M . En intégrant le long de la trajectoire parcourue entre les temps (de R) t_1 et t_2 , on voit que

$$\Delta\tau = \tau_2 - \tau_1 = \int_{t_1}^{t_2} \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} dt \leq (t_2 - t_1) .$$

On a donc obtenu le résultat que l'intervalle de temps propre d'un mobile est toujours inférieur à l'intervalle de temps mesuré par un observateur au repos.

Si le mobile parcourt une trajectoire et recroise celle de l'observateur au repos, on peut comparer les laps de temps écoulés entre les deux croisements, et on trouve $\Delta\tau < t_2 - t_1$. Par conséquent, dans la géométrie hyperbolique de l'espace-temps, la trajectoire rectiligne

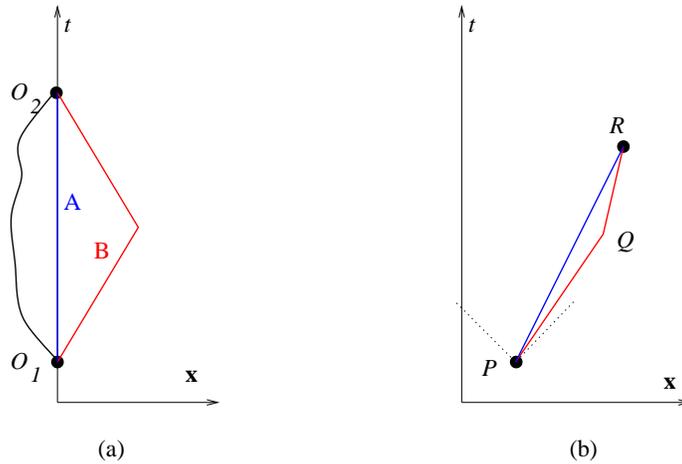


Fig. 1: (a) : trois trajectoires de genre temps entre O_1 et O_2 : la ligne droite donne le chemin le plus long en temps ! (b) : l'inégalité triangulaire inversée pour des points P, Q, R de séparations mutuelles de genre temps.

de l'observateur au repos fournit le temps le plus long entre deux “points” (événements), supérieur à celui de toute trajectoire curviligne, voir figure 1. Autrement dit, parmi toutes les trajectoires entre deux points O_1 et O_2 , le mouvement uniforme donne celle qui maximise $\int_1^2 d\tau$ (ou encore $\int_1^2 ds$) (cf ci-après le principe variationnel d'action).

- Le “paradoxe” des jumeaux de Langevin.

Ce fameux paradoxe se rattache au même ordre d'idées. Le jumeau A est au repos, à l'origine du référentiel R . À l'instant $t = 0$, son jumeau B le quitte à bord d'un engin spatial se déplaçant à une vitesse v d'ordre c . Au bout d'un temps t' (mesuré dans son référentiel), B fait “demi-tour” et repart immédiatement à la même vitesse vers la Terre, qu'il atteint au bout du temps $2t'$ (temps de R'). Quel est le temps de l'aller-retour du point de vue de A ? Pour A, l'horloge de B retarde d'un facteur γ et le temps écoulé mesuré dans R est donc $\Delta t = 2t'\gamma$. A a vieilli plus vite que son jumeau astronaute ! Pour $v = 0.6 c$, donc $\gamma = 1.25$, et $t' = 2$ ans, on trouve $\Delta t' = 4$ ans, $\Delta t = 2\gamma t' = 5$ ans.

Jusque là, tout est en accord avec la discussion précédente de dilatation du temps. Le paradoxe (apparent !) consiste à inverser les rôles de A et B. Dans le référentiel de B, A s'éloigne à la vitesse $-\mathbf{v}$, etc. Pourquoi devrait-il y avoir une dissymétrie entre A et B?

Il n'y a pas symétrie parce que A est un observateur inertiel tandis que B ne l'est pas. Comme on l'a discuté plus haut, la ligne brisée parcourue par B entre ses deux rencontres avec A fournit un intervalle de temps inférieur à celui de la ligne droite suivie par A (voir figure 1(a)).

- *Contraction des distances.*

On parle aussi de contraction des règles, ou de contraction de Lorentz. C'est la contrepartie spatiale du phénomène de dilatation du temps. Une "règle" de longueur L au repos dans le référentiel R' a pour coordonnées de ses extrémités x'_A et $x'_B = x'_A + L$, mesurées au même temps $t'_A = t'_B$. Dans le référentiel R , on mesure en un même temps $t_A = t_B$ des coordonnées x_A, x_B , et selon la première équation (1.5), $x' = \gamma(x - vt)$, donc

$$L = |x'_A - x'_B| = \gamma|x_A - x_B| \quad \text{donc} \quad |x_A - x_B| = L/\gamma < L. \quad (1.8)$$

Il y a donc *contraction* de la règle, quand elle est observée dans le référentiel R .

En physique des particules, le phénomène se manifeste dans les expériences de très haute énergie. Des particules "ultrarelativistes" (c'est-à-dire accélérées à des vitesses proches de c), qu'on décrirait au repos comme de petites sphères, apparaissent plutôt comme des galettes aplaties dans leurs collisions...

La description de certains phénomènes impliquant des hadrons (protons, neutrons, ...) de très haute énergie fait appel à la notion de *partons*, sous-constituants de ces hadrons, dotés de mouvements internes à ce hadron. Dans l'image où le hadron est aplati par l'effet précédent, seules comptent les impulsions *transverses* des partons par rapport à celle du hadron.

- *Composition des vitesses.*

En différentiant (1.6) on trouve immédiatement la loi de transformation des vitesses $\mathbf{V} = \frac{d\mathbf{x}}{dt} \mapsto \mathbf{V}' = \frac{d\mathbf{x}'}{dt'}$

$$\mathbf{V}'_{\parallel} = \frac{\mathbf{V}_{\parallel} - \mathbf{v}}{1 - \frac{\mathbf{v} \cdot \mathbf{V}}{c^2}} \quad , \quad \mathbf{V}'_{\perp} = \sqrt{1 - \frac{v^2}{c^2}} \frac{\mathbf{V}_{\perp}}{1 - \frac{\mathbf{v} \cdot \mathbf{V}}{c^2}}. \quad (1.9)$$

On en déduit que $\mathbf{V}^2 = \mathbf{V}'^2_{\perp} + \mathbf{V}'^2_{\parallel}$ est tel que

$$\left(1 - \frac{\mathbf{V}'^2}{c^2}\right) \left(1 - \frac{\mathbf{v} \cdot \mathbf{V}}{c^2}\right)^2 = \left(1 - \frac{v^2}{c^2}\right) \left(1 - \frac{\mathbf{V}^2}{c^2}\right),$$

une relation compatible avec le fait que c est la limite supérieure des vitesses observables, et que la vitesse de la lumière est c quel que soit le référentiel d'observation (Postulat 2).

Application. Considérons l'expérience de Fizeau où la lumière se propage dans un milieu d'indice n , au repos dans un référentiel R' . La vitesse de la lumière observée dans R' est donc $V' = c/n$. Si R' se déplace avec une vitesse \mathbf{v} (disons le long des x) par rapport à R , la vitesse observée dans R est (composante x)

$$V = \frac{(V' + v)}{\left(1 + \frac{v}{cn}\right)} \approx \frac{c}{n} + v\left(1 - \frac{1}{n^2}\right) + O(v^2).$$

C'est là une formule trouvée (par un raisonnement tout différent !) par Fresnel (1818) et vérifiée expérimentalement par Fizeau (1851).

2. Espace de Minkowski. Groupe de Lorentz $O(3,1)$.

2.1. Métrique de l'espace de Minkowski. Groupes de Lorentz et de Poincaré

On peut résumer ce qui précède en disant que l'espace à trois dimensions de la physique classique (augmenté d'un temps universel) a été remplacé par un *espace-temps*, espace réel de dimension 4, doté d'une métrique indéfinie de signature $(+, -, -, -)$, l'espace de Minkowski $\mathcal{M} = \mathbb{R}^4$. Les vecteurs V de cet espace sont appelés *quadrivecteurs*. Dans un repère orthonormé, le produit scalaire de deux quadrivecteurs V et W de coordonnées $V = (V^0, \mathbf{V})$ et $W = (W^0, \mathbf{W})$ est donc ⁵

$$(V, W) = V.W = V^0W^0 - \mathbf{V}.\mathbf{W} = g_{\mu\nu}V^\mu W^\nu \quad (2.1)$$

avec le tenseur métrique ⁶

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (2.2)$$

Un cas particulièrement important de quadrivecteur est le vecteur position x avec des coordonnées $x^0 = ct$ et \mathbf{x} . En prenant la norme carrée de sa différentielle dx , on obtient le carré de l'élément de longueur

$$ds^2 = (dx, dx) = c^2 dt^2 - d\mathbf{x}^2 = g_{\mu\nu} dx^\mu dx^\nu. \quad (2.3)$$

La signature indéfinie de cette métrique a des conséquences géométriques inhabituelles, comme l'inégalité triangulaire inversée déjà mentionnée plus haut (§1.4). Anticipant un peu sur une terminologie précisée au § 2.3, si P , Q et R sont trois points d'espace-temps séparés par des intervalles de genre temps, avec Q dans le "cône futur" de P , et R dans

⁵ On utilisera la notation abrégée $V.W$ toutes les fois qu'elle est non ambiguë, sinon (V, W) . La composante en caractère gras \mathbf{V} ou surmontée d'une flèche, \vec{V} , désigne la partie spatiale du vecteur, de coordonnées cartésiennes V^i , $i = 1, 2, 3$. Par convention, les indices grecs se réfèrent aux coordonnées minkowskienne, $\mu, \nu, \dots = 0, 1, 2, 3$, les indices latins aux coordonnées spatiales $i, j, \dots = 1, 2, 3$.

⁶ La notation $g_{\mu\nu}$ pour ce tenseur est remplacée par $\eta_{\mu\nu}$ dès qu'on aborde la relativité générale : g désigne alors la métrique non triviale couplée à la matière, tandis que η désigne la métrique minkowskienne de référence.

ceux de P et Q , voir figure 1(b), alors leurs distances minkowskiennes satisfont $d(P, R) \geq d(P, Q) + d(Q, R)$.

Le groupe d'invariance du produit scalaire (2.1) est le groupe de Lorentz $\mathcal{L} = \text{O}(1, 3)$, c'est-à-dire le sous-groupe du groupe linéaire, $V \mapsto \Lambda V$, constitué des matrices réelles $\Lambda = (\Lambda^\mu_\rho)$ satisfaisant

$$\Lambda^T g \Lambda = g, \quad (2.4)$$

ou, en composantes, $V^\mu \mapsto \Lambda^\mu_\rho V^\rho$,

$$\Lambda^\mu_\rho \Lambda^\nu_\sigma g_{\mu\nu} = g_{\rho\sigma}. \quad (2.5)$$

Il contient comme sous-groupe le groupe de rotations $\text{O}(3)$ de l'espace \mathbb{R}^3 . Il contient aussi toute transformation de Lorentz spéciale de rapidité β le long d'une direction \mathbf{n} (\mathbf{n} un vecteur unitaire) (on rappelle que $\mathbf{x}_\parallel := \mathbf{x} \cdot \mathbf{n} \mathbf{n}$, $\mathbf{x}_\perp = \mathbf{x} - \mathbf{x}_\parallel$)

$$L_{\mathbf{n}}(\beta) \begin{cases} \mathbf{x}'_\parallel = \mathbf{x}_\parallel \cosh \beta - x_0 \mathbf{n} \sinh \beta \\ x'_0 = -\mathbf{x}_\parallel \cdot \mathbf{n} \sinh \beta + x_0 \cosh \beta \\ \mathbf{x}'_\perp = \mathbf{x}_\perp, \end{cases} \quad (2.6)$$

une formule qui généralise (1.3).

Les transformations de Lorentz spéciales forment-elles un groupe? Montrer que les transformations de Lorentz spéciales de vecteur directeur \mathbf{n} fixé forment un groupe dont les rapidités se composent de façon additive : $L_{\mathbf{n}}(\beta)L_{\mathbf{n}}(\beta') = L_{\mathbf{n}}(\beta')L_{\mathbf{n}}(\beta) = L_{\mathbf{n}}(\beta + \beta')$.

Le groupe de Lorentz a quatre composantes connexes (on dit aussi "nappes"). En effet, en prenant le déterminant de (2.4), on apprend que $\det \Lambda = \pm 1$ et en prenant l'élément de matrice 00 de (2.5), on lit $1 = g_{\mu\nu} \Lambda^\mu_0 \Lambda^\nu_0 = (\Lambda^0_0)^2 - \sum_i (\Lambda^i_0)^2$, donc $|\Lambda^0_0| > 1$, Λ^0_0 positif ou négatif. Les quatre nappes correspondent aux quatre choix de signes de $(\det \Lambda, \Lambda^0_0)$, la nappe $(+, +)$ constituant le sous-groupe \mathcal{L}_+^\uparrow des transformations de Lorentz

"propres, orthochrones". La parité d'espace $\Lambda_P^\mu_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$, le renversement du temps $\Lambda_T = -\Lambda_P$ et leur produit $\Lambda_P \Lambda_T = -I$ fournissent des représentants des trois autres nappes ; en les composant avec un élément quelconque de \mathcal{L}_+^\uparrow , on obtient l'élément générique de chacune de ces nappes. Toute transformation de \mathcal{L}_+^\uparrow s'écrit comme produit d'une rotation "ordinaire" de $\text{SO}(3)$ par une transformation de Lorentz spéciale.

Le groupe de Poincaré \mathcal{P} , appelé aussi groupe de Lorentz inhomogène, est le groupe engendré par \mathcal{L} et les translations d'espace-temps.

Plus précisément, c'est le "produit semi-direct" $\mathbb{R}^4 \rtimes \mathcal{L}$ du groupe \mathbb{R}^4 des translations par \mathcal{L} , puisque si on note (a, Λ) la transformation $x' = \Lambda \cdot x + a$, la loi de composition est (le vérifier)

$$(a', \Lambda') \cdot (a, \Lambda) = (a' + \Lambda' \cdot a, \Lambda' \Lambda) ,$$

avec une action de \mathcal{L} dans le groupe des translations \mathbb{R}^4 . Quel est l'inverse de (a, Λ) ?

Nous reviendrons par la suite sur ces groupes, leurs propriétés et leurs "représentations".

2.2. Vecteurs et tenseurs contravariants et covariants.

La métrique indéfinie impose un peu de soin dans la manipulation des quadrivecteurs et de leurs composantes. Les objets se transformant selon

$$V^\mu \mapsto \Lambda^\mu_\nu V^\nu \tag{2.7}$$

sont appelés vecteurs *contravariants*, tandis que

$$U_\mu \mapsto (\Lambda^{-1})^\nu_\mu U_\nu = ((\Lambda^T)^{-1})^\nu_\mu U_\nu \tag{2.8}$$

définit un vecteur *covariant*. Un objet tensoriel peut être doté d'indices des deux types, covariant et contravariant. (Exemples : le tenseur de Kronecker δ_μ^ν , égal à 1 si $\mu = \nu$, à 0 sinon, est un tenseur invariant (pourquoi?). Le tenseur métrique $g_{\mu\nu}$ est un tenseur deux fois covariant, en fait invariant par définition de $O(1, 3)$ (cf (2.5)).

Si V est un vecteur contravariant de composantes V^μ , $V_\mu := g_{\mu\nu} V^\nu$ est sa version covariante, et on peut écrire le produit scalaire $V \cdot W = V_\mu W^\mu = V_0 W^0 + V_1 W^1 + V_2 W^2 + V_3 W^3$. Attention au changement de signe des composantes spatiales $V_i = -V^i$, $i = 1, 2, 3$, tandis que $V_0 = V^0$. Inversement on passe des composantes covariantes aux contravariantes par multiplication par l'inverse de g , noté $g^{\mu\nu}$:

$$\begin{aligned} g_{\mu\nu} g^{\nu\lambda} &= \delta_\mu^\lambda , \\ V^\mu &= g^{\mu\nu} V_\nu . \end{aligned} \tag{2.9}$$

On vérifie immédiatement que $g^{\mu\nu}$ est un tenseur deux fois contravariant, égal dans une base orthonormée à $g_{\mu\nu}$.

Exemples de vecteur covariant : le gradient

$$\partial = (\partial_\mu \equiv \frac{\partial}{\partial x^\mu}) = (\frac{\partial}{\partial x^0}, \vec{\nabla}) \tag{2.10}$$

ou, (cf infra) le potentiel électromagnétique A .

La notation avec indices covariants et contravariants permet d'identifier aisément les invariants (sous l'action des transformations (2.7) et (2.8)). Ainsi si $A(x)$ se transforme comme un quadrivecteur, $\partial A \equiv \partial_\mu A^\mu = \frac{1}{c} \frac{\partial A^0}{\partial t} + \vec{\nabla} \cdot \vec{A}$ est invariant, de même que $A_\mu(x) dx^\mu$, et donc aussi $\int_{s_1}^{s_2} A_\mu(x(s)) \frac{dx^\mu}{ds} ds$.

• *Tenseur complètement antisymétrique $\epsilon^{\mu\nu\rho\sigma}$ de Levi-Civita.*

Le tenseur complètement antisymétrique à trois indices ϵ^{ijk} , invariant par l'action du groupe des rotations $SO(3)$, a un analogue à 4 dimensions, noté $\epsilon^{\mu\nu\rho\sigma}$. Il est défini par sa propriété de complète antisymétrie (donc toute composante avec deux indices identiques s'annule) et par la donnée de $\epsilon^{0123} = 1$. Noter que sa version covariante est telle que $\epsilon_{0123} = -1$, et que $\epsilon^{0ijk} = \epsilon^{ijk} = \epsilon_{ijk}$.

Exercices

1. Montrer que le tenseur $\epsilon^{\mu\nu\rho\sigma}$ est en fait *invariant* par \mathcal{L}_+^\uparrow , c'est-à-dire que $\Lambda_\mu^\mu, \Lambda_{\nu'}^{\nu'}, \Lambda_\rho^\rho, \Lambda_{\sigma'}^{\sigma'}, \epsilon^{\mu'\nu'\rho'\sigma'} = \epsilon^{\mu\nu\rho\sigma}$ si $\Lambda \in \mathcal{L}_+^\uparrow$. (Pour cela, se rappeler la définition du déterminant).
2. Soit q un quadrivecteur. Montrer que le tenseur

$$q_\mu \epsilon_{\nu\rho\sigma\tau} + q_\nu \epsilon_{\rho\sigma\tau\mu} + q_\rho \epsilon_{\sigma\tau\mu\nu} + q_\sigma \epsilon_{\tau\mu\nu\rho} + q_\tau \epsilon_{\mu\nu\rho\sigma}$$

est complètement antisymétrique en $(\mu, \nu, \rho, \sigma, \tau)$. Que peut-on dire d'un tel tenseur dans un espace à quatre dimensions ? Les identités de ce type, parfois appelées "circulantes", sont utiles quand on cherche à construire une base d'invariants.

3. De même qu'on a les identités $\epsilon^{ijk} \epsilon_{i'j'k} = \delta_i^{i'} \delta_j^{j'} - \delta_i^{j'} \delta_j^{i'}$, $\epsilon^{ijk} \epsilon_{i'jk} = 2\delta_i^{i'}$ et $\epsilon^{ijk} \epsilon_{ijk} = 6$, on a $\epsilon^{\mu\nu\rho\sigma} \epsilon_{\mu\nu\rho\sigma} = -24$, $\epsilon^{\mu\nu\rho\sigma} \epsilon_{\mu'\nu\rho\sigma} = -6\delta_{\mu'}^\mu$. Que valent $\epsilon^{\mu\nu\rho\sigma} \epsilon_{\mu'\nu'\rho\sigma}$? $\epsilon^{\mu\nu\rho\sigma} \epsilon_{\mu'\nu'\rho'\sigma'}$? Trouver une expression déterminantale pour $\epsilon^{\mu\nu\rho\sigma} \epsilon_{\mu'\nu'\rho'\sigma'}$.

2.3. *Géométrie de l'espace de Minkowski, temps propre, lignes d'univers.*

Dans la suite "vecteur" signifie quadrivecteur.

Un vecteur de norme carrée $> 0, = 0, < 0$ est dit respectivement de genre temps, lumière et espace.

Tout vecteur non nul W orthogonal à un vecteur V de genre temps est de genre espace. Si $V \cdot W = 0$ et $V^2 > 0$, on peut en effet trouver un repère où $V = (V^0, \mathbf{0})$, et alors nécessairement $W^0 = 0$, donc $W^2 = -\mathbf{W}^2 < 0$.

La distance minkowskienne entre deux points A et B de coordonnées x et y est définie par $d^2(A, B) = (x - y)^2 = (x^0 - y^0)^2 - (\mathbf{x} - \mathbf{y})^2$. Plaçons un de ces points à l'origine. Le *cône de lumière* (light cone) sépare l'ensemble des points séparés de l'origine par un intervalle de genre espace de ceux séparés par un intervalle de genre temps. Ces notions sont invariantes par transformation de Lorentz. On distingue le cône futur du cône passé.

Une notion importante est celle de *causalité*. Selon le principe que la vitesse de la lumière est la plus grande qu'on puisse communiquer à un corps matériel ou à un signal,

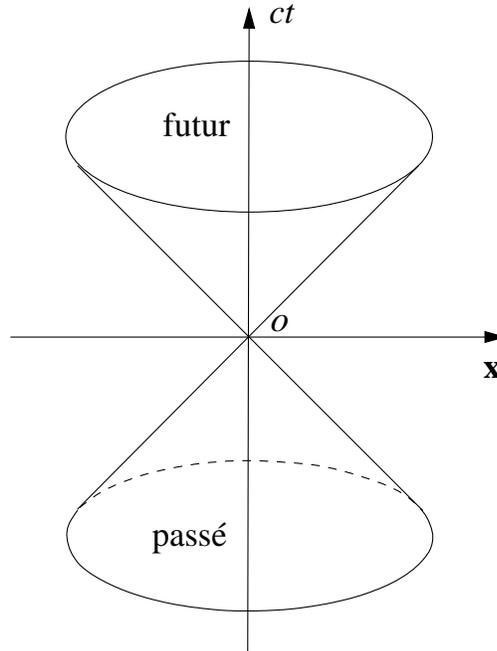


Fig. 2: Cône de lumière

un événement ayant lieu par exemple à l'origine O de l'espace-temps ne peut influencer causalement que les points de l'espace-temps situés dans le cône futur de O , et ne peut avoir reçu d'influence causale que des points situés dans son cône passé (voir figure 2). Autrement dit, il est certain que deux événements x et y séparés par un intervalle de genre espace ne peuvent pas s'être influencés causalement dans le passé, ni s'influencer dans le futur, puisque cela nécessiterait un signal se propageant à une vitesse $|\mathbf{x} - \mathbf{y}|/(x_0 - y_0) > c$.

Attention, si le cône est globalement invariant sous l'effet des transformations de Lorentz, seules les transformations telles que $\Lambda_0^0 > 0$ préservant le sens du temps conservent le cône futur...

Dans tout espace (pseudo-)riemannien, donc doté d'une métrique $ds^2 = g_{\mu\nu} dx^\mu dx^\nu$ dans des coordonnées x^μ , avec $g := \det g_{\mu\nu}$, et $g^{\lambda\mu}$ le tenseur inverse de $g_{\mu\nu}$, on peut construire un laplacien et un élément de volume invariants par changement de coordonnées

$$\Delta = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^\mu} g^{\mu\nu} \sqrt{|g|} \frac{\partial}{\partial x^\nu}$$

$$dV = \sqrt{|g|} \prod_{\mu} dx^\mu .$$
(2.11)

C'est le cas de l'espace de Minkowski, pour lequel le laplacien, aussi appelé le d'Alembertien et traditionnellement noté \square , est simplement

$$\square = g^{\mu\nu} \partial_\mu \partial_\nu = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \quad (2.12)$$

avec Δ le laplacien usuel à 3 dimensions. C'est bien sûr l'opérateur différentiel familier dans l'équation des ondes. Le d'Alembertien et l'élément de volume $dV = c dt d^3\mathbf{x}$ sont invariants par les transformations de Lorentz.

3. Formalisme lagrangien et relativité. Cinématique relativiste

3.1. Rappels : Principe de moindre action. Équations d'Euler-Lagrange. Formalisme hamiltonien, etc.

On rappelle la formulation lagrangienne de la dynamique d'un système comme principe variationnel appliqué à une fonctionnelle d'action dépendant de la trajectoire $q(t)$ et de sa dérivée première $\dot{q}(t)$

$$S = \int_{t_1}^{t_2} dt L(q(t), \dot{q}(t)) . \quad (3.1)$$

Ici $q = (q_1, q_2, \dots, q_p)$ désigne l'ensemble des coordonnées des degrés de liberté ; $L(\cdot)$ est une fonctionnelle "locale" (ne dépendant de q et \dot{q} qu'à une certaine valeur de t). Le principe variationnel conduit aux équations d'Euler-Lagrange

$$\frac{\delta S}{\delta q_i(t)} = \frac{\partial L}{\partial q_i(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i(t)} = 0 , \quad i = 1, 2, \dots, p . \quad (3.2)$$

Les variables d'impulsion p sont les variables conjuguées au sens que

$$p_i = \frac{\partial L}{\partial \dot{q}_i}(q, \dot{q}) . \quad (3.3)$$

Le hamiltonien se construit alors par une transformée de Legendre de L par rapport aux variables \dot{q} . C'est la fonction de q et p donnée par

$$H(p, q) = \sum_i p_i \dot{q}_i(p, q) - L(q, \dot{q}(p, q)) \quad (3.4)$$

où la fonction $\dot{q}(p, q)$ est obtenue en inversant (3.3). La fonctionnelle $H(p, q)$ mesure l'énergie du système. Les équations de Hamilton

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad (3.5)$$

réexpriment pour la première la relation (3.3) entre p et \dot{q} et pour la seconde une équation dynamique, l'équation du mouvement équivalente à (3.2). On peut encore les récrire à l'aide des crochets de Poisson

$$\dot{q}_i = \{H, q_i\} \quad \dot{p}_i = \{H, p_i\} \quad (3.6)$$

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

$$\{f(p, q), g(p, q)\} := \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} . \quad (3.7)$$

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

$$m\ddot{\mathbf{x}} = -\nabla V(\mathbf{x}) =: \mathbf{F} . \quad (3.8)$$

L'impulsion est $\mathbf{p} = m\dot{\mathbf{x}}$. Pour une particule libre, le membre de droite s'annule, et l'impulsion est conservée $d\mathbf{p}/dt = 0$.

3.2. Particule libre relativiste

Revenant aux notations antérieures pour les coordonnées x , cherchons quelle action ne dépendant que de x et de sa dérivée première et invariante relativiste décrit une particule libre. Il est clair que le choix de lagrangien $L = \text{const.} \frac{d\tau}{dt}$ où la constante est fixée par la limite non-relativiste, remplit les deux conditions précédentes :

$$L = -mc^2 \frac{d\tau}{dt} = -mc^2 \sqrt{1 - \frac{\mathbf{v}^2}{c^2}} \quad (3.9)$$

dont le développement à petit v donne bien

$$L \approx -mc^2 + \frac{1}{2}m\mathbf{v}^2 + \dots$$

c'est-à-dire le lagrangien classique non-relativiste, à une constante additive sans importance près. L'action entre les temps t_1 et t_2 est donc $I = \int_{t_1}^{t_2} dt L = -mc \int_{s_1}^{s_2} ds$. Cette action est invariante de Lorentz, ds l'étant. Le signe $-$ garantit que cette action est *minimale* pour un mouvement uniforme entre deux points (cf discussion plus haut au sujet du temps propre). Sa variation conduit à l'équation du mouvement

$$\frac{d^2 x^\mu}{ds^2} = 0 , \quad (3.10)$$

qui est bien satisfaite par un mouvement uniforme.

Remarques

- Montrer que cette action peut se mettre sous la forme $\int mc ds = \int d^4x \frac{ds}{dt} \mu(\mathbf{x})$ à condition d'écrire $m = \int d^3\mathbf{x} \mu(x)$ en termes d'une densité de masse. En discuter l'invariance de Lorentz.
- Montrer que l'action $S = -\frac{1}{2}mc \int_1^2 ds(dx/ds)^2$, où s est un paramètre le long de la trajectoire, conduit à la bonne équation du mouvement, et que s s'identifie alors à la longueur invariante le long de la trajectoire.

À partir de (3.9), l'impulsion relativiste et l'énergie se calculent comme en (3.3)-(3.4) :

$$\mathbf{p} = \frac{\partial L}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \quad E = \mathbf{p} \cdot \mathbf{v} - L = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} . \quad (3.11)$$

Le développement de E dans la limite non relativiste est important

$$E = mc^2 + \frac{1}{2}m\mathbf{v}^2 + \dots \quad (3.12)$$

puisque'il présente le terme d'énergie de repos plus le terme classique d'énergie cinétique.

Les expressions (3.11) coïncident avec celles qu'on obtient à partir du *quadrivecteur vitesse* u^μ . Ce dernier est défini par dérivation de x^μ par rapport au temps propre

$$u^\mu = \frac{dx^\mu}{d\tau} = \left(c \frac{dt}{d\tau}, \frac{d\mathbf{x}}{d\tau} \right) = \frac{dt}{d\tau} (c, \mathbf{v}) . \quad (3.13)$$

Calculons sa norme carrée : $u^2 = c^2$ puisque $(d\tau/dt)^2 = 1 - v^2/c^2$. Le quadrivecteur vitesse u est de genre temps et de longueur constante. Le quadrivecteur accélération défini par $du^\mu/d\tau$ lui est donc orthogonal, et par conséquent de genre espace (cf début du § 2.3). On définit finalement la quadri-impulsion

$$p^\mu = mu^\mu = \left(\frac{1}{c}E, \mathbf{p} \right) \quad (3.14)$$

en accord avec les expressions ci-dessus de \mathbf{p} et E . Cette quantité est de grande importance physique, puisque l'énergie et l'impulsion totales d'un système isolé sont conservées (dans un référentiel donné).

Récapitulons : pour une particule libre relativiste

$$p^\mu = mu^\mu = \left(\frac{1}{c}E, \mathbf{p} \right) = \frac{(mc, m\mathbf{v})}{\sqrt{1 - \mathbf{v}^2/c^2}}$$

$$\mathbf{p} = \frac{E\mathbf{v}}{c^2}$$

$$p^2 = p_\mu p^\mu = m^2 c^2 = \frac{E^2}{c^2} - \mathbf{p}^2 .$$

(3.15)

La dernière relation exprime que la quadri-impulsion est sur la “couche de masse” (*mass shell*), c’est-à-dire sur l’hyperboloïde $p^2 = m^2 c^2$. On peut encore relire cette dernière relation comme $E^2 = (mc^2)^2 + (c\mathbf{p})^2$, une autre manifestation du bilan d’énergie entre énergie de masse et énergie cinétique.

En physique des particules, les énergies sont couramment mesurées en unités d’électrons-volts (rappelez qu’1 eV = 1.6 10⁻¹⁹ Joule) et dans ses multiples keV, MeV, GeV, TeV, ... en attendant le PeV (Peta-eV = 10¹⁵ eV) ... Selon les expressions ci-dessus, une impulsion peut alors se mesurer en XeV/c, une masse en XeV/c². Par abus de langage, on oublie souvent le facteur c et on parle par exemple de masse de ≈ 91 GeV pour le boson intermédiaire Z ...

• *Élément de volume sur la couche de masse.*

La couche de masse $p^2 = m^2 c^2$ est elle-même une variété à 3 dimensions dans l’espace des impulsions. Géométriquement, c’est un hyperboloïde à deux nappes, $p^0 = \pm \sqrt{\mathbf{p}^2 + m^2 c^2}$. Une particule physique est dotée d’une énergie positive, donc on se restreint en général à la nappe $p^0 > 0$. Il est important pour la suite de connaître l’élément de volume naturel (c’est-à-dire invariant de Lorentz) sur cette nappe. On peut soit réduire la forme $d^4 p \delta_+(p^2 - m^2 c^2)$, où δ_+ ne garde que la racine $p^0 > 0$, soit procéder comme plus haut, en calculant l’élément de longueur $(dp)^2$ sur la couche de masse et l’élément de volume associé. Montrer que l’une et l’autre méthode conduisent à

$$d\tilde{p} := \frac{mc}{2p^0} d^3 \mathbf{p} = \frac{mc}{2\sqrt{\mathbf{p}^2 + m^2 c^2}} d^3 \mathbf{p}, \quad (3.16)$$

à une constante multiplicative près, choisie ici pour respecter la dimension $[p]^3$. On peut ensuite aussi écrire $d^3 \mathbf{p} = d^2 \Omega_p |\mathbf{p}|^2 d|\mathbf{p}|$, réexprimer $d|\mathbf{p}|$ en termes de dE , etc.

• *Cas des particules de masse nulle.*

Si $m \rightarrow 0$, les situations physiquement intéressantes correspondent à $v \rightarrow c$, \mathbf{p} et E fixés. Ce qui remplace (3.15) est donc

$$\begin{aligned}
 p^\mu &= \left(\frac{1}{c} E, \mathbf{p} \right) \\
 m = 0 \quad \mathbf{p} &= \frac{E\mathbf{v}}{c^2}, \quad |\mathbf{v}| = c \\
 p^2 = p_\mu p^\mu &= 0 = \frac{E^2}{c^2} - \mathbf{p}^2.
 \end{aligned}$$

(3.17)

3.3. Cinématique relativiste : collisions et désintégration de particules

- *Centre de masses.*

Considérons un ensemble de points matériels (“particules”) de masses m_i , $i = 1, \dots, N$, dotés d’énergies et d’impulsions (E_i, \mathbf{p}_i) dans un certain référentiel, dit repère du laboratoire (*laboratory frame*). Notons $P^\mu = \sum_i p^\mu$ la (quadri-)impulsion totale. C’est un vecteur de genre temps futur si tous les p_i le sont (pourquoi ?). On appelle repère du centre de masses un référentiel inertiel dans lequel les nouvelles énergies et impulsions (E'_i, \mathbf{p}'_i) sont telles que $\sum_i \mathbf{p}'_i = 0$. L’énergie $E' := \sum_i E'_i$ est l’“énergie totale dans le centre de masses”. Pour la calculer, faisons appel au fait que par changement de référentiel, $P^\mu \rightarrow P'^\mu$ de même norme carrée.

$$\frac{1}{c^2} E'^2 = P'^2 = P^2 = \frac{1}{c^2} \left(\sum_i E_i \right)^2 - \left(\sum_i \mathbf{p}_i \right)^2 =: s \quad (3.18)$$

(la notation $s = (\sum_i p_i)^2$ est traditionnelle). L’énergie totale dans le centre de masses est donc $c\sqrt{s}$.

- *Désintégration de particules*

Considérons le cas d’une particule de masse M se désintégrant en deux particules de masses m_1 et m_2 . Si P, p_1, p_2 désignent les quadri-impulsions des trois particules, on a $P = p_1 + p_2$.

On se place d’abord dans le repère où la particule initiale est au repos, qui est aussi le repère du centre de masse du système final. **Attention**, pour simplifier les expressions, on va faire $c = 1$ dans le reste de ce paragraphe.

État initial $P^\mu = (M, \mathbf{0})$, état final $p_1^\mu = (E_1, \mathbf{p})$, $p_2^\mu = (E_2, -\mathbf{p})$. Conservation de l’énergie-impulsion $E_1 + E_2 = M$, avec $\mathbf{p}^2 = E_1^2 - m_1^2 = E_2^2 - m_2^2$. Éliminant E_2 entre ces deux équations conduit à

$$\begin{aligned} E_1 &= \frac{M^2 + m_1^2 - m_2^2}{2M} & E_2 &= \frac{M^2 + m_2^2 - m_1^2}{2M} \\ \mathbf{p}^2 &= \frac{M^4 + m_1^4 + m_2^4 - 2m_1^2 m_2^2 - 2M^2 m_1^2 - 2M^2 m_2^2}{4M^2} \end{aligned} \quad (3.19)$$

et donc tous les éléments cinématiques en dehors de la direction de \mathbf{p} sont déterminés.

Si maintenant la particule se désintègre “en vol”, avec une certaine vitesse \mathbf{v} et une énergie E dans le “repère du laboratoire”, on a $|\mathbf{v}| = \sqrt{1 - M^2/E^2}$ selon les expressions (3.15). Choissant l’axe des x le long de \mathbf{v} , on applique les formules (1.5) avec $\gamma = E/M$ pour déterminer l’expression de E_i^{lab} et $\mathbf{p}_i^{\text{lab}}$ dans le laboratoire à partir de celles de E et \mathbf{p} dans le centre de masses

$$\begin{aligned} E_i^{\text{lab}} &= \frac{E}{M} E_i + p_{i,x} \sqrt{\frac{E^2}{M^2} - 1} \\ p_{1,x}^{\text{lab}} &= \frac{E}{M} p_x + E_1 \sqrt{\frac{E^2}{M^2} - 1} & p_{1,y}^{\text{lab}} &= p_y \\ p_{2,x}^{\text{lab}} &= -\frac{E}{M} p_x + E_2 \sqrt{\frac{E^2}{M^2} - 1} & p_{2,y}^{\text{lab}} &= -p_y \end{aligned} \quad (3.20)$$

Supposons que la particule se désintègre dans le plan x, y selon la direction θ , donc $p_z = 0$, et donc aussi $p_{i,z}^{\text{lab}} = 0$, $i = 1, 2$, et $p_x = p \cos \theta$, $p_y = p \sin \theta$; soit θ^{lab} l’angle d’émission de la particule 1 par rapport à

la particule incidente, donc $p_{1,x}^{\text{lab}} = p^{\text{lab}} \cos \theta^{\text{lab}}$, $p_{1,y}^{\text{lab}} = p^{\text{lab}} \sin \theta^{\text{lab}}$. Les formules précédentes fournissent θ^{lab} en termes de θ et de E

$$\text{tg } \theta^{\text{lab}} = \frac{p \sin \theta}{\frac{E}{M} p \cos \theta + E_1 \sqrt{\frac{E^2}{M^2} - 1}} .$$

Si la distribution angulaire est uniforme en θ (dans le centre de masses), $dn/d \cos \theta = \frac{1}{2} N$, N le nombre total de désintégrations observées, la distribution angulaire en θ^{lab} se ramène à un “simple” jacobien

$$\frac{dn}{d \cos \theta^{\text{lab}}} = \frac{d \cos \theta}{d \cos \theta^{\text{lab}}} \frac{dn}{d \cos \theta} = \frac{1}{2} N \frac{d \cos \theta}{d \cos \theta^{\text{lab}}} .$$

- *Collisions à deux corps*

On considère un processus de diffusion

$$1 + 2 \rightarrow 3 + 4$$

avec quatre particules de masses m_i et de quadri-impulsions p_i , $i = 1, \dots, 4$. Soit $P = p_1 + p_2 = p_3 + p_4$ l’impulsion totale conservée. Si on se place dans le référentiel du centre de masses, où $P = (E = \sqrt{s}, \mathbf{0})$, on est ramené à la discussion de la désintégration.

Combien d’invariants relativistes indépendants peut-on construire avec les quatre quadri-impulsions p_i^μ , compte tenu des conditions $p_i^2 = m_i^2$ (“la particule i est sur sa couche de masse”)? On a déjà vu $s = (p_1 + p_2)^2 = (p_3 + p_4)^2$. On définit les deux autres “variables de Mandelstam”

$$\begin{aligned} t &= (p_1 - p_3)^2 = (p_2 - p_4)^2 \\ u &= (p_1 - p_4)^2 = (p_2 - p_3)^2 . \end{aligned} \tag{3.21}$$

Montrer que $s+t+u = \sum_i m_i^2$ et que l’on ne peut construire d’autre invariant indépendant. Relier t à l’angle de diffusion (entre les particules 1 et 3) dans le repère du centre de masses.

4. Théorie classique des champs. Équations de Maxwell et relativité

Nous allons maintenant généraliser les considérations précédentes à des systèmes à nombre infini de degrés de liberté, les théories de champs.

4.1. Théorie des champs lagrangienne

En théorie des champs, classique pour le moment (telle l’électrodynamique), mais aussi quantique par la suite, on cherche à écrire le lagrangien comme une intégrale d’espace d’une *densité de lagrangien* \mathcal{L} , dépendant des champs (et autres variables dynamiques) ϕ et de

leurs dérivées premières, et éventuellement d'autres fonctions des coordonnées (sources extérieures, etc). L'action est donc de la forme générale

$$S = \int Ldt = \int dt \int d^3\mathbf{x} \mathcal{L}(\phi(\mathbf{x}, t), \partial_\mu\phi(\mathbf{x}, t); x) . \quad (4.1)$$

Dans une approche relativiste, l'action doit être invariante de Lorentz. La mesure $d^4x = cdt d^3\mathbf{x}$ l'étant, on va s'efforcer de construire une densité de lagrangien \mathcal{L} elle-même invariante de Lorentz.

N.B. Un nouvel abus de langage conduit alors à omettre le "densité de" et à appeler \mathcal{L} le lagrangien...

On peut alors écrire les équations d'Euler-Lagrange (ou équations du mouvement) qui expriment la stationnarité de cette action⁷

$$(E - L) \quad \frac{\delta S}{\delta\phi(x)} := \frac{\partial\mathcal{L}}{\partial\phi(x)} - \partial_\mu \left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi(x))} \right) = 0 . \quad (4.2)$$

Exemple. Pour un champ $\phi(x)$, un lagrangien simple peut être écrit sous la forme $\mathcal{L}(\phi) = \frac{1}{2}(\hbar\partial\phi)^2 - \frac{1}{2}m^2c^2\phi^2 - V(\phi)$ où $(\partial\phi)^2$ doit être compris au sens minkowskien $(\partial\phi)^2 = (\frac{1}{c}\dot{\phi})^2 - (\nabla\phi)^2$. Pour $V = 0$, l'équation du mouvement s'écrit

$$(\partial_0^2 - \Delta + \frac{m^2c^2}{\hbar^2})\phi(x) = 0 ,$$

c'est l'équation de Klein-Gordon sur laquelle on reviendra par la suite.

• *Formalisme hamiltonien*

On peut à nouveau définir le moment conjugué du champ $\phi(x)$

$$\pi(\mathbf{x}, t) = \frac{\delta L(\phi, \partial\phi)}{\delta(\partial_t\phi(\mathbf{x}, t))} , \quad (4.3)$$

avec une dérivée fonctionnelle par rapport à $\phi(\mathbf{x}, t)$ à temps fixé t , puis la densité d'hamiltonien

$$\mathcal{H}(\phi, \pi) = \pi(x)\partial_t\phi(x) - \mathcal{L}(\phi, \partial\phi) \quad (4.4)$$

et l'hamiltonien $H = \int d^3\mathbf{x}\mathcal{H}$. Les équations du mouvement peuvent s'écrire

$$\partial_t\phi(\mathbf{x}, t) = \frac{\delta H}{\delta\pi(\mathbf{x}, t)} \quad \partial_t\pi(\mathbf{x}, t) = -\frac{\delta H}{\delta\phi(\mathbf{x}, t)} \quad (4.5)$$

⁷ Bien noter la distinction de notations entre la *dérivée fonctionnelle* $\frac{\delta S}{\delta\phi(x)}$ et la dérivée partielle "ordinaire" $\frac{\partial\mathcal{L}}{\partial\phi(x)}$.

ou encore sous forme de crochets de Poisson,

$$\partial_t \pi = \{H, \pi\} \quad , \quad \partial_t \phi = \{H, \phi\} \quad (4.6)$$

avec le crochet de Poisson “canonique” à temps égaux

$$\{\pi(\mathbf{x}, t), \phi(\mathbf{y}, t)\} = \delta^3(\mathbf{x} - \mathbf{y}) \quad , \quad (4.7)$$

les autres crochets, toujours à temps égaux, s’annulant

$$\{\pi(\mathbf{x}, t), \pi(\mathbf{y}, t)\} = \{\phi(\mathbf{x}, t), \phi(\mathbf{y}, t)\} = 0 \quad .$$

4.2. Symétries, lois de conservation. Courant de Noether.

Montrons maintenant comment l’existence de symétries continues, c’est-à-dire de transformations des variables laissant l’action inchangée et dépendant de paramètres continus, implique celle de quantités conservées (énergie, impulsion, charge, ...). Cela apparaît simplement en formalisme lagrangien, pour un nombre fini de degrés de liberté q_i , cf § 3.1 ci-dessus. Pour une variation infinitésimale des coordonnées q_i

$$\begin{aligned} \delta S &= \int_{t_1}^{t_2} dt \sum_i \left(\delta q_i(t) \frac{\partial L}{\partial q_i} + \delta \dot{q}_i(t) \frac{\partial L}{\partial \dot{q}_i} \right) = \int_{t_1}^{t_2} dt (\dot{p}_i \delta q_i + p_i \delta \dot{q}_i) = \int_{t_1}^{t_2} dt \frac{d}{dt} \sum_i p_i \delta q_i \\ &= \sum_i p_i(t_2) \delta q_i(t_2) - \sum_i p_i(t_1) \delta q_i(t_1) \end{aligned} \quad (4.8)$$

(où on a utilisé la définition (3.3) de la variable conjuguée p_i et l’équation du mouvement (3.2)) et l’invariance $\delta S = 0$ implique que $\sum_i p_i(t_2) \delta q_i(t_2) = \sum_i p_i(t_1) \delta q_i(t_1)$, autrement dit que la quantité $\sum p_i \delta q_i$ est conservée entre les temps t_1 et t_2 . Des exemples classiques de cette situation sont donnés par la conservation de l’impulsion $\sum_i p_i$ s’il y a invariance par translation des coordonnées d’espace, de l’énergie H s’il y a invariance par translation du temps, du moment angulaire s’il y a invariance par rotation, etc.

En théorie lagrangienne des champs, toute invariance de l’action se traduit encore par une loi de conservation, mais le théorème de Noether nous dit en outre que cela s’exprime localement par la conservation d’un courant.

Anticipant un peu sur ce qui va être discuté dans la suite du cours, considérons la situation où un groupe G de transformations continues (plus précisément un “groupe de Lie”) agit sur les champs ϕ d’une théorie décrite par un lagrangien $\mathcal{L}(\phi)$. Dans un premier temps, nous supposons que ces transformations n’agissent pas sur les coordonnées

d'espace-temps et qu'elles agissent "localement" sur ϕ (c'est-à-dire en un même point x) : $g \in G \quad \phi(x) \mapsto \phi'(x) = {}^g\phi(x)$. Sous forme infinitésimale, on écrit $g = e^{t_s \alpha^s}$, où t_s sont les générateurs infinitésimaux (dans l'"algèbre de Lie" \mathfrak{g}) de G , et $\phi' = \phi + \delta\phi(x)$, avec

$$\delta\phi(x) = \alpha^s \mathcal{T}_s(\phi(x)) . \quad (4.9)$$

On peut alors répéter le calcul du nombre fini de degrés de liberté

$$\delta\mathcal{L} = \alpha^s \partial_\mu \left(\frac{\partial\mathcal{L}(\phi, \partial\phi)}{\partial\partial_\mu\phi} \mathcal{T}_s(\phi(x)) \right) , \quad (4.10)$$

où on a utilisé les équations d'Euler-Lagrange (4.2). Si le lagrangien est invariant $\delta\mathcal{L} = 0$ pour tout α^s , on a

$$\begin{aligned} \partial_\mu j_s^\mu(x) &= 0 , \\ \text{où } j_s^\mu(x) &= \frac{\partial\mathcal{L}(\phi, \partial\phi)}{\partial\partial_\mu\phi} \mathcal{T}_s(\phi(x)) . \end{aligned} \quad (4.11)$$

Si l'action est invariante sans que \mathcal{L} le soit, on suppose qu'on peut écrire $\delta\mathcal{L} = \alpha^s \partial_\mu J_s^\mu$ sous forme d'une dérivée totale, impliquant bien l'annulation de son intégrale δS (on suppose qu'il n'y a pas de termes de bord...). Avec la définition précédente de j_s^μ , c'est alors par $\partial_\mu(j_s^\mu - J_s^\mu) = 0$ qu'on exprime la loi de conservation du courant.

Une autre méthode consiste à effectuer dans l'action, au voisinage d'une configuration de champs stationnaire (une solution des équations d'Euler-Lagrange), une transformation de la forme (4.9), mais où α est maintenant "locale"⁸, c'est-à-dire dépend de x

$$\delta\phi(x) = \alpha^s(x) \mathcal{T}_s(\phi(x)) . \quad (4.12)$$

On suppose que $\alpha(x)$ s'annule aux bords du domaine d'intégration (à l'infini), pour pouvoir intégrer par parties. Écrivons $\widehat{\mathcal{L}}(\phi, \partial\phi, \alpha, \partial\alpha) := \mathcal{L}(\phi + \delta\phi, \partial(\phi + \delta\phi))$. L'action étant stationnaire, on a (après une intégration par parties)

$$0 = \int d^4x \alpha^s(x) \left(\frac{\partial\widehat{\mathcal{L}}}{\partial\alpha^s(x)} - \partial_\mu \frac{\partial\widehat{\mathcal{L}}}{\partial(\partial_\mu\alpha^s(x))} \right) .$$

Comme α^s est quelconque, on en tire

$$\frac{\partial\widehat{\mathcal{L}}}{\partial\alpha^s(x)} = \partial_\mu \frac{\partial\widehat{\mathcal{L}}}{\partial(\partial_\mu\alpha^s(x))}$$

⁸ Noter les multiples utilisations du mot "local" dans des sens légèrement différents...

qui n'est autre que l'équation du mouvement (E-L) pour les variations (4.12) des champs. Si on définit le courant de Noether par

$$\boxed{j_s^\mu(x) := \frac{\partial \widehat{\mathcal{L}}}{\partial(\partial_\mu \alpha^s(x))}} \quad (4.13)$$

nous venons de montrer que

$$\partial_\mu j_s^\mu(x) = \frac{\partial \widehat{\mathcal{L}}}{\partial \alpha^s(x)} \quad (4.14)$$

en vertu de l'équation du mouvement. En fait $\frac{\partial \widehat{\mathcal{L}}}{\partial \alpha^s(x)} \equiv \frac{\partial \widehat{\mathcal{L}}}{\partial \alpha^s}$, la dérivée par rapport à des α^s indépendants de x , et on peut récrire l'identité précédente comme

$$\boxed{\partial_\mu j_s^\mu(x) = \frac{\partial \widehat{\mathcal{L}}}{\partial \alpha^s}} \quad (4.15)$$

L'avantage de cette deuxième méthode par rapport à la précédente est qu'elle nous fournit une information supplémentaire sur la divergence du courant quand elle n'est pas nulle. Discuter l'équivalence entre les deux définitions du courant de Noether (4.11) et (4.13).

Si le lagrangien est invariant sous l'action des transformations (4.12) "globales", c'est-à-dire avec α^s indépendant de x , alors le membre de droite s'annule, $\partial \widehat{\mathcal{L}} / \partial \alpha^s = 0$, et

$$\partial_\mu j_s^\mu(x) = 0 \quad (4.16)$$

Par abus de langage, on dit que le courant est "conservé". En fait la loi de conservation se lit sur la charge

$$Q_s = \int d^3x j_s^0(\mathbf{x}, t) \quad (4.17)$$

qui *a priori* pourrait dépendre du temps t mais qui en fait est conservée

$$\frac{d}{dt} Q_s = c \int d^3x \partial_0 j_s^0(\mathbf{x}, t) = -c \int d^3x \nabla \cdot \mathbf{j}_s(\mathbf{x}, t) = 0 \quad (4.18)$$

Si l'action est invariante mais le lagrangien ne l'est qu'à une dérivée totale près : $\delta \mathcal{L} = \alpha^s \partial_\mu J_s^\mu$, on écrit plutôt $\partial_\mu j_s^\mu = \partial_\mu J_s^\mu$ et c'est $j_s^\mu - J_s^\mu$ qui est le courant conservé.

Récapitulons : Nous nous sommes intéressés aux transformations du champ de la forme (4.9). L'artifice de rendre ces transformations locales, selon (4.12), a suggéré une

définition naturelle d'un "courant" $j_s^\mu(x)$. Alors, l'invariance de \mathcal{L} sous (4.9) implique une loi de conservation (grâce aux équations du mouvement), $\partial_\mu j_s^\mu = 0$, qui implique elle-même que les charges $Q_s = \int d^3x j_s^0(\mathbf{x}, t)$ ne dépendent pas du temps. Il y a autant de telles lois de conservation indépendantes qu'il y a de transformations infinitésimales indépendantes, (c'est-à-dire leur nombre égale la "dimension de l'algèbre de Lie \mathfrak{g} ").

Exercices.

1. Champ scalaire complexe, de lagrangien $\mathcal{L} = \hbar^2(\partial\phi)^*(\partial\phi) - m^2c^2\phi^*\phi - \frac{1}{2}g(\phi^*\phi)^2$. Montrer que le groupe $U(1)$ agissant par $\phi(x) \rightarrow e^{-iq\alpha}\phi(x)$ est une invariance de \mathcal{L} . Calculer le courant de Noether et en vérifier la conservation.

2. Vérifier que dans le formalisme hamiltonien, $Q_s(\pi, \phi) = \int d^3x \pi(\mathbf{x}, t)\mathcal{T}_s(\phi(\mathbf{x}, t))$ engendre au sens des crochets de Poisson la variation (4.9) $\delta\phi(x) = \alpha_s\{Q_s, \phi(x)\}$. En Mécanique Quantique (et en théorie quantique des champs), ces relations impliquent des commutateurs d'opérateurs à la place des crochets de Poisson : $-i\hbar\{.,.\} \rightarrow [.,.]$.

• *Tenseur d'énergie-impulsion* (en anglais *energy-momentum tensor* ou *stress-energy tensor*)

Examinons maintenant l'effet de transformations agissant sur les coordonnées des champs. Considérons d'abord l'effet d'une translation infinitésimale : $\delta x^\mu = a^\mu$. Dans l'esprit de la dérivation précédente du théorème de Noether, nous la prenons "locale", $\delta x^\mu = a^\mu(x)$, ce qui induit une transformation infinitésimale de ϕ (cf ci-dessous (4.25))

$$\delta\phi(x) = \phi(x) - \phi(x-a) = -a^\mu(x)\partial_\mu\phi(x) \quad \delta\partial_\mu\phi(x) = -a^\nu\partial_\mu\partial_\nu\phi(x) - \partial_\mu(a^\nu(x))\partial_\nu\phi(x).$$

Sous l'effet de cette variation, l'action $S = \int d^4x \mathcal{L}(\phi(x), \partial\phi(x))$, où toute la dépendance de l'intégrand en x est celle de ϕ et $\partial\phi$, (ce qui exclut la présence d'une source extérieure), varie de

$$\begin{aligned} \delta S &= - \int d^4x \left\{ a^\nu \underbrace{\left(\frac{\partial\mathcal{L}}{\partial\phi(x)}\partial_\nu\phi(x) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi(x))}\partial_\nu\partial_\mu\phi(x) \right)} + (\partial_\mu a^\nu) \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi(x))}\partial_\nu\phi(x) \right\} \\ &= - \int d^4x \left\{ \begin{array}{ccc} a^\nu & \partial_\nu\mathcal{L} & + (\partial_\mu a^\nu) \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi(x))}\partial_\nu\phi(x) \end{array} \right\} \\ &= - \int d^4x \left\{ \partial_\mu a^\nu(x) \left(-\delta^\mu_\nu\mathcal{L} + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\partial_\nu\phi(x) \right) + \partial_\nu(a^\nu(x)\mathcal{L}) \right\} \end{aligned} \quad (4.19)$$

Dans la dernière ligne de (4.19), on néglige le dernier terme (dérivée totale), et on lit alors la variation de l'action sous l'effet de la transformation locale $\xi^\mu = a^\mu(x)$

$$\delta S = - \int d^4x (\partial_\mu a_\nu)\tilde{\Theta}^{\mu\nu}(x) \quad (4.20)$$

où on a identifié selon la méthode précédente du théorème de Noether le “courant”

$$\boxed{\tilde{\Theta}^{\mu\nu} = \frac{\delta S}{\delta \partial_\mu a_\nu(x)} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial^\nu \phi - \mathcal{L} g^{\mu\nu} .}$$
(4.21)

Selon le principe d’action, (donc en utilisant les équations du mouvement, ce que nous n’avons pas encore fait à ce stade), la variation δS est nulle pour tout $a^\nu(x)$ et il en découle une loi de conservation : $\partial_\mu \tilde{\Theta}^{\mu\nu} = 0$. Le tilde signale que ce tenseur énergie-impulsion n’a peut-être pas encore sa forme optimale, en particulier, il n’est pas toujours symétrique en μ, ν contrairement à ce qu’on souhaiterait. (Il pourra par la suite être éventuellement modifié, “amélioré”, de façon à réaliser cette symétrie sans affecter sa conservation.) Les intégrales d’espace de $\tilde{\Theta}$ donnent

$$P^\nu = \int d^3x \tilde{\Theta}^{0\nu}(\mathbf{x}, t)$$
(4.22)

indépendant de t . C’est la loi de conservation du quadrivecteur énergie-impulsion, générateur des translations d’espace-temps. L’intégrale (4.22) est faite sur une surface de temps constant, ou bien plus généralement sur toute surface de genre-espace. La conservation de l’énergie-impulsion est la conséquence de l’invariance par translation (d’espace-temps) de la théorie, implicite dans l’hypothèse que \mathcal{L} n’a pas de dépendance explicite en x .

Examinons maintenant l’invariance de Lorentz. Une transformation de Lorentz infinitésimale est de la forme

$$x'^\mu = \Lambda^\mu_\nu x^\nu = x^\mu + \delta x^\mu = x^\mu + \omega^{\mu\nu} x_\nu ,$$
(4.23)

avec $\omega^{\mu\nu}$ antisymétrique, indépendant de x et infinitésimal. Pour toute grandeur physique f dépendant localement de la coordonnée, on note $f'(x')$ ou ${}^\Lambda f(\Lambda x)$ la fonction transformée décrivant cette même grandeur dans les nouvelles coordonnées. Pour un champ “scalaire” ϕ , toute la dépendance en ω vient du changement d’argument :

$$\phi'(x') = \phi(x) \Leftrightarrow \phi'(x) \equiv {}^\Lambda \phi(x) = \phi(\Lambda^{-1}x) .$$

(Bien noter l’argument $\Lambda^{-1}x$ et non pas Λx , ce qui est compatible avec la composition $\phi \xrightarrow{\Lambda} \phi' = \Lambda \phi \xrightarrow{\Lambda'} \phi'' = \Lambda' \phi' = \Lambda' \Lambda \phi$.) Mais en général, ϕ subit aussi une transformation

linéaire non triviale due à son spin, (on dira plus tard que “ ϕ porte une représentation non triviale du groupe de Lorentz”), et on écrit

$$\Lambda\phi(x) = S(\Lambda)\phi(\Lambda^{-1}x) , \quad (4.24)$$

ou de façon plus explicite, en composantes

$$\Lambda\phi_\alpha(x) = S(\Lambda)_\alpha^\beta\phi_\beta(\Lambda^{-1}x) ,$$

avec $S(\Lambda)$ une matrice donnée. Sous forme infinitésimale, $(\Lambda^{-1}x)^\mu = x^\mu - \omega^{\mu\nu}x_\nu$, $S(\Lambda) = I + S_{\mu\nu}\omega^{\mu\nu}$ et on peut choisir $S_{\mu\nu}$ antisymétrique en μ, ν , puisque $\omega^{\mu\nu}$ l’est. On écrit

$$\begin{aligned} \delta\phi(x) &:= \Lambda\phi(x) - \phi(x) = \omega^{\mu\nu}(S_{\mu\nu} - x_\nu\partial_\mu)\phi(x) \\ &= \omega^{\mu\nu} \left(S_{\mu\nu} - \frac{1}{2}(x_\nu\partial_\mu - x_\mu\partial_\nu) \right) \phi(x) . \end{aligned} \quad (4.25)$$

Calculons alors la variation du lagrangien de deux manières différentes. D’une part, si la théorie est invariante de Lorentz, le lagrangien est scalaire $\mathcal{L}'(x') = \mathcal{L}(x)$, soit

$$\begin{aligned} \delta\mathcal{L} &= \mathcal{L}(\Lambda^{-1}x) - \mathcal{L}(x) = -\omega^{\mu\nu}x_\nu\partial_\mu\mathcal{L} = -\frac{1}{2}\omega^{\mu\nu}(x_\nu\partial_\mu - x_\mu\partial_\nu)\mathcal{L} \\ &= -\frac{1}{2}\omega^{\mu\nu}\partial^\rho(x_\nu g_{\mu\rho}\mathcal{L} - x_\mu g_{\nu\rho}\mathcal{L}) ; \end{aligned} \quad (4.26)$$

d’autre part, toute la variation de \mathcal{L} venant de sa dépendance dans le champ ϕ ,

$$\begin{aligned} \delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial\partial_\rho\phi}\partial_\rho\delta\phi = \partial_\rho \left(\frac{\partial\mathcal{L}}{\partial\partial_\rho\phi}\delta\phi \right) \quad \Leftarrow \text{(E-L)} \\ &= \frac{1}{2}\omega^{\mu\nu}\partial_\rho \left\{ \frac{\partial\mathcal{L}}{\partial\partial_\rho\phi} \left(2S_{\mu\nu} + (x_\mu\partial_\nu - x_\nu\partial_\mu) \right) \phi \right\} \quad \Leftarrow \text{(4.25)} \end{aligned} \quad (4.27)$$

si bien qu’en rapprochant les deux expressions de $\delta\mathcal{L}$, on obtient que

$$\omega^{\mu\nu}\partial^\rho \left\{ 2S_{\mu\nu}\phi \frac{\partial\mathcal{L}}{\partial\partial^\rho\phi} + \left[x_\mu \underbrace{\left(\frac{\partial\mathcal{L}}{\partial\partial^\rho\phi}\partial_\nu\phi - g_{\nu\rho}\mathcal{L} \right)}_{\tilde{\Theta}_{\rho\nu}} - (\mu \leftrightarrow \nu) \right] \right\} = 0 .$$

Par conséquent, si on définit le tenseur à trois indices, antisymétrique en μ, ν ,

$$\boxed{M^{\rho, \mu\nu} = x^\mu \tilde{\Theta}^{\rho\nu} - x^\nu \tilde{\Theta}^{\rho\mu} + 2(S^{\mu\nu}\phi) \frac{\partial\mathcal{L}}{\partial(\partial_\rho\phi)}} \quad (4.28)$$

alors l'invariance de Lorentz se traduit par

$$\partial_\rho M^{\rho, \mu\nu} = 0 \quad (4.29)$$

et la charge correspondante est

$$J^{\mu\nu} = \int d^3x M^{0, \mu\nu} = \int d^3x \left\{ x^\mu \tilde{\Theta}^{0\nu} - x^\nu \tilde{\Theta}^{0\mu} + 2(S^{\mu\nu} \phi) \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi)} \right\}. \quad (4.30)$$

C'est le moment angulaire généralisé, générateur des transformations de Lorentz infinitésimales. Le premier terme représente la partie "orbitale", le second le moment angulaire intrinsèque ou spin.

$J^{\mu\nu}$ n'est pas invariant sous l'effet des translations (par exemple, il ne commute pas au sens du crochet de Poisson avec P^ρ). Il est approprié de construire la combinaison suivante de ces opérateurs

$$W_\mu = -\frac{1}{2} \epsilon_{\mu\nu\rho\sigma} J^{\nu\rho} P^\sigma, \quad (4.31)$$

appelé vecteur de Pauli-Lubanski. Noter que ce vecteur, orthogonal à P^μ de genre temps, est de genre espace. Dans un référentiel de repos où $P^\mu = (m, \mathbf{0})$, ce vecteur n'a pas de composante temporelle et se réduit à $W^i = \frac{1}{2} m \epsilon_{ijk} J^{jk}$, c'est le moment angulaire (total) usuel (au facteur m près).

Calculant explicitement (4.29), on trouve

$$\tilde{\Theta}^{\mu\nu} - \tilde{\Theta}^{\nu\mu} + 2\partial_\rho \left((S^{\mu\nu} \phi) \frac{\partial \mathcal{L}}{\partial(\partial_\rho \phi)} \right) = 0. \quad (4.32)$$

Construisons alors le nouveau tenseur

$$\Theta^{\mu\nu} = \tilde{\Theta}^{\mu\nu} + \partial_\rho \left((S^{\mu\nu} \phi) \frac{\partial \mathcal{L}}{\partial(\partial_\rho \phi)} - (S^{\rho\nu} \phi) \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} - (S^{\rho\mu} \phi) \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)} \right). \quad (4.33)$$

Observons que

- (i) l'expression entre parenthèses dans (4.32) est antisymétrique en μ, ρ , et donc $\Theta^{\mu\nu}$ est aussi conservé $\partial_\mu \Theta^{\mu\nu} = \partial_\rho \tilde{\Theta}^{\mu\nu} = 0$;
- (ii) $\Theta^{0\nu}$ diffère de $\tilde{\Theta}^{0\nu}$ par une dérivée spatiale totale, et conduit donc à la même intégrale (4.22) ;
- (iii) la partie antisymétrique de $\Theta^{\mu\nu}$ s'annule en vertu de (4.32) :

$$\Theta^{\mu\nu} - \Theta^{\nu\mu} = \tilde{\Theta}^{\mu\nu} - \tilde{\Theta}^{\nu\mu} + 2\partial_\rho \left((S^{\mu\nu} \phi) \frac{\partial \mathcal{L}}{\partial(\partial_\rho \phi)} \right) = 0.$$

Le nouveau tenseur $\Theta^{\mu\nu}$ se qualifie donc parfaitement comme nouveau tenseur énergie-impulsion. C'est la *tenseur de Belinfante*.

4.3. Invariance de Lorentz des équations de Maxwell

Un mot sur le système d'unités : on choisit le système tel que la loi de Coulomb s'écrive $F = qq'/4\pi r^2$ (donc $\epsilon_0 = 1$). Nous gardons dans le reste de ce chapitre la vitesse de la lumière explicite, avant de faire par la suite $c = 1$.

En présence de sources extérieures décrites par des densités de charge $\rho(t, \mathbf{x})$ et de courant $\mathbf{j}(t, \mathbf{x})$, les équations de Maxwell satisfaites par le champ électrique \mathbf{E} et l'induction magnétique \mathbf{B} s'écrivent

$$\begin{aligned} \text{(a)} \quad \operatorname{div} \mathbf{E} &= \rho & \text{(b)} \quad \operatorname{rot} \mathbf{B} - \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} &= \frac{1}{c^2} \mathbf{j} \\ \text{(c)} \quad \operatorname{div} \mathbf{B} &= 0 & \text{(d)} \quad \operatorname{rot} \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} &= 0 \end{aligned} \quad (4.34)$$

avec la loi de conservation de la charge électrique

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \mathbf{j} = 0 . \quad (4.35)$$

Rappelons que ces équations expriment de façon locale les grandes lois de l'électromagnétisme :

- (a) $\int_S \mathbf{E} \cdot d\mathbf{S} = \int \rho(x) d^3x$: le flux de \mathbf{E} à travers la surface fermée S égale la charge intérieure (loi de Gauss) ;
- (b) $\oint_C \mathbf{B} \cdot d\mathbf{x} = \frac{1}{c^2} \int_S (\mathbf{j} + \frac{\partial \mathbf{E}}{\partial t}) \cdot d\mathbf{S}$: la circulation de \mathbf{B} le long d'une courbe fermée C qui borde une surface S égale (au facteur c^2 près) le flux à travers S de la somme du courant \mathbf{j} et du "courant de déplacement de Maxwell" $\frac{\partial \mathbf{E}}{\partial t}$;
- (c) $\int_S \mathbf{B} \cdot d\mathbf{S} = 0$: le flux de \mathbf{B} à travers toute surface fermée est nul, il n'y a pas de charges magnétiques isolées (monopoles) ;
- (d) $\int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} = - \oint_C \mathbf{E} \cdot d\mathbf{x}$: un flux magnétique variable engendre une force électromotrice (loi de Faraday).

Selon l'observation de Lorentz, éclairée par le principe de relativité d'Einstein, ces équations de Maxwell sont invariantes par l'action du groupe $O(3,1)$, pourvu que le quadricourant défini comme

$$j^\mu = (c\rho, \mathbf{j})$$

se transforme comme un quadrivecteur, et que les champs \mathbf{E} et \mathbf{B} se transforment ... de façon adéquate mais un peu compliquée à mémoriser ! Les choses deviennent beaucoup

plus simples si on utilise des notations minkowskienne compactes pour ces champs. On définit le tenseur de champ électromagnétique

$$F^{\mu\nu} = \begin{pmatrix} 0 & -\frac{1}{c}E^1 & -\frac{1}{c}E^2 & -\frac{1}{c}E^3 \\ \frac{1}{c}E^1 & 0 & -B^3 & B^2 \\ \frac{1}{c}E^2 & B^3 & 0 & -B^1 \\ \frac{1}{c}E^3 & -B^2 & B^1 & 0 \end{pmatrix}. \quad (4.36)$$

Il s'agit d'un tenseur antisymétrique, $F^{\mu\nu} = -F^{\nu\mu}$. Le tenseur dual, lui aussi antisymétrique, est défini via le tenseur complètement antisymétrique ϵ

$$\tilde{F}^{\mu\nu} = -\tilde{F}^{\nu\mu} := \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}F_{\rho\sigma} = \begin{pmatrix} 0 & -B^1 & -B^2 & -B^3 \\ B^1 & 0 & \frac{1}{c}E^3 & -\frac{1}{c}E^2 \\ B^2 & -\frac{1}{c}E^3 & 0 & \frac{1}{c}E^1 \\ B^3 & \frac{1}{c}E^2 & -\frac{1}{c}E^1 & 0 \end{pmatrix}. \quad (4.37)$$

Il est donc obtenu à partir de F en changeant $\frac{1}{c}\mathbf{E} \mapsto \mathbf{B}$, $\mathbf{B} \mapsto -\frac{1}{c}\mathbf{E}$.

C'est un exercice aisé de vérifier que les équations de Maxwell prennent la forme simple ⁹

$$\partial_\mu F^{\mu\nu} = \frac{1}{c^2}j^\nu \quad \partial_\mu \tilde{F}^{\mu\nu} = 0 \quad (4.38)$$

et la compatibilité de la première avec l'antisymétrie est assurée par la loi de conservation de j qui s'écrit

$$\partial_\mu j^\mu = 0. \quad (4.39)$$

Sous la forme (4.38), l'invariance de Lorentz des équations de Maxwell devient manifeste, à condition de supposer que le tenseur de champ F (et donc aussi \tilde{F}) et le courant j se transforment comme un tenseur (resp. un vecteur).

Exercice : à partir de la loi de transformation de F , retrouver celle de \mathbf{E} et \mathbf{B} sous l'effet des transformations de $O(3,1)$. Vérifier qu'ils se transforment bien comme des vecteurs sous l'effet des rotations de $O(3)$, et examiner l'effet des transformations de Lorentz spéciales (2.6).

• *Invariants du champ électromagnétique*

Par contraction des indices, on construit divers invariants

$$\begin{aligned} F_{\mu\nu}F^{\mu\nu} &= -2\left(\frac{1}{c^2}\mathbf{E}^2 - \mathbf{B}^2\right) \\ cF_{\mu\nu}\tilde{F}^{\mu\nu} &= -4\mathbf{E}\cdot\mathbf{B} \end{aligned} \quad (4.40)$$

ce qui indique par exemple que si $\mathbf{E} \perp \mathbf{B}$ ou si $|\mathbf{E}| = c|\mathbf{B}|$ dans un référentiel, il en est de même dans tout autre référentiel.

⁹ On rappelle que la coordonnée $x^0 = ct$ et donc $\partial_0 = \frac{1}{c}\partial_t$.

La deuxième des lois de Maxwell, $\partial \tilde{F} = 0$ autorise à écrire le tenseur de champ en termes d'un nouveau champ $A^\mu(x)$, le *potentiel vecteur*

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu, \quad (4.41)$$

soit

$$\mathbf{E} = -c\nabla A^0 - \frac{\partial \mathbf{A}}{\partial t} \quad \mathbf{B} = \mathbf{rot} \mathbf{A} .$$

(Noter que le potentiel scalaire usuel est $\phi = cA^0$.) Si on suppose que $A = (A^\mu)$ est un vecteur de Lorentz, il en découle le caractère tensoriel de F . Ce potentiel A n'est pas défini de façon unique. La transformation

$$A^\mu(x) \rightarrow A^\mu(x) + \partial^\mu \phi(x), \quad (4.42)$$

dite transformation de jauge, laisse le tenseur F de (4.41) inchangé pour toute fonction $\phi(x)$. Ces transformations de jauge vont jouer un rôle considérable par la suite, puisque c'est le champ A qui se prête bien au formalisme lagrangien et à la quantification.

4.4. Lagrangien de l'électrodynamique

L'action

$$S = \int d^4x \left(-\frac{c}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{c} j_\mu A^\mu \right) = \int d^4x \left(-\frac{c}{4} (\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) - \frac{1}{c} j_\mu A^\mu \right)$$

conduit aux bonnes équations (4.38). Pour une particule de charge q dont la trajectoire d'espace-temps $x^\mu(\tau)$ est donnée, la densité de charge s'écrit $\rho(y) = \rho(\mathbf{y}, ct) = q\delta^3(\mathbf{y} - \mathbf{x}(\tau))|_{ct=x^0(\tau)}$ et la densité de courant $j^\mu(y) = q\frac{dx^\mu}{dt}\delta^3(\mathbf{y} - \mathbf{x}(\tau))|_{ct=x^0(\tau)} = qc \int d\tau \frac{dx^\mu}{d\tau} \delta^4(y - x(\tau))$, donc $\frac{1}{c} \int d^4y j_\mu(y) A^\mu(y) = q \int dx^\mu A_\mu(x(\tau))$, intégrale curviligne le long de la trajectoire d'espace-temps. En ajoutant l'action écrite plus haut ((3.9) et sq), on a finalement l'action du système couplé champ-particule chargée

$$\begin{aligned} S &= -c \int d^4x \frac{1}{4} F^2 - q \int dx^\mu A_\mu(x(\tau)) - mc \int ds \\ &= -c \int d^4x \frac{1}{4} F^2 + \int dt \left(-mc^2 \sqrt{1 - \mathbf{v}^2/c^2} + q\mathbf{A} \cdot \mathbf{v} - qcA^0 \right). \end{aligned} \quad (4.43)$$

L'intégrand du deuxième terme est le lagrangien L_{ch} de la particule chargée en présence du champ, et on en déduit le moment conjugué \mathbf{p}

$$\mathbf{p} = \frac{\partial L_{\text{ch}}}{\partial \mathbf{v}} = \frac{m\mathbf{v}}{\sqrt{1 - \mathbf{v}^2/c^2}} + q\mathbf{A},$$

et l'hamiltonien

$$H = \mathbf{v} \cdot \frac{\partial L_{\text{ch}}}{\partial \mathbf{v}} - L_{\text{ch}} = \frac{mc^2}{\sqrt{1 - \mathbf{v}^2/c^2}} + qcA^0 = (m^2c^4 + c^2(\mathbf{p} - q\mathbf{A})^2)^{\frac{1}{2}} + qcA^0 .$$

Noter que la relation $(H - qcA^0)^2 = m^2c^4 + c^2(\mathbf{p} - q\mathbf{A})^2$ s'obtient à partir de la relation correspondante pour la particule libre, $E^2 = (mc^2)^2 + (c\mathbf{p})^2$ (cf §3.2) via le “couplage minimal”

$$p^\mu = \left(\frac{E}{c}, \mathbf{p}\right) \rightarrow (p - qA)^\mu . \quad (4.44)$$

Quant à l'équation du mouvement, on peut lui donner la forme covariante

$$m \frac{du^\mu}{d\tau} = qF^{\mu\nu} u_\nu$$

en termes du quadrivecteur vitesse u^μ , ou encore la forme plus classique ¹⁰

$$\begin{aligned} \frac{d}{dt} \frac{m\mathbf{v}}{\sqrt{1 - \mathbf{v}^2/c^2}} &= q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \\ \frac{dE}{dt} &= q\mathbf{E} \cdot \mathbf{v} \end{aligned}$$

où la première équation est la force de Lorentz, et la seconde rappelle que seul le champ électrique fournit du travail !

5. Rotation du minkowskien vers l'euclidien...

Il est souvent souhaitable d'effectuer une transformation des coordonnées qui ramène la forme indéfinie minkowskienne à une forme définie positive (ou négative) euclidienne. Cela est souhaitable par exemple pour la comparaison entre les deux situations physiques, celle –minkowskienne– rencontrée en physique des particules, et celle euclidienne, dans un contexte de mécanique statistique ou physique des solides. C'est aussi utile pour des raisons techniques, car il est souvent plus facile, ou mieux défini, de travailler “dans l'euclidien” plutôt que dans la métrique minkowskienne.

Cette transformation impose de complexifier le temps : la variable $x_0 = ct$ est complexifiée en $x_0 = \pm ix_4$, ou de façon duale, pour la variable conjuguée d'énergie, $k_0 \rightarrow \pm ik_4$. De savoir s'il est légitime d'effectuer ce changement de variable (par exemple une rotation de contour d'intégration), et dans quel sens l'effectuer (choix du \pm !), nécessite de connaître

¹⁰ Ne pas confondre l'énergie E sans indice avec E_i composante du champ \mathbf{E} .

une certaine information sur les propriétés d'analyticité de la fonction dans laquelle on l'effectue. De tout cela, il sera question plus en détail dans les cours de théorie des champs. On donne le nom de *rotation de Wick* à cette rotation du minkowskien vers l'eulidien.

Bibliographie sommaire

Les deux articles originaux d'A. Einstein, *Annalen der Physik*, vol XVII (1905), pages 891-921, et vol XVIII (1905) p. 639-641 se lisent très bien ! Ils sont traduits et annotés dans les *Œuvres choisies d'Albert Einstein* éditées par F. Balibar, O. Darrigol et B. Jech, Seuil/CNRS 1994. Ils sont également reproduits dans le fascicule *Einstein 1905-2005*, Séminaire Poincaré 9/04/2005, qui contient aussi des articles variés sur l'histoire (O. Darrigol) et les tests de la relativité (C.M. Will, gr-qc/0504085), la géométrie de l'espace de Minkovski, (J. Bros et U. Moschella), etc.

Les textes classiques, tels W. Panofsky et M. Phillips, *Classical Electricity and Magnetism*, Addison-Wesley ; L. Landau et E. Lifschitz [LL] ; J.D. Jackson, *Classical Electrodynamics*, John Wiley & Sons, contiennent des exposés très détaillés sur l'électrodynamique classique et la relativité, . . . mais attention aux conventions, facteurs c , etc ! . . .

Il est bon de comparer les discussions du théorème de Noether et du tenseur énergie-impulsion par les différents auteurs [BC, Bdf, IZ, LL, Wf . . .].

Chapitre 00

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

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

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 sont représentées dans une base orthonormée par des matrices 3×3 orthogonales réelles, de déterminant 1 : elles forment le groupe $\text{SO}(3)$.

Formule d'O. 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

$$\mathbf{x}' = R_{\mathbf{n}}(\psi)\mathbf{x} = \cos \psi \mathbf{x} + (1 - \cos \psi)(\mathbf{x} \cdot \mathbf{n})\mathbf{n} + \sin \psi (\mathbf{n} \wedge \mathbf{x}) . \quad (1.1)$$

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 figure 1) un élément de $\text{SO}(3)$ est paramétrisé par 3 variables continues. On prendra ainsi

$$0 \leq \theta \leq \pi, \quad 0 \leq \phi < 2\pi, \quad 0 \leq \psi \leq \pi . \quad (1.2)$$

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

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

$$\mathcal{R}_z(\psi) = \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (1.3)$$

tandis qu'autour des axes Ox et Oy

$$\mathcal{R}_x(\psi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & -\sin \psi \\ 0 & \sin \psi & \cos \psi \end{pmatrix} \quad \mathcal{R}_y(\psi) = \begin{pmatrix} \cos \psi & 0 & \sin \psi \\ 0 & 1 & 0 \\ -\sin \psi & 0 & \cos \psi \end{pmatrix} . \quad (1.3)'$$

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.

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

Une relation que nous allons abondamment utiliser est que

$$RR_{\mathbf{n}}(\psi)R^{-1} = R_{\mathbf{n}'}(\psi) \quad (1.4)$$

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 (1.4) : on dira plus tard que les "classes de conjugaison" du groupe $SO(3)$ sont caractérisées par l'angle ψ .

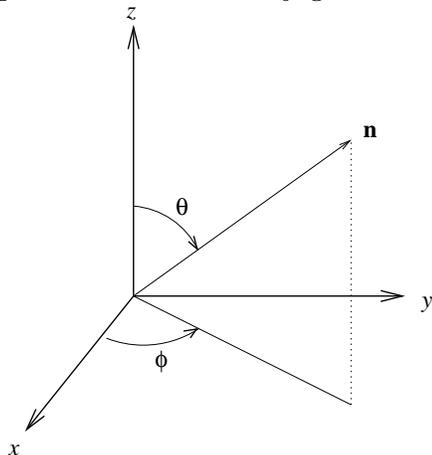


Fig. 1

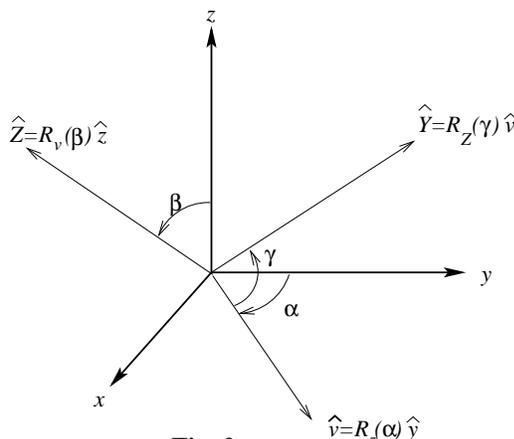


Fig. 2

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 Figure 2). On prend donc $0 \leq \alpha < 2\pi$, $0 \leq \beta \leq \pi$, $0 \leq \gamma < 2\pi$ et on écrit

$$R(\alpha, \beta, \gamma) = R_Z(\gamma)R_v(\beta)R_z(\alpha) \quad (1.5)$$

mais selon (1.4)

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

d'où en reportant dans (1.5)

$$R(\alpha, \beta, \gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma) . \quad (1.6)$$

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 (1.4), écrire l'expression d'une matrice R qui amène le vecteur unitaire \hat{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 (1.6) et en déduire les relations entre θ, ϕ, ψ et les angles d'Euler. (Voir aussi plus bas, équ. (3.19).)

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 $\sigma_i, i = 1, 2, 3$.

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{1.7}$$

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

$$\sigma_i \sigma_j = \delta_{ij} I + i \epsilon_{ijk} \sigma_k , \tag{1.8}$$

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

$$U = u_0 I - i \mathbf{u} \cdot \boldsymbol{\sigma} \tag{1.9}$$

qui est unitaire et de déterminant 1 (le vérifier et montrer aussi la réciproque : toute matrice unitaire 2×2 est de la forme (1.9), 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 (1.8), on peut vérifier que

$$e^{-i \frac{\psi}{2} \mathbf{n} \cdot \boldsymbol{\sigma}} = \cos \frac{\psi}{2} - i \sin \frac{\psi}{2} \mathbf{n} \cdot \boldsymbol{\sigma} . \tag{1.10}$$

Il est suggéré que la multiplication des matrices

$$U_n(\psi) = e^{-i\frac{\psi}{2}\mathbf{n}\cdot\boldsymbol{\sigma}} = \cos\frac{\psi}{2} - i\sin\frac{\psi}{2}\mathbf{n}\cdot\boldsymbol{\sigma}, \quad 0 \leq \psi \leq 2\pi, \quad \mathbf{n} \in S^2 \quad (1.11)$$

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

$$X = \mathbf{x}\cdot\boldsymbol{\sigma} = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix}, \quad (1.12)$$

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

$$X \mapsto X' = UXU^\dagger, \quad (1.13)$$

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

$$\det X = -(x_1^2 + x_2^2 + x_3^2) \quad (1.14)$$

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 $U = I$ 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 (1.8), le calcul explicite de X' conduit après un peu d'algèbre à

$$\begin{aligned} X' &= \left(\cos\frac{\psi}{2} - i\mathbf{n}\cdot\boldsymbol{\sigma}\sin\frac{\psi}{2}\right)X\left(\cos\frac{\psi}{2} + i\mathbf{n}\cdot\boldsymbol{\sigma}\sin\frac{\psi}{2}\right) \\ &= (\cos\psi\mathbf{x} + (1 - \cos\psi)(\mathbf{x}\cdot\mathbf{n})\mathbf{n} + \sin\psi(\mathbf{n} \wedge \hat{x}))\cdot\boldsymbol{\sigma} \end{aligned} \quad (1.15)$$

sur lequel on reconnaît la formule (1.1). On en conclut que la transformation $x \rightarrow x'$ effectuée par les matrices de $SU(2)$ dans (1.13) 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 (1.11)-(1.13).

2. Générateurs infinitésimaux. L'algèbre de Lie $su(2)$

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 en écrivant les générateurs infinitésimaux des rotations comme des opérateurs hermitiens $J = J^\dagger$. Ainsi on écrit

$$R_{\mathbf{n}}(d\psi) = (I - i d\psi J_{\mathbf{n}}) \quad (2.1)$$

où $J_{\mathbf{n}}$ est le générateur de ces rotations, 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,

$$R_{\mathbf{n}}(\psi + d\psi) = R_{\mathbf{n}}(d\psi)R_{\mathbf{n}}(\psi) = (I - i d\psi J_{\mathbf{n}})R_{\mathbf{n}}(\psi) , \quad (2.2)$$

ou encore

$$\frac{\partial R_{\mathbf{n}}(\psi)}{\partial \psi} = -i J_{\mathbf{n}} R_{\mathbf{n}}(\psi) \quad (2.3)$$

qui, compte tenu de $R(0) = I$, s'intègre en

$$\boxed{R_{\mathbf{n}}(\psi) = e^{-i\psi J_{\mathbf{n}}} .} \quad (2.4)$$

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 (1.3) on tire

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad J_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \quad J_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.5)$$

ce que l'on peut exprimer par une formule unique

$$\boxed{(J_k)_{ij} = -i\epsilon_{ijk}} \quad (2.6)$$

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

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.

Montrons que les trois matrices (2.5) forment une base des générateurs infinitésimaux et que $J_{\mathbf{n}}$ s'exprime simplement comme

$$J_{\mathbf{n}} = \sum_k J_k n_k \quad (2.7)$$

ce qui permet d'écrire (2.4) sous la forme

$$R_{\mathbf{n}}(\psi) = e^{-i\psi \sum_k n_k J_k} . \quad (2.8)$$

L'expression (2.7) découle simplement de la version infinitésimale de la formule d'O. Rodrigues, $R_{\mathbf{n}}(d\psi) = (I + d\psi \mathbf{n} \wedge)$ donc $-iJ_{\mathbf{n}} = \mathbf{n} \wedge$ ou si on préfère $-i(J_{\mathbf{n}})_{ij} = \epsilon_{ijk} n_k = n_k (-iJ_k)_{ij}$, cqfd.

Un commentaire sur (2.8) : 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 (1.6), on peut écrire toute rotation de $SO(3)$ sous la forme

$$R(\alpha, \beta, \gamma) = e^{-i\alpha J_3} e^{-i\beta J_2} e^{-i\gamma J_3} . \quad (2.9)$$

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

$$[J_i, J_j] = i\epsilon_{ijk} J_k \quad (2.10)$$

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

$$\epsilon_{iab}\epsilon_{bjc} + \epsilon_{icb}\epsilon_{baj} + \epsilon_{ijb}\epsilon_{bca} = 0 . \quad (2.11)$$

Exercice : bien comprendre la structure de cette identité et vérifier qu'elle implique (2.10).

Au vu de l'importance des relations (2.7)–(2.10), il est utile de les retrouver par une autre route. Notons d'abord que l'équation (1.4) implique que pour tout R

$$R e^{-i\psi J_{\mathbf{n}}} R^{-1} = e^{-i\psi R J_{\mathbf{n}} R^{-1}} = e^{-i\psi J_{\mathbf{n}'}} \quad (2.12)$$

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

$$R J_{\mathbf{n}} R^{-1} = J_{\mathbf{n}'}, \quad (2.13)$$

c'est-à-dire $J_{\mathbf{n}}$ se transforme comme le vecteur \mathbf{n} .

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

$$\epsilon_{lmn} \mathcal{R}_{il} \mathcal{R}_{jm} \mathcal{R}_{kn} = \epsilon_{ijk} \det \mathcal{R} = \epsilon_{ijk} \quad (2.14)$$

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

$$\epsilon_{lmn} \mathcal{R}_{jm} \mathcal{R}_{kn} = \epsilon_{ijk} \mathcal{R}_{il} \quad (2.15)$$

ce qui au vu de (2.6) exprime que

$$\mathcal{R}_{jm} (J_l)_{mn} \mathcal{R}_{nk}^{-1} = (J_i)_{jk} \mathcal{R}_{il} \quad (2.16)$$

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

$$R J_l R^{-1} = J_i \mathcal{R}_{il} . \quad (2.17)$$

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

$$J_{\mathbf{n}} \stackrel{(2.13)}{=} R J_3 R^{-1} \stackrel{(2.17)}{=} J_k \mathcal{R}_{k3} = J_k n_k , \quad (2.18)$$

qui n'est autre que (2.7). Noter que les équations (2.17) et (2.18) sont bien compatibles avec (2.13)

$$J_{\mathbf{n}'} \stackrel{(2.13)}{=} R J_{\mathbf{n}} R^{-1} \stackrel{(2.18)}{=} R J_k n_k R^{-1} \stackrel{(2.17)}{=} J_l \mathcal{R}_{lk} n_k = J_l n'_l .$$

Comme on le verra de façon plus systématique par la suite, la relation (2.10) 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

$$R_2(d\psi) J_1 R_2^{-1}(d\psi) \stackrel{(2.17)}{=} J_k [\mathcal{R}_2(d\psi)]_{k1} \quad (2.19)$$

mais au premier ordre, $R_2(d\psi) = I - id\psi J_2$, donc le membre de gauche de (2.19) 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 (2.6), d'où $i[J_1, J_2] = -J_3$, qui est l'une des relations (2.10).

2.2. Générateurs infinitésimaux dans $SU(2)$

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 $U = V \exp\{i \text{diag}(\lambda_k)\} V^\dagger$ et donc s'écrire

$$U = \exp iH = \sum_0^{\infty} \frac{(iH)^n}{n!} \quad (2.20)$$

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

$$H = \sum_{i=1}^3 \eta_k \frac{\sigma_k}{2} , \quad (2.21)$$

ce que l'on peut reporter dans (2.20). On a en fait déjà observé plus haut que toute matrice unitaire 2×2 peut s'écrire sous la forme (1.10). En comparant cette forme avec celle

obtenue en (2.8), ou encore en comparant sa version infinitésimale $U_n(d\psi) = (I - i d\psi \mathbf{n} \cdot \frac{\boldsymbol{\sigma}}{2})$ avec (2.1)-(2.18), on voit que les matrices $\frac{1}{2}\sigma_j$ jouent ici dans SU(2) le rôle joué par les générateurs infinitésimaux J_j dans SO(3). Or ces matrices $\frac{1}{2}\sigma$ vérifient les relations de commutation

$$\left[\frac{\sigma_i}{2}, \frac{\sigma_j}{2} \right] = i\epsilon_{ijk} \frac{\sigma_k}{2} . \quad (2.22)$$

avec les mêmes *constantas de structure* ϵ_{ijk} que dans (2.10). Autrement dit, nous venons de découvrir que les générateurs infinitésimaux J_i (éq. (2.5)) de SO(3) et $\frac{1}{2}\sigma_i$ de SU(2) satisfont aux mêmes relations de commutation (on dira plus tard qu'ils forment deux représentations de la même algèbre de Lie $su(2) = so(3)$). 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. Par exemple, des relations (2.17), il découle sans aucun calcul supplémentaire que pour les matrices de Pauli, on a

$$e^{-i\frac{\beta}{2}\sigma_2}\sigma_k e^{i\frac{\beta}{2}\sigma_2} = D_2(\beta)\sigma_k D_2^{-1}(\beta) = \sigma_l \mathcal{R}_y(\beta)_{lk} \quad (2.23)$$

où on lit les éléments de matrice \mathcal{R}_y en (1.3)'. 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, (3.15). Au contraire, 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 $su(2)$.

2.3. Algèbre de Lie $su(2)$

Récapitulons : nous venons d'introduire l'algèbre de commutation des générateurs infinitésimaux (ou algèbre de Lie) du groupe SU(2) (ou SO(3)), notée $su(2)$ ou $so(3)$. Elle est définie par les relations (2.10), que nous récrivons

$$[J_i, J_j] = i\epsilon_{ijk} J_k . \quad (2.10)$$

On utilise aussi beaucoup les trois combinaisons

$$J_z \equiv J_3, \quad J_+ = J_1 + iJ_2, \quad J_- = J_1 - iJ_2 . \quad (2.24)$$

Il est immédiat de calculer

$$[J_3, J_+] = J_+ \quad (2.25a)$$

$$[J_3, J_-] = -J_- \quad (2.25b)$$

$$[J_+, J_-] = 2J_3 . \quad (2.25c)$$

On vérifie aussi que l'opérateur

$$\mathbf{J}^2 = J_1^2 + J_2^2 + J_3^2 = J_3^2 + J_+ J_- \quad (2.26)$$

commute avec tous les J

$$[\mathbf{J}^2, J_i] = 0 . \quad (2.27)$$

\mathbf{J}^2 est l'*opérateur de Casimir*. Anticipant un peu sur la suite, nous serons souvent intéressés aux “représentations unitaires”, où les générateurs J_i , $i = 1, 2, 3$ sont hermitiques, donc

$$J_i^\dagger = J_i, \quad i = 1, 2, 3 \quad J_\pm^\dagger = J_\mp . \quad (2.28)$$

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

$$\begin{aligned} f'(\mathbf{x}) &= f(R^{-1}\mathbf{x}) = f(\mathbf{x} - \delta\psi \mathbf{n} \wedge \mathbf{x}) \\ &= (1 - \delta\psi \mathbf{n} \cdot \mathbf{x} \wedge \nabla) f(\mathbf{x}) \\ &= (1 - i\delta\psi \mathbf{n} \cdot \mathbf{J}) f(\mathbf{x}) . \end{aligned} \quad (2.29)$$

On identifie donc

$$\mathbf{J} = -i\mathbf{x} \wedge \nabla, \quad J_i = -i\epsilon_{ijk} x_j \frac{\partial}{\partial x_k} \quad (2.30)$$

ce qui permet de le calculer dans des coordonnées quelconques, par exemple sphériques (Appendice A). (Comparer aussi (2.30) 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 (2.10).

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 (§ 2.3 du Chap. 0 et Appendice A). Il doit en particulier être invariant par rotation, être de degré 2 dans les J , 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

$$\begin{aligned} \Delta_3 &= \frac{1}{r} \frac{\partial^2}{\partial r^2} r - \frac{\mathbf{J}^2}{r^2} \\ &= \frac{1}{r} \frac{\partial^2}{\partial r^2} r + \frac{\Delta_{\text{sphère } S^2}}{r^2} . \end{aligned} \quad (2.31)$$

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

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

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 ?

3. Représentations de $SU(2)$

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 (2.10)).

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 $su(2)$.

3.2. Représentations de l'algèbre $su(2)$

Procédons à la construction classique des représentations de l'algèbre $su(2)$. Comme précédemment, J_{\pm} et J_z désignent les représentants des générateurs infinitésimaux dans une certaine représentation. Ils satisfont aux relations de commutation (2.25) et d'hermiticité (2.28). 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 $|j m\rangle$

$$\begin{aligned}\mathbf{J}^2|j m\rangle &= j(j+1)|j m\rangle \\ J_z|j m\rangle &= m|j m\rangle .\end{aligned}\tag{3.1}$$

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 (2.25)), on calcule la norme carrée de $J_{\pm}|j m\rangle$:

$$\begin{aligned}\langle j m|J_-J_+|j m\rangle &= (j(j+1) - m(m+1)) \langle j m|j m\rangle \\ &= (j-m)(j+m+1) \langle j m|j m\rangle \\ \langle j m|J_+J_-|j m\rangle &= (j(j+1) - m(m-1)) \langle j m|j m\rangle \\ &= (j+m)(j-m+1) \langle j m|j m\rangle .\end{aligned}\tag{3.2}$$

Ces normes carrées ne peuvent être négatives donc

$$\begin{aligned}(j-m)(j+m+1) &\geq 0 & : & \quad -j-1 \leq m \leq j \\ (j+m)(j-m+1) &\geq 0 & : & \quad -j \leq m \leq j+1\end{aligned}\tag{3.3}$$

qui impliquent

$$-j \leq m \leq j .\tag{3.4}$$

En outre $J_+|j m\rangle = 0$ si et seulement si $m = j$ et $J_-|j m\rangle = 0$ si et seulement si $m = -j$

$$J_+|j j\rangle = 0 \quad J_-|j -j\rangle = 0 .\tag{3.5}$$

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

$$\begin{aligned} \mathbf{J}^2 J_+|j m\rangle &= J_+ \mathbf{J}^2|j m\rangle = j(j+1)J_+|j m\rangle \\ J_z J_+|j m\rangle &= J_+(J_z+1)|j m\rangle = (m+1)J_+|j m\rangle . \end{aligned} \quad (3.6)$$

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 (3.4), 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 (3.5), $J_-^p|j m\rangle$ est un vecteur propre de valeurs propres $(j, -j)$ donc $m-p = -j$ c'est-à-dire

$$(j+m) \text{ est un entier positif ou nul.} \quad (3.7)$$

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

$$(j-m) \text{ est un entier positif ou nul} \quad (3.8)$$

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 $2j+1$ valeurs ³

$$m = -j, -j+1, \dots, j-1, j . \quad (3.9)$$

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

$$J_+|j m\rangle = \sqrt{j(j+1) - m(m+1)}|j m+1\rangle \quad (3.10a)$$

$$J_-|j m\rangle = \sqrt{j(j+1) - m(m-1)}|j m-1\rangle \quad (3.10b)$$

$$J_z|j m\rangle = m|j m\rangle . \quad (3.10c)$$

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.

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

En fait, cette représentation de l’algèbre $su(2)$ 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 : $\|J_{\pm}|jm\rangle\|^2 \geq 0 \implies -j \leq m \leq j$, 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 $|jj\rangle$, notation que nous adoptons donc dans la suite de cette discussion. A partir de ce vecteur, les $J_-^p|jj\rangle$ forment une suite qui doit elle aussi être finie

$$\exists q \quad J_-^{q-1}|jj\rangle \neq 0 \quad J_-^q|jj\rangle = 0. \quad (3.11)$$

On démontre aisément par récurrence que

$$J_+J_-^q|jj\rangle = [J_+, J_-^q]|jj\rangle = q(2j + 1 - q)J_-^{q-1}|jj\rangle = 0 \quad (3.12)$$

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 (3.9). 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.

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

$$|jm\rangle \mapsto D^j(U)|jm\rangle = |jm'\rangle \mathcal{D}_{m'm}^j(U). \quad (3.13)$$

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 (1.6), on a donc

$$\begin{aligned} \mathcal{D}_{m'm}^j(\alpha, \beta, \gamma) &= \langle jm'|D(\alpha, \beta, \gamma)|jm\rangle \\ &= \langle jm'|e^{-i\alpha J_z} e^{-i\beta J_y} e^{-i\gamma J_z}|jm\rangle \\ &= e^{-i\alpha m'} d_{m'm}^j(\beta) e^{-i\gamma m} \end{aligned} \quad (3.14)$$

où la matrice d^j est définie par

$$d_{m'm}^j(\beta) = \langle jm'|e^{-i\beta J_y}|jm\rangle. \quad (3.15)$$

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

$$\begin{aligned} \mathcal{D}_{m'm}^j(\hat{z}, \psi) &= e^{-i\psi m} \delta_{mm'} \\ \mathcal{D}_{m'm}^j(\hat{y}, \psi) &= d_{m'm}^j(\psi) \end{aligned} \quad (3.16)$$

Exercice : Calculer $\mathcal{D}^j(\hat{x}, \psi)$. (On pourra utiliser (1.4).)

On note que $\mathcal{D}^j(\hat{z}, 2\pi) = (-1)^{2j}I$, puisque $(-1)^{2m} = (-1)^{2j}$, compte tenu de (3.8), et la propriété est vraie pour tout axe \mathbf{n} par conjugaison (1.4)

$$\mathcal{D}^j(\mathbf{n}, 2\pi) = (-1)^{2j}I . \quad (3.17)$$

Cela montre qu'une rotation de 2π dans $\text{SO}(3)$ est représentée par $-I$ dans une représentation de spin demi-entier de $\text{SU}(2)$. Les représentations de spin demi-entier de $\text{SU}(2)$ sont des représentations "projectives", (c'est-à-dire ici à un signe près), de $\text{SO}(3)$; 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 $\mathbf{n} = R\hat{z}$, $\mathcal{D}(\mathbf{n}, \psi) = \mathcal{D}(R)\mathcal{D}(\hat{z}, \psi)\mathcal{D}^{-1}(R)$, donc

$$\det \mathcal{D}(\mathbf{n}, \psi) = \det \mathcal{D}(\hat{z}, \psi) = \det e^{-i\psi J_z} = \prod_{m=-j}^j e^{-im\psi} = 1 . \quad (3.18)$$

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

$$\begin{aligned} \mathcal{D}^{\frac{1}{2}}(U) = U = e^{-i\frac{1}{2}\psi\mathbf{n}\cdot\boldsymbol{\sigma}} &= \begin{pmatrix} \cos \frac{\psi}{2} - i \cos \theta \sin \frac{\psi}{2} & -i \sin \frac{\psi}{2} \sin \theta e^{-i\phi} \\ -i \sin \frac{\psi}{2} \sin \theta e^{i\phi} & \cos \frac{\psi}{2} + i \cos \theta \sin \frac{\psi}{2} \end{pmatrix} \\ &= e^{-i\frac{\alpha}{2}\sigma_3} e^{-i\frac{\beta}{2}\sigma_2} e^{-i\frac{\gamma}{2}\sigma_3} = \begin{pmatrix} \cos \frac{\beta}{2} e^{-\frac{i}{2}(\alpha+\gamma)} & -\sin \frac{\beta}{2} e^{-\frac{i}{2}(\alpha-\gamma)} \\ \sin \frac{\beta}{2} e^{\frac{i}{2}(\alpha-\gamma)} & \cos \frac{\beta}{2} e^{\frac{i}{2}(\alpha+\gamma)} \end{pmatrix} \end{aligned} \quad (3.19)$$

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 (2.5) !)

$$J_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad J_+ = \sqrt{2} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad J_- = \sqrt{2} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad (3.20)$$

d'où

$$d^1(\beta) = e^{-i\beta J_y} = \begin{pmatrix} \frac{1+\cos \beta}{2} & -\frac{\sin \beta}{\sqrt{2}} & \frac{1-\cos \beta}{2} \\ \frac{\sin \beta}{\sqrt{2}} & \cos \beta & -\frac{\sin \beta}{\sqrt{2}} \\ \frac{1-\cos \beta}{2} & \frac{\sin \beta}{\sqrt{2}} & \frac{1+\cos \beta}{2} \end{pmatrix} \quad (3.21)$$

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 B du chapitre 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 Chapitre 2 que la représentation étant unitaire, elle est soit irréductible soit complètement réductible ; 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 (3.10) toute matrice M commutant avec J_z est diagonale, $M_{mm'} = \mu_m \delta_{mm'}$, (le vérifier !), et la commutation avec J_+ force tous les μ_m à être égaux : la matrice M est multiple de l'identité et la représentation est bien irréductible.

On peut aussi se demander pourquoi l'étude des représentations de dimension finie que vous venons de construire suffit aux besoins du physicien, par exemple en mécanique quantique, où la scène se passe en général dans un espace de Hilbert de dimension infinie. On 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 donc est complètement réductible en une somme (finie ou infinie) de représentations irréductibles de dimension finie.

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

$$P_{jm} = \frac{\xi^{j+m} \eta^{j-m}}{\sqrt{(j+m)!(j-m)!}} \quad m = -j, \dots, j \quad (3.22)$$

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 $GL(2, \mathbb{C})$), qui n'est autre que la représentation de spin j précédente. Cela permet d'écrire des formules très explicites pour les \mathcal{D}^j .

$$P_{jm}(\xi', \eta') = \sum_{m'} P_{jm'}(\xi, \eta) \mathcal{D}_{m'm}^j(U). \quad (3.23)$$

On obtient

$$\mathcal{D}_{m'm}^j(U) = ((j+m)!(j-m)!(j+m')!(j-m')!)^{\frac{1}{2}} \sum_{\substack{n_1, n_2, n_3, n_4 \geq 0 \\ n_1+n_2=j+m'; \quad n_3+n_4=j-m' \\ n_1+n_3=j+m; \quad n_2+n_4=j-m}} \frac{a^{n_1} b^{n_2} c^{n_3} d^{n_4}}{n_1! n_2! n_3! n_4!}. \quad (3.24)$$

Pour $U = -I$, on vérifie à nouveau que $\mathcal{D}^j(-I) = (-1)^{2j} I$.

Dans le cas particulier de $U = e^{-i\psi \frac{\sigma_2}{2}} = \cos \frac{\psi}{2} I - i \sin \frac{\psi}{2} \sigma_2$, on a donc

$$d_{m'm}^j(\psi) = ((j+m)!(j-m)!(j+m')!(j-m')!)^{\frac{1}{2}} \sum_{k \geq 0} \frac{(-1)^{k+j-m} \cos \frac{\psi}{2}^{2k+m+m'} \sin \frac{\psi}{2}^{2j-2k-m-m'}}{(m+m'+k)!(j-m-k)!(j-m'-k)!k!}. \quad (3.25)$$

L'expression des générateurs infinitésimaux sur les polynômes P_{jm} s'obtient en considérant des U proches de l'identité. On trouve

$$J_+ = \xi \frac{\partial}{\partial \eta} \quad J_- = \eta \frac{\partial}{\partial \xi} \quad J_z = \frac{1}{2} \left(\xi \frac{\partial}{\partial \xi} - \eta \frac{\partial}{\partial \eta} \right) \quad (3.26)$$

dont il est immédiat de vérifier les relations de commutation ainsi que l'action sur les P_{jm} en accord avec (3.10). Cela achève l'identification de (3.22) avec la représentation de spin j .

Remarques

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.

4. Produit direct de représentations de $SU(2)$

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

$$|j_1 m_1\rangle \otimes |j_2 m_2\rangle \equiv |j_1 m_1; j_2 m_2\rangle \quad \text{abrégé en} \quad |m_1 m_2\rangle \quad (4.1)$$

sur lesquels agissent les générateurs infinitésimaux sous la forme

$$\mathbf{J} = \mathbf{J}^{(1)} \otimes I^{(2)} + I^{(1)} \otimes \mathbf{J}^{(2)} . \quad (4.2)$$

L'indice supérieur indique sur quel espace agissent les opérateurs. Par abus de notation, on écrit souvent au lieu de (4.2)

$$\mathbf{J} = \mathbf{J}^{(1)} + \mathbf{J}^{(2)} \quad (4.2)'$$

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 (4.1) 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

$$|(j_1 j_2) J M\rangle \quad \text{ou plus simplement} \quad |J M\rangle \quad (4.3)$$

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

$$\begin{aligned} \langle m_1 m_2 | J_z | J M \rangle &= (m_1 + m_2) \langle m_1 m_2 | J M \rangle \\ &= M \langle m_1 m_2 | J M \rangle \end{aligned} \quad (4.4)$$

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

$$M = m_1 + m_2 . \quad (4.5)$$

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 (4.5), soit

$$n(M) = \begin{cases} 0 & \text{si } |M| > j_1 + j_2 \\ j_1 + j_2 + 1 - |M| & \text{si } |j_1 - j_2| \leq |M| \leq j_1 + j_2 \\ 2 \inf(j_1, j_2) + 1 & \text{si } 0 \leq |M| \leq |j_1 - j_2| \end{cases} \quad (4.6)$$

(voir Fig. 3 pour laquelle $j_1 = 5/2$ and $j_2 = 1$). Soit N_J le nombre de fois où la représentation de spin J apparaît dans la décomposition du produit des représentations de spin j_1 et j_2 . Les $n(M)$ vecteurs de valeur propre M pour J_z peuvent aussi s'interpréter

comme provenant des N_J vecteurs $|JM\rangle$ pour les différentes valeurs de J compatibles avec cette valeur de M

$$n(M) = \sum_{J \geq |M|} N_J \quad (4.7)$$

soit en retranchant membre à membre deux telles relations

$$\begin{aligned} N_J &= n(J) - n(J+1) \\ &= 1 \quad \text{si et seulement si} \quad |j_1 - j_2| \leq J \leq j_1 + j_2 \\ &= 0 \quad \text{sinon.} \end{aligned} \quad (4.8)$$

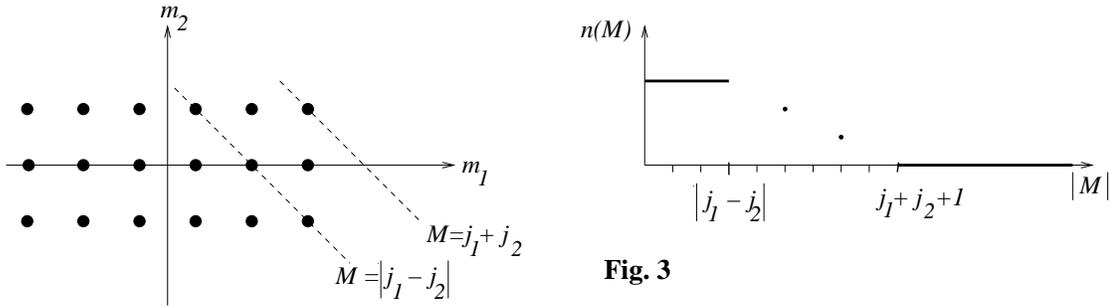


Fig. 3

En conclusion, nous venons de démontrer que les $(2j_1 + 1)(2j_2 + 1)$ vecteurs (4.1) (à j_1 et j_2 fixés) peuvent se réexprimer en fonction des vecteurs $|JM\rangle$ où

$$\begin{aligned} J &= |j_1 - j_2|, |j_1 - j_2| + 1, \dots, j_1 + j_2 \\ M &= -J, -J + 1, \dots, J. \end{aligned} \quad (4.9)$$

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.

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

$$|j_1 m_1; j_2 m_2\rangle = \sum_{J=|j_1-j_2|}^{j_1+j_2} \sum_{M=-J}^J \langle (j_1 j_2) JM | j_1 m_1; j_2 m_2 \rangle |(j_1 j_2) JM\rangle \quad (4.10a)$$

$$|(j_1 j_2); JM\rangle = \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} \langle (j_1 j_2) JM | j_1 m_1; j_2 m_2 \rangle^* |j_1 m_1; j_2 m_2\rangle. \quad (4.10b)$$

Leur valeur dépend en fait d'un choix de phase relative entre les vecteurs (4.1) et (4.3) ; la convention habituelle est que pour chaque valeur de J , on choisit

$$\langle j_1 m_1 = j_1; j_2 m_2 = J - j_1 | J M = J \rangle \quad \text{réel.} \quad (4.11)$$

Les autres vecteurs sont alors définis sans ambiguïté par (3.10) et on va montrer que tous les C.G. sont réels. Les C.G. satisfont des relations de récurrence conséquences de (3.10). Appliquant en effet J_{\pm} aux deux membres de (4.10a), on obtient

$$\begin{aligned} \sqrt{J(J+1) - M(M \pm 1)} \langle (j_1 j_2) J M | j_1 m_1; j_2 m_2 \rangle \\ = \sqrt{j_1(j_1+1) - m_1(m_1 \pm 1)} \langle (j_1 j_2) J M \pm 1 | j_1 m_1 \pm 1; j_2 m_2 \rangle \\ + \sqrt{j_2(j_2+1) - m_2(m_2 \pm 1)} \langle (j_1 j_2) J M \pm 1 | j_1 m_1; j_2 m_2 \pm 1 \rangle \end{aligned} \quad (4.12)$$

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 (4.11) 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

$$\begin{aligned} \sum_{m_1=-j_1}^{j_1} \langle j_1 m_1; j_2 m_2 | (j_1 j_2) J M \rangle \langle j_1 m_1; j_2 m_2 | (j_1 j_2) J' M' \rangle = \delta_{JJ'} \delta_{MM'} \\ \text{si } |j_1 - j_2| \leq J \leq j_1 + j_2 \\ \sum_{J=|j_1-j_2|}^{j_1+j_2} \langle j_1 m_1; j_2 m_2 | (j_1 j_2) J M \rangle \langle j_1 m'_1; j_2 m'_2 | (j_1 j_2) J M \rangle = \delta_{m_1 m'_1} \delta_{m_2 m'_2} \\ \text{si } |m_1| \leq j_1, |m_2| \leq j_2. \end{aligned} \quad (4.13)$$

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

$$\begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & -M \end{pmatrix} = \frac{(-1)^{j_1-j_2+M}}{\sqrt{2J+1}} \langle j_1 m_1; j_2 m_2 | (j_1 j_2) J M \rangle \quad (4.14)$$

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

$$\frac{1}{2} \otimes \frac{1}{2} : \begin{aligned} |(\frac{1}{2}, \frac{1}{2})1, 1\rangle &= |\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, \frac{1}{2}\rangle \\ |(\frac{1}{2}, \frac{1}{2})1, 0\rangle &= \frac{1}{\sqrt{2}} \left(|\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2}\rangle + |\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}, \frac{1}{2}\rangle \right) \\ |(\frac{1}{2}, \frac{1}{2})0, 0\rangle &= \frac{1}{\sqrt{2}} \left(|\frac{1}{2}, \frac{1}{2}; \frac{1}{2}, -\frac{1}{2}\rangle - |\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}, \frac{1}{2}\rangle \right) \\ |(\frac{1}{2}, \frac{1}{2})1, -1\rangle &= |\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}, -\frac{1}{2}\rangle \end{aligned} \quad (4.15)$$

et

$$\frac{1}{2} \otimes 1 : \begin{aligned} |(\frac{1}{2}, 1)\frac{3}{2}, \frac{3}{2}\rangle &= |\frac{1}{2}, \frac{1}{2}; 1, 1\rangle \\ |(\frac{1}{2}, 1)\frac{3}{2}, \frac{1}{2}\rangle &= \frac{1}{\sqrt{3}} \left(\sqrt{2}|\frac{1}{2}, \frac{1}{2}; 1, 0\rangle + |\frac{1}{2}, -\frac{1}{2}; 1, 1\rangle \right) \\ |(\frac{1}{2}, 1)\frac{3}{2}, -\frac{1}{2}\rangle &= \frac{1}{\sqrt{3}} \left(|\frac{1}{2}, \frac{1}{2}; 1, -1\rangle + \sqrt{2}|\frac{1}{2}, -\frac{1}{2}; 1, 0\rangle \right) \\ |(\frac{1}{2}, 1)\frac{3}{2}, -\frac{3}{2}\rangle &= |\frac{1}{2}, -\frac{1}{2}; 1, -1\rangle \\ |(\frac{1}{2}, 1)\frac{1}{2}, \frac{1}{2}\rangle &= \frac{1}{\sqrt{3}} \left(-|\frac{1}{2}, \frac{1}{2}; 1, 0\rangle + \sqrt{2}|\frac{1}{2}, -\frac{1}{2}; 1, 1\rangle \right) \\ |(\frac{1}{2}, 1)\frac{1}{2}, -\frac{1}{2}\rangle &= \frac{1}{\sqrt{3}} \left(-\sqrt{2}|\frac{1}{2}, \frac{1}{2}; 1, -1\rangle + |\frac{1}{2}, -\frac{1}{2}; 1, 0\rangle \right) \end{aligned} \quad (4.16)$$

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 (4.14) 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

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

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

$$\langle j_1 (j_2 j_3); J M | (j_1 j_2) j_3; J' M' \rangle = \delta_{JJ'} \delta_{MM'} \sqrt{(2J_1 + 1)(2J_2 + 1)} (-1)^{j_1 + j_2 + j_3 + J} \begin{Bmatrix} j_1 & j_2 & J_1 \\ j_3 & J & J_2 \end{Bmatrix}. \tag{4.18}$$

et les $\begin{Bmatrix} \end{Bmatrix}$ 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.

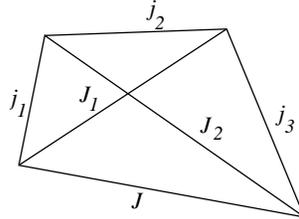


Fig. 4

5. Une application physique : l'isospin

Le groupe SU(2) n'intervient pas en Physique qu'en tant que (lié au) groupe de rotation de l'espace euclidien. Illustrons une autre de ses apparitions par la symétrie d'isospin. Il existe dans la nature un certain nombre de particules élémentaires présentant des propriétés voisines, mais différant par leur charge électrique. C'est le cas du proton et du neutron, de masses $938,28 \text{ MeV}/c^2$ 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 mésons K etc. Il a été proposé que ceci est la manifestation d'une symétrie brisée par les effets électromagnétiques. En l'absence d'interactions électromagnétiques, le proton et le neutron d'une part, les trois mésons π de l'autre seraient des particules de même nature, de même masse, différant seulement par un nombre quantique "interne", à la façon de deux électrons dotés de spins différents. En fait le groupe régissant cette symétrie est aussi SU(2), mais un SU(2) agissant dans un espace abstrait autre que l'espace usuel. On a donné le nom d'*isospin* ou *spin*

isotopique au nombre quantique correspondant. Pour résumer, la proposition est donc qu'il existe un groupe SU(2) de symétrie de l'Hamiltonien des interactions fortes, et que les différentes particules sujettes à ces interactions forment des représentations de SU(2) : représentation d'isospin $I = \frac{1}{2}$ pour le nucléon (proton $I_z = +\frac{1}{2}$, neutron $I_z = -\frac{1}{2}$), isospin $I = 1$ pour les pions ($\pi^\pm : I_z = \pm 1$, $\pi^0 : I_z = 0$) etc. L'isospin est donc un "bon nombre quantique", conservé dans ces interactions. Ainsi la réaction $N \rightarrow N + \pi$, (N pour nucléon) importante 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

$$\begin{aligned}
 p + \pi^+ &\rightarrow p + \pi^+ & I_z &= \frac{3}{2} \\
 p + \pi^0 &\rightarrow p + \pi^0 & I_z &= \frac{1}{2} \\
 &\rightarrow n + \pi^+ & &'' \\
 p + \pi^- &\rightarrow p + \pi^- & I_z &= -\frac{1}{2} \\
 &\rightarrow n + \pi^0 & &'' \\
 n + \pi^- &\rightarrow n + \pi^- & I_z &= -\frac{3}{2}
 \end{aligned}$$

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 (4.16), 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 $\langle I I_z | \mathcal{T} | I' I'_z \rangle = \mathcal{T}_I \delta_{II'} \delta_{I_z I'_z}$, comme on le justifiera plus tard (lemme de Schur). En calculant alors les éléments de matrice de l'opérateur de transition \mathcal{T} entre ces différents états,

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

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

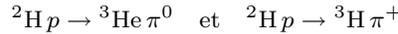
$$\sigma(\pi^+ p \rightarrow \pi^+ p) : \sigma(\pi^- p \rightarrow \pi^0 n) : \sigma(\pi^- p \rightarrow \pi^- p) = 9 : 2 : 1$$

ce qui indique qu'à cette énergie, la diffusion dans le canal d'isospin 3/2 est prédominante et signale en fait l'existence d'un état intermédiaire, particule très instable ou "résonance", notée Δ , d'isospin 3/2 donc avec quatre états de charge

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

Cette particule a un spin 3/2 et une masse $M(\Delta) \approx 1230 \text{ MeV}/c^2$.

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 deuteron qui est formé d'un proton et d'un neutron dans un état antisymétrique de leurs isospins (pour que la fonction d'onde, 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 $\frac{1}{2}$. 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 $\frac{1}{2}$.

6. Représentations de $SO(3,1)$ et $SL(2, \mathbb{C})$

6.1. Algèbre de Lie des groupes de Lorentz et Poincaré

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

$$\begin{aligned} (a, \Lambda) \quad : \quad x &\mapsto x' = \Lambda x + a \\ (a', \Lambda')(a, \Lambda) &= (a' + \Lambda'a, \Lambda'\Lambda) ; \end{aligned} \tag{6.1}$$

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

Une transformation infinitésimale de Poincaré s'écrit $(\alpha^\mu, \Lambda^\mu_\nu = \delta^\mu_\nu + \omega^\mu_\nu)$. L'algèbre de Lie est engendrée par des opérateurs différentiels agissant sur les fonctions des coordonnées de telle sorte que si $x'^\lambda = x^\lambda + \delta x^\lambda = x^\lambda + \alpha^\lambda + \omega^{\lambda\nu}x_\nu$, $\delta f(x) = f(x^\mu - \alpha^\mu - \omega^{\lambda\nu}x_\nu) - f(x) = (I - i\alpha^\mu P_\mu - \frac{i}{2}\omega^{\mu\nu} J_{\mu\nu})f(x)$, (cf (2.29) ou chap. 1 (4.25)), donc

$$J_{\mu\nu} = i(x_\mu \partial_\nu - x_\nu \partial_\mu) \quad P_\mu = -i\partial_\mu \quad (6.2)$$

dont on calcule aisément les commutateurs

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

(6.3)

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

$$J_{ij} = \epsilon_{ijk} J^k \quad K^i = J_{0i} . \quad (6.4)$$

On a alors

$$\begin{aligned} [J^i, J^j] &= i\epsilon_{ijk} J^k \\ [J^i, K^j] &= i\epsilon_{ijk} K^k \\ [K^i, K^j] &= -i\epsilon_{ijk} J^k \end{aligned} \quad (6.5)$$

et aussi

$$\begin{aligned} [J^i, P^j] &= i\epsilon_{ijk} P^k & [K^i, P^j] &= iP^0 \delta_{ij} \\ [J^i, P^0] &= 0 & [K^i, P^0] &= iP^i . \end{aligned} \quad (6.6)$$

N.B. Les deux premières des relations (6.5) et la première de (6.6) expriment bien, comme attendu, que $\mathbf{J} = \{J^i\}$, $\mathbf{K} = \{K^i\}$ et $\mathbf{P} = \{P^i\}$ se transforment comme des vecteurs sous l'action des rotations de \mathbb{R}^3 . Formons les combinaisons

$$M^j = \frac{1}{2}(J^j + iK^j) \quad N^j = \frac{1}{2}(J^j - iK^j) \quad (6.7)$$

elles satisfont

$$\begin{aligned} [M^i, M^j] &= i\epsilon_{ijk} M^k \\ [N^i, N^j] &= i\epsilon_{ijk} N^k \\ [M^i, N^j] &= 0 . \end{aligned} \quad (6.8)$$

On voit donc que, si on considère les combinaisons complexes M et N de ses générateurs, l'algèbre de Lie de $\mathcal{L} = O(3, 1)$ 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.

6.2. Groupes de recouvrement de \mathcal{L}_+^\dagger et \mathcal{P}_+^\dagger

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}_+^\dagger sont reliés, qui est une extension quadri-dimensionnelle de la méthode suivie au §1.2. On utilise les matrices σ_μ constituées de $\sigma_0 = I$ et des trois matrices de Pauli familières. Notons que l'on a

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

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}_+^\dagger et que $A \rightarrow \Lambda$ est un homomorphisme de $SL(2, \mathbb{C})$ dans \mathcal{L}_+^\dagger . 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}_+^\dagger : $SL(2, \mathbb{C})$ est un recouvrement d'ordre 2 de \mathcal{L}_+^\dagger . 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 $ISL(2, \mathbb{C})$).

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

La construction du § 3.3 fournit une représentation explicite de $\mathrm{GL}(2, \mathbb{C})$ et donc de $\mathrm{SL}(2, \mathbb{C})$. Pour $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{C})$, (3.24) donne l'expression de $\mathcal{D}_{mm'}^j(A)$:

$$\mathcal{D}_{mm'}^j(A) = [(j+m)!(j-m)!(j+m')!(j-m')!]^{\frac{1}{2}} \sum_{\substack{n_1, n_2, n_3, n_4 \geq 0 \\ n_1+n_2=j+m; \quad n_3+n_4=j-m' \\ n_1+n_3=j+m; \quad n_2+n_4=j-m}} \frac{a^{n_1} b^{n_2} c^{n_3} d^{n_4}}{n_1! n_2! n_3! n_4!} \quad (3.24)$$

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 (3.24) 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. § 1.3) $\mathcal{D}^j(A^{\dagger -1})$. Le remplacement de A par $A^{\dagger -1}$ s'interprète dans la construction du § 6.2 si au lieu d'associer $X = x^\mu \sigma_\mu$ 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 $\mathrm{SL}(2, \mathbb{C})$ 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 $\mathrm{SL}(2, \mathbb{C})$), on vérifie aisément que $A^{-1} = a^0 \sigma_0 - \mathbf{a} \cdot \boldsymbol{\sigma}$, donc

$$\sigma_2 A^T \sigma_2 = A^{-1}. \quad (6.9)$$

Finalement

$$X' = AXA^\dagger \iff \tilde{X}' = (A^{-1})^\dagger \tilde{X} A^{-1}. \quad (6.10)$$

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

Finalement, on démontre que toute représentation de dimension finie de $\mathrm{SL}(2, \mathbb{C})$ est complètement réductible et peut donc s'écrire comme somme directe de représentations irréductibles. La représentation irréductible de dimension finie la plus générale de $\mathrm{SL}(2, \mathbb{C})$ 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}_\uparrow^+ . 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 $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{C})$

$$\begin{aligned} (\frac{1}{2}, 0) \quad \xi = (\xi^\alpha) &\mapsto \xi' = A\xi = \begin{pmatrix} a\xi^1 + b\xi^2 \\ c\xi^1 + d\xi^2 \end{pmatrix} \\ (0, \frac{1}{2}) \quad \xi = (\xi^{\dot{\alpha}}) &\mapsto \xi' = A^*\xi = \begin{pmatrix} a^*\xi^{\dot{1}} + b^*\xi^{\dot{2}} \\ c^*\xi^{\dot{1}} + d^*\xi^{\dot{2}} \end{pmatrix} \end{aligned} \tag{6.11}$$

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 (6.9)

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

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

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

• *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 :

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

c'est-à-dire

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

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

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

6.4. 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" $\text{SL}(2, \mathbb{C})$ et $\text{ISL}(2, \mathbb{C})$. Comme on le verra plus bas (chap. 2), les représentations unitaires (de classe L^2) du groupe non compact $\text{SL}(2, \mathbb{C})$ 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 (6.3). 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

$$P_\mu |p\rangle = p_\mu |p\rangle . \quad (6.12)$$

On considère aussi le tenseur de Pauli-Lubanski introduit au chapitre 0

$$W^\lambda = \frac{1}{2} \epsilon^{\lambda\mu\nu\rho} J_{\mu\nu} P_\rho \quad (6.13)$$

et on vérifie (exercice !) que (6.3) implique

$$[W_\mu, P_\nu] = 0 \quad (6.14a)$$

$$[W^\mu, W^\nu] = -i \epsilon^{\mu\nu\rho\sigma} W_\rho P_\sigma \quad (6.14b)$$

$$[J_{\mu\nu}, W_\lambda] = i(g_{\nu\lambda} W_\mu - g_{\mu\lambda} W_\nu) . \quad (6.14c)$$

La dernière relation signifie que W est un 4-vecteur de Lorentz. On note aussi que $W \cdot P = 0$ 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, § 1.4) 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 $P^2 = 0$, $W^2 = 0$.

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

• *Représentations* $P^2 = p^2 = M^2 > 0$

Les représentations où $P^2 = p^2 = M^2 > 0$, $p^0 > 0$ et $W^2 < 0$ décrivent des particules de masse M . Les vecteurs de la représentation peuvent être choisis vecteurs propres des P_μ de valeur propre p_μ et d'une composante de W . On peut écrire

$$\frac{W_\mu}{M} = \sum_{i=1}^3 S_i n_\mu^{(i)} \quad , \quad \frac{W^2}{M^2} = -\mathbf{S}^2 = -(S_1^2 + S_2^2 + S_3^2)$$

où $n^{(i)}$ sont trois quadrivecteurs orthogonaux à p et entre eux, donc de genre espace, normés à $(n^{(i)})^2 = -1$. Les S_i satisfont

$$[S_i, S_j] = i\epsilon_{ijk} S_k \quad ,$$

comme conséquence de la relation (6.14b). Nous sommes donc de retour sur le terrain familier des représentations de l'algèbre $\mathfrak{su}(2)$. Les représentations unitaires sont telles que \mathbf{S}^2 a pour valeur propre $s(s+1)$, où s est entier ou demi-entier. Les vecteurs $n^{(i)}$ forment avec p un repère orthonormé (ou "tétrade"), noté $[p]$. On suppose cette tétrade $\{p, n^{(1)}, n^{(2)}, n^{(3)}\}$ orientée, $\det[p] = 1$. Les vecteurs de la représentation sont alors indexés par $[p]$ et par la valeur propre $s_3 \in \{-s, -s+1, \dots, +s\}$ de S_3 .

Ces vecteurs forment une représentation irréductible unitaire (cf plus bas) de $ISL(2, \mathbb{C})$. Cette représentation est de **dimension infinie**, puisque p peut prendre toute valeur sur la couche de masse $p^2 = M^2$, $p^0 > 0$. L'action des générateurs infinitésimaux sur le vecteur $|[p], s_3\rangle$ est donnée par

$$\begin{aligned} P_\mu |[p], s_3\rangle &= p_\mu |[p], s_3\rangle \\ S_3 |[p], s_3\rangle &= s_3 |[p], s_3\rangle \\ (S_1 \pm iS_2) |[p], s_3\rangle &= \sqrt{s(s+1) - s_3(s_3 \pm 1)} |[p], s_3 \pm 1\rangle \\ \frac{W^2}{M^2} |[p], s_3\rangle &= -s(s+1) |[p], s_3\rangle \end{aligned} \quad (6.15)$$

et celle des transformations finies de $ISL(2, \mathbb{C})$ par

$$U(a, A) |[p], s_3\rangle = U(a, I) U(0, A) |[p], s_3\rangle = e^{i(A \cdot p) \cdot a} |[A \cdot p], s'_3\rangle \mathcal{D}_{s'_3 s_3}^s ([Ap]^{-1} A [p]) \quad . \quad (6.16)$$

Par un petit abus de notation, $[p]$ désigne ici un élément de $SL(2, \mathbb{C})$ correspondant à la transformation de Lorentz qui envoie $\overset{0}{[p]} = \{(M, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\}$ sur $[p] = \{p, n^{(1)}, n^{(2)}, n^{(3)}\}$. Noter que $[Ap]^{-1} A [p]$ envoie $\overset{0}{[p]}$ sur lui-même : c'est un élément du *stabilisateur* (les physiciens disent "petit groupe") de $\overset{0}{[p]}$, qui est ici $SU(2)$. $\mathcal{D}^s([Ap]^{-1} A [p])$ est donc une matrice unitaire.

Si on change de tétrade, $[p] \rightarrow [p]', |p], s_3\rangle = [p]', s'_3\rangle \mathcal{D}_{s'_3 s_3}^s ([p]'^{-1} [p])$.

Cette propriété d'unitarité fait que le produit scalaire

$$\langle [p]', s'_3 | [p], s_3\rangle = 2p^0 (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}') \delta_{s_3 s'_3} \quad (6.17)$$

où $p^0 := \sqrt{\mathbf{p}^2 + M^2}$, est invariant par $SL(2, \mathbb{C})$. La normalisation globale de (6.17) est bien sûr affaire de convention. (Rappelons aussi le calcul au chapitre 1 de la mesure invariante de Lorentz sur la couche de masse).

Une dernière remarque : ces vecteurs $|[p], s_3\rangle$ sont *non normalisables*, en raison de la divergence de δ^3 à argument nul. On peut leur préférer des superpositions linéaires

$$|\phi\rangle = \int \frac{d^3 p}{2p^0 (2\pi)^3} \sum_{s_3} \varphi_{s_3}([p]) |[p], s_3\rangle$$

avec maintenant

$$\langle \phi | \psi \rangle = \int \frac{d^3 p}{2p^0 (2\pi)^3} \sum_{s_3} \varphi_{s_3}^*([p]) \psi_{s_3}([p]) \quad ,$$

qui sont normalisables si les fonctions d'onde φ et ψ sont de carré sommable.

• *Représentations où $P^2 = 0$, $W^2 = 0$*

De telles représentations décrivent des particules de masse nulle. Dans l'espace propre de P_μ , $p^2 = 0$, on a $W.P = 0$.

Soient $n_1(p), n_2(p)$ deux quadrivecteurs de genre espace orthogonaux à p , tels que $\det(\hat{t}, n^{(1)}, n^{(2)}, p) > 0$ (ici \hat{t} est le vecteur unitaire porté par l'axe du temps futur.) Comme situation de référence, on prendra $\overset{0}{p} = (1, 0, 0, 1)$, $n^{(1)}(\overset{0}{p}) = (0, 1, 0, 0)$, $n^{(2)}(\overset{0}{p}) = (0, 0, 1, 0)$ complétés par $\hat{t} = (1, 0, 0, 0)$. On décompose W sur cette base selon

$$W_\mu = \lambda(p)p_\mu + W_1 n_\mu^{(1)} + W_2 n_\mu^{(2)} \quad (6.18)$$

avec des relations de commutation

$$[W_1, \lambda] = -iW_2 \quad [W_2, \lambda] = iW_1 \quad [W_1, W_2] = 0. \quad (6.19)$$

L'opérateur $\lambda(p)$ est appelé "hélicité".

Soit à nouveau $[p]$ la transformation de $SL(2, \mathbb{C})$ qui envoie $(\overset{0}{p}, n^{(1)}(\overset{0}{p}), n^{(2)}(\overset{0}{p}))$ sur $(p, n^{(1)}, n^{(2)})$. Les états à une particule $|p_\mu, \lambda(p)\rangle$ que nous considérons sont vecteurs propres de P_μ et de $\lambda(p)$

$$\begin{aligned} P_\mu |[p], \lambda\rangle &= p_\mu |[p], \lambda\rangle \\ \lambda(p) |[p], \lambda\rangle &= \lambda |[p], \lambda\rangle. \end{aligned} \quad (6.20)$$

Seules les représentations où $W^2 = 0$, c'est-à-dire où W_1 et W_2 sont représentés par 0, sont réalisées dans la nature.

Cherchons quel est le stabilisateur de $\overset{0}{p}$, c'est-à-dire, une fois transcrit en termes de matrices 2×2 , de $\overset{0}{p} \sigma_0 + \overset{0}{p}^3 \sigma_3 = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$. Ce stabilisateur est constitué des matrices unimodulaires A telles que $A \overset{0}{p} A^\dagger = \overset{0}{p}$, soit (petit calcul simple)

$$A = \begin{pmatrix} e^{i\phi/2} & ze^{-i\phi/2} \\ 0 & e^{-i\phi/2} \end{pmatrix}.$$

Le stabilisateur est donc maintenant $ISO(2)$, le groupe de rotations et translations à deux dimensions ("I" pour inhomogène : translation de z et rotation de ϕ). La rotation est représentée par $e^{i\lambda\phi}$, donc λ doit être entier ou demi-entier. Pour une transformation finie

$$U(a, A) |[p], \lambda\rangle = e^{i(A.p).a} |[A.p], \lambda\rangle \mathcal{D}_{\lambda\lambda}^{| \lambda |} ([Ap]^{-1} A[p]). \quad (6.21)$$

Noter que selon (6.20) ou (6.21), l'hélicité n'est pas affectée par le groupe de Lorentz \mathcal{L}_+^\dagger . C'est l'introduction (éventuelle) de la parité qui oblige à considérer simultanément les hélicités $\pm\lambda$, un fait que nous rencontrerons dans l'étude des spineurs de Dirac de masse nulle. Pour un photon, $\alpha_+ |p, +1\rangle + \alpha_- |p, -1\rangle$ décrit un état de polarisation "elliptique" ; $\alpha_+ = 0$ ou $\alpha_- = 0$, un état de polarisation circulaire ; $|\alpha_+| = |\alpha_-| = 1/\sqrt{2}$ un état de polarisation linéaire.

Il faudrait maintenant compléter cette discussion des représentations du groupe orthochrone propre de Lorentz (ou de Poincaré) par l'étude des représentations des transformations discrètes de renversement du temps et de la parité...

Bibliographie

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

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 and Breach 1968.

Problème

1. 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. (4.2), (4.2)'.

- a) Que peut-on dire des opérateurs $\mathbf{J}^{(1)2}$, $\mathbf{J}^{(2)2}$ et \mathbf{J}^2 et de leurs valeurs propres ?
- b) 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.

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

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

Appendix A. Calcul du laplacien sur les sphères S^2 et S^3 .

La méthode proposée au chapitre 0, § 2.3, pour calculer l'élément de volume $d\mu(x)$ et le laplacien Δ sur une variété à partir de son tenseur métrique s'applique bien au cas des sphères S^2 et S^3 . Considérons la sphère S^2 de rayon r fixé avec les coordonnées sphériques $0 \leq \theta \leq \pi$, $0 \leq \phi \leq 2\pi$ (Figure 1) ⁵. On a

$$\begin{aligned} ds^2 &= r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \\ \sqrt{g} &= r^2 \sin \theta \\ d\mu(x) &= r^2 \sin \theta d\theta d\phi \end{aligned} \tag{A.1}$$

$$\Delta_{\text{sphère } S^2} = \frac{1}{r^2} \Delta_{\text{sphère unité } S^2} = \frac{1}{r^2} \left(\frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} \right) .$$

⁵ Rappelons que dans l'espace \mathbb{R}^n , le laplacien peut s'écrire

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

Les générateurs J_i s'écrivent

$$\begin{aligned} J_3 &= -i \frac{\partial}{\partial \phi} \\ J_1 &= -i \left[-\cos \phi \cotg \theta \frac{\partial}{\partial \phi} - \sin \phi \frac{\partial}{\partial \theta} \right] \\ J_2 &= -i \left[-\sin \phi \cotg \theta \frac{\partial}{\partial \phi} + \cos \phi \frac{\partial}{\partial \theta} \right] \end{aligned} \quad (\text{A.2})$$

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

Pour la sphère S^3 on calcule les formules analogues. Dans la paramétrisation (1.11), on prendra par exemple

$$ds^2 = \frac{1}{2} \text{tr} dU dU^\dagger = \left(d\frac{\psi}{2} \right)^2 + \sin^2 \frac{\psi}{2} (d\theta^2 + \sin^2 \theta d\phi^2) \quad (\text{A.3})$$

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

$$d\mu(U) = \frac{1}{2} \left(\sin \frac{\psi}{2} \right)^2 \sin \theta d\psi d\theta d\phi. \quad (\text{A.4})$$

Dans la paramétrisation des angles d'Euler,

$$U = e^{-i\alpha \frac{\sigma_3}{2}} e^{-i\beta \frac{\sigma_2}{2}} e^{-i\gamma \frac{\sigma_3}{2}} \quad (\text{A.5})$$

d'où

$$ds^2 = \frac{1}{2} \text{tr} dU dU^\dagger = \frac{1}{4} (d\alpha^2 + 2d\alpha d\gamma \cos \beta + d\gamma^2 + d\beta^2) \quad (\text{A.6})$$

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

$$d\mu(U) = \frac{1}{8} \sin \beta d\alpha d\beta d\gamma \quad (\text{A.7})$$

$$\Delta_{\text{sphère } S^3} = \frac{4}{\sin^2 \beta} \left[\frac{\partial^2}{\partial \alpha^2} + \frac{\partial^2}{\partial \gamma^2} + \frac{\partial^2}{\partial \alpha \partial \gamma} \right] + \frac{4}{\sin \beta} \frac{\partial}{\partial \beta} \sin \beta \frac{\partial}{\partial \sin \beta}. \quad (\text{A.8})$$

Chapitre 000

Équations de Klein-Gordon et de Dirac

(Notes de cours de 2007)

1. Rappels de Mécanique Quantique

Les phénomènes de la physique quantique exhibent deux aspects fondamentaux, en rupture avec ceux de la physique classique :

- le caractère *discret* de certaines quantités physiques (énergie, moment cinétique, ...) lié à l'existence d'une nouvelle grandeur fondamentale, ayant les dimensions d'une action

$$\hbar = 1,055 \cdot 10^{-34} \text{ kg.m}^2/\text{s}$$

(Exemples : spectre d'énergie de "quanta" $\hbar\omega$, ω une fréquence caractéristique du système ; moment cinétique = $\hbar \times$ générateur infinitésimal des rotations \Leftrightarrow spectre = $\hbar \times$ entier ou demi-entier...)

- le caractère *probabiliste* des observations, qui force à abandonner le déterminisme de la physique classique.

1.1. Postulats

Le formalisme de la Mécanique Quantique rend compte de ces phénomènes en se basant sur les postulats suivants :

- a) les états ("purs") d'un système sont décrits par les *rayons* d'un espace de Hilbert \mathcal{H} , c'est-à-dire par les vecteurs de \mathcal{H} à un facteur non nul près (ou encore par les vecteurs normés à une phase près). Ces vecteurs sont notés ψ , ou encore, selon la notation de Dirac $|\psi\rangle$ ("ket") et leur conjugué, $\langle\psi|$ ("bra"), avec un produit scalaire $\langle\psi|\phi\rangle$. Un tel vecteur dépend du temps : $|\psi(t)\rangle$. Ainsi, pour une particule sans degré de liberté supplémentaire (spin etc), dans l'espace à trois dimensions, $\mathcal{H} = \mathcal{L}^2(\mathbb{R}^3)$.

Par opposition à un état pur, un mélange est représenté par une matrice densité...

- b) Les quantités physiquement observables (ou simplement "observables") sont des opérateurs auto-adjoints sur \mathcal{H} , $A = A^\dagger$.

En général, l'opérateur est non borné, (cf $A = \hat{x}$ dans $\mathcal{H} = \mathcal{L}^2(\mathbb{R})$), et défini seulement sur un sous-ensemble $D(A)$ dense dans \mathcal{H} .

- c) Dans une mesure de A , on n'observe que les valeurs $\alpha \in \text{Spec}(A)$. On notera P_I le projecteur spectral sur l'intervalle $I = [\alpha_1, \alpha_2]$.

- d) La probabilité de mesurer $\alpha \in I$ dans l'état $|\psi\rangle$ (c'est-à-dire le pourcentage d'occurrence de la valeur α parmi celles obtenues dans un grand nombre de mesures répétées sur des systèmes identiques, préparés dans le même état) est

$$p(I) = \frac{\langle \psi | P_I | \psi \rangle}{\langle \psi | \psi \rangle}$$

La valeur moyenne de A est donc $\langle A \rangle = \frac{\langle \psi | A | \psi \rangle}{\langle \psi | \psi \rangle}$.

Après mesure de la valeur $a \in I$ pour A , il y a "réduction du paquet d'onde" $\psi \mapsto P_I \psi$.

- e) Parmi les observables, l'opérateur hamiltonien H joue un rôle privilégié : il régit l'évolution dans le temps selon l'équation de Schrödinger

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle .$$

1.2. Principe de correspondance

Ces postulats sont complétés par le "principe de correspondance" qui suggère ce que doit être l'opérateur A du formalisme quantique, étant donné son analogue classique. Ce principe fait aussi appel à l'observation que dans le passage mécanique classique \rightarrow mécanique quantique, on passe du crochet de Poisson au commutateur

$$\{f, g\} \rightarrow [f_{\text{op}}, g_{\text{op}}] = -i\hbar \{f, g\} .$$

En particulier on a le commutateur canonique entre les opérateurs q_{op} et p_{op} de position et d'impulsion, soit

$$[q_{\text{op}}, p_{\text{op}}] = i\hbar$$

(opérateurs conjugués). Noter cependant que ce principe de correspondance n'est pas sans ambiguïté : il ne dit rien sur l'ordre des opérateurs p et q à adopter pour passer d'une fonction $f(p, q)$ à sa version quantique $f_{\text{op}}(p_{\text{op}}, q_{\text{op}})$. On omettra l'indice "op" dans la suite chaque fois que cela ne prêterait pas à confusion.

Notons $|q\rangle$ l'état propre de l'opérateur q_{op} de valeur propre q

$$q_{\text{op}} |q\rangle = q |q\rangle , \quad \langle q | q' \rangle = \delta(q - q') \quad \int dq |q\rangle \langle q| = Id .$$

À la description d'un état par un vecteur normalisé ψ (à une phase près), on peut préférer celle par sa *fonction d'onde* $\psi(q)$, obtenue par produit scalaire

$$\psi(q) = \langle q | \psi \rangle .$$

L'action de p_{op} sur ψ se traduit en un opérateur différentiel sur $\psi(q)$

$$p_{\text{op}} = \frac{\hbar}{i} \frac{\partial}{\partial q} \quad \text{i.e.} \quad \langle q | p_{\text{op}} \psi \rangle = -i\hbar \frac{\partial}{\partial q} \psi(q)$$

de telle sorte que $[q_{\text{op}}, p_{\text{op}}] = i\hbar$ est bien vérifiée. En particulier, pour les états propres notés $|p\rangle$ de p_{op}

$$p_{\text{op}}|p\rangle = p|p\rangle$$

on a les fonctions d'onde $\psi_p(q) = \langle q | p \rangle$ satisfaisant

$$\langle q | p_{\text{op}} | p \rangle = -i\hbar \frac{\partial}{\partial q} \psi_p(q) = p \psi_p(q)$$

d'où

$$\langle q | p \rangle = \psi_p(q) = \frac{1}{\sqrt{2\pi\hbar}} \exp \frac{i}{\hbar} p \cdot q ,$$

(avec une normalisation conventionnelle).

Toutes les considérations précédentes s'étendent bien sûr à des vecteurs positions et impulsions à d dimensions euclidiennes ou minkovskiennes. Résumons donc le principe de correspondance, en notant désormais les coordonnées de positions par $\mathbf{x} = (x^1, x^2, x^3)$ ou (x, y, z)

$$\begin{aligned} \text{Energie } E &\mapsto i\hbar \frac{\partial}{\partial t} \\ \text{Impulsion } p^i &\mapsto -i\hbar \frac{\partial}{\partial x^i} \end{aligned}$$

ou en notations minkovskiennes,

$$p^\mu \mapsto i\hbar \partial^\mu = (i\partial^0, -i\nabla) .$$

2. Équation de Klein-Gordon

Le principe de correspondance appliqué à l'expression de l'énergie d'une particule massive non relativiste, $E = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x})$ conduit à l'équation de Schrödinger pour la fonction d'onde notée maintenant $\psi(\mathbf{x}, t)$

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{x}, t) = \left(-\frac{\Delta}{2m} + V(\mathbf{x}) \right) \psi(\mathbf{x}, t) , \quad (2.1)$$

(avec $\Delta = \nabla^2$ le laplacien).

De la même façon, il est suggéré d'appliquer le principe de correspondance à l'expression de l'énergie d'une particule relativiste libre, $E^2 = \mathbf{p}^2 c^2 + m^2 c^4$, ce qui conduit à l'équation de Klein-Gordon

$$\left(\frac{\hbar^2}{c^2} \frac{\partial^2}{\partial t^2} - \hbar^2 \Delta + m^2 c^2 \right) \psi(\mathbf{x}, t) = 0 . \quad (2.2)$$

On a déjà rencontré au chapitre 0 le d'Alembertien $\square = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta$: c'est le laplacien dans la géométrie minkovskienne $\square = g^{\mu\nu} \partial_\mu \partial_\nu$. À partir de maintenant, nous adoptons des unités telles que $\hbar = c = 1$, si bien que nous écrirons l'équation de Klein-Gordon libre sous la forme

$$(\square + m^2) \psi = 0 . \quad (2.2)'$$

Couplage minimal au champ électromagnétique

On peut écrire aisément une équation de Klein-Gordon couplée au champ électromagnétique si on admet le principe de couplage minimal, qui consiste à remplacer les dérivées ∂_μ par une dérivée covariante (covariante vis à vis des transformations de jauge)

$$\partial_\mu \rightarrow \partial_\mu + iqA_\mu(x) . \quad (2.3)$$

On écrira donc

$$((\partial + iqA)^\mu (\partial + iqA)_\mu + m^2) \psi = 0 \quad (2.4)$$

Exercice : cette équation découle-t-elle d'un lagrangien ? Si oui, lequel ? De quelle transformation sur ψ faut-il accompagner une transformation de jauge $A_\mu \rightarrow A_\mu + \partial\phi(x)$?

L'équation (2.2), qui est bien invariante relativiste (et découle d'un principe d'action comme on a vu au chapitre précédent), constitue-t-elle une équation décente pour une fonction d'onde ? Il y a plusieurs manifestations des difficultés rencontrées dans cette interprétation :

- 1) Les niveaux d'énergie qui découlent de (2.2), sont solutions de $E^2 = \mathbf{p}^2 + m^2$ avec les deux signes $E = \pm \sqrt{\mathbf{p}^2 + m^2}$. L'existence de niveaux d'énergie négative non bornée inférieurement, donc arbitrairement basse, est un problème majeur, qui semble indiquer une instabilité de la théorie.
- 2) Peut-on arbitrairement se restreindre aux racines positives, en ne gardant que $H = +\sqrt{\mathbf{p}^2 + m^2}$? L'évolution d'une particule libre décrite par cet hamiltonien semble violer la causalité. Pour deux points séparés x, y par un intervalle de genre espace,

donc en principe sans communication causale, on trouve que l'élément de matrice de l'opérateur d'évolution

$$\langle \mathbf{x}, t | e^{-iHt} | \mathbf{y}, 0 \rangle = \int d^3p e^{-it\sqrt{\mathbf{p}^2+m^2}} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}$$

ne s'annule pas pour $|\mathbf{x} - \mathbf{y}|^2 > t^2$, cf cours de C. Bachas. et [PS, p 14].

- 3) Si la théorie décrite par (2.2) admet bien un courant conservé dans l'évolution temporelle, à savoir

$$j_\mu = \frac{i}{2m} (\psi^* \partial_\mu \psi - (\partial_\mu \psi)^* \psi) \quad (2.5)$$

satisfaisant $\partial^\mu j_\mu = 0$, on constate que la densité $\rho(x) = j_0$ n'est pas définie positive. Elle ne saurait représenter la densité de probabilité de présence de la particule décrite par ψ , qu'on attend dans une théorie sensée.

Toutes ces incohérences vont être levées par l'introduction de la théorie quantique des champs, une théorie décrivant un système à nombre arbitraire de particules, dans laquelle l'objet ψ est promu du rôle de fonction d'onde à celui d'opérateur capable de créer et annihiler ces particules. C'est le changement de point de vue traditionnellement appelé *seconde quantification*.

Nous examinerons maintenant une autre équation d'onde relativiste, originellement proposée par Dirac pour pallier aux insuffisances de l'équation de Klein-Gordon. L'équation de Dirac conduit elle aussi à des incohérences, également levées par la théorie quantique. Néanmoins son étude se révèle très profitable, sur le plan physique –elle offre une excellente description de phénomènes de basse énergie, atome d'hydrogène, etc– et sur le plan technique –représentations spinorielles, etc– qui seront utiles dans le contexte de la théorie quantique des champs.

3. Équation de Dirac

3.1. Les matrices γ de Dirac

Selon l'idée de Dirac, essayons de trouver une racine carrée à l'opérateur $\mathbf{p}^2 + m^2 = -\Delta + m^2$ en utilisant une représentation matricielle

$$i \frac{\partial \psi}{\partial t} = \left(\frac{1}{i} \boldsymbol{\alpha} \cdot \nabla + \beta m \right) \psi =: H \psi \quad (3.1)$$

où la fonction d'onde ψ a maintenant plusieurs composantes, supposées indépendantes, et $\boldsymbol{\alpha}$ et β sont des matrices hermitiennes, de telle façon que l'opérateur H défini par le

membre de droite soit lui même hermitien. On demande aussi que (3.1) soit compatible avec (2.2), chaque composante de ψ satisfaisant cette équation.

Si les matrices satisfont les relations d'*anticommutation*¹

$$\begin{aligned}\{\alpha_i, \alpha_j\} &= 0 \quad \text{si } i \neq j \\ \{\alpha_i, \beta\} &= 0 \\ \alpha_i^2 &= \beta^2 = I\end{aligned}\tag{3.2}$$

on calcule

$$\left(\frac{1}{i}\boldsymbol{\alpha}\cdot\nabla + \beta m\right)^2 = -\nabla^2 + m^2 = -\Delta + m^2$$

et l'on retrouve bien l'équation de Klein-Gordon en calculant $(i\partial_t)^2\psi$. Mettons l'équation de Dirac (3.1) sous forme covariante en introduisant le nouvel ensemble de matrices

$$\begin{aligned}\gamma^0 &= \gamma_0 = \beta \\ \gamma^i &= -\gamma_i = \beta\alpha^i\end{aligned}\tag{3.3}$$

qui satisfont donc

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}I.\tag{3.4}$$

Dans la suite, la matrice identité I du membre de droite de (3.4) (d'une dimension pas encore fixée !) sera souvent considérée comme implicite et omise. Noter que puisque les matrices α^i et β sont hermitiennes et anticommulent, la matrice γ^0 est hermitienne mais les γ^i , $i = 1, 2, 3$ sont *antihermitiennes*. Il sera utile d'utiliser aussi les matrices γ avec indices covariants

$$\gamma_\mu = g_{\mu\nu}\gamma^\nu = (\gamma_0, \gamma_1, \gamma_2, \gamma_3) = (\gamma^0, -\gamma^1, -\gamma^2, -\gamma^3).\tag{3.5}$$

ainsi que la notation

$$\not{a} := a^\mu\gamma_\mu.\tag{3.6}$$

On vérifie immédiatement comme conséquence de (3.4) la relation d'usage constant

$$\not{a}^2 = a^2 = a_\mu a^\mu.\tag{3.7}$$

On peut maintenant récrire l'équation de Dirac sous la forme

$$(i\not{\partial} - m)\psi \equiv (i\gamma^\mu\partial_\mu - m)\psi = 0,\tag{3.8}$$

et on se réfère à l'opérateur différentiel $(i\not{\partial} - m)$ comme l'opérateur de Dirac. On vérifie que

$$(i\not{\partial} - m)(i\not{\partial} + m) = (i\not{\partial} + m)(i\not{\partial} - m) = -(\square + m^2).$$

¹ L'anticommutateur $\{, \}$ est quelquefois noté $[,]_+$ par certains auteurs.

3.2. Une représentation explicite des matrices γ

Notons d'abord qu'à deux ou trois dimensions d'espace-temps, avec la métrique "minkowskienne" $(+, -)$ ou $(+, -, -)$, on connaît une solution de (3.4) en termes des matrices de Pauli²

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (3.9)$$

qui satisfont les relations bien connues

$$\sigma_i \sigma_j = \delta_{ij} + i \epsilon_{ijk} \sigma_k \quad \text{donc} \quad \{\sigma_i, \sigma_j\} = 2\delta_{ij}.$$

Cela permet de choisir à deux ou trois dimensions,

$$\beta = \sigma_1 \quad \alpha_1 = \sigma_2 \quad \text{et éventuellement} \quad \alpha_2 = \sigma_3.$$

À quatre dimensions, cela suggère de construire les matrices cherchées comme blocs de matrices de Pauli. On peut prendre par exemple

$$\begin{aligned} \gamma^0 &= \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} & \gamma^i &= \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \\ \beta &= \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} & \alpha^i &= \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \end{aligned} \quad (3.10)$$

ce qui constitue la "représentation de Dirac", mais il existe d'autres choix utiles, comme on le verra par la suite.

Noter que les matrices (3.10) satisfont aussi la propriété de conjugaison suivante

$$\gamma^{\mu\dagger} = \gamma^0 \gamma^\mu \gamma^0. \quad (3.11)$$

3.3. L'algèbre de Clifford des matrices γ . La "γ-gymnastique".

L'algèbre engendrée par l'identité I et des générateurs γ^μ satisfaisant (3.4) est ce qu'on appelle une algèbre de Clifford. Cherchons en une base linéaire. On introduit la notation

$$\gamma_5 \equiv \gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3. \quad (3.12)$$

On vérifie que $\gamma^{5\dagger} = -i\gamma_3^\dagger\gamma_2^\dagger\gamma_1^\dagger\gamma_0^\dagger = \dots = i\gamma^0\gamma^1\gamma^2\gamma^3 = \gamma^5$ est hermitienne et que

$$\{\gamma^5, \gamma^\mu\} = 0 \quad \text{et} \quad (\gamma^5)^2 = I. \quad (3.13)$$

² N.B. Pour les matrices de Pauli, on ne distinguera pas les positions supérieure ou inférieure de l'indice : $\sigma^i \equiv \sigma_i$.

Introduisons encore une autre notation très utile dans la suite :

$$\sigma_{\mu\nu} = -\sigma_{\nu\mu} := \frac{i}{2}[\gamma_\mu, \gamma_\nu] \quad (3.14)$$

avec cette fois les commutateurs des matrices γ . Le facteur i rend les matrices σ_{ij} hermitiennes, et les σ_{0i} antihermitiennes. Noter que γ^5 commute avec les $\sigma_{\mu\nu}$.

Vérifier que dans la représentation (3.10)

$$\gamma^5 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \quad \sigma^{0j} = i \begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix} \quad \sigma^{ij} = \epsilon_{ijk} \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix}. \quad (3.15)$$

Avec ces notations, nous pouvons énoncer le résultat

Proposition L'algèbre (de Clifford) des matrices γ est engendrée linéairement par les $4^2 = 16$ matrices suivantes $I, \gamma_\mu, \sigma_{\mu\nu}, \gamma^5\gamma_\mu, \gamma^5$.

Cela permet de démontrer le théorème important suivant

Théorème Toutes les solutions de (3.4) sont équivalentes, et toutes les solutions de (3.4)-(3.11) (en dimension 4) sont unitairement équivalentes. Autrement dit, pour toute paire de matrices γ_μ, γ'_μ satisfaisant (3.4), il existe une matrice inversible S telle que

$$\exists S \text{ inversible} \quad \gamma'_\mu = S\gamma_\mu S^{-1};$$

si en outre γ_μ et γ'_μ satisfont (3.11)

$$\exists U \text{ unitaire} \quad \gamma'_\mu = U\gamma_\mu U^\dagger. \quad (3.16)$$

La démonstration est laissée en exercice.

Dans le reste de ce paragraphe, toutes les identités énoncées ne sont conséquences que de (3.4) et éventuellement de (3.11), et sont donc indépendantes de la représentation choisie pour ces matrices.

Exercices Montrer que chacune des seize matrices de la liste précédente a un carré égal à $\pm I$ et que la trace de chacune (sauf I) s'annule. Vérifier leur indépendance linéaire. On en déduit alors la proposition précédente (cf [M], tome 2, p. 773). Pour chacune des matrices Γ de cette liste, comparer $\gamma^0\Gamma^\dagger\gamma^0$ et Γ .

Que peut-on dire de la matrice $i\epsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma$?

Vérifier la commutation de γ_5 et de $\sigma^{\mu\nu}$.

Montrer la décomposition de $\gamma_5\sigma_{\mu\nu} = \frac{i}{2}\epsilon_{\mu\nu\rho\sigma}\sigma^{\rho\sigma}$.

[En dimension paire quelconque $d = 2n$, une base linéaire de l'algèbre de Clifford est donnée par les produits antisymétrisés $1, \gamma_\mu, \gamma_{[\mu}\gamma_{\nu]}, \dots, \gamma_{[\mu_1} \dots \gamma_{\mu_d]}$. Quelle est la dimension de cette algèbre ? Quelle est la dimension minimale à donner aux matrices γ ?]

Des identités d'usage très courant dans les calculs sont les suivantes

$$\begin{aligned}
 \not{a}\not{b} &= a.b - i\sigma_{\mu\nu}a^\mu b^\nu \\
 \gamma^\mu\gamma_\mu &= 4 & \gamma^\mu\gamma^\nu\gamma_\mu &= -2\gamma^\nu \\
 \gamma^\mu\gamma^\nu\gamma^\rho\gamma_\mu &= 4g^{\nu\rho} & \gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\gamma_\mu &= -2\gamma^\sigma\gamma^\rho\gamma^\nu \\
 \gamma^\mu\sigma^{\rho\sigma}\gamma_\mu &= 0 & \gamma^\mu\sigma^{\nu\rho}\gamma^\sigma\gamma_\mu &= 2\gamma^\sigma\sigma^{\nu\rho}
 \end{aligned} \tag{3.17}$$

etc, etc. C'est évidemment un bon exercice de les vérifier.

Finalement, des traces des matrices de l'algèbre interviennent aussi souvent à la fin des calculs. Il est bon de savoir que

$$\text{tr}I = 4 \quad \text{tr}\gamma^\mu = 0 \quad \text{tr}\gamma_5 = 0 , \tag{3.18}$$

de se rappeler que la trace d'un nombre impair de matrices γ^μ , ($\mu = 0, 1, 2, 3$), s'annule (pourquoi ?), par exemple

$$\text{tr}\gamma_5\gamma^\mu = 0 \tag{3.19}$$

et de savoir calculer les autres, telles

$$\begin{aligned}
 \text{tr}\gamma^\mu\gamma^\nu &= 4g^{\mu\nu} & \text{tr}\sigma^{\mu\nu} &= 0 \\
 \text{tr}\gamma_5\gamma^\mu\gamma^\nu &= 0 & \text{tr}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma &= 4(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma} - g^{\mu\sigma}g^{\nu\rho}) \\
 \text{tr}\gamma_5\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma &= -4i\epsilon^{\mu\nu\rho\sigma} .
 \end{aligned} \tag{3.20}$$

3.4. Covariance relativiste. Représentation spinorielle. Parité

Pour démontrer l'invariance relativiste de l'équation de Dirac, il nous faut trouver la transformation de $\psi(x)$ sous l'effet d'une transformation de Lorentz de son argument x . Soit $x' = \Lambda x$ une transformation de Lorentz faisant passer des coordonnées dans un premier référentiel à celles dans un second³. On cherche une transformation $\psi \rightarrow \psi'$ telle que l'équation de Dirac prenne la même forme dans les deux référentiels,

$$(i\gamma^\mu \frac{\partial}{\partial x^\mu} - m)\psi(x) = 0 \tag{3.21a}$$

$$(i\gamma^\mu \frac{\partial}{\partial x'^\mu} - m)\psi'(x') = 0 . \tag{3.21b}$$

Supposons la transformation $\psi \rightarrow \psi'$ linéaire et écrivons-la sous la forme

$$\psi'(x') = S(\Lambda)\psi(x) . \tag{3.22}$$

³ Il s'agit là du point de vue *passif*. Un autre point de vue, dit actif, totalement équivalent au passif dans ses implications, consiste à considérer que x' est la coordonnée décrivant dans le même référentiel un autre système, obtenu à partir de l'original par la transformation Λ .

La matrice cherchée $S(\Lambda)$ est de dimension 4, elle est donc une combinaison linéaire des 16 matrices de la base de l'algèbre de Clifford du paragraphe précédent. Cette matrice doit aussi être inversible (transformation inverse $x' \rightarrow x$). L'équation (3.21b) se réécrit donc

$$(i\gamma^\mu \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} - m)S(\Lambda)\psi(x) = 0$$

et ceci doit être une conséquence de (3.21a), c'est-à-dire

$$S^{-1}(\Lambda)\gamma^\mu S(\Lambda) = \Lambda^\mu_\nu \gamma^\nu . \quad (3.23)$$

Cherchons $S(\Lambda)$ pour une transformation infinitésimale. On écrit

$$\Lambda^\mu_\nu = \delta^\mu_\nu + \omega^\mu_\nu \quad (\Lambda^{-1})^\mu_\nu = \delta^\mu_\nu - \omega^\mu_\nu + O(\omega^2)$$

avec $\omega_{\mu\nu}$ antisymétrique (cf. chapitre 0, § 4.2). Montrons que

$$\begin{aligned} S(\Lambda) &= I - \frac{i}{4}\sigma_{\mu\nu}\omega^{\mu\nu} + \dots \\ S^{-1}(\Lambda) &= I + \frac{i}{4}\sigma_{\mu\nu}\omega^{\mu\nu} + \dots \end{aligned} \quad (3.24)$$

où les points de suspension représentent des termes d'ordre supérieur en ω et les matrices σ sont celles qui ont été introduites plus haut en (3.14),

$$\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu] \quad (3.14)$$

satisfont bien (3.23). Cela est bien le cas puisque

$$\frac{i}{4}[\sigma_{\rho\sigma}, \gamma^\mu]\omega^{\rho\sigma} = \omega^{\mu\nu}\gamma_\nu ,$$

comme l'identité

$$[\sigma_{\rho\sigma}, \gamma^\mu] = -2i(\delta^\mu_\rho \gamma_\sigma - \delta^\mu_\sigma \gamma_\rho)$$

le montre, ce qui achève la démonstration. Pour une transformation finie, $S(\Lambda)$ est donc de la forme

$$S(\Lambda) = e^{-(i/4)\sigma_{\mu\nu}\omega^{\mu\nu}} . \quad (3.25)$$

On rappelle de la discussion du § 3.3 que les σ_{ij} sont hermitiennes, les σ_{0i} antihermitiennes, donc les $S(\Lambda)$ sont unitaires pour les rotations usuelles de \mathbb{R}^3 , mais hermitiennes pour un "boost".

Remarque. On vient de montrer que les matrices $\sigma_{\mu\nu}$ fournissent une solution à (3.23). Est-ce la seule ? Pourquoi est-elle naturelle ? Ces questions seront réexaminées plus bas, au chap. ?.

Moment angulaire

Pour terminer cette discussion, nous revenons à l'opérateur de spin covariant W^μ , construit à partir de P^μ et de l'opérateur de moment angulaire $J^{\nu\rho}$, générateur infinitésimal des transformations de Lorentz (cf chap. 0, § 4.2)

$$W_\mu = -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}J^{\nu\rho}P^\sigma . \quad (3.26)$$

Comme on le verra plus bas au chapitre ? dans l'étude des représentations du groupe de Poincaré, si M^2 désigne la valeur propre de l'opérateur P^2 , les valeurs propres possibles de W^2 sont

$$W^2 = -M^2S(S+1) \quad (3.27)$$

où S est le spin, entier ou demi-entier, de la représentation. Pour l'équation de Dirac (ou de Klein-Gordon) libre, $P^\mu = i\partial^\mu$ donc $P^2 = -\partial^2 = m^2$. Selon la discussion du chapitre 1, on identifie $J_{\mu\nu}$ en calculant

$$\begin{aligned} \psi'(x) &= \left(I - \frac{i}{2}J_{\mu\nu}\omega^{\mu\nu}\right)\psi(x) \\ &= \left(I - \frac{i}{4}\sigma_{\mu\nu}\omega^{\mu\nu}\right)\psi(x^\rho - \omega^\rho_\sigma x^\sigma) \\ &= \left(I - \frac{i}{4}\sigma_{\mu\nu}\omega^{\mu\nu} + x_\mu\omega^{\mu\nu}\partial_\nu\right)\psi(x) \end{aligned}$$

d'où

$$J_{\mu\nu} = \frac{1}{2}\sigma_{\mu\nu} + i(x_\mu\partial_\nu - x_\nu\partial_\mu) . \quad (3.28)$$

On calcule alors

$$W_\mu = \frac{i}{4}\epsilon_{\mu\nu\rho\sigma}\sigma^{\nu\rho}\partial^\sigma \quad (3.29)$$

et un dernier calcul, faisant appel à l'une des identités sur les produits de tenseurs ϵ du chapitre 1 (§ 2.2), conduit à

$$W^2 = -\frac{1}{2}\left(\frac{1}{2} + 1\right)m^2 \quad (3.30)$$

d'où il découle que le spin de la particule décrite par l'équation de Dirac est $\frac{1}{2}$.

Transformations discrètes P et T

Rappelons que le groupe de Lorentz a quatre *nappes* disconnexes, cf chap. 0, § 2.1. Nous avons examiné la covariance de l'équation de Dirac sous l'effet des transformations du groupe de Lorentz infinitésimales. Par intégration, les transformations finies sont celles qui sont dans la même composante connexe que l'identité. Il nous faut encore examiner le cas des transformations discrètes. Considérons d'abord la parité d'espace

$$x = (x^0, \mathbf{x}) \mapsto x' = \Lambda_P x = (x^0, -\mathbf{x}) , \quad (3.31)$$

avec

$$(\Lambda_P)^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} . \quad (3.32)$$

Selon le même argument qu'en (3.23), l'équation de Dirac garde la même forme dans les coordonnées x et x' si l'on peut trouver une matrice S_P , $\psi'(x') = S_P \psi(x)$, telle que S_P commute avec γ_0 et anticommute avec les γ^i , $i = 1, 2, 3$. Il est clair que tout multiple de la matrice γ^0 fait l'affaire et si on veut une transformation unitaire, il reste une phase arbitraire

$$\psi'(x') = \eta_P \gamma^0 \psi(x) . \quad (3.33)$$

L'opération \mathcal{T} de renversement du temps $t \rightarrow -t$ devrait en principe pouvoir se décrire de manière analogue. Il y existe en fait une différence importante, liée au caractère nécessairement *anti*-linéaire de la transformation, dans un contexte de mécanique quantique.

Rappelons d'abord ce qu'on entend par opérateur antilinéaire. Un tel opérateur satisfait

$$U(\lambda|\phi\rangle + \mu|\psi\rangle) = \lambda^* U|\phi\rangle + \mu^* U|\psi\rangle$$

et son adjoint est défini par

$$\langle\phi|U^\dagger|\psi\rangle = \langle U\phi|\psi\rangle^* = \langle\psi|U\phi\rangle .$$

S'il est en outre unitaire, on a

$$\langle\psi|\phi\rangle^* = \langle\phi|\psi\rangle = \langle\phi|U^\dagger U|\psi\rangle = \langle U\phi|U\psi\rangle^* ,$$

donc $\langle U\phi|U\psi\rangle = \langle\psi|\phi\rangle$.

En mécanique quantique, la transformation \mathcal{T} supposée unitaire (cf Théorème de Wigner au chapitre 3) laisse l'opérateur position \mathbf{x} inchangé, mais change le signe des vitesses, donc de l'impulsion \mathbf{p}

$$\begin{aligned} \mathbf{x}' &= \mathcal{T}\mathbf{x}\mathcal{T}^\dagger = \mathbf{x} \\ \mathbf{p}' &= \mathcal{T}\mathbf{p}\mathcal{T}^\dagger = -\mathbf{p} . \end{aligned} \quad (3.34)$$

Les relations de commutation canoniques ne sont compatibles avec cette transformation que si \mathcal{T} est antilinéaire

$$\begin{aligned} [x'_j, p'_k] &= -[x_j, p_k] = -i\hbar\delta_{jk} \\ &= \mathcal{T}[x_j, p_k]\mathcal{T}^\dagger = \mathcal{T}i\hbar\delta_{jk}\mathcal{T}^\dagger \end{aligned}$$

Autre argument : \mathcal{T} commute avec les translations dans le temps dont le générateur est l'hamiltonien : $\mathcal{T}iHT^\dagger = -iH$ (puisque $t \rightarrow -t$). Si \mathcal{T} était linéaire, on conclurait que $\mathcal{T}HT^\dagger = -H$, ce qui est gênant si on veut que $\text{Spec}(H) \geq 0$!

Revenant à l'équation de Dirac, on cherche donc une transformation antilinéaire $\psi(x) \rightarrow \psi'(x')$ où $x' = (-t, \mathbf{x})$. La transformation est de la forme

$$\psi'(x') = \eta_T S_T \psi^*(x) , \quad (3.35)$$

où S_T est une matrice 4×4 telle que l'équation pour ψ' obtenue par conjugaison de (3.1) ait la forme (3.1)

$$i\frac{\partial\psi'}{\partial t'} = \left(-\frac{1}{i}(S_T\boldsymbol{\alpha}^*S_T^{-1})\cdot\nabla + S_T\beta^*S_T^{-1}m \right) \psi'(x') \quad (3.36)$$

Cela est bien le cas si

$$S_T\boldsymbol{\alpha}^*S_T^{-1} = -\boldsymbol{\alpha} \quad S_T\beta^*S_T^{-1} = \beta .$$

Dans la représentation usuelle des matrices γ , où $\alpha^{2*} = -\alpha^2$ tandis que les autres α et β sont réelles, il faut trouver S_T qui anticommute avec α^1 et α^3 et commute avec α^2 et β . On prend

$$S_T = -i\alpha^1\alpha^3 = i\gamma_1\gamma_3 \quad (3.37)$$

le facteur i étant choisi de façon à avoir $S_T^2 = I$.

3.5. Combinaisons bilinéaires de ψ et $\bar{\psi}$

Par conjugaison hermitique de (3.8), on a

$$i\partial_\mu\psi^\dagger\gamma^{\mu\dagger} + m\psi^\dagger = 0$$

ce qu'on note aussi

$$\psi^\dagger(i\overleftarrow{\partial}_\mu\gamma^{\mu\dagger} + m) = 0 .$$

On introduit la notation

$$\bar{\psi} := \psi^\dagger\gamma^0 \quad (3.38)$$

qui sera d'usage constant dans la suite. En utilisant (3.11) on a donc

$$\bar{\psi}(i\overleftarrow{\not{\partial}} + m) = 0 . \quad (3.39)$$

Par transformation de Lorentz, $\bar{\psi}$ se transforme selon

$$\bar{\psi}'(x') = \bar{\psi}(x)\gamma^0 S(\Lambda)^\dagger \gamma^0 = \bar{\psi}(x)S^{-1}(\Lambda) \quad (3.40)$$

y compris pour la parité, pour laquelle $\bar{\psi}(x') = \eta_P^* \bar{\psi}(x)\gamma^0$.

Nous allons maintenant étudier les différentes combinaisons bilinéaires $\bar{\psi}(x)\gamma^A\psi(x)$, où γ^A parcourt la base de l'algèbre de Clifford considérée plus haut. En utilisant (3.22), (3.40) et (3.23), on note que $\bar{\psi}\psi$ est un scalaire de Lorentz, $\bar{\psi}\gamma_\mu\psi$ un vecteur, $\bar{\psi}\sigma_{\mu\nu}\psi$ un tenseur antisymétrique d'ordre 2, et $\bar{\psi}\gamma_5\psi$ et $\bar{\psi}\gamma_5\gamma_\mu\psi$ respectivement un *pseudo*-scalaire et un *pseudo*-vecteur, ce qui veut dire que sous l'effet de la parité, ils reçoivent un signe opposé à celui d'un scalaire ou d'un vecteur.

Physiquement, le vecteur $j_\mu = \bar{\psi}\gamma_\mu\psi$ joue un rôle important. D'abord, en vertu des équations de Dirac satisfaites par ψ et $\bar{\psi}$, il est conservé

$$\partial^\mu j_\mu = \bar{\psi}(\overleftarrow{\not{\partial}} + \overrightarrow{\not{\partial}})\psi = 0. \quad (3.41)$$

En outre, sa composante temporelle $\psi^\dagger\psi$ est définie positive, et il paraît donc qualifié pour le rôle de courant de probabilité, dans l'interprétation de ψ comme une fonction d'onde. (Rappelons notre échec à ce stade avec l'équation de Klein-Gordon, cf fin du § 2.) En fait, par la suite, en théorie quantique des champs, $\bar{\psi}\gamma_\mu\psi$ sera interprété comme le courant électromagnétique.

Les considérations précédentes nous permettent aussi de répondre à la question : l'équation de Dirac découle-t-elle d'un principe d'action ? La réponse est oui, l'action de Dirac libre s'écrit

$$S = \int d^4x \bar{\psi}(x)(i\not{\partial} - m)\psi(x). \quad (3.42)$$

Dans le calcul des variations, on doit considérer les parties réelle et imaginaire de ψ comme indépendantes, ou ce qui revient au même, ψ et $\bar{\psi}$ comme indépendants, et on retrouve immédiatement (3.8) et (3.39). L'action est bien invariante relativiste.

Il est enfin important de noter que j_μ est le courant de Noether pour la transformation des champs

$$\psi(x) \rightarrow e^{i\alpha}\psi(x) \quad \bar{\psi}(x) \rightarrow e^{-i\alpha}\bar{\psi}(x). \quad (3.43)$$

Sa conservation reflète l'invariance de l'action (3.42) pour de telles transformations indépendantes de x ("globales").

On peut aussi étudier l'effet d'une transformation *chirale*

$$\psi(x) \rightarrow e^{i\alpha\gamma_5}\psi(x) \quad \bar{\psi}(x) \rightarrow \bar{\psi}(x)e^{i\alpha\gamma_5}, \quad (3.44)$$

d'une grande importance en physique hadronique. Montrer que le courant de Noether correspondant est $j_\mu^5(x) = \bar{\psi}(x)\gamma_\mu\gamma_5\psi(x)$ et que sa divergence est

$$\partial_\mu j^{\mu 5}(x) = 2im\bar{\psi}\gamma^5\psi. \quad (3.45)$$

Exercice. Identités de Fierz [...]

3.6. Représentations de Weyl et de Majorana.

Selon une observation déjà faite, il est quelquefois utile de considérer d'autres représentations des matrices γ que celle donnée plus haut. Deux autres représentations sont souvent considérées :

(*) la représentation de Weyl, ou représentation chirale, consiste à choisir γ^0 anti-bloc-diagonale, au contraire de (3.10)⁴.

$$\gamma^0 = -\sigma_1 \otimes I = \begin{pmatrix} 0 & -I \\ -I & 0 \end{pmatrix} \quad \gamma^i = i\sigma_2 \otimes \sigma_i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}. \quad (3.46)$$

On calcule alors

$$\gamma_5 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \quad \sigma^{0i} = i \begin{pmatrix} \sigma_i & 0 \\ 0 & -\sigma_i \end{pmatrix} \quad \sigma^{ij} = \epsilon_{ijk} \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k \end{pmatrix}. \quad (3.47)$$

On constate que les matrices $\sigma^{\mu\nu}$ sont diagonales par blocs. Puisque les transformations $S(\Lambda)$ de (3.25) en sont des exponentielles, on en conclut que dans cette représentation, si on écrit

$$\psi = \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix} \quad (3.48)$$

les composantes ψ_L et ψ_R , spineurs à deux composantes, se transforment indépendamment sous l'effet des rotations d'espace et des "boosts"⁵. (Dans le langage de la théorie des

⁴ Attention que d'autres auteurs peuvent avoir des conventions légèrement différentes, par exemple [PS].

⁵ Cette conclusion est bien sûr indépendante de la représentation choisie. Si on définit les combinaisons $\psi_{L,R} = \frac{1}{2}(I \pm \gamma^5)\psi$, vecteurs propres de la chiralité γ^5 : $\gamma^5\psi_{L,R} = \pm\psi_{L,R}$, la commutation de γ^5 avec $\sigma_{\mu\nu}$ déjà notée plus haut, donc avec $S(\Lambda)$, implique que la chiralité est préservée par les transformations de Lorentz.

groupes qu'on étudiera plus tard, la représentation du sous-groupe de Lorentz propre \mathcal{L}_+^\uparrow est manifestement "réductible".) Il est légitime de se demander pourquoi on ne s'est donc pas contenté de spineurs à deux composantes. La raison tient à la transformation de parité, effectuée par la matrice γ^0 . C'est la parité qui mélange ψ_L et ψ_R (et rend la représentation du groupe entier de Lorentz irréductible).

Noter que dans cette représentation, l'équation de Dirac s'écrit

$$(i\cancel{\partial} - m)\psi = \begin{pmatrix} -m & -i(\partial_0 - \vec{\sigma} \cdot \vec{\nabla}) \\ -i(\partial_0 + \vec{\sigma} \cdot \vec{\nabla}) & -m \end{pmatrix} \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix} = 0, \quad (3.49)$$

et donc que le terme de masse couple les deux *spineurs de Weyl*.

Dit autrement, le terme de masse $\bar{\psi}\psi$ se réécrit en termes de $\psi_L = \frac{1}{2}(I \pm \gamma^5)\psi$ et de $\bar{\psi}_L = \frac{1}{2}\bar{\psi}(I \mp \gamma^5)$ comme $\bar{\psi}\psi = \bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L$. Il couple les deux chiralités.

En l'absence de terme de masse, si $m = 0$, les deux équations se découplent en

$$i(\partial_0 \pm \vec{\sigma} \cdot \vec{\nabla})\psi_L = 0$$

(cf plus bas au § 6, la discussion des particules de masse nulle).

(*) La représentation de Majorana permet de bien prendre en compte les propriétés de *réalité* des spineurs. Si on choisit

$$\begin{aligned} \gamma^0 &= \begin{pmatrix} 0 & \sigma_2 \\ \sigma_2 & 0 \end{pmatrix} & \gamma^1 &= \begin{pmatrix} i\sigma_3 & 0 \\ 0 & i\sigma_3 \end{pmatrix} \\ \gamma^2 &= \begin{pmatrix} 0 & -\sigma_2 \\ \sigma_2 & 0 \end{pmatrix} & \gamma^3 &= \begin{pmatrix} -i\sigma_1 & 0 \\ 0 & -i\sigma_1 \end{pmatrix}, \end{aligned} \quad (3.50)$$

c'est-à-dire toutes les matrices γ purement imaginaires, l'opérateur différentiel $(i\cancel{\partial} - m)$ est purement réel et on peut en chercher des solutions réelles.

Exercice. Pour chacune de ces deux représentations, trouver quelle est la matrice unitaire U qui l'"entrelace" avec la représentation de Dirac, au sens de (3.16).

3.7. Matrices de Dirac en métrique euclidienne

Dans le cas euclidien à d dimensions, on cherche des matrices γ^μ , $\mu = 1, \dots, d$, hermitiennes et satisfaisant

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}I. \quad (3.51)$$

Ces matrices engendrent une algèbre de Clifford de dimension (comme espace vectoriel) 2^d . Supposons d pair. La dimension de ces matrices est donc au moins $2^{d/2}$. On en donnera une construction explicite plus bas. Étant de telles matrices, la matrice

$$\gamma_{d+1} = i^{-d/2}\gamma_1\gamma_2 \cdots \gamma_d$$

satisfait

$$\gamma_{d+1}^2 = I \quad \{\gamma_{d+1}, \gamma_\mu\} = 0 \quad (3.52)$$

(le vérifier : d est pair !). On peut donc considérer que (3.51) s'applique à $\mu, \nu = 1, \dots, d+1$.

Si d est impair, on peut prendre les $(d-1)$ matrices $\gamma_1, \dots, \gamma_{d-1}$ de la dimension $d-1$ et leur adjoindre la matrice $\gamma_d = i^{-\frac{d-1}{2}}\gamma_1\gamma_2 \cdots \gamma_{d-1}$.

Construction explicite par produit tensoriel du cas $d = 2$.

À $d = 2$, on prend $\gamma_1 = \sigma_1, \gamma_2 = \sigma_2, \gamma_3 = \sigma_3$ (comparer avec le cas minkowskien au § 3.2). Ces matrices sont hermitiennes et de trace nulle. Puis on construit par récurrence, pour toute dimension d paire

$$\begin{aligned} \gamma_\mu^{(d+2)} &= \sigma_1 \otimes \gamma_\mu^{(d)} = \begin{pmatrix} 0 & \gamma_\mu^{(d)} \\ \gamma_\mu^{(d)} & 0 \end{pmatrix} \quad \text{pour } \mu = 1, \dots, d+1 \\ \gamma_{d+2}^{(d+2)} &= \sigma_2 \otimes I_d = \begin{pmatrix} 0 & -iI_d \\ iI_d & 0 \end{pmatrix} \\ \text{d'où } \gamma_{d+3}^{(d+2)} &= \sigma_3 \otimes I_d = \begin{pmatrix} I_d & 0 \\ 0 & -I_d \end{pmatrix}. \end{aligned} \quad (3.53)$$

(I_d est la matrice identité de dimensions $2^{d/2} \times 2^{d/2}$.) Ces matrices γ sont hermitiennes en toute dimension paire. Vérifier qu'elles satisfont bien les relations d'anticommutation. Il est aussi intéressant de savoir lesquelles sont des matrices symétriques. Par construction, si $\gamma_\mu^{(d)}$ est symétrique, resp. antisymétrique, alors $\gamma_\mu^{(d+2)}$ l'est aussi pour $\mu = 1, \dots, d+1$ tandis que $\gamma_{d+2}^{(d+2)}$ est antisymétrique et $\gamma_{d+3}^{(d+2)}$ est symétrique. Donc en toute dimension paire,

$$\begin{aligned} \gamma_1, \gamma_3, \dots, \gamma_{d+3} &\text{ sont symétriques} \\ \gamma_2, \gamma_4, \dots, \gamma_{d+2} &\text{ sont antisymétriques.} \end{aligned}$$

Identités de trace dans le cas euclidien

Si d est pair, $\text{tr}\gamma_\mu = 0$, $\mu = 1, \dots, d+1$ et toute trace d'un nombre pair de matrices se calcule explicitement

$$\text{tr}\gamma_{\mu_1}\gamma_{\mu_2}\cdots\gamma_{\mu_{2n}} = \left(\sum_P \epsilon_P \delta_{\mu_{p_1}\mu_{p_2}} \cdots \delta_{\mu_{p_{2n-1}}\mu_{p_{2n}}} \right) \text{tr}I_d \quad \text{avec } \mu_i = 1, \dots, d+1$$

où la somme porte sur tous les *appariements* d'indices, et ϵ_P désigne la signature de la permutation correspondante. Le nombre de ces appariements est le nombre de Catalan $C_n = (2n)!/(n!(n+1)!)$. Pour un nombre impair de matrices γ , si une matrice γ_μ donnée apparaît un nombre de fois pair (ou 0 fois), la trace s'annule. Si toutes apparaissent une fois et une seule, et donc en particulier γ_{d+1} , on place cette dernière en dernière position et on a

$$\text{tr}\gamma_{\mu_1}\gamma_{\mu_2}\cdots\gamma_{\mu_d}\gamma_{d+1} = i^{d/2}\epsilon_{\mu_1\mu_2\cdots\mu_d}\text{tr}I$$

(noter que dans l'euclidien, $\epsilon_{12\dots d} = 1$.)

Exercice : vérifier toutes les assertions de ce paragraphe !

4. Contenu physique de l'équation de Dirac

4.1. Ondes planes et projecteurs

Comme $e^{\pm ik \cdot x}$ est fonction propre de l'opérateur différentiel $(i\partial - m)$, il est naturel de chercher des solutions de (3.8) sous la forme d'ondes planes. La notation traditionnelle consiste à écrire

$$\begin{aligned} \psi^{(+)}(x) &= e^{-ik \cdot x} u(k) && \text{solution d'énergie positive} \\ \psi^{(-)}(x) &= e^{ik \cdot x} v(k) && \text{solution d'énergie négative} \end{aligned} \quad (4.1)$$

étant entendu que $k^0 > 0$, et que $k^2 = m^2$. Le quadrivecteur k^μ est en effet l'énergie-impulsion de la particule. Les spineurs u et v satisfont

$$(\not{k} - m)u(k) = 0 \quad (\not{k} + m)v(k) = 0. \quad (4.2)$$

Pour construire explicitement $u(k)$ et $v(k)$, on se place dans le référentiel où la particule (de masse $m > 0$) est au repos : $k = (m, \mathbf{0})$. Les équations ci-dessus se réduisent à

$$(\gamma^0 - 1)u(m, \mathbf{0}) = 0 \quad (\gamma^0 + 1)v(m, \mathbf{0}) = 0$$

et une base de solutions, dans la représentation de Dirac (3.10), est donnée par

$$u^{(1)}(m, \mathbf{0}) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad u^{(2)}(m, \mathbf{0}) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad v^{(1)}(m, \mathbf{0}) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad v^{(2)}(m, \mathbf{0}) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (4.3)$$

Les solutions à k quelconque s'obtiennent alors en effectuant un boost sur les solutions au repos, ou plus simplement en écrivant

$$u^{(\alpha)}(k) = \frac{\not{k} + m}{\sqrt{2m(m + k^0)}} u^{(\alpha)}(m, \mathbf{0}) \quad v^{(\alpha)}(k) = \frac{-\not{k} + m}{\sqrt{2m(m + k^0)}} v^{(\alpha)}(m, \mathbf{0}), \quad (4.4)$$

qui sont bien solutions de (4.2) puisque $(\not{k} - m)(\not{k} + m) = k^2 - m^2 = 0$. Leur normalisation est choisie de façon que (avec $\alpha, \beta = 1, 2$)

$$\begin{aligned} \bar{u}^{(\alpha)}(k)u^{(\beta)}(k) &= \delta_{\alpha\beta} & \bar{u}^{(\alpha)}(k)v^{(\beta)}(k) &= 0 \\ \bar{v}^{(\alpha)}(k)v^{(\beta)}(k) &= -\delta_{\alpha\beta} & \bar{v}^{(\alpha)}(k)u^{(\beta)}(k) &= 0, \end{aligned} \quad (4.5)$$

comme on va le montrer ci-dessous. En effet, les équations (4.5) sont équivalentes à dire que les matrices

$$\Lambda_+(k) := \sum_{\alpha=1,2} u^{(\alpha)}(k) \otimes \bar{u}^{(\alpha)}(k) \quad \Lambda_-(k) := - \sum_{\alpha=1,2} v^{(\alpha)}(k) \otimes \bar{v}^{(\alpha)}(k) \quad (4.6)$$

sont les projecteurs sur les états d'énergie positive, resp. négative. Un petit calcul auxiliaire consiste à vérifier que $(\not{k} + m)^2 = 2m(\not{k} + m)$, $(\not{k} + m)(I + \gamma^0)(\not{k} + m) = 2(m + k^0)(\not{k} + m)$ ou encore $(I + \gamma^0)(\not{k} + m)(I + \gamma^0) = 2(m + k^0)(I + \gamma^0)$. Compte tenu de la normalisation dans (4.4), cela conduit à

$$\Lambda_{\pm}(k) = \frac{1}{2m(m + k^0)} (\not{k} \pm m) \frac{I \pm \gamma^0}{2} (\not{k} \pm m) = \frac{\pm \not{k} + m}{2m}. \quad (4.7)$$

et à

$$\Lambda_{\pm}^2(k) = \Lambda_{\pm}(k) \quad \Lambda_{\pm}(k)\Lambda_{\mp}(k) = 0 \quad \text{tr}\Lambda_{\pm} = 2 \quad \Lambda_+(k) + \Lambda_-(k) = I \quad (4.8)$$

comme on s'y attend pour des projecteurs orthogonaux sur des espaces de dimension 2.

Cette dimensionnalité 2 reflète l'existence des degrés de liberté de spin. On désire maintenant construire les projecteurs sur les états propres de spin, analogues covariants de $\frac{1}{2}(1 \pm \sigma_3)$ pour un spin $\frac{1}{2}$ non-relativiste. Les solutions au repos $u^{(\alpha)}(m, \mathbf{0})$ et $v^{(\alpha)}(m, \mathbf{0})$ de (4.3) sont clairement vecteurs propres de $\Sigma^3 := \sigma_{12} = \begin{pmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{pmatrix}$ pour les valeurs propres 1 pour $\alpha = 1$ et -1 pour $\alpha = 2$. (En général, on définit $\Sigma^i = \frac{1}{2}\epsilon_{ijk}\sigma_{jk}$ et dans la base usuelle de matrices γ , $\Sigma^i = \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix}$.) Les projecteurs sur les états propres de spin des solutions au repos sont donc $\frac{1}{2}(I \pm \Sigma^3)$. Si \hat{z} est le vecteur unitaire dans la direction 3 (axe des z), on définit

$$P(\hat{z}) = \frac{1}{2}(I + \gamma^5 \not{\hat{z}}) = \frac{1}{2} \begin{pmatrix} I + \sigma_3 & 0 \\ 0 & I - \sigma_3 \end{pmatrix}$$

puisque $\gamma^5 \not{\hat{z}} = -\gamma^5 \gamma^3 = i\gamma^0 \gamma^1 \gamma^2 = \gamma^0 \sigma_{12} = \gamma^0 \Sigma^3$. P projette donc sur la solution $u^{(1)}$ d'énergie positive de spin $\frac{1}{2}$ et sur celle $v^{(2)}$ d'énergie négative de spin $-\frac{1}{2}$. Attention au signe ! Ce choix peut paraître un peu pervers, mais en fait se prête bien à l'application d'une transformation de Lorentz. Un boost de Lorentz Λ amène en effet les solutions au repos sur les solutions (4.4), amène \hat{z} sur le (quadri)vecteur $n^{(z)} := \Lambda \hat{z}$ et transforme $P(\hat{z})$ en le projecteur cherché

$$P(n^{(z)}) = \frac{1}{2}(I + \gamma_5 \not{n}^{(z)}), \quad (4.9)$$

le projecteur sur $u^{(1)}(k)$ et $v^{(2)}(k)$. (On rappelle que les matrices $S(\Lambda)$ de Lorentz commutent avec γ^5 et transforment $\not{\hat{z}}$ en $\not{n}^{(z)}$.) On peut généraliser cette expression à tout n orthogonal à k , donc de genre espace, et normalisé à $n^2 = -1$ et considérer donc $P(n) = \frac{1}{2}(I + \gamma^5 \not{n})$. Si Λ est le boost qui amène $(m, \mathbf{0})$ sur k (ne pas confondre ce Λ avec les projecteurs $\Lambda_{\pm}(k)$!), l'opérateur $P(n)$ projette sur l'état de spin qui dans le référentiel au repos est polarisé dans la direction $\Lambda^{-1}n$. Par abus de langage, on dira que $P(n)$ est le projecteur sur l'état polarisé le long de n . Les relations

$$\begin{aligned} [\Lambda_{\pm}(k), P(n)] &= 0 \\ (\Lambda_+(k) + \Lambda_-(k))(P(n) + P(-n)) &= I \\ \text{tr} \Lambda_{\pm}(k) P(n) &= 1, \end{aligned} \quad (4.10)$$

qui sont aisément vérifiées, montrent que les deux types de projecteurs Λ_{\pm} et P sont compatibles (ils commutent), que leurs produits $\Lambda_{\pm}P(\pm n)$ sont des projecteurs sur des

espaces complémentaires de dimension 1, toutes choses compatibles avec leur interprétation comme projecteurs sur les états d'énergie-impulsion et de spin fixés.

Remarque. Une remarque importante. Tout, dans la discussion de ce paragraphe, a été restreint au cas où $m \neq 0$: usage répété du référentiel de repos, normalisations avec dénominateur m ... La situation de masse nulle va nécessiter un traitement séparé, cf plus bas § 6.

Remarques et exercices

- De quel opérateur de spin généralisant Σ_3 les solutions (4.4) sont-elles vecteurs propres ? Montrer que $u^{(\alpha)}, v^{(\alpha)}$ sont vecteurs propres de $W.n^{(z)}$, où W est l'opérateur de spin de Pauli-Lubanski, cf (3.26).
- Montrer que (4.6) peut être raffiné en $u^{(1)}(k) \otimes \bar{u}^{(1)}(k) = \frac{k+m}{2m} \frac{1+\gamma^5 \not{p}^{(z)}}{2}$. Écrire les relations analogues pour les autres spineurs (4.4).
- **Hélicité** Il existe un choix privilégié de n tel que \mathbf{n} soit proportionnel à \mathbf{k} dans le référentiel considéré (ou dans tout autre référentiel s'en déduisant par un boost dans la direction de \mathbf{k}). Soit

$$n_k := \left(\frac{|\mathbf{k}|}{m}, \frac{k^0}{m} \frac{\mathbf{k}}{|\mathbf{k}|} \right) \quad (4.11)$$

On a bien $n_k^2 = -1$, $n_k.k = 0$, donc $\Lambda^{-1}n_k = (0, \mathbf{k}/|\mathbf{k}|)$ dans le référentiel de repos. Alors, le projecteur $P(n_k)$ agissant sur les solutions d'énergie positive, resp. négative, donne

$$P(n_k)\Lambda_{\pm}(k) = \frac{1}{2} \left(1 \pm \frac{\boldsymbol{\Sigma}.\mathbf{k}}{|\mathbf{k}|} \right) \Lambda_{\pm}(k)$$

et donc projette sur les états propres de spin le long de la direction de \mathbf{k} , ce qu'on appelle les états d'hélicité : il projette sur les états d'énergie et d'hélicité positives, et sur les états d'énergie et d'hélicité négatives.

Un état propre d'hélicité positive le reste-t-il par transformation de Lorentz ? La réponse est non, expliquer pourquoi.

- Démontrer l'identité de Gordon

$$\bar{u}^{(\alpha)}(p)\gamma^{\mu}u^{(\beta)}(q) = \frac{1}{2m}\bar{u}^{(\alpha)}(p)\left((p+q)^{\mu} + i\sigma^{\mu\nu}(p-q)_{\nu}\right)u^{(\beta)}(q). \quad (4.12)$$

(On pourra noter que, sandwiché entre $\bar{u}(p)$ et $u(q)$, γ^{μ} peut être remplacé par $(\not{p}\gamma^{\mu} + \gamma^{\mu}\not{q})/2m$.) Quelle serait la formule analogue pour des solutions d'énergie négative ? pour l'élément de matrice de $\gamma_5\gamma^{\mu}$?

4.2. Couplage minimal au champ électromagnétique

Supposons maintenant que la particule sujette à l'équation de Dirac porte une charge électrique q . (Pour un électron cette charge serait négative $q = e = -|e|$.) Selon un principe maintenant bien éprouvé, nous construisons l'équation de Dirac en présence d'un champ électromagnétique par le couplage minimal $i\partial_{\mu} \rightarrow i\partial_{\mu} - qA_{\mu}$, soit

$$(i\cancel{\partial} - q\cancel{A} - m)\psi = 0. \quad (4.13)$$

Notant $D^\mu = \partial^\mu + iqA^\mu$ (dérivée covariante), donc $\mathcal{D} = \not{\partial} + iq\not{A}$, calculons $(\mathcal{D})^2$. On ne peut utiliser brutalement l'identité $\not{a}^2 = a^2$ valable pour un quadrivecteur, en raison de la non-commutation des dérivées covariantes. On écrit

$$\begin{aligned}
 \mathcal{D}^2 &= D^\mu D^\nu \gamma_\mu \gamma_\nu \\
 &= \frac{1}{2} D^\mu D^\nu \{\gamma_\mu, \gamma_\nu\} + \frac{1}{2} [D^\mu, D^\nu] \gamma_\mu \gamma_\nu \\
 &= D_\mu D^\mu + \frac{iq}{4} (\partial^\mu A^\nu - \partial^\nu A^\mu) [\gamma_\mu, \gamma_\nu] \\
 &= D^2 + \frac{1}{2} q F^{\mu\nu} \sigma_{\mu\nu} .
 \end{aligned} \tag{4.14}$$

En formant

$$(i\not{\partial} - q\not{A} + m)(i\not{\partial} - q\not{A} - m)\psi = (-D^2 - \frac{1}{2} q F^{\mu\nu} \sigma_{\mu\nu} - m^2)\psi = 0$$

on trouve une équation qui diffère de celle obtenue à partir de l'équation de Klein-Gordon par le couplage minimal. Le terme supplémentaire $-\frac{q}{2} F \cdot \sigma$ décrit un couplage du champ à un moment dipolaire électrique et à un moment magnétique portés par la particule. En effet, dans la représentation de Dirac où $\sigma_{0j} = i\alpha_j$ et $\sigma_{ij} = \epsilon_{ijk} \text{diag}(\sigma_k)$, on trouve

$$-\frac{1}{2} q F^{\mu\nu} \sigma_{\mu\nu} = -q(i\boldsymbol{\alpha} \cdot \mathbf{E} + \boldsymbol{\sigma} \cdot \mathbf{B}) .$$

L'interprétation va en être plus claire dans la limite non-relativiste.

4.3. Limite non relativiste

Examinons la limite non relativiste de l'équation de Dirac couplée à un champ, (4.13). Autrement dit, on suppose les énergies en jeu, cinétique ou électromagnétique, petites par rapport à l'énergie au repos m (fois c^2 implicite). La représentation de Dirac des matrices se prête bien à cet exercice, car les deux composantes supérieures, resp. inférieures de ψ dans cette base sont d'ordre différent dans cette limite. Écrivant $\psi = \begin{pmatrix} \varphi \\ \chi \end{pmatrix}$, $\boldsymbol{\pi} := (\mathbf{p} - q\mathbf{A})$, et se rappelant la forme (3.10) des matrices $\boldsymbol{\alpha}$ et β , on trouve

$$\begin{aligned}
 i \frac{\partial \varphi}{\partial t} &= \boldsymbol{\sigma} \cdot \boldsymbol{\pi} \chi + qA^0 \varphi + m\varphi \\
 i \frac{\partial \chi}{\partial t} &= \boldsymbol{\sigma} \cdot \boldsymbol{\pi} \varphi + qA^0 \chi - m\chi .
 \end{aligned} \tag{4.15}$$

Dans la limite relativiste, le terme de masse m est le terme dominant dans la première équation, et il est cohérent de supposer que χ est sous-dominant par rapport à φ . Si on pose

$$\varphi = e^{-imt} \Phi \quad \chi = e^{-imt} X$$

avec Φ et X lentement variables, (4.15) devient

$$\begin{aligned} i\frac{\partial\Phi}{\partial t} &= \boldsymbol{\sigma}\cdot\boldsymbol{\pi}X + qA^0\Phi \\ i\frac{\partial X}{\partial t} &= \boldsymbol{\sigma}\cdot\boldsymbol{\pi}\Phi + qA^0X - 2mX . \end{aligned} \quad (4.16)$$

Sous l'hypothèse que $|qA^0| \ll 2m$ et que X est lentement variable, on peut approximer la solution de la deuxième équation par

$$X \approx \frac{\boldsymbol{\sigma}\cdot\boldsymbol{\pi}}{2m}\Phi \ll \Phi$$

et en reportant dans la première, on trouve

$$i\frac{\partial\Phi}{\partial t} = \left[\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2}{2m} + qA^0 \right] \Phi , \quad (4.17)$$

qui n'est autre que l'équation de Pauli, équation non-relativiste satisfaite par la fonction d'onde d'une particule de spin $\frac{1}{2}$ dans un champ électromagnétique. On calcule sans peine, en utilisant les relations de commutation des matrices de Pauli

$$\begin{aligned} (\boldsymbol{\sigma}\cdot\boldsymbol{\pi})^2 &= \frac{1}{2}\{\sigma_i, \sigma_j\}\pi^i\pi^j + \frac{1}{4}[\sigma_i, \sigma_j][\pi^i, \pi^j] = \boldsymbol{\pi}^2 - \frac{1}{2}i\epsilon_{ijk}\sigma_k \left(\frac{-q}{i} \right) (\nabla_i A^j - \nabla_j A^i) \\ &= \boldsymbol{\pi}^2 - q\boldsymbol{\sigma}\cdot\mathbf{B} , \end{aligned}$$

ce qui permet de récrire l'hamiltonien (le membre de droite de (4.17)) comme

$$H_{\text{Pauli}} = \frac{(\mathbf{p} - q\mathbf{A})^2}{2m} - \frac{q}{2m}\boldsymbol{\sigma}\cdot\mathbf{B} + qA^0 .$$

Le premier et le dernier termes sont ceux attendus du couplage minimal appliqué à l'équation de Schrödinger. Le deuxième terme décrit un couplage entre le spin et le champ magnétique, de la forme (en rétablissant les \hbar et c)

$$H_{\text{magnet}} = -\frac{q\hbar}{2m}\boldsymbol{\sigma}\cdot\mathbf{B} = -\boldsymbol{\mu}\cdot\mathbf{B}$$

où le moment magnétique $\boldsymbol{\mu}$ est

$$\boldsymbol{\mu} = 2\frac{q}{2m}\frac{\hbar\boldsymbol{\sigma}}{2} . \quad (4.18)$$

En général, pour un objet chargé de moment cinétique \mathbf{J} et de moment magnétique $\boldsymbol{\mu}$, on définit le *rapport gyromagnétique* g par

$$\boldsymbol{\mu} = g\frac{q}{2m}\mathbf{J} . \quad (4.19)$$

Ici $\mathbf{J} = \frac{1}{2}\hbar\boldsymbol{\sigma}$. Pour la particule décrite par l'équation de Dirac (et son approximation non-relativiste de Pauli), le rapport gyromagnétique est donc 2.

La théorie quantique des champs va apporter des corrections à cette valeur. Ces corrections sont infimes mais calculables et mesurables avec une précision inouïe pour une particule comme l'électron ou le muon qui ne sont soumis qu'aux interactions électrofaibles : l'accord entre expérience et théorie sur la valeur de $g_e - 2$ (ou de $g_\mu - 2$) est un des grands triomphes de la TQC. Les corrections sont par contre grandes et non calculables systématiquement pour le proton ou le neutron, soumis aussi aux interactions fortes : $g_p \approx 5.585$.

4.4. Spectre relativiste de l'atome d'hydrogène

cf TD, [BDm] et [IZ]

5. Particules et trous. La mer de Dirac.

5.1. La mer de Dirac. Particules et trous, antiparticules.

La présence d'états d'énergie négative non bornée inférieurement est très préoccupante puisqu'elle semble indiquer une instabilité de la théorie. La solution proposée par Dirac est élégante et instructive –on en retrouve des avatars dans des contextes très divers, depuis la physique des solides (théorie de Fermi) jusqu'à des problèmes de mathématiques (représentations de l'algèbre de Virasoro, par exemple). Elle consiste à supposer que dans le vide, tous les états d'énergie négative sont occupés –c'est la "mer de Dirac". Cela, joint au principe d'exclusion de Pauli, interdit une dangereuse transition d'un état physique d'énergie positive vers un de ces états. Inversement un état de cette mer peut être excité vers un état d'énergie positive, laissant derrière lui un "trou", de charge opposée, qui va être interprété comme une anti-particule, un positron si la particule de Dirac est l'électron. Il faut pour cela fournir une énergie E supérieure à l'intervalle d'énergie (*gap*) entre l'état initial et l'état final, soit $E \geq 2mc^2$. L'excédent $E - 2mc^2$ est emporté comme énergie cinétique par l'électron et le trou-positron. C'est le phénomène de création de paire.

La théorie de Dirac prédit donc des phénomènes nouveaux (nouveaux pour l'époque où elle a été proposée, et pour le cadre de la mécanique quantique habituelle) : création et annihilation de paires électron–positron. Cette théorie contient donc ce qui fait l'essence d'une théorie quantique de champs : l'existence d'un nombre de particules arbitraire.

5.2. Conjugaison de charge

Si ψ solution de l'équation de Dirac (4.13) décrit la particule de charge q , mettons l'électron, on voudrait trouver quel ψ^c décrit son antiparticule, de charge $-q$. On veut avoir simultanément

$$(i\cancel{\partial} - q\cancel{A} - m)\psi = 0 \quad (5.1a)$$

$$(i\cancel{\partial} + q\cancel{A} - m)\psi^c = 0 . \quad (5.1b)$$

En conjugant et transposant la première équation, on a

$$[\gamma_\mu^T(-i\partial^\mu - qA^\mu) - m]\bar{\psi}^T = 0 .$$

Si on peut trouver une matrice C telle que

$$C\gamma_\mu^T C^{-1} = -\gamma_\mu \quad (5.2)$$

alors $C\bar{\psi}^T$ satisfait (5.1b). Or il existe une telle matrice C dans toute représentation des matrices γ . Ainsi dans la représentation de Dirac, où γ^1 et γ^3 sont antisymétriques, et γ^0 et γ^2 symétriques, il suffit de prendre un multiple de $\gamma^0\gamma^2$. On choisit

$$\begin{aligned} C &= i\gamma^2\gamma^0 = \begin{pmatrix} 0 & -i\sigma_2 \\ -i\sigma_2 & 0 \end{pmatrix} \\ C^{-1} &= C^T = C^\dagger = -C \end{aligned} \quad (5.3)$$

et on prend

$$\psi^c = \eta_c C\bar{\psi}^T \quad (5.4)$$

avec une phase arbitraire η_c qu'on prendra égale à 1 dans la suite. L'opération de conjugaison de charge qui est antilinéaire dans le présent contexte, reprendra un caractère linéaire dans le cadre de la TQC.

Exercice : montrer que ψ^c se transforme par Lorentz comme ψ et donc que $\bar{\psi}\psi^c$ est un invariant de Lorentz. Un tel terme est appelé terme de masse de Majorana, quelle symétrie brise-t-il ? [Voir plus bas et en TD de TQC la discussion des champs de Majorana.]

Action de la conjugaison de charge sur le spin.

Soit ψ une solution de Dirac, état propre des projecteurs du § 4.1

$$\psi = \frac{\epsilon\not{p} + m}{2m} \frac{1 + \gamma^5\not{n}}{2} \psi \quad p^0 > 0 ,$$

donc d'énergie de signe ϵ et de polarisation "le long du vecteur n ". On calcule alors sa conjuguée de charge

$$\begin{aligned} \psi^c &= C\bar{\psi}^T = C\gamma^0 \left(\frac{\epsilon\not{p} + m}{2m} \right)^* \left(\frac{1 + \gamma^5\not{n}}{2} \right)^* \psi^* \\ &= \frac{-\epsilon\not{p} + m}{2m} \frac{1 + \gamma^5\not{n}}{2} \psi^c \end{aligned} \quad (5.5)$$

La conjugaison de charge fait donc passer –comme prévu– d'une solution d'énergie positive à une d'énergie négative. Le spin est renversé, ce qui est en accord avec notre discussion du § 5.1.

6. Particules de masse nulle

6.1. Chiralité et hélicité

Considérons l'équation de Dirac de masse nulle.

$$i\rlap{/}\partial\psi = 0 \text{ ou encore } \rlap{/}\mathbf{p}\psi = 0 . \quad (6.1)$$

On remarque est que l'opérateur de Dirac $\rlap{/}\partial$ anticommute avec l'opérateur de "chiralité" γ^5 . Ou dit autrement, puisque

$$\rlap{/}\partial(1 \pm \gamma^5) = (1 \mp \gamma^5)\rlap{/}\partial ,$$

l'opérateur de Dirac $\rlap{/}\partial$ fait passer de spineurs de chiralité positive (vecteurs propres de $\frac{1}{2}(1 + \gamma^5)$) à ceux de chiralité négative.

Plus tard dans le cours de TQC (discussion des "anomalies chirales"), on rencontrera le déterminant de l'opérateur de Dirac en présence d'un champ de jauge $\det \mathcal{D}(A)$. La remarque précédente, qui s'étend à $\mathcal{D}(A)$, montre que l'opérateur $\mathcal{D}(A)$ fait passer d'un espace à un autre, et que son déterminant ne peut avoir un sens intrinsèque, une source potentielle de difficultés . . .

On note aussi que $\gamma^5\gamma^0 = -i\gamma^1\gamma^2\gamma^3$, et donc $\gamma^5\gamma^0\gamma^i = \Sigma^i$, donc (6.1) multiplié par $\gamma^5\gamma^0$ implique

$$\gamma^5\mathbf{p}^0\psi = \mathbf{\Sigma}\cdot\mathbf{p}\psi \quad (6.2)$$

Pour une solution d'énergie positive, $\psi(x) = e^{-ik\cdot x}\tilde{\psi}(k)$, avec $k^2 = 0$, $k^0 = |\mathbf{k}|$, $\mathbf{k} = k^0\hat{\mathbf{k}}$ et (6.2) se réécrit

$$\mathbf{\Sigma}\cdot\hat{\mathbf{k}}\tilde{\psi}(k) = \gamma^5\tilde{\psi}(k) , \quad (6.3)$$

ce qui signifie que l'hélicité égale la chiralité. Pour une solution d'énergie négative, c'est l'opposé.

Remarque. On a déjà souligné que la chiralité est conservée par les transformations du groupe de Lorentz. Ce n'est en général pas le cas de l'hélicité, sauf dans le cas de masse nulle. Si $m \neq 0$, un spineur d'hélicité positive acquiert sous l'effet d'une transformations du groupe de Lorentz une composante d'hélicité opposée. Si $m = 0$, ce n'est pas le cas, comme on vient de le voir.

Exercices

– En indexant les solutions d'énergie positive et négative par leur chiralité, on écrit

$$\psi(x) = \begin{cases} e^{-ik\cdot x}u_{\pm}(k) \\ e^{ik\cdot x}v_{\pm}(k) \end{cases} \quad \text{avec } k^2 = 0, k^0 = |\mathbf{k}| > 0 . \quad (6.4)$$

$$\gamma^5u_{\pm}(k) = \pm u_{\pm}(k) \quad \gamma^5v_{\pm}(k) = \pm v_{\pm}(k) .$$

Trouver l'expression explicite des spineurs à quatre composantes $u_{\pm}(k)$ et $v_{\pm}(k)$ comme fonctions des coordonnées θ et φ de $\hat{\mathbf{k}}$.

– Montrer que le conjugué de charge d'un spineur de chiralité positive est de chiralité négative.

6.2. Spineurs de Weyl

Comme on l'a noté plus haut au § 3.6, pour des particules de masse nulle, les deux composantes de chiralité obéissent à des équations découplées. Si on n'admet qu'une seule chiralité et qu'on renonce à l'invariance par parité, on peut se contenter d'une réalisation des matrices de Dirac de dimension 2. On parle de spineurs de Weyl pour ces spineurs à deux composantes, de chiralité donnée

$$\begin{aligned} \text{chiralité} = 1, & & (p^0 - \mathbf{p}\cdot\boldsymbol{\sigma})\psi_R(p) = 0 \\ \text{chiralité} = -1, & & (p^0 + \mathbf{p}\cdot\boldsymbol{\sigma})\psi_L(p) = 0 \end{aligned} \tag{6.5}$$

6.3. Petite incursion dans le monde des neutrinos

Pendant longtemps, on a cru que les neutrinos des trois espèces (“familles” ou “générations”) connues, ν_e , ν_μ et ν_τ , étaient de masse nulle, et n'existaient dans la nature que dans leur chiralité (et hélicité) négative, (neutrinos “gauches”). Leurs antiparticules, les antineutrinos $\bar{\nu}_e$, $\bar{\nu}_\mu$ et $\bar{\nu}_\tau$, n'existant, eux, que dans l'hélicité et la chiralité positives.

Les résultats observationnels accumulés au cours des dix ou quinze dernières années (oscillations dans les neutrinos solaires et autres) ne semblent plus laisser place au doute : les neutrinos ont une masse petite mais non nulle, les deux chiralités existent même si leur physique est très différente.

Plusieurs scénarii sont actuellement envisagés. Ou bien les neutrinos sont des spineurs de Dirac, dotés d'une petite masse de Dirac couplant leurs composantes gauche et droite. La composante droite du ν ou la gauche du $\bar{\nu}$ ne participent pas aux interactions faibles usuelles et ne sont couplés que par leur terme de masse (et leur couplage au boson de Higgs). On parle de neutrinos “stériles”. Ce scénario est compatible avec la conservation du nombre leptonique, la quantité conservée associée par le théorème de Noether à l'invariance par $\psi \rightarrow e^{i\alpha}\psi$, $\bar{\psi} \rightarrow e^{-i\alpha}\bar{\psi}$ pour les différents champs de leptons.

Ou bien ils sont identiques à leur antiparticule, c'est-à-dire qu'ils sont ce qu'on appelle des particules de Majorana, ce qui autorise l'existence d'un terme de masse de Majorana (cd § 5.2). Ce qu'on appelle usuellement le ν et le $\bar{\nu}$ sont alors les deux états d'hélicité de cette particule de Majorana. Le terme de masse de Majorana viole la conservation du nombre leptonique. Un test de ce scénario consisterait en l'observation de la désintégration β double

$$(Z, A) \rightarrow (Z + 2, A) + e^- + e^-$$

interdite si le neutrino est une particule de Dirac, puisque violant la conservation du nombre leptonique.

En fait, la situation est plus compliquée que ce qui vient d'être esquissé. L'existence des trois familles de neutrinos autorise des termes de masse de type Dirac ou Majorana, bilinéaires dans ces différents champs et leurs conjugués. On parle alors d'une matrice de masse...

Ce sujet est d'une grande actualité, pour ses implications cosmologiques etc.

Les spineurs de Majorana jouent aussi un rôle central dans la construction de théories supersymétriques...

Pour plus de détails sur les spineurs et champs de Majorana, voir les TD ou [PS].

Problème

On démontre dans le cours de TQC que l'élément de matrice du courant électromagnétique entre deux états d'une particule de spin $\frac{1}{2}$ peut s'écrire

$$\langle p', \beta | j^\mu(x) | p, \alpha \rangle = e^{iq \cdot x} \bar{u}^{(\beta)}(p') O^\mu(p', p) u^{(\alpha)}(p)$$

où $q = p' - p$ et $O^\mu(p', p)$ est une matrice de l'algèbre de Clifford.

1. Dresser la liste des $O^\mu(p', p)$ compatibles avec l'invariance sous le groupe de Lorentz propre orthochrone et montrer que six d'entre eux sont linéairement indépendants, compte tenu de l'équation de Dirac satisfaite par $u^{(\alpha)}(p)$ et $\bar{u}^{(\beta)}(p')$.
2. Montrer que la contrainte de conservation du courant $\partial_\mu j^\mu(x) = 0$, donc de "transversité" $q_\mu O^\mu(p', p) = 0$, restreint cette liste à quatre termes.
3. En conclure que l'expression donnée au chapitre 3 d'un livre de TQC bien connu est incomplète et qu'on doit par exemple écrire

$$O^\mu(p', p) = \gamma^\mu F_1(q^2) + i \frac{\sigma^{\mu\nu} q_\nu}{2m} F_2(q^2) + \gamma^5 \frac{\sigma^{\mu\nu} q_\nu}{2m} F_3(q^2) + \gamma^5 (q^2 \gamma^\mu - \not{q} q^\mu) F_4(q^2) .$$

4. Montrer que l'hermiticité de l'opérateur courant j^μ se traduit par la condition $O^\mu(p, p')^\dagger = \gamma^0 O^\mu(p', p) \gamma^0$ et discuter quelles conditions de réalité cela fournit sur les fonctions $F_i(q^2)$.
5. Montrer que si j^μ se transforme comme $\bar{\psi} \gamma^\mu \psi$ sous l'effet des transformations discrètes, l'invariance par parité implique que $F_3(q^2) = F_4(q^2) = 0$, celle par renversement du temps $F_3(q^2) = 0$.
6. $j^0(x)$ étant la densité de charge électrique, montrer qu'on a

$$\int d^3x \langle p', \beta | j^0(x) | p, \alpha \rangle = q \frac{p^0}{2m} (2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}') \delta_{\alpha\beta}$$

(avec une normalisation conventionnelle) et en déduire que $F_1(0) = q$.

7. Pour l'interprétation physique des autres termes $F_i(q^2)$, il est utile de considérer la limite de particules presque au repos, $|\mathbf{p}|, |\mathbf{p}'| \ll m$. En étudiant dans cette limite $\int d^4x A_\mu(x) \langle p', \beta | j^\mu(x) | p, \alpha \rangle$, montrer que $F_1(0) + F_2(0)$ et $F_3(0)$ donnent les coefficients des termes $\boldsymbol{\mu} \cdot \mathbf{B}$ et $\boldsymbol{\sigma} \cdot \mathbf{E}$ dans le lagrangien d'interaction et décrivent les moments magnétique et dipolaire électrique de la particule considérée.
8. Le dernier coefficient, qui décrirait dans cette même limite le couplage du champ électromagnétique à un moment multipolaire d'ordre plus élevé de la distribution de charges est le seul à survivre pour des particules de Majorana, identiques à leur antiparticule.

Bibliographie sommaire

Pour la discussion générale de l'équation de Dirac, les identités sur les matrices γ etc, on a suivi [IZ], qui contient aussi une discussion détaillée de la limite non relativiste, du spectre de l'atome d'hydrogène, etc.

[BDm] est une autre référence très utile.

Pour la version "euclidienne" de l'équation de Dirac et des matrices γ , voir [Z-J].

La démonstration de la proposition et du théorème du § 3.3 peut se trouver dans A. Messiah, [M], t. 2, p 773.

Sur la physique des neutrinos, voir Cecilia Jarlskog, cours à l'École de Gif 92.

Chapitre 1

Groupes. Groupes et algèbres de Lie

1. Généralités sur les groupes

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...)

- 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 $n!$;
 - * les groupes d'homotopie, que nous allons rencontrer bientôt, sont d'autres exemples, etc, etc.
- Groupes infinis discrets. L'exemple le plus simple est le groupe additif \mathbb{Z} . Citons aussi les groupes de translations des réseaux réguliers.

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 ($\Im m \omega_2 \omega_1^* = \Im m \omega'_2 \omega_1'^*$) et leur effet sur $\tau = \omega_2/\omega_1$: $\tau \rightarrow (a\tau + b)/(c\tau + d)$. 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...

- 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 $((+)^p, (-)^q)$.

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 déplacement dans \mathbb{R}^3 , et les groupes obtenus en lui ajoutant les dilatations, puis les inversions par rapport à un point ;
- * le groupe de Galilée des transformations $\mathbf{x}' = R\mathbf{x} + \mathbf{v}t + \mathbf{x}_0$, $t' = t + t_0$,
- * $O(3, 1)$, le groupe de Lorentz, et le groupe de Poincaré,
etc etc.

1.2. Classes d'un groupe

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

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

et on dit aussi 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 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 $SU(n)$, où une classe est caractérisée par un n -tuple non ordonné de valeurs propres $(\lambda_1, \dots, \lambda_n)$. La notion de classe joue un rôle important dans la discussion des représentations des groupes et sera abondamment illustrée par la suite.

Ce type de classe d'équivalence ne doit pas être confondu avec celui impliquant un sous-groupe, que nous examinerons au § 1.5.

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 $a\mathbf{1} + b\sigma_1$, $a^2 - b^2 \neq 0$.

* 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

$$\begin{aligned} Z(S) &= \{y : \forall s \in S \quad y.s = s.y\} \\ N(S) &= \{x : x^{-1}.S.x = S\} . \end{aligned}$$

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:

$$\forall g, h \in G, \quad \rho(g.h) = \rho(g).\rho(h) \quad (1.6)$$

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 00 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 de l'élément neutre e' de G' . 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 $n!/2$.

1.5. Classes par rapport à un sous-groupe

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

$$g \sim g' \iff g.g'^{-1} \in H, \quad (1.7)$$

ce qu'on peut encore écrire comme

$$g \sim g' \iff \exists h \in H : g = h.g' \quad \text{ou encore} \quad g \in H.g' . \quad (1.8)$$

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

$$g \sim_g g' \iff g^{-1}.g' \in H \iff g \in g'.H. \tag{1.9}$$

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. 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 $|G|/|H|$ classes, et on obtient comme corollaire le théorème de Lagrange : l'ordre $|H|$ de tout sous-groupe H divise celui de G , et le rapport $|G|/|H|$ est l'ordre de l'ensemble quotient G/H .

L'équivalence à gauche donne en général une partition différente. 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.

1.6. *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 assertions.

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, 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 .

1.7. 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, § 2.2). Le groupe $SU(2)$ n'est ni simple, ni semi-simple, il contient en effet le sous-groupe invariant $\{I, -I\}$. Le groupe S_n n'est pas simple, pour $n > 2$ (pourquoi?).

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 $(g, h) \mapsto g.h$ 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 $g'.h'$ est proche de $g.h$ 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.

Commençons par étudier quelques propriétés topologiques de tels groupes continus.

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 est connexe s'il est à la fois ouvert et fermé), 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. [Po].

Exemples. $O(3)$ est disconnexe et la composante connexe de l'identité est $SO(3)$; pour le groupe de Lorentz $\mathcal{L}=O(3,1)$ 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 $PT\dots$

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

2.2. Simple connexité. Groupe d'homotopie. Recouvrement universel

Cette notion ne doit pas être confondue avec la précédente. On considère les chemins fermés tracés dans le groupe ou *lacets* à extrémité fixée, c'est-à-dire les applications continues $g(t)$ de $[0, 1]$ dans G telles que $g(0) = g(1) = g_0$. Étant donnés deux tels chemins $g_1(\cdot)$ et $g_2(\cdot)$ de g_0 à g_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 le groupe, telle que

$$\begin{aligned} \forall \xi \in [0, 1] \quad f(0, \xi) = f(1, \xi) = g_0 & \quad : \text{courbes fermées} \\ \forall t \in [0, 1] \quad f(t, 0) = g_1(t) \quad f(t, 1) = g_2(t) & \quad : \text{interpolation .} \end{aligned} \quad (2.1)$$

Si c'est le cas, on dit que les lacets g_1 et g_2 sont *homotopes* (c'est une relation d'équivalence), ou encore qu'ils appartiennent à la même *classe d'homotopie*, voir figure 1.

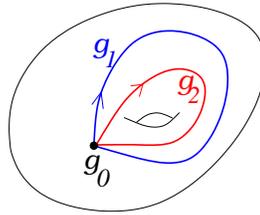


Fig. 1: Les lacets g_1 et g_2 sont homotopes. Mais aucun d'eux n'est homotope au lacet "trivial" qui reste en g_0 . L'espace (le groupe) n'est pas simplement connexe.

On peut aussi *composer* les chemins : Si $g_1(\cdot)$ et $g_2(\cdot)$ sont deux lacets de g_0 à g_0 , le chemin $g_2 \circ g_1$ va aussi de g_0 à g_0 en parcourant d'abord g_1 puis g_2 . Le lacet inverse de $g_1(\cdot)$ pour cette composition est le lacet parcouru en sens inverse : $g_1^{-1}(t) := g_1(1 - t)$. La composition et le passage à l'inverse sont compatibles avec l'homotopie : si $g_1 \sim g'_1$ et $g_2 \sim g'_2$, alors $g_2 \circ g_1 \sim g'_2 \circ g'_1$ et $g_1^{-1} \sim g'^{-1}_1$. 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 le *groupe d'homotopie* $\pi_1(G, g_0)$. Ainsi, un représentant de la classe identité est fourni par le lacet "trivial", $g(t) = g_0, \forall t$. On montre enfin que les groupes relatifs à des extrémités g_0 différentes sont isomorphes (dans un groupe connexe), et on peut se ramener au choix du point de base à l'identité $g_0 = e$. On parle donc **du** groupe d'homotopie (ou groupe fondamental) $\pi_1(G)$. Pour plus de détails, voir par exemple [Po], [DNF].

Si tous les lacets de g_0 à g_0 peuvent être contractés en le lacet trivial $\{g_0\}$, on dit que G est *simplement connexe*. Dans le cas contraire, on démontre, et nous admettrons,

que l'on peut construire un groupe \tilde{G} , dit *groupe de recouvrement universel* de G , tel que \tilde{G} est simplement connexe et que **localement**, G et \tilde{G} sont *homéomorphes*. Cela signifie qu'il existe une application continue surjective p de \tilde{G} dans G tel que tout point g de \tilde{G} ait un voisinage V_g et que $V_g \mapsto p(V_g)$ soit un homéomorphisme, c'est-à-dire une application bijective et bicontinue². Le groupe \tilde{G} de recouvrement universel de G est unique (à un homéomorphisme près). Dans le cas qui nous occupe d'une variété de groupe, l'application p est en fait 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

$$\tilde{G}/\pi_1(G) \simeq G, \quad (2.2)$$

(selon une propriété générale du groupe quotient par le noyau d'un homomorphisme, cf. § 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 $\pi_1(G)$.

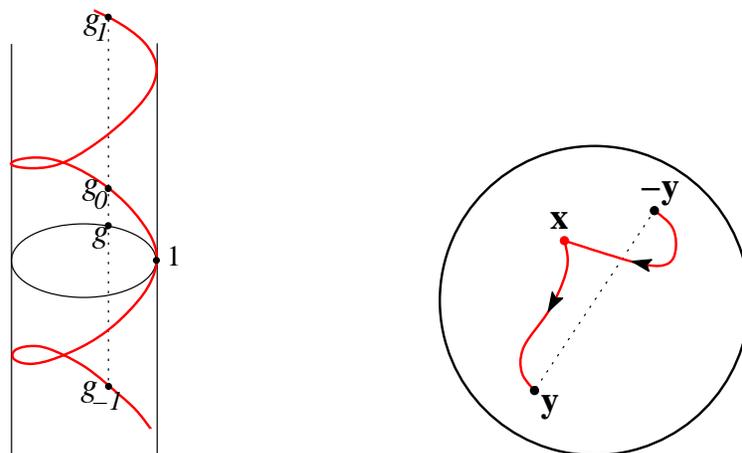


Fig. 2: Le groupe $U(1)$, identifié au cercle et son groupe de recouvrement universel \mathbb{R} , identifié à l'hélice. Un élément $g \in U(1)$ se relève en des points $\dots, g_{-1}, g_0, g_1, \dots$ sur l'hélice.

Fig. 3: 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.

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

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 $\pi_1(S^1) = \mathbb{Z}$. En fait la notion de groupe d'homotopie s'applique à tout espace topologique, et en particulier on se convainc aisément que pour les sphères, $\pi_1(S^n)$ est trivial (tous les lacets sont contractibles) dès que $n > 1$.

Autre exemple fondamental : Le groupe des rotations $SO(3)$ n'est pas simplement connexe, comme cela a été pressenti au Chapitre 00. Pour nous en convaincre, visualisons la rotation $R_{\mathbf{n}}(\psi)$ 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 00, §1.2.

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 $\pi_n(S^n) = \mathbb{Z}$. Cette notion est importante pour le physicien pour décrire des défauts topologiques, solitons, instantons, monopoles, etc. Voir [DNF] pour plus de détails et des calculs de ces groupes π_n .

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 quelques-unes des nombreuses caractérisations équivalentes d'un ensemble compact E . Toute suite infinie y admet un sous-suite convergente. Étant donné un recouvrement de E par un ensemble d'ouverts U_i , E peut être recouvert par un nombre fini d'entre eux. Toute fonction 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 ?

2.4. Mesure invariante

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

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

(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, § 2.3 : on définit d'abord une métrique sur $U(n)$ en écrivant $ds^2 = \text{tr } dU \cdot dU^\dagger$ 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 $d\mu_L(g) = y^{-2} dx dy$, $d\mu_R(g) = y^{-1} dx dy$ sont les mesures invariantes à gauche et à droite, respectivement, et que leurs intégrales divergent. Réf. [Bu].

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 !³ Ceci laisse entrevoir que la structure de groupe de Lie est très puissante et très rigide.

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

³ 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 !

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 $g(\xi'') = g(\xi').g(\xi)$, 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 $\dim(Sp(2n, \mathbb{R})) = \dim(USp(n)) = \dim(SO(2n+1))$, et 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.

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

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

$$\begin{aligned} (\lambda_1 X_1 + \lambda_2 X_2) * Y &= \lambda_1 X_1 * Y + \lambda_2 X_2 * Y \\ X * (\mu_1 Y_1 + \mu_2 Y_2) &= \mu_1 X * Y_1 + \mu_2 X * Y_2 . \end{aligned} \tag{3.1}$$

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.

Une *algèbre de Lie* est une algèbre dont le produit noté $[X, Y]$ a la propriété supplémentaire d'être antisymétrique et de satisfaire l'identité de Jacobi

$$\begin{aligned} [X, Y] &= -[Y, X] \\ [X_1, [X_2, X_3]] + [X_2, [X_3, X_1]] + [X_3, [X_1, X_2]] &= 0 . \end{aligned} \tag{3.2}$$

Exemples : Toute algèbre associative pour un produit noté $*$, en particulier toute algèbre de matrices, est une algèbre de Lie pour le produit (ou 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", $\mathbf{u} \wedge (\mathbf{v} \wedge \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$.

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),

$$g(t_1)g(t_2) = g(t_1 + t_2) \quad g^{-1}(t) = g(-t) . \quad (3.3)$$

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

$$g(t + \delta t) = g(t)g(\delta t) \quad \Leftrightarrow \quad g^{-1}(t)g(t + \delta t) = g(\delta t) . \quad (3.4)$$

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

$$X = \left. \frac{d}{dt} g(t) \right|_{t=0} , \quad (3.5)$$

c'est le vecteur vitesse en $t = 0$ (ou en $g = e$) le long de la courbe. L'équation (3.4) se réécrit donc

$$g'(t) = g(t)X . \quad (3.6)$$

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 $X = X^\alpha \frac{\partial}{\partial \xi^\alpha}$.

Dans le cas auquel nous nous restreignons d'un groupe $G \subset \text{GL}(n, \mathbb{R})$, $X \in \mathfrak{g} \subset M(n, \mathbb{R})$, l'ensemble des matrices réelles $n \times n$, et on peut effectuer tous les calculs dans cette algèbre. En particulier, on peut intégrer (3.6) selon

$$g(t) = \exp tX = \sum_{n=0}^{\infty} \frac{t^n}{n!} X^n, \quad (3.7)$$

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 $\exp X$, application dotée des propriétés usuelles de l'exponentielle, cf Appendice B4.)

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

★ Si G est le groupe linéaire $\text{GL}(n, \mathbb{R})$, \mathfrak{g} est l'ensemble des matrices réelles $n \times n$, noté $M(n, \mathbb{R})$. Si G est le groupe de matrices unitaires $\text{U}(n)$, \mathfrak{g} est l'ensemble des matrices antihermitiennes $n \times n$. Elles sont en outre de trace nulle si $G = \text{SU}(n)$. De même, pour le groupe orthogonal $\text{O}(n)$, \mathfrak{g} est constitué des matrices antisymétriques, et donc de trace nulle. Pour le groupe symplectique $G = \text{USp}(n)$, \mathfrak{g} est engendré par les matrices quaternioniques "antiselfduales", cf. Appendice A.

★ 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 (=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 $G = \text{U}(1)$, 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 $\text{SU}(2)$ qui est simplement connexe, si \mathbf{n} est de norme 1, $e^{i\pi \mathbf{n} \cdot \boldsymbol{\sigma}} = -I$, donc tous les éléments $\pi i \mathbf{n} \cdot \boldsymbol{\sigma}$ de $\mathfrak{g} = \text{su}(2)$ ont même image !

★ Exemple de groupe non compact pour lequel l'application \exp n'est pas surjective : $G = \text{SL}(2, \mathbb{R})$, pour lequel $\mathfrak{g} = \text{sl}(2, \mathbb{R})$, 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.

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

$$\begin{aligned} e^{tX} e^{uY} e^{-tX} e^{-uY} &= (I + tX + \frac{1}{2}t^2 X^2)(I + uY + \frac{1}{2}u^2 Y^2)(I - tX + \frac{1}{2}t^2 X^2)(I - uY + \frac{1}{2}u^2 Y^2) \\ &= I + (XY - YX)tu + O(t^3). \end{aligned} \quad (3.8)$$

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

$$e^{tX} e^{uY} e^{-tX} e^{-uY} = e^Z \quad , \quad Z = tu[X, Y] + O(t^3) \quad (3.9)$$

et on démontre que ce crochet a les propriétés (3.2) d'un crochet de Lie.

◦ *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

$$Y \mapsto (\text{ad } X)Y := [X, Y] \quad , \quad (3.10)$$

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 $e^X e^Y = e^Z$? 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 .

$$\begin{aligned} e^X e^Y &= e^Z \\ Z &= X + \int_0^1 dt \psi(\exp \text{ad } X \exp t \text{ad } Y)Y \end{aligned} \quad (3.11)$$

où $\psi(\cdot)$ est la fonction

$$\psi(u) = \frac{u \ln u}{u-1} = 1 + \frac{1}{2}(u-1) - \frac{1}{6}(u-1)^2 + \dots \quad (3.12)$$

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

$$Z = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}([X, [X, Y]] + [Y, [Y, X]]) + \dots \quad (3.13)$$

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

$$e^X e^Y = e^{X+Y + \frac{1}{2}[X, Y]} = e^{X+Y} e^{\frac{1}{2}[X, Y]}, \quad (3.14)$$

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

$$e^X Y e^{-X} = \sum_0^{\infty} \frac{1}{n!} \text{ad}^n X Y \quad (3.15)$$

(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 $f(t) = e^{tX} e^{tY}$, $f(0) = 1$

$$\begin{aligned} f'(t) &= (X + e^{tX} Y e^{-tX}) f(t) \\ &= (X + Y + t[X, Y]) f(t). \end{aligned} \quad (3.16)$$

Par ailleurs, au premier ordre en Y , on peut remplacer l'argument de ψ dans (3.11) par $\exp \text{ad} X$ et on voit qu'on a

$$Z = X + \sum_{n=0}^{\infty} \frac{B_n}{n!} (-1)^n (\text{ad} X)^n Y + O(Y^2) \quad (3.17)$$

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.

3.5. Un exemple explicite : l'algèbre de Lie de $SO(n)$

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 00, 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 $x \rightarrow x' = Ox$, 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

$$\begin{aligned} \delta x^i &= \omega^i_j x^j = -\frac{1}{2} \omega^{kl} J_{kl} x^i \\ J_{kl} &= x^k \partial_l - x^l \partial_k \quad : \quad J_{kl} x^i = x^k \delta_{il} - x^l \delta_{ik} \end{aligned} \quad (3.18)$$

(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 $\mathfrak{so}(n)$. C'est alors un calcul simple de calculer les relations de commutation⁴

$$[J_{ij}, J_{kl}] = \delta_{il} J_{jk} - \delta_{ik} J_{jl} - \delta_{jl} J_{ik} + \delta_{jk} J_{il} . \quad (3.19)$$

(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

⁴ Noter que par rapport au calcul mené pour le groupe $O(3,1)$ au chapitre 00, § 6.1, nous avons changé nos conventions et adopté ici des générateurs infinitésimaux antihermitiens.

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 (3.19).

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 $g = \text{diag}((+1)^p, (-1)^q)$.

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

$$\ell_n = -z^{n+1} \frac{\partial}{\partial z} , \quad n \in \mathbb{Z} \quad (3.20)$$

qui satisfont

$$[\ell_n, \ell_m] = (n - m)\ell_{n+m} \quad (3.21)$$

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

$$[L_n, L_m] = (n - m) + \frac{c}{12}n(n^2 - 1)\delta_{n,-m} \quad [c, L_n] = 0 . \quad (3.22)$$

(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 (3.21) 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].

4. Des propriétés de l'algèbre de Lie à celles du groupe

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

4.1. Simplicité, semi-simplicité

On va définir 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 $[\mathfrak{J}, \mathfrak{g}] \subset \mathfrak{J}$. L'idéal est dit *abélien* si $[\mathfrak{J}, \mathfrak{J}] = \{0\}$.

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 $\text{SO}(4)$, notée $\text{so}(4)$, cf les formules données en (3.19) pour $\text{so}(n)$. 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 $\text{so}(3)$. On écrit $\text{so}(4) = \text{so}(3) \oplus \text{so}(3)$.

À l'évidence l'algèbre $\text{so}(4)$ n'est pas simple, mais elle est semi-simple.

Bien noter la différence entre ce cas de $\text{so}(4)$ et le cas de l'algèbre $\text{so}(3,1)$ étudiée au chapitre 00, § 6.1. Là, la signature indéfinie nous a obligés à complexifier l'algèbre pour "découpler" les deux copies de l'algèbre $\text{so}(3)$.

On a les relations suivantes

$$G \text{ simple} \implies \mathfrak{g} \text{ simple}$$

$$G \text{ semi-simple} \implies \mathfrak{g} \text{ semi-simple}$$

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 § 1.7. ⁵

⁵ 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 qui n'est pas de Lie) ...

4.2. Compacité. Complexifiée

On dit de l'algèbre de Lie d'un groupe compact qu'elle est compacte.

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 $O(3)$ et $O(2,1)$, l'un compact, l'autre non, ont pour algèbres de Lie

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

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. 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.

Exercice. Étudier les algèbres réelles de dimension 3 : $\mathfrak{so}(3)=\mathfrak{su}(2)$, $\mathfrak{so}(2,1)$, $\mathfrak{su}(1,1)$, $\mathfrak{sp}(2, \mathbb{R})$, $\mathfrak{usp}(1)$ et $\mathfrak{sl}(2, \mathbb{R})$. En trouver les isomorphismes et montrer qu'à isomorphisme près, deux seulement sont indépendantes. (Voir [DNF] vol. 1, §13 et 24 pour plus de détails sur ces isomorphismes et leur interprétation géométrique.)

Les algèbres $\mathfrak{so}(4)$ et $\mathfrak{so}(3,1)$ étudiées plus haut et au chapitre 1 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.

L'algèbre de Lie ne capte cependant pas les propriétés topologiques globales du groupe.

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 $\mathfrak{g} = \mathfrak{g}'$.

– 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 $SU(2)$; $SO(3,1)$ 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 Cartan affirme que toute algèbre de Lie est l'algèbre de Lie d'un certain groupe de Lie [Ki, p.99]. 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.

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

Choisissant 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

$$[t_\alpha, t_\beta] = C_{\alpha\beta}^\gamma t_\gamma, \quad (4.1)$$

qui sont évidemment antisymétriques dans leurs deux indices inférieurs, $C_{\alpha\beta}^\gamma = -C_{\beta\alpha}^\gamma$. Considérons alors

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

pour l'opérateur ad défini plus haut. Considérons l'opérateur linéaire $\text{ad } X \text{ad } Y$ qui agit dans l'algèbre de Lie selon

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

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

$$\sum_{\delta} (C_{\alpha\delta}^\epsilon C_{\beta\gamma}^\delta + C_{\beta\delta}^\epsilon C_{\gamma\alpha}^\delta + C_{\gamma\delta}^\epsilon C_{\alpha\beta}^\delta) = 0 \quad (4.2)$$

(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

$$[\text{ad } X, \text{ad } Y]Z = \text{ad } [X, Y]Z. \quad (4.3)$$

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

$$(X, Y) := \text{tr}(\text{ad } X \text{ ad } Y) = \sum_{\gamma, \delta} C_{\alpha\delta}{}^{\gamma} C_{\beta\gamma}{}^{\delta} x^{\alpha} y^{\beta} =: g_{\alpha\beta} x^{\alpha} y^{\beta}, \quad (4.4)$$

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 $\text{ad } Z$:

$$\forall X, Y, Z \in \mathfrak{g} \quad ([Z, X], Y) + (X, [Z, Y]) = 0 \quad (4.5)$$

(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 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 $C_{\alpha\beta\gamma} + C_{\gamma\beta\alpha} = 0$. Exercice : le vérifier, en utilisant l'identité de Jacobi. Cette complète antisymétrie de $C_{\alpha\beta\gamma}$ équivaut à la propriété de cyclicité

$$(X, [Y, Z]) = (Y, [Z, X]) = (Z, [X, Y]) = C_{\alpha\beta\gamma} x^{\alpha} y^{\beta} z^{\gamma} \quad (4.6)$$

qui découle de la cyclicité de la trace et qui en offre une autre preuve.

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 , $g_{ij} = \sum_{1 \leq k, l \leq r} C_{i,k}^l C_{jl}^k$. 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 ((4.5))). 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

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

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 $\mathfrak{so}(2, 1)$.

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 ((4.6)). 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 ((4.6)), $([\mathfrak{C}, \mathfrak{J}], \mathfrak{J}) = (\mathfrak{C}, [\mathfrak{J}, \mathfrak{J}]) = (\mathfrak{C}, \mathfrak{J}) = 0$ (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.

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

$$C_2 = \sum_{\alpha,\beta} g^{\alpha\beta} t_\alpha t_\beta \quad (4.7)$$

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 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.

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

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.

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

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

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 (cf Chap. 00, § 2.3 pour le cas de $SO(3)$).

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.
- C. Chevalley, *Theory of Lie groups*, Princeton University Press.
- 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. (il y a aussi un troisième tome qui traite d'homologie ...)
- [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

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

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

[Ra] P. Ramond, *Group Theory, A Physicist's Survey*, Cambridge 2010.

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 $e_i = -i\sigma_i$ en termes des matrices de Pauli.

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

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

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

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

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

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

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 $Sp(2n, \mathbb{R})$ et $USp(n)$, algèbres de Lie $sp(2n)$ et $usp(n)$

Soit la matrice $2n \times 2n$

$$S = \begin{pmatrix} 0 & \mathbf{1}_N \\ -\mathbf{1}_N & 0 \end{pmatrix} \quad (\text{A.10})$$

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 $Sp(2n, \mathbb{R})$ 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 $S = \text{diag} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \text{diag}(-e_2)$ 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 $sp(2n, \mathbb{R})$ 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 $n = 1$, $Sp(2, \mathbb{R}) = SL(2, \mathbb{R})$.

Un groupe relié est le groupe $USp(n)$ 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 $U(2n)$. Son algèbre de Lie $usp(n)$ est engendrée par les matrices quaternioniques réelles antiselfduales $A = -A^R = -A^\dagger$ (le vérifier). Elle a aussi $n(2n + 1)$ pour dimension. Pour $n = 1$, $USp(1) = SU(2)$.

En exprimant la condition sur les matrices A de $sp(n, \mathbb{R})$ en termes de quaternions, on constate que les deux algèbres $sp(2n, \mathbb{R})$ et $usp(n)$ ont la même algèbre complexifiée, qui n'est autre que $sp(2n, \mathbb{C})$. Seule $usp(n)$ 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 : – toute suite infinie de points de E admet un point d'accumulation ;

– Si E est compact et $f : E \mapsto F$ est continue, $f(E)$ est compact ;

– toute fonction continue sur E compact est bornée.

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

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.

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 ,

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

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.

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

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

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 $\frac{\partial}{\partial x^i} \Big|_{x_0}$: 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} .

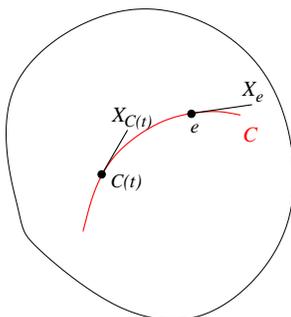


Fig. 4: Le champ de vecteurs tangents à la courbe $C(t)$ est un champ invariant à gauche.

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

$$C(t) = \exp tX_e \quad (\text{B.1})$$

l'unique solution de l'équation différentielle

$$C'(t) = X_{C(t)} \quad (\text{B.2})$$

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.

Cette fonction satisfait la propriété (3.3). En effet, si $C(t)$ satisfait (B.2), $C(t+t')$ 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 $C(-t) = C(t)^{-1}$.

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 (3.7).

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 (00-1.8) (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 $v_0^2 + \mathbf{v}^2 = 1$), ceci 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 rotation de l'espace \mathbb{R}^4 qui préserve la mesure naturelle $d^4u \delta(u^2 - 1)$ 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

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

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 $v(SO(3)) = \int_{SO(3)} d\mu(g) = \pi^2$.

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

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

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

Une autre méthode pour obtenir ces résultats passe par l’introduction d’une métrique invariante sur le groupe ; on définit une distance carrée entre deux éléments U et $U + dU$ par $ds^2 = \frac{1}{2} \text{tr} dU dU^\dagger$, invariante par $U \rightarrow UV$, $U \rightarrow VU$ ou $U \rightarrow U^{-1}$, et on en déduit une mesure d’intégration invariante (cf Chapitre 0, §2.3). Avec la paramétrisation ($\mathbf{n} = (\theta, \phi, \psi)$), on a

$$ds^2 = \frac{1}{2} \text{tr} dU dU^\dagger = \left(d\frac{\psi}{2} \right)^2 + \sin^2 \frac{\psi}{2} (d\theta^2 + \sin^2 \theta d\phi^2) , \quad (C.4)$$

qui conduit bien à (C.1). Dans la paramétrisation des angles d’Euler,

$$U = e^{-i\alpha \frac{\sigma_3}{2}} e^{-i\beta \frac{\sigma_2}{2}} e^{-i\gamma \frac{\sigma_3}{2}} \quad (C.5)$$

d’où

$$ds^2 = \frac{1}{2} \text{tr} dU dU^\dagger = \frac{1}{4} (d\alpha^2 + 2d\alpha d\gamma \cos \beta + d\gamma^2 + d\beta^2) \quad (C.6)$$

et avec $\sqrt{g} = \sin \beta$ on retrouve bien (C.3) (le vérifier).

Cas de $U(n)$

Examinons finalement rapidement le cas du groupe $U(n)$. Toute matrice unitaire $U \in U(n)$ peut se diagonaliser sous la forme

$$U = V \Lambda V^\dagger , \quad (C.7)$$

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 $U \mapsto UU'$, s’écrit $\text{tr}(dU dU^\dagger)$. Or $dU = V(d\Lambda + [dX, \Lambda])V^\dagger$, où $dX := V^\dagger dV$ est antihermitienne (et sans termes diagonaux, pourquoi ?). On a donc $\text{tr}(dU dU^\dagger) = \sum_i |d\alpha_i|^2 + 2 \sum_{i < j} |dX_{ij}|^2 |\lambda_i - \lambda_j|^2$ 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

$$d\mu(U) = \sqrt{\det g} \prod d\xi^\alpha = \text{const.} |\Delta(e^{i\alpha})|^2 \prod d\alpha_i d\mu(V) . \quad (C.8)$$

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

$$\Delta(\lambda) := \prod_{i < j} (\lambda_i - \lambda_j) = \begin{vmatrix} \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \\ \vdots & \vdots & & \vdots \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ 1 & 1 & \dots & 1 \end{vmatrix} . \quad (C.9)$$

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

$$d\mu(U) = \text{const.} \prod_{i < j} \sin^2 \left(\frac{\alpha_i - \alpha_j}{2} \right) \prod d\alpha_i \times \text{partie angulaire} . \quad (C.10)$$

Exercices pour le 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. Écrire précisément les conditions requises. On définit l'orbite $O(x)$ d'un point $x \in E$ comme l'ensemble des images $\beta(g)x$ pour $g \in G$.

1. Montrer que l'appartenance à une même orbite est une relation d'équivalence.
2. Exemple : action du groupe $O(n)$ sur l'espace \mathbb{R}^n . Que sont les orbites ?
3. 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 $\mathcal{L} = O(3,1)$.

4. 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 :

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

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 ?

5. 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.

(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 $(t_{(ij)})_{ab} = \delta_{ia} \delta_{jb}$ en 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 $C^{(r)} = \sum_{1 \leq i_1, i_2, \dots, i_r \leq n} t_{(i_1 i_2)} t_{(i_2 i_3)} \cdots t_{(i_r i_1)}$ 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

A-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

$$\delta S = \int d^d x (\partial_\mu a_\nu) \Theta^{\mu\nu}(x); \quad (\text{C.11})$$

- $\Theta_{\mu\nu}$ 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

$$ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu \rightarrow ds'^2 = g_{\mu\nu}(x')dx'^\mu dx'^\nu = \alpha(x)ds^2 \quad (\text{C.12})$$

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

$$\partial_\mu a_\nu + \partial_\nu a_\mu = \frac{2}{d}g_{\mu\nu}\partial_\rho a^\rho. \quad (\text{C.13})$$

3. Montrer en utilisant (C.11) et (C.13) 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 (C.13). On pose $D := \frac{1}{d}\partial_\rho a^\rho$.

a) En dérivant (C.13) par rapport à x^ν , montrer que

$$\partial^2 a_\mu = (2 - d)\partial_\mu D. \quad (\text{C.14})$$

b) En dérivant (C.14) 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 (C.14) par rapport à x^ν , en symétrisant en μ et ν et en utilisant (C.13), montrer que si $d > 2$, alors $\partial_\mu \partial_\nu D = 0$.

En conclure qu'il existe une constante h et un vecteur k tel que $D = k_\mu x^\mu + h$.

d) En dérivant (C.13) par rapport à x^σ et en antisymétrisant en ν et σ , montrer que

$$\partial_\mu(\partial_\sigma a_\nu - \partial_\sigma a_\nu) = 2(g_{\mu\nu}k_\sigma - g_{\mu\sigma}k_\nu) = \partial_\mu(2k_\sigma x_\nu - 2k_\nu x_\sigma). \quad (\text{C.15})$$

e) En déduire qu'il existe un tenseur antisymétrique constant $l_{\sigma\nu}$ tel que

$$\partial_\sigma a_\nu - \partial_\nu a_\sigma = (2k_\sigma x_\nu - 2k_\nu x_\sigma) + 2l_{\sigma\nu}, \quad (\text{C.16})$$

qui combiné avec (C.13) donne

$$\partial_\sigma a_\nu = x_\nu k_\sigma - x_\sigma k_\nu + l_{\sigma\nu} + g_{\nu\sigma}k_\rho x^\rho + h.$$

f) En conclure que l'expression générale d'une transformation conforme infinitésimale en dimension $d > 2$ s'écrit

$$a_\nu = k_\sigma x^\sigma x_\nu - \frac{1}{2} x_\sigma x^\sigma k_\nu + l_{\sigma\nu} x^\sigma + h x_\nu + c_\nu \quad (\text{C.17})$$

avec c un vecteur constant⁶. De combien de paramètres réels indépendants dépend une telle transformation en dimension d ?

B-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 (C.17), 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 ?

C-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 $\vec{x}^2 = x_1^2 + \dots + x_d^2$, sur la sphère S^d . Cette sphère est définie par l'équation $\vec{r}^2 + r_{d+1}^2 = 1$ dans l'espace \mathbb{R}^{d+1} , et l'application est réalisée grâce à la projection stéréographique à partir du “pôle Nord” $\vec{r} = 0, r_{d+1} = 1$ (voir figure). Montrer que l'on a

$$\vec{r} = \frac{2\vec{x}}{\vec{x}^2 + 1} \quad r_{d+1} = \frac{\vec{x}^2 - 1}{\vec{x}^2 + 1}.$$

Quelle est l'image du point à l'infini ? Quel est l'effet de l'inversion dans \mathbb{R}^d sur le point $r = (\vec{r}, r_{d+1}) \in S^d$?

⁶ Ce joli raisonnement est dû à Michel Bauer.

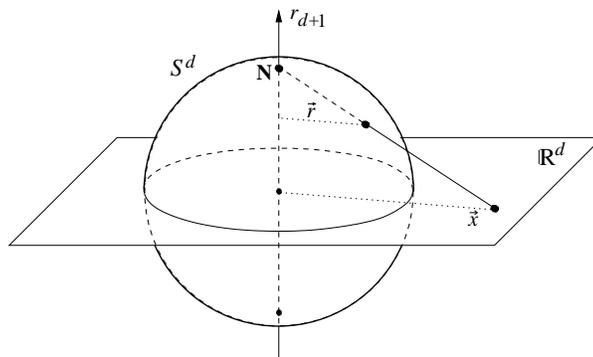


Fig. 5: Projection stéréographique

2. La sphère précédente est à son tour considérée comme la section du cône de lumière \mathcal{C} dans l'espace de Minkowski $\mathcal{M}_{d+1,1}$ de métrique $z_0^2 - \vec{z}^2 - z_{d+1}^2 = 0$ 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 $\vec{x} \in \mathbb{R}^d$ en fonction de $z = (z_0, \vec{z}, z_{d+1}) \in \mathcal{C}$ est

$$\vec{x} = \frac{\vec{z}}{z_0 - z_{d+1}} .$$

3. On va montrer maintenant que l'action du groupe conforme dans \mathbb{R}^d découle de transformations *linéaires* dans $\mathcal{M}_{d+1,1}$ préservant le cône de lumière. Sans aucun calcul, montrer que ces transformations doivent alors appartenir au groupe de Lorentz de $\mathcal{M}_{d+1,1}$, soit $O(d+1, 1)$.

a) Identifier les transformations linéaires de z correspondant aux rotations de \vec{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 $O(d+1, 1)$ qui préservent $z_0 - z_{d+1}$. Écrire la matrice T_a d'une telle transformation infinitésimale agissant sur les coordonnées (z_0, \vec{z}, z_{d+1}) telle que $\delta \vec{z} = \vec{a}(z_0 - z_{d+1})$ (au premier ordre en \vec{a}). A quelle transformation de $\vec{x} \in \mathbb{R}^d$ 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 $\mathcal{M}_{d+1,1}$? Que dire des transformations conformes spéciales ? Quelle est la dimension du groupe $O(d+1, 1)$? Qu'en conclure sur la relation entre le groupe de Lorentz dans l'espace de Minkowski $\mathcal{M}_{d+1,1}$ et le groupe conforme dans \mathbb{R}^d ?

D. Question subsidiaire : Connaissez-vous des transformations conformes de l'espace \mathbb{R}^2 autres que celles mentionnées au B.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.

1. Définition. Représentations équivalentes. Représentations unitaires

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 $GL(E)$:

$$\begin{aligned}\forall g \in G \quad g &\mapsto D(g) \in GL(E) \\ \forall g, g' \in G \quad D(g.g') &= D(g).D(g') \\ D(e) &= I \\ \forall g \in G \quad D(g^{-1}) &= (D(g))^{-1}\end{aligned}\tag{1.1}$$

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.

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 $D(g)$:

$$D(g)e_j = e_i \mathcal{D}_{ij}(g) \quad (1.2)$$

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 (1.1). En effet, on a bien

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

donc $\mathcal{D}_{ik}(g.g') = \mathcal{D}_{ij}(g) \mathcal{D}_{jk}(g')$. (1.3)

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

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (1.4)$$

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 ?

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

$$\forall g \in G \quad VD(g) = D'(g)V . \quad (1.5)$$

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.

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

$$\chi(g) = \text{tr}D(g) . \quad (1.6)$$

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 §3.1.

On notera encore que le caractère, évalué pour l'élément identité du groupe, fournit la dimension de la représentation

$$\chi(e) = \dim D . \quad (1.7)$$

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.

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

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.

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 .

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 de dimensions $\dim E_1$ et $\dim E_2$

$$\forall g \in G \quad \mathcal{D}(g) = \begin{pmatrix} \mathcal{D}_1(g) & 0 \\ 0 & \mathcal{D}_2(g) \end{pmatrix}. \quad (1.8)$$

Si la représentation est réductible sans être complètement réductible, sa matrice prend la forme, dans une base constituée par une base de E_1 et d'une base d'un sous-espace supplémentaire

$$\mathcal{D}(g) = \begin{pmatrix} \mathcal{D}_1(g) & \mathcal{D}'(g) \\ 0 & \mathcal{D}_2(g) \end{pmatrix}. \quad (1.9)$$

C'est le cas des représentations du groupe des translations à une dimension. La représentation

$$\mathcal{D}(a) = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \quad (1.10)$$

est réductible, puisqu'elle laisse invariants les vecteurs $(X, 0)$ mais n'a pas de sous-espace supplémentaire invariant.

Par contre, 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. (1.9), 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 (1.4) 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

$$\begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix}.$$

◦ *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 (1.3)

$$\mathcal{D}_{ik}^*(g \cdot g') = \mathcal{D}_{ij}^*(g) \mathcal{D}_{jk}^*(g') .$$

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 $SU(2)$.)

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

$$\overline{\mathcal{D}}(g) = D^{-1T}(g)$$

c'est-à-dire $\overline{\mathcal{D}}_{ij}(g) = \mathcal{D}_{ji}(g^{-1})$, qui satisfait bien aussi (1.3). Pour une représentation unitaire, cf alinéa suivant, $\overline{\mathcal{D}}_{ij}(g) = \mathcal{D}_{ij}^*(g)$, la contragrédiente s'identifie à la conjuguée. Les représentations D , D^* et \overline{D} sont simultanément réductibles ou irréductibles.

◦ *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 $x, y \in E$

$$\begin{aligned} \langle x|y \rangle &= \langle D(g)x|D(g)y \rangle \\ \text{donc } D(g)^\dagger D(g) &= I \end{aligned} \tag{1.11}$$

² Ceci est vrai au moins pour les représentations irréductibles des groupes finis ou continus compacts pour lesquelles on verra plus bas (§3) que deux représentations irréductibles non équivalentes ne peuvent avoir le même caractère.

³ Un espace est dit *séparable* s'il contient un sous-ensemble dénombrable dense.

et

$$\boxed{D(g^{-1}) = D^{-1}(g) = D^\dagger(g) \ .} \quad (1.12)$$

On a les propriétés importantes suivantes :

(i) *Toute représentation unitaire réductible est complètement réductible.*

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

$$\langle x | D(g)y \rangle = \langle D(g^{-1})x | y \rangle = 0 \quad (1.13)$$

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

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

Considérons d'abord le cas d'un groupe fini et formons

$$Q = \sum_{g' \in G} D^\dagger(g')D(g') \quad (1.14)$$

qui satisfait

$$D^\dagger(g)QD(g) = \sum_{g' \in G} D^\dagger(g'.g)D(g'.g) = Q \quad (1.15)$$

où on a remplacé $\sum_{g'}$ par $\sum_{g'.g}$ ("lemme de réarrangement"). L'opérateur auto-adjoint Q est défini positif, (pourquoi ?), on peut donc l'écrire sous la forme

$$Q = V^\dagger V \quad (1.16)$$

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 :

$$\begin{aligned} D'(g) &= VD(g)V^{-1} \\ D'^\dagger(g)D'(g) &= V^{\dagger-1}D^\dagger(g)V^\dagger VD(g)V^{-1} \\ &= V^{\dagger-1}D^\dagger(g)QD(g)V^{-1} = V^{\dagger-1}QV^{-1} = I \ . \end{aligned} \quad (1.17)$$

Dans le cas d'un groupe continu compact, l'existence de la mesure invariante de Haar (cf chap. 1, §2.4) permet de répéter l'argument avec $Q = \int d\mu(g') D^\dagger(g') D(g')$. \square

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 classifier 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).

1.4. Lemme de Schur

Soit deux représentations irréductibles D dans E et D' dans E' et un opérateur d'entrelacement entre elles, comme défini en (1.5). 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 (1.10) 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 contre-exemples aux deux propositions précédentes : toute matrice $\mathcal{D}(\alpha)$ commute avec $\mathcal{D}(\theta)$ mais n'a pas de valeur propre réelle 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 les choisir 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 .

1.5. Produit tensoriel de représentations ; décomposition de Clebsch-Gordan

o 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 tensoriel ⁴ $E = E_1 \otimes E_2$ 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 . Sur l'élément z ci-dessus

$$D(g)z = \sum_i D_1(g)x^{(i)} \otimes D_2(g)y^{(i)}. \quad (1.18)$$

⁴ On trouvera à l'Appendice A 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

$$\chi(g) = \chi_1(g)\chi_2(g) \quad (1.19)$$

En particulier en évaluant cette relation pour $g = e$, on a pour des représentations de dimension finie

$$\dim D = \dim (E_1 \otimes E_2) = \dim E_1 \cdot \dim E_2 = \dim D_1 \cdot \dim D_2 \quad (1.20)$$

comme il est bien connu pour un produit tensoriel.

◦ *Décomposition de Clebsch-Gordan*

La représentation produit direct de deux représentations irréductibles D et D' n'est en général pas irréductible. Si elle est complètement réductible (comme c'est le cas pour les représentations unitaires qui vont nous intéresser au premier chef), on effectue la *décomposition de Clebsch-Gordan* en représentations irréductibles

$$D \otimes D' = \oplus_j D_j \quad (1.21)$$

où au second membre apparaissent un certain nombre de représentations irréductibles D_1, \dots . La notation \oplus_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 (1.21) une autre écriture qui indique lesquelles de ces représentations inéquivalentes apparaissent, *et avec quelle multiplicité*

$$D \otimes D' = \oplus_{\rho} m_{\rho} D^{(\rho)} . \quad (1.22)$$

Une écriture plus correcte serait $E \otimes E' = \oplus_{\rho} F_{\rho} \otimes E^{(\rho)}$ où F_{ρ} est un espace vectoriel de dimension m_{ρ} , l'espace de multiplicité.

Les entiers m_{ρ} sont non négatifs. Les équations (1.21) et (1.22) impliquent des règles simples sur les caractères et les dimensions

$$\begin{aligned} \chi_D \cdot \chi_{D'} &= \sum_j \chi_j = \sum_{\rho} m_{\rho} \chi^{(\rho)} \\ \dim D \cdot \dim D' &= \sum_j \dim D_j = \sum_{\rho} m_{\rho} \dim D^{(\rho)} . \end{aligned} \quad (1.23)$$

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) ; 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

$$D^{(3)} \otimes D^{(3)} = D^{(1)} \oplus D^{(3)} \oplus D^{(5)}. \quad (1.24)$$

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

$$(j) \otimes (j') = \bigoplus_{j''=|j-j'|}^{j+j'} (j''). \quad (1.25)$$

En itérant, on trouve de même

$$D^{(3)} \otimes D^{(3)} \otimes D^{(3)} = D^{(1)} \oplus 3D^{(3)} \oplus 2D^{(5)} \oplus D^{(7)}, \quad (1.26)$$

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 (1.22), (1.24) ou (1.26), 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 $SO(3)$.

⁵ (un tel tenseur est “dual” d'un vecteur : $A_{ij} = \epsilon_{ijk} z_k$)

◦ *Coefficients de Clebsch-Gordan*

La formule (1.21) 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

$$e_\alpha^{(\rho)} \otimes e_\beta^{(\sigma)} = \sum_{\tau, \gamma, i} C_{\rho, \alpha; \sigma, \beta | \tau_i, \gamma} e_\gamma^{(\tau_i)}. \quad (1.27)$$

ou encore avec des notations plus familières en Mécanique Quantique

$$|\rho, \alpha; \sigma, \beta\rangle \equiv |\rho\alpha\rangle|\sigma\beta\rangle = \sum_{\tau, \gamma, i} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle |\tau_i \gamma\rangle. \quad (1.28)$$

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 (1.22), 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

$$\begin{aligned} \sum_{\tau, \gamma, i} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle \langle \tau_i \gamma | \rho, \alpha'; \sigma, \beta' \rangle^* &= \delta_{\alpha, \alpha'} \delta_{\beta, \beta'} \\ \sum_{\alpha, \beta} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle \langle \tau'_j \gamma' | \rho, \alpha; \sigma, \beta \rangle^* &= \delta_{\tau, \tau'} \delta_{\gamma, \gamma'} \delta_{i, j}. \end{aligned} \quad (1.29)$$

Ceci permet d'inverser la relation (1.28) en

$$|\tau_i \gamma\rangle = \sum_{\alpha, \beta} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle^* |\rho, \alpha; \sigma, \beta\rangle \quad (1.30)$$

et justifie la notation

$$\langle \rho, \alpha; \sigma, \beta | \tau_i \gamma \rangle = \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle^* \quad (1.31)$$

$$|\tau_i \gamma\rangle = \sum_{\alpha, \beta} \langle \rho, \alpha; \sigma, \beta | \tau_i \gamma \rangle |\rho, \alpha; \sigma, \beta\rangle. \quad (1.32)$$

Finalement en appliquant une opération du groupe aux deux membres de (1.28) et en utilisant ces relations, on décompose le produit des matrices $\mathcal{D}^{(\rho)}$ et $\mathcal{D}^{(\sigma)}$ de façon tout à fait explicite

$$\mathcal{D}_{\alpha\alpha'}^{(\rho)} \mathcal{D}_{\beta\beta'}^{(\sigma)} = \sum_{\tau, \gamma, \gamma', i} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle^* \langle \tau_i \gamma' | \rho, \alpha'; \sigma, \beta' \rangle \mathcal{D}_{\gamma\gamma'}^{(\tau_i)}. \quad (1.33)$$

On verra plus bas (§4.3) une application de ces formules au théorème de Wigner-Eckart.

1.6. Décomposition en représentations irréductibles d'un sous-groupe d'une représentation d'un 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

$$\forall h \in H \quad D'(h) = D(h) .$$

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, D' ne l'est pas, 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. Représentations des groupes et représentations des algèbres de Lie

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 chapitre 00, § 2.2, une illustration de cette propriété d'universalité.

En parallèle avec les définitions du § 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. 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 $g = e^{tX}$)

$$d(X) := \left. \frac{d}{dt} \right|_{t=0} D(g(t)) , \quad (2.1)$$

ou encore, pour t infinitésimal,

$$D(e^{tX}) = e^{td(X)} . \quad (2.2)$$

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

$$\begin{aligned} e^{d(Z)} &= D(e^Z) = D(e^{tX} e^{uY} e^{-tX} e^{-uY}) = D(e^{tX}) D(e^{uY}) D(e^{-tX}) D(e^{-uY}) \\ &= e^{td(X)} e^{ud(Y)} e^{-td(X)} e^{-ud(Y)} \\ &= e^{ut[d(X), d(Y)] + \dots} , \end{aligned} \quad (2.3)$$

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

$$X \in \mathfrak{g} \quad D^{\text{adj}}(g)(X) = gXg^{-1} , \quad (2.4)$$

ce qu'on note $\text{Ad } g X$. (Il faut comprendre le membre de droite de (2.4) 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, 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 (2.4), 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

$$\left. \frac{d}{dt} \text{Ad } g(t)X \right|_{t=0} = [Y, X] = \text{ad } Y X . \quad (2.5)$$

où on retrouve (et justifie) notre notation ad du chap. 1.

Exercice : montrer que les matrices T_i définies par $(T_i)_j^k = -C_{ij}^k$ 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 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. 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' . Dans notre cas, 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 chapitre 00 sur les deux cas concrets de $\text{SU}(2)$ et de $\text{SL}(2, \mathbb{C})$.

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 $U(1) = \{e^{ix}\}$ 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.

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 1.2. (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*

$$\int \frac{d\mu(g)}{v(G)} \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\alpha'\beta'}^{(\rho')*}(g) = \frac{1}{n_\rho} \delta_{\rho\rho'} \delta_{\alpha\alpha'} \delta_{\beta\beta'} \quad (3.1)$$

et leurs caractères satisfont donc

$$\int \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g) \chi^{(\rho')*}(g) = \delta_{\rho\rho'} . \quad (3.2)$$

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

$$V = \int d\mu(g') \mathcal{D}^{(\rho)}(g') M \mathcal{D}^{(\rho')\dagger}(g') . \quad (3.3)$$

Le membre de gauche de (3.1) 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

$$V \mathcal{D}^{(\rho')}(g) = \mathcal{D}^{(\rho)}(g) V \quad (3.4)$$

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 (3.1).

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 (3.1), cqfd.

La proposition (3.2) découle simplement de la précédente en prenant la trace sur $\alpha = \beta$ et $\alpha' = \beta'$.

Soulignons deux conséquences importantes de cette discussion :

1. on vient de voir que toute représentation irréductible (et unitaire) d'un groupe compact est de dimension finie ;
2. la relation (3.2) implique que deux représentations $D^{(\rho)}$ et $D^{(\sigma)}$ sont équivalentes (en fait identiques, compte tenu de notre convention d'indexation) sissi leurs caractères sont égaux : $\chi^{(\rho)} = \chi^{(\sigma)}$.

o *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 §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 (3.1) (ou (3.2), 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.

o *Complétude.*

Revenons au cas d'un groupe compact. On démontre que les matrices $\mathcal{D}_{\alpha\beta}^{(\rho)}(g)$ satisfont aussi une propriété de complétude

$$\sum_{\rho, \alpha, \beta} n_{\rho} \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\alpha\beta}^{(\rho)*}(g') = v(G) \delta(g, g') .$$

(3.5)

ou encore si on préfère

$$\sum_{\rho, \alpha, \beta} n_{\rho} \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\beta\alpha}^{(\rho)\dagger}(g') = \sum_{\rho} n_{\rho} \chi^{(\rho)}(g.g'^{-1}) = v(G) \delta(g, g') , \quad (3.5)'$$

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 $\mathcal{D}_{\alpha\beta}^{(\rho)}(g)$

$$f(g) = \int d\mu(g') \delta(g, g') f(g') = \sum_{\rho, \alpha, \beta} n_\rho \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \int \frac{d\mu(g')}{v(G)} \mathcal{D}_{\beta\alpha}^{(\rho)\dagger}(g') f(g') =: \sum_{\rho, \alpha, \beta} n_\rho \mathcal{D}_{\alpha\beta}^{(\rho)}(g) f_{\alpha\beta}^{(\rho)}. \quad (3.6)$$

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 (3.6) :

$$\begin{aligned} f_{\alpha\beta}^{(\rho)} &= \int \frac{d\mu(g')}{v(G)} f(g') \mathcal{D}_{\beta\alpha}^{(\rho)\dagger}(g') = \int \frac{d\mu(g')}{v(G)} f(hg'h^{-1}) \mathcal{D}_{\beta\alpha}^{(\rho)\dagger}(hg'h^{-1}) \quad \forall h \\ &= \int \frac{d\mu(h)}{v(G)} \frac{d\mu(g')}{v(G)} f(g') \mathcal{D}_{\beta\gamma}^{(\rho)\dagger}(h) \mathcal{D}_{\gamma\delta}^{(\rho)\dagger}(g') \mathcal{D}_{\delta\alpha}^{(\rho)\dagger}(h^{-1}) \quad (3.7) \\ &= \int \frac{d\mu(g')}{v(G)} f(g') \mathcal{D}_{\gamma\delta}^{(\rho)\dagger}(g') \frac{1}{n_\rho} \delta_{\alpha\beta} \delta_{\gamma\delta} \quad \text{par (3.1)} \\ &= \frac{1}{n_\rho} \int \frac{d\mu(g')}{v(G)} f(g') \chi^{(\rho)*}(g') \delta_{\alpha\beta} \end{aligned}$$

d'où il découle que (3.6) se réduit bien à un développement sur les caractères.

Testons à nouveau ces relations de complétude sur le cas $U(1)$. Elles expriment dans ce cas

$$\sum_{k=-\infty}^{\infty} e^{ikx} e^{-ikx'} = 2\pi \delta_P(x - x') \quad (3.8)$$

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 (3.6) signifie que toute fonction périodique de période 2π (et avec des conditions adéquates de régularité) peut être représentée par sa *série* de Fourier

$$f(x) = \sum_{k=-\infty}^{\infty} e^{ikx} f_k \quad f_k = \int_{-\pi}^{\pi} \frac{dx}{2\pi} f(x) e^{-ikx}. \quad (3.9)$$

Pour le groupe non compact \mathbb{R} , la relation de complétude (qui est encore vraie dans ce cas) équivaut à la *transformation* de Fourier

$$f(x) = \int_{-\infty}^{\infty} dk \tilde{f}(k) e^{ikx} \quad \tilde{f}(k) = \int_{-\infty}^{\infty} \frac{dx}{2\pi} f(x) e^{-ikx}. \quad (3.10)$$

Les théorèmes de décomposition pour un groupe quelconque sont donc des généralisations 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

$$\begin{aligned} \mathcal{D}^{(k)}(\alpha) &= \begin{pmatrix} \cos k\alpha & -\sin k\alpha \\ \sin k\alpha & \cos k\alpha \end{pmatrix} & k \in \mathbb{N}^* \\ \chi^{(k)}(\alpha) &= 2 \cos k\alpha \end{aligned} \quad (3.11)$$

Que deviennent les relations d'orthogonalité et de complétude?

3.2. Conséquences

Pour un groupe compact,

(i) toute représentation étant complètement réductible, son caractère s'écrit

$$\chi = \sum_{\rho} m_{\rho} \chi^{(\rho)} \quad (3.12)$$

et les multiplicités peuvent se calculer par la formule

$$m_{\rho} = \int \frac{d\mu(g)}{v(G)} \chi(g) \chi^{(\rho)*}(g) . \quad (3.13)$$

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 sissi son caractère satisfait la condition $\int \frac{d\mu(g)}{v(G)} |\chi(g)|^2 = 1$. Et 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 §1.5 et 1.6.

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

$$D^{(\rho)} \otimes D^{(\sigma)} = \oplus_{\tau} m_{\tau} D^{(\tau)} \quad (3.14a)$$

$$\chi^{(\rho)} \chi^{(\sigma)} = \sum_{\tau} m_{\tau} \chi^{(\tau)} \quad (3.14b)$$

$$m_{\tau} = \int \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g) \chi^{(\sigma)}(g) \chi^{(\tau)*}(g) , \quad (3.14c)$$

donc la représentation τ apparaît dans le produit $\rho \otimes \sigma$ avec la même multiplicité que σ^* dans $\rho \otimes \tau^*$.

◦ *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 dans l'Appendice B.

3.3. Cas des groupes finis

Nous n'évoquerons que très brièvement le cas des groupes finis. En fait les théorèmes (3.1), (3.2) et (3.5) et leurs conséquences (3.12), (3.13), (3.14), 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 $\sum_{g \in G}$:

$$\begin{aligned} \frac{1}{|G|} \sum_{g \in G} \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\alpha'\beta'}^{(\rho')*}(g) &= \frac{1}{n_\rho} \delta_{\rho\rho'} \delta_{\alpha\alpha'} \delta_{\beta\beta'} \\ \sum_{\rho, \alpha, \beta} \frac{n_\rho}{|G|} \mathcal{D}_{\alpha\beta}^{(\rho)}(g) \mathcal{D}_{\alpha\beta}^{(\rho)*}(g') &= \delta_{g, g'} . \end{aligned} \quad (3.15)$$

Mais les représentations des groupes finis jouissent de propriétés supplémentaires. Montrons ainsi que les dimensions des représentations irréductibles non équivalentes vérifient

$$\sum_{\rho} n_{\rho}^2 = |G| . \quad (3.16)$$

Cela découle du fait que le système d'équations (3.15) 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 $\mathcal{U}\mathcal{U}^{\dagger} = \mathbb{I}$, $\mathcal{U}^{\dagger}\mathcal{U} = \mathbb{I}$, ce qui n'est possible que si c'est une matrice carrée, cqfd.

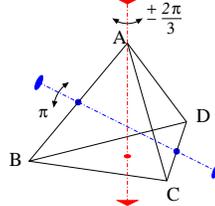
Cela implique en particulier que le nombre r de représentations irréductibles inéquivalentes est fini, et nous allons montrer que *ce nombre r est égal au nombre m des classes \mathcal{C}_i dans le groupe*. En notant $\chi_j^{(\rho)}$ la valeur du caractère $\chi^{(\rho)}$ dans la classe \mathcal{C}_i , on peut récrire les relations d'orthogonalité et de complétude des caractères selon

$$\begin{aligned} \frac{1}{|G|} \sum_{i=1}^m |\mathcal{C}_i| \chi_i^{(\rho)} \chi_i^{(\rho')*} &= \delta_{\rho\rho'} \\ \frac{|\mathcal{C}_i|}{|G|} \sum_{\rho=1}^r \chi_i^{(\rho)} \chi_j^{(\rho)*} &= \delta_{ij} . \end{aligned} \quad (3.17)$$

(Exercice : déduire la deuxième relation de (3.6) et (3.7), appliquées à un groupe fini.)
 Mais à nouveau, ces relations expriment que la matrice $\mathcal{K}_{\rho i} := \left(\frac{|C_i|}{|G|}\right)^{\frac{1}{2}} \chi_i^{(\rho)}$ de dimensions $r \times m$ satisfait $\mathcal{K}\mathcal{K}^\dagger = \mathbb{I}$, $\mathcal{K}^\dagger\mathcal{K} = \mathbb{I}$, 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) $\chi_i^{(\rho)}$, $\rho = 1, \dots, r$, $i = 1, \dots, m = r$. Ses lignes et colonnes satisfont les propriétés d'orthogonalité (3.17).

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 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 figure.



Il a donc 4 représentations irréductibles, dont on vérifie aisément à l'aide de (3.16) 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 $SO(3)$ fournit une représentation de dimension 3 de T dont le caractère χ prend les valeurs $\chi_i = 1 + 2 \cos \theta_i = (1, -1, 0, 0)$ dans les quatre classes ; selon le critère du § 3.2, $\|\chi\|^2 = \sum_i \frac{|C_i|}{|G|} |\chi_i|^2 = 1$ donc ce caractère est irréductible. Cela nous fournit une deuxième ligne (celle notée D_4). La représentation de spin 2 de $SO(3)$ 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 $j + j^2 = -1$.

↓ irreps \ classes →	$C(0)$	$C(\pi)$	$C(\frac{2\pi}{3})$	$C(-\frac{2\pi}{3})$
$ C_i $	1	3	4	4
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

Vérifier que les relations (3.17) sont bien satisfaites. Expliquer aussi pourquoi le groupe T n'est autre que le groupe A_4 des permutations paires de 4 objets.

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 $SL(2, \mathbb{C})$). L'ensemble des représentations irréductibles est indexé par des paramètres discrets et *continus*.

4. Représentations projectives. Théorème de Wigner. Cocycles et termes centraux. Théorème de Bargmann

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

$$U(g_1)U(g_2) = e^{i\zeta(g_1, g_2)}U(g_1g_2) . \quad (4.1)$$

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

$$\zeta(g_1, g_2) \rightarrow \zeta'(g_1, g_2) = \zeta(g_1, g_2) + \alpha(g_1) + \alpha(g_2) - \alpha(g_1g_2) . \quad (4.2)$$

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 :

$$\forall g_1, g_2, g_3 \quad (\partial\zeta)(g_1, g_2, g_3) := \zeta(g_1, g_2) + \zeta(g_1g_2, g_3) - \zeta(g_2, g_3) - \zeta(g_1, g_2g_3) = 0 \quad (4.3)$$

(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 :

$$(\partial\varphi)(g_1, \dots, g_{n+1}) = \sum_{i=1}^n \varphi(g_1, g_2, \dots, (g_i g_{i+1}), \dots, g_{n+1}) - \varphi(g_2, \dots, g_{n+1}) + (-1)^n \varphi(g_1, \dots, g_n) .$$

Pour une 1-cochaîne $\alpha(g)$, $\partial\alpha(g_1, g_2) = \alpha(g_1g_2) - \alpha(g_1) - \alpha(g_2)$, et donc (4.2) 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 (4.2), $\zeta' = 0$.

Autrement dit, le 2-cocycle ζ , qui est fermé ($\partial\zeta = 0$) par (4.3), 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 § 4.4) 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 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.

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} , et dans lequel les 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 observables $|\langle \phi | A | \psi \rangle|^2$, c'est-à-dire

$$|\psi\rangle \rightarrow |{}^g\psi\rangle, \quad A \rightarrow {}^gA \quad \text{t.q.} \quad |\langle \phi | A | \psi \rangle| = |\langle {}^g\phi | {}^gA | {}^g\psi \rangle|. \quad (4.4)$$

On démontre alors le théorème suivant

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*

$$|\langle \phi | A | \psi \rangle| = |\langle {}^g\phi | {}^gA | {}^g\psi \rangle|, \quad (4.5)$$

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

$$|{}^g\psi\rangle = U(g)|\phi\rangle, \quad {}^gA = U(g)AU^\dagger(g); \quad U(g)U^\dagger(g) = U(g)^\dagger U(g) = I. \quad (4.6)$$

Rappelons d'abord ce qu'on entend par opérateur antilinéaire. Un tel opérateur satisfait

$$U(\lambda|\phi\rangle + \mu|\psi\rangle) = \lambda^*U|\phi\rangle + \mu^*U|\psi\rangle \quad (4.7)$$

et son adjoint est défini par

$$\langle \phi | U^\dagger | \psi \rangle = \langle U\phi | \psi \rangle^* = \langle \psi | U\phi \rangle, \quad (4.8)$$

de façon à être compatible avec la linéarité :

$$\langle \lambda \phi | U^\dagger | \psi \rangle = \lambda^* \langle \phi | U^\dagger | \psi \rangle . \quad (4.9)$$

S'il est en outre unitaire,

$$\langle \psi | \phi \rangle^* = \langle \phi | \psi \rangle = \langle \phi | U^\dagger U | \psi \rangle = \langle U \phi | U \psi \rangle^* , \quad (4.10)$$

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 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

$$\begin{aligned} \langle {}^g\phi | {}^gA | {}^g\psi \rangle &= \langle U\phi | UAU^\dagger | U\psi \rangle \\ &= \langle \phi | U^\dagger UAU^\dagger U | \psi \rangle^\# \\ &= \langle \phi | A | \psi \rangle^\# \end{aligned} \quad (4.11)$$

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}

$$\begin{aligned} \mathbf{x}' &= U(T)\mathbf{x}U^\dagger(T) = \mathbf{x} \\ \mathbf{p}' &= U(T)\mathbf{p}U^\dagger(T) = -\mathbf{p} . \end{aligned} \quad (4.12)$$

Les relations de commutation canoniques ne sont compatibles avec cette transformation que si $U(T)$ est antilinéaire

$$\begin{aligned} [x'_j, p'_k] &= -[x_j, p_k] = -i\hbar\delta_{jk} \\ &= U(T)[x_j, p_k]U^\dagger(T) = U(T)i\hbar\delta_{jk}U^\dagger(T) \end{aligned}$$

Autre argument : $U(T)$ commute avec les translations dans le temps dont le générateur est l'hamiltonien : $U(T)iHU^\dagger(T) = -iH$ (puisque $t \rightarrow -t$). Si U était linéaire, on conclurait que $UHU^\dagger = -H$, 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} etc. 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.

4.3. Invariance d'un système quantique

Supposons maintenant que sous l'action d'un certain groupe de transformations G , le système est invariant, en ce sens que sa dynamique, contrôlée par son hamiltonien H , est inchangée. On va écrire

$$H = U(g)HU^\dagger(g)$$

ou encore

$$[H, U(g)] = 0 . \quad (4.13)$$

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

$$i\hbar \frac{\partial \mathcal{F}(U(g))}{\partial t} = [\mathcal{F}(U(g)), H] = 0 \quad (4.14)$$

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.

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

$$[T_i, T_j] = iC_{ij}^k T_k \quad (4.15)$$

(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 de H égales à $\dim E^{(\rho)}$ ⁶. 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 $\oplus_i E^{(\rho,i)}$, 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 ...

4.4. Transformations des observables. Théorème de Wigner-Eckart

Selon (4.6), 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 :

$$A_\alpha \rightarrow U(g)A_\alpha U(g)^\dagger = \sum_{\alpha'} A_{\alpha'} \mathcal{D}_{\alpha'\alpha}(g) . \quad (4.16)$$

⁶ 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

$$U(g)A_\alpha|\sigma\beta\rangle = U(g)A_\alpha U(g)^\dagger U(g)|\sigma\beta\rangle = \mathcal{D}_{\alpha'\alpha}^{(\rho)}(g)\mathcal{D}_{\beta'\beta}^{(\sigma)}(g)A_{\alpha'}|\sigma\beta'\rangle \quad (4.17)$$

c'est-à-dire selon le produit tensoriel des représentations $D^{(\rho)}$ et $D^{(\sigma)}$. Comme on l'a fait en (1.33), on peut développer sur des représentations irréductibles

$$\mathcal{D}_{\alpha'\alpha}^{(\rho)}(g)\mathcal{D}_{\beta'\beta}^{(\sigma)}(g) = \sum_{\tau, \gamma, \gamma', i} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle \langle \tau_i \gamma' | \rho, \alpha'; \sigma, \beta' \rangle^* \mathcal{D}_{\gamma' \gamma}^{(\tau)}(g). \quad (4.18)$$

Supposons maintenant que le groupe G est compact (ou fini). Les matrices des représentations satisfont les propositions d'orthogonalité (3.1). On peut alors écrire

$$\begin{aligned} \langle \tau \gamma | A_\alpha | \sigma \beta \rangle &= \langle \tau \gamma | U(g)^\dagger U(g) A_\alpha | \sigma \beta \rangle && \forall g \in G \\ &= \int \frac{d\mu(g)}{v(G)} \langle \tau \gamma | U(g)^\dagger U(g) A_\alpha | \sigma \beta \rangle \\ &= \int \frac{d\mu(g)}{v(G)} \sum_{\alpha', \beta', \gamma'} \mathcal{D}_{\gamma' \gamma}^{(\tau)*}(g) \langle \tau \gamma' | A_{\alpha'} | \sigma \beta' \rangle \mathcal{D}_{\alpha' \alpha}^{(\rho)}(g) \mathcal{D}_{\beta' \beta}^{(\sigma)}(g) \\ &= \frac{1}{n_\tau} \sum_{\alpha', \beta', \gamma', i} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle \langle \tau_i \gamma' | \rho, \alpha'; \sigma, \beta' \rangle^* \langle \tau_i \gamma' | A_{\alpha'} | \sigma \beta' \rangle. \end{aligned} \quad (4.19)$$

Notons

$$\langle \tau \parallel A \parallel \sigma \rangle_i = \frac{1}{n_\tau} \sum_{\alpha', \beta', \gamma'} \langle \tau_i \gamma' | \rho, \alpha'; \sigma, \beta' \rangle^* \langle \tau_i \gamma' | A_{\alpha'} | \sigma \beta' \rangle. \quad (4.20)$$

Il en découle que (théorème de Wigner–Eckart) :

$$\langle \tau \gamma | A_\alpha | \sigma \beta \rangle = \sum_{i=1}^{m_\tau} \langle \tau \parallel A \parallel \sigma \rangle_i \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle \quad (4.21)$$

dans laquelle les éléments de matrice “réduits” $\langle \cdot \parallel A \parallel \cdot \rangle_i$ sont indépendants de α, β, γ . L'élément de matrice du membre de gauche dans (4.21) 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 B.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 § 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 (4.21)

$$\langle \rho\alpha i | H | \rho\alpha' i' \rangle = \delta_{\alpha\alpha'} \langle \rho i | H | \rho i' \rangle$$

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

$$\langle j, m_1 | K_1^m | j, m_2 \rangle = \langle j, m_1 | J^m | j, m_2 \rangle \frac{\langle \vec{J} \cdot \vec{K} \rangle}{j(j+1)}$$

où $\langle \vec{J} \cdot \vec{K} \rangle$ 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 $\vec{J} \frac{\langle \vec{J} \cdot \vec{K} \rangle}{j(j+1)}$.

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 $\alpha\beta$

$$i\zeta(I + \alpha t_a, I + \beta t_b) = \alpha\beta z_{ab} . \tag{4.22}$$

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)

$$[T_a, T_b] = C_{ab}{}^c T_c + C_{ab} \quad [C_{ab}, T_c] = 0 . \tag{4.23}$$

(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

$$T_a \rightarrow \tilde{T}_a = T_a + X_a \quad [\tilde{T}_a, \tilde{T}_b] = C_{ab}{}^c \tilde{T}_c, \quad (4.24)$$

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 (4.24) est alors possible.

On démontre (Bargmann) que pour un groupe de Lie connexe G , les cocycles sont triviaux si

- a) il n'existe pas d'extension centrale non triviale de \mathfrak{g} ;
- b) G est simplement connexe.

En ce qui concerne le point a), un théorème de Bargmann nous assure qu'il n'existe pas d'extension centrale non triviale pour tout groupe semi-simple, comme les groupes classiques $SU(n)$, $SO(n)$, $Sp(2n)$. C'est donc plutôt le point b) 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

Outre les références déjà données au chap. 1,

Théorie générale des représentations

A.A. Kirillov, *Elements of the theory of representations*, Springer.

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 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 V. Bargmann, *Ann. Math.* **59** (1954) 1-46, ou S. Weinberg [Wf] Chap 2.7.

★

Appendix A. 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...

◦ 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

$$z = \sum_{\alpha} x^{(\alpha)} \otimes y^{(\alpha)} \quad (\text{A.1})$$

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

$$\begin{aligned} A \otimes B(x \otimes y) &= Ax \otimes By \\ A \otimes B \sum_{\alpha} (x^{(\alpha)} \otimes y^{(\alpha)}) &= \sum_{\alpha} Ax^{(\alpha)} \otimes By^{(\alpha)} \end{aligned} \quad (\text{A.2})$$

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

$$A \otimes Bz = \sum_{i,i',j,j'} A_{ii'} B_{jj'} x^{i'} y^{j'} e_i f_j =: (A \otimes B)_{ii';jj'} z^{i'j'} e_i \otimes f_j \quad (\text{A.3})$$

soit

$$(A \otimes B)_{ij;i'j'} = A_{ii'} B_{jj'} \quad (\text{A.4})$$

formule qui est parfois prise comme définition...

◦ 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

$$D(g) \otimes D'(g)(x \otimes y) = D(g)x \otimes D'(g)y \quad (\text{A.5})$$

en accord avec (A.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 « 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

$$(x' \otimes y')^{ij} = x^i y^j = \mathcal{D}_{i'i'} \mathcal{D}'_{j'j} x^i y^j \quad (\text{A.6})$$

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 Chapitre 00, § 3.3, dans la construction des représentations de $\text{SU}(2)$ par produits tensoriels symétrisés de la représentation de spin $\frac{1}{2}$, ou au § 6.2 pour celles de $\text{SL}(2, \mathbb{C})$ par produits tensoriels symétrisés des deux représentations à indices pointés $(0, \frac{1}{2})$ ou non pointés $(\frac{1}{2}, 0)$.

Appendix B. Compléments sur les matrices de représentation de $\text{SU}(2)$

On revient ici sur les matrices \mathcal{D}^j des représentations de $\text{SU}(2)$ définies au § 3.2 du chapitre 00.

B.1. Orthogonalité, complétude, caractères

Nous faisons appel à la mesure invariante sur le groupe $\text{SU}(2)$ introduite au chapitre 1 (§ 2.4 et Appendice C) pour énoncer des propriétés d'orthogonalité et de complétude des matrices \mathcal{D}^j . Nous avons construit au chapitre 00 *toutes* les représentations unitaires de $\text{SU}(2)$. Selon la discussion du § 3, les éléments des matrices \mathcal{D}^j satisfont des propriétés d'orthogonalité et de complétude

$$(2j+1) \int \frac{d\mu(U)}{2\pi^2} \mathcal{D}_{mn}^j(U) \mathcal{D}_{m'n'}^{j'*}(U) = \delta_{jj'} \delta_{mm'} \delta_{nn'} \quad (\text{B.1a})$$

$$\sum_{jmn} (2j+1) \mathcal{D}_{mn}^j(U) \mathcal{D}_{mn}^{j'*}(U') = 2\pi^2 \delta(U, U') \quad (\text{B.1b})$$

La fonction $\delta(U, U')$ qui apparaît dans le second membre de (B.1b) 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,

$$\delta(U, U') = 8\delta(\alpha - \alpha') \delta(\cos \beta - \cos \beta') \delta(\gamma - \gamma'), \quad (\text{B.2})$$

(cf Appendice C du chapitre 1). La signification de l'équation (B.1b) 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

$$\begin{aligned}\chi_j(U) = \chi_j(\psi) &= \text{tr} \mathcal{D}^j(\mathbf{n}, \psi) = \sum_{m=-j}^j e^{im\psi} \\ &= \frac{\sin\left(\frac{2j+1}{2}\psi\right)}{\sin\frac{\psi}{2}}.\end{aligned}\tag{B.3}$$

Noter que ces expressions sont des polynômes (dits de Tchebichev de 2ème espèce) de la variable $2 \cos \frac{\psi}{2}$ (voir l'exercice D en fin de chapitre). En particulier

$$\chi_0(\psi) = 1 \quad \chi_{\frac{1}{2}}(\psi) = 2 \cos \frac{\psi}{2} \quad \chi_1(\psi) = 1 + 2 \cos \psi \quad \text{etc}.\tag{B.4}$$

On est alors en mesure de vérifier toutes les propriétés attendues

$$\begin{aligned}\text{unitarité et réalité} \quad & \chi_j(U^{-1}) = \chi_j^*(U) = \chi_j(U) \\ & \chi_j(-U) = \chi_j(2\pi + \psi) = (-1)^{2j} \chi_j(U) \\ \text{orthogonalité} \quad & \int_0^{2\pi} d\psi \sin^2 \frac{\psi}{2} \chi_j(\psi) \chi_{j'}(\psi) = \pi \delta_{jj'} \\ \text{complétude} \quad & \sum_{j=0, \frac{1}{2}, \dots} \chi_j(\psi) \chi_j(\psi') = \frac{\pi}{\sin^2 \frac{\psi}{2}} \delta(\psi - \psi') = \frac{\pi}{2 \sin \frac{\psi}{2}} \delta\left(\cos \frac{\psi}{2} - \cos \frac{\psi'}{2}\right)\end{aligned}\tag{B.5}$$

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.

La formule de multiplicité (3.14c) conduit-elle bien aux formules connues (1.25) ?

B.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 aussi noté que la discussion du chapitre 00, §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 (3.23) du chapitre 00 est donc toujours vraie dans ce cas

$$P_{jm}(\xi', \eta') = \sum_{m'} P_{jm'}(\xi, \eta) \mathcal{D}_{m'm}^j(A). \quad (00.3.23)$$

La combinaison $(a\xi + c\eta)^{j+m}(b\xi + d\eta)^{j-m}$ satisfait évidemment

$$\left(\frac{\partial^2}{\partial a \partial d} - \frac{\partial^2}{\partial b \partial c} \right) (a\xi + c\eta)^{j+m}(b\xi + d\eta)^{j-m} = 0 \quad (B.6)$$

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 $A = \rho U$, $U \in SU(2)$, et (B.6) conduit à

$$\Delta_4 \mathcal{D}_{m'm}^j(A) = 4 \left(\frac{\partial^2}{\partial a \partial a^*} + \frac{\partial^2}{\partial b \partial b^*} \right) \mathcal{D}_{m'm}^j(A) = 0 \quad (B.7)$$

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,

$$\Delta_4 = \frac{\partial^2}{\partial \rho^2} + \frac{3}{\rho} \frac{\partial}{\partial \rho} + \frac{1}{\rho^2} \Delta_{\text{sphère } S^3} \quad (B.8)$$

où le dernier terme, laplacien sur la sphère S^3 , n’agit que sur les “variables angulaires” $U \in SU(2)$. Les fonctions \mathcal{D}^j étant homogènes de degré $2j$ en a, b, c, d donc en ρ , on a finalement

$$-\frac{1}{4} \Delta_{\text{sphère } S^3} \mathcal{D}_{m'm}^j(U) = j(j+1) \mathcal{D}_{m'm}^j(U). \quad (B.9)$$

Par exemple, en utilisant la paramétrisation des angles d’Euler, on trouve que

$$\left\{ \frac{1}{\sin \beta} \frac{\partial}{\partial \beta} \sin \beta \frac{\partial}{\partial \beta} + \frac{1}{\sin^2 \beta} \left[\frac{\partial^2}{\partial \alpha^2} + \frac{\partial^2}{\partial \gamma^2} - 2 \cos \beta \frac{\partial^2}{\partial \alpha \partial \gamma} \right] + j(j+1) \right\} \mathcal{D}^j(\alpha, \beta, \gamma)_{m'm} = 0. \quad (B.10)$$

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

$$\left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + l(l+1) \right] \mathcal{D}_{m0}^l(\phi, \theta, 0) = 0. \quad (B.11)$$

L'opérateur différentiel constitué des deux premiers termes est le laplacien Δ_{S^2} sur la sphère unité S^2 . L'équation (B.11) définit donc les *harmoniques sphériques* $Y_l^m(\theta, \phi)$ comme vecteurs propres du laplacien Δ_{S^2} . La normalisation correcte est que

$$\left[\frac{2l+1}{4\pi} \right]^{\frac{1}{2}} \mathcal{D}_{m0}^l(\phi, \theta, 0) = Y_l^{m*}(\theta, \phi) . \quad (\text{B.12})$$

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

$$\begin{aligned} P_l(u) &= \frac{1}{2^l l!} \frac{d^l}{du^l} (u^2 - 1)^l \\ P_l^m(u) &= (1 - u^2)^{\frac{1}{2}m} \frac{d^m}{du^m} P_l(u) \quad \text{pour } 0 \leq m \leq l . \end{aligned} \quad (\text{B.13})$$

Les polynômes de Legendre $P_l(u)$ sont des polynômes orthogonaux sur l'intervalle $[-1, 1]$ avec le poids 1 : $\int_{-1}^1 P_l(u) P_{l'}(u) du = \frac{2}{2l+1} \delta_{ll'}$. Les premiers P_l sont

$$P_0 = 1 \quad P_1 = u \quad P_2 = \frac{1}{2}(3u^2 - 1) \quad P_3 = \frac{1}{2}(5u^3 - 3u) , \dots \quad (\text{B.14})$$

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

$$Y_l^m(\theta, \phi) = (-1)^m \left[\frac{(2l+1)}{4\pi} \frac{(l-m)}{(l+m)} \right]^{\frac{1}{2}} P_l^m(\cos \theta) e^{im\phi} \quad (\text{B.15})$$

donc

$$\mathcal{D}_{m0}^l(0, \theta, 0) = d_{m0}^l(\theta) = (-1)^m \left[\frac{(l-m)}{(l+m)} \right]^{\frac{1}{2}} P_l^m(\cos \theta) = \left(\frac{4\pi}{2l+1} \right)^{\frac{1}{2}} Y_l^{m*}(\theta, 0) . \quad (\text{B.16})$$

En particulier, $d_{00}^l(\theta) = P_l(\cos \theta)$. En général, $d_{m'm}^l(\theta)$ est relié au polynôme de Jacobi

$$P_l^{(\alpha, \beta)}(u) = \frac{(-1)^l}{2^l l!} (1-u)^{-\alpha} (1+u)^{-\beta} \frac{d^l}{du^l} [(1-u)^{\alpha+l} (1+u)^{\beta+l}] \quad (\text{B.17})$$

par

$$d_{m'm}^j(\theta) = \left[\frac{(j+m')!(j-m')!}{(j+m)!(j-m)!} \right]^{\frac{1}{2}} \left(\cos \frac{\theta}{2} \right)^{m+m'} \left(\sin \frac{\theta}{2} \right)^{m-m'} P_{j-m'}^{(m'-m, m'+m)}(\cos \theta) . \quad (\text{B.18})$$

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

$$\int_{-1}^1 du (1-u)^\alpha (1+u)^\beta P_j^{(\alpha,\beta)}(u) P_{j'}^{(\alpha,\beta)}(u) = \delta_{jj'} \frac{2^{\alpha+\beta+1} \Gamma(l+\alpha+1) \Gamma(l+\beta+1)}{(2l+\alpha+\beta+1)! \Gamma(l+\alpha+\beta+1)} \quad (\text{B.19})$$

et satisfont la relation de récurrence

$$\begin{aligned} & 2(l+1)(l+\alpha+\beta+1)(2l+\alpha+\beta)P_{l+1}^{(\alpha,\beta)}(u) \\ &= (2l+\alpha+\beta+1)[(2l+\alpha+\beta)(2l+\alpha+\beta+2)u + \alpha^2 - \beta^2]P_l^{(\alpha,\beta)}(u) \\ & - 2(l+\alpha)(l+\beta)(2l+\alpha+\beta+2)P_{l-1}^{(\alpha,\beta)}. \end{aligned} \quad (\text{B.20})$$

Le polynôme de Jacobi $P_l^{(\alpha,\beta)}(u)$ est solution de l'équation différentielle

$$\{(1-u^2)\frac{d^2}{du^2} + [\beta - \alpha - (2+\alpha+\beta)u] \frac{d}{du} + l(l+\alpha+\beta+1)\} P_l^{(\alpha,\beta)}(u) = 0. \quad (\text{B.21})$$

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

$$\begin{aligned} & (\Delta_{S^2} + l(l+1)) Y_l^m = 0 \\ & J_z Y_l^m = -i \frac{\partial}{\partial \phi} Y_l^m = m Y_l^m \end{aligned} \quad (\text{B.22})$$

et peuvent s'écrire

$$Y_l^m(\theta, \phi) = \frac{(-1)^l}{2^l l!} \sqrt{\frac{(2l+1)(l+m)!}{4\pi(l-m)!}} e^{im\phi} \sin^{-m} \theta \left(\frac{d}{d \cos \theta} \right)^{l-m} \sin^{2l} \theta. \quad (\text{B.23})$$

(ii) Elles sont normalisées à 1 sur la sphère unité et plus généralement y satisfont des propriétés d'orthogonalité et de complétude

$$\begin{aligned} & \int d\Omega Y_l^{m*} Y_{l'}^{m'} = \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta Y_l^{m*} Y_{l'}^{m'} = \delta_{ll'} \delta_{mm'} \\ & \sum_{l=0}^{\infty} \sum_{m=-l}^l Y_l^{m*}(\theta, \phi) Y_l^m(\theta', \phi') = \delta(\Omega - \Omega') = \frac{\delta(\theta - \theta') \delta(\phi - \phi')}{\sin \theta} \\ & = \delta(\cos \theta - \cos \theta') \delta(\phi - \phi') \end{aligned} \quad (\text{B.24})$$

(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

$$Y_l^m(\mathbf{n}') = Y_l^{m'}(\mathbf{n}) \mathcal{D}^l(R)_{m'm} \quad (\text{B.25})$$

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

$$Y_l^{m*}(\theta, \phi) = (-1)^m Y_l^{-m}(\theta, \phi) \quad (\text{B.26})$$

et de parité

$$Y_l^m(\pi - \theta, \phi + \pi) = (-1)^l Y_l^m(\theta, \phi). \quad (\text{B.27})$$

Noter qu'à $\theta = 0$, $Y_l^m(0, \phi)$ s'annule sauf pour $m = 0$, cf. (B.13), (B.15).

(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 (00.3.10)

$$e^{\pm i\phi} \left[\pm \frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right] Y_l^m = \sqrt{l(l+1) - m(m \pm 1)} Y_l^{m \pm 1} \quad (\text{B.28})$$

et celles provenant de la multiplication des représentations,

$$\sqrt{2l+1} \cos \theta Y_l^m = \left(\frac{(l+m)(l-m)}{2l-1} \right)^{\frac{1}{2}} Y_{l-1}^m + \left(\frac{(l+m+1)(l-m+1)}{2l+3} \right)^{\frac{1}{2}} Y_{l+1}^m. \quad (\text{B.29})$$

On a plus généralement la formule de produit

$$Y_l^m(\theta, \phi) Y_{l'}^{m'}(\theta, \phi) = \sum_L \langle lm; l'm' | L, m+m' \rangle \left[\frac{(2l+1)(2l'+1)}{4\pi(2L+1)} \right]^{\frac{1}{2}} Y_L^{m+m'}(\theta, \phi). \quad (\text{B.30})$$

(vi) Finalement citons la très utile "formule d'addition"

$$\frac{2l+1}{4\pi} P_l(\cos \theta) = \sum_{m=-l}^l Y_l^m(\mathbf{n}) Y_l^{m*}(\mathbf{n}') \quad (\text{B.31})$$

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 (B.31) (On rappelle que $P_l(1) = 1$).

2. Montrer qu'une fonction génératrice des polynômes de Legendre est

$$\frac{1}{\sqrt{1-2ut+t^2}} = \sum_{l=0}^{\infty} t^l P_l(u)$$

On pourra vérifier que l'équation différentielle des P_l (cas particulier de (B.21) 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 $r' < r$),

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \sum_{l=0}^{\infty} \frac{r'^l}{r^{l+1}} P_l(\cos \theta) = \sum_{l,m} \frac{4\pi}{2l+1} \frac{r'^l}{r^{l+1}} Y_l^{m*}(\mathbf{n}) Y_l^m(\mathbf{n}'). \quad (\text{B.32})$$

Les expressions des premiers Y_l^m peuvent être utiles

$$\begin{aligned} Y_0^0 &= \frac{1}{\sqrt{4\pi}} \\ Y_1^0 &= \sqrt{\frac{3}{4\pi}} \cos \theta & Y_1^{\pm 1} &= \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi} \\ Y_2^0 &= \sqrt{\frac{5}{16\pi}} (3 \cos^3 \theta - 1) & Y_2^{\pm 1} &= \mp \sqrt{\frac{15}{8\pi}} \cos \theta \sin \theta e^{\pm i\phi} & Y_2^{\pm 2} &= \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi}. \end{aligned}$$

B.3. Applications physiques

o Moments multipolaires

On considère un potentiel créé par une distribution de charge statique $\rho(\mathbf{r})$

$$\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{d^3r' \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$$

et on le développe sur les harmoniques sphériques selon (B.32). Il vient

$$\phi(\mathbf{r}) = \frac{1}{\epsilon_0} \sum_{l,m} \frac{1}{2l+1} \frac{Y_l^{m*}(\mathbf{n})}{r^{l+1}} Q_{lm} \quad (\text{B.33})$$

où les Q_{lm} , définis par

$$Q_{lm} = \int d^3r' \rho(\mathbf{r}') r'^l Y_l^m(\mathbf{n}') \quad (\text{B.34})$$

sont les *moments multipolaires* de la distribution de charge ρ . Par exemple, si $\rho(\mathbf{r}) = \rho(r)$ 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 $\rho(\mathbf{r})$ quelconque, les trois composantes de Q_{1m} reconstruisent le moment dipolaire $\int d^3r' \rho(\mathbf{r}') \mathbf{r}'$. 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. (B.27), 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 $l \leq 2j$.

o *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 $\vec{L} = \hbar \vec{J} = \hbar \mathbf{r} \wedge \vec{\nabla}$

$$Y_l^m(\theta, \phi) = \langle \theta, \phi | l, m \rangle$$

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

$$\begin{aligned} \langle p, \theta, \phi | \mathcal{T} | p, 0, 0 \rangle &= \sum_{l' m m'} Y_l^m(\theta, \phi) \langle p, l, m | \mathcal{T} | p, l', m' \rangle Y_{l'}^{m'*}(0, 0) \\ &= \sum_{lm} Y_l^m(\theta, \phi) \langle p, l, m | \mathcal{T} | p, l, m \rangle Y_l^{m*}(0, 0) \\ &= \sum_l \frac{2l+1}{4\pi} \mathcal{T}_l(p) P_l(\cos \theta) \end{aligned}$$

selon à nouveau la formule d'addition et $\langle plm | \mathcal{T} | pl'm' \rangle = \delta_{l'l} \delta_{m'm} \mathcal{T}_l(p)$ exprimant l'invariance par rotation. C'est le très utile développement en ondes partielles de l'amplitude de diffusion.

★

Exercices pour le 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 .

- a) Par un calcul trigonométrique élémentaire, exprimer $U_{l-1} + U_{l+1}$ en fonction de U_l .
- b) En déduire que U_l est un polynôme en $z = 2 \cos \theta$ de degré l , qu'on notera $U_l(z)$.
- c) Quelle est l'interprétation groupiste du résultat obtenu en a) ?
- d) 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 de

$$\frac{2}{\pi} \int_{-1}^1 dz (1-z^2)^{\frac{1}{2}} U_l(z) U_{l'}(z) U_{l''}(z) ?$$

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 d)) et satisfont une relation de récurrence à trois termes (question a)), 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 1. 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 **anti-symétriques** réelles et satisfont donc $\text{tr} T^a T^b = -\delta_{ab}$. On va s'intéresser à la quantité

$$X_{\alpha\beta; \alpha'\beta'} := \sum_{a=1}^d T_{\alpha\alpha'}^{(\rho)a} T_{\beta\beta'}^{(\sigma)a}. \quad (\text{B.35})$$

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

$$\left(\mathcal{M}^{(\tau\gamma)} \right)_{\alpha\beta} = \langle \tau\gamma | \rho\alpha; \sigma\beta \rangle. \quad (\text{B.36})$$

- 1) Rappeler pourquoi ces coefficients satisfont des propriétés d'orthogonalité et de complétude qu'on écrira.

2) En déduire qu'on peut écrire

$$X_{\alpha\beta;\alpha'\beta'} = - \sum_{\tau\gamma} \left(\mathcal{M}^{(\tau\gamma)} \right)_{\alpha\beta} \left(T^{(\rho)a} \mathcal{M}^{(\tau\gamma)} T^{(\sigma)a} \right)_{\alpha'\beta'} . \quad (\text{B.37})$$

3) En faisant agir le générateur infinitésimal T^a sur les deux membres de la relation

$$|\rho\alpha; \sigma\beta\rangle = \sum_{\tau,\gamma} \left(\mathcal{M}^{(\tau\gamma)} \right)_{\alpha\beta} |\tau\gamma\rangle \quad (\text{B.38})$$

montrer qu'on obtient

$$\sum_{\gamma'} T_{\gamma\gamma'}^{(\tau)a} \left(\mathcal{M}^{(\tau\gamma')} \right)_{\alpha\beta} = \sum_{\alpha'} \left(\mathcal{M}^{(\tau\gamma)} \right)_{\alpha'\beta} (T^{(\rho)a})_{\alpha'\alpha} + \sum_{m'_2} \left(\mathcal{M}^{(\tau\gamma)} \right)_{\alpha\beta'} (T^{(\sigma)a})_{\beta'\beta} \quad (\text{B.39})$$

ou encore en termes de matrices de dimensions $\dim(\rho) \times \dim(\sigma)$

$$\sum_{\gamma'} T_{\gamma\gamma'}^{(\tau)a} \mathcal{M}^{(\tau\gamma')} = -T^{(\rho)a} \mathcal{M}^{(\tau\gamma)} + \mathcal{M}^{(\tau\gamma)} T^{(\sigma)a} . \quad (\text{B.40})$$

4) En utilisant de façon répétée cette relation (B.40) dans (B.37), montrer qu'on a

$$X_{\alpha\beta;\alpha'\beta'} = \frac{1}{2} \sum_{\tau\gamma} (C_\rho + C_\sigma - C_\tau) \left(\mathcal{M}^{(\tau\gamma)} \right)_{\alpha\beta} \left(\mathcal{M}^{(\tau\gamma)} \right)_{\alpha'\beta'} \quad (\text{B.41})$$

où les C sont les opérateurs de Casimir, par exemple

$$C_\rho = - \sum_a (T^{(\rho)a})^2 . \quad (\text{B.42})$$

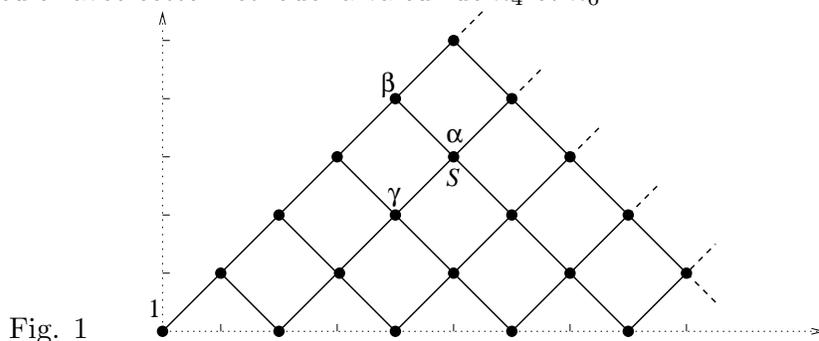
5) Pourquoi peut-on dire que les “grandes représentations” τ tendent à rendre le coefficient $(C_\rho + C_\sigma - C_\tau)$ de plus en plus négatif ? On pourra prendre l'exemple de $\text{SU}(2)$ avec ρ et σ deux représentations de spin (entier) égal à j .

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 ?

Problème 2. Produit tensoriel dans $\text{SU}(2)$

1. On considère la représentation $R_{\frac{1}{2}}$ de spin $\frac{1}{2}$ de $\text{SU}(2)$ et on veut calculer le nombre n_r de fois où la représentation identité apparaît dans la décomposition en représentations irréductibles du produit tensoriel de r copies de $R_{\frac{1}{2}}$.

- a) Interpréter n_r en termes du nombre d'invariants linéairement indépendants multilinéaires en ξ_1, \dots, ξ_r où les ξ_i sont des spineurs se transformant selon la représentation $R_{\frac{1}{2}}$.
- b) Par convention $n_0 = 1$. Sans aucun calcul, que valent n_1 et n_2 ?
- c) Montrer que l'on peut exprimer simplement n_r à l'aide d'une intégrale impliquant les caractères $\chi_j(\psi)$ de $SU(2)$, cf les formules (A.3-A.5) du chapitre 2 du cours. (On ne cherchera pas à calculer explicitement cette intégrale pour r arbitraire.)
- d) Vérifier que cette formule donne bien les valeurs de n_1 et n_2 obtenues au b).
- e) On va montrer que l'on peut aussi obtenir les n_r par la méthode graphique et récursive suivante. Sur le graphe de la Fig. 1, on attache $n_0 = 1$ au sommet le plus à gauche, puis à chaque sommet S , la somme $\alpha = \beta + \gamma$ des nombres situés aux sommets situés à sa gauche et directement reliés à S .
 - Montrer que les n_r sont les nombres figurant sur l'axe horizontal. Quelle est l'interprétation des axes horizontal et vertical ?
 - Calculer avec cette méthode la valeur de n_4 et n_6 .



2. On cherche à répéter ce calcul pour la représentation R_1 de spin 1, et donc à déterminer le nombre N_r de fois où la représentation identité apparaît dans le produit tensoriel de r copies de R_1 .

- a) Comment le graphe de la figure 1 doit-il être modifié pour obtenir les N_r ?
- b) Calculer ainsi N_2 , N_3 et N_4 .
- c) Que représentent ces nombres en termes de vecteurs V_1, \dots, V_r se transformant selon la représentation R_1 ?

Chapitre 3

Algèbres de Lie simples, classification et représentations.

Racines et poids

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 $SU(2)$

$$[J_z, J_{\pm}] = \pm J_{\pm} \quad [J_+, J_-] = 2J_z . \quad (1.1)$$

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.

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.

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 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 engendrée ¹ par les matrices diagonales de trace nulle. Une base en est donnée par les $n - 1$ matrices

$$H_1 = \text{diag}(1, -1, 0, \dots, 0), H_2 = \text{diag}(0, 1, -1, 0, \dots, 0), \dots, H_{n-1} = \text{diag}(0, \dots, 0, 1, -1). \quad (1.2)$$

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 [Mehta, *Elements of Matrix Theory*, p 41].

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 H_i hermitiens. Par définition $[H_i, H_j] = 0$, (sous-algèbre abélienne) ou plus précisément puisqu'on est dans la représentation adjointe,

$$[\text{ad } H_i, \text{ad } H_j] = 0. \quad (1.3)$$

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

$$\text{ad } H_i E_\alpha = \alpha_{(i)} E_\alpha \quad (1.4)$$

¹ Nous nous plaçons momentanément dans la représentation "de définition" (matrices $n \times n$) et non plus dans la représentation adjointe.

c'est-à-dire un ensemble d'éléments de \mathfrak{g} tels que

$$[H_i, E_\alpha] = \alpha_{(i)} E_\alpha, \quad (1.5)$$

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 H_i hermitiens, leurs valeurs propres $\alpha_{(i)}$ sont réelles. Par combinaison linéaire, pour un élément arbitraire de \mathfrak{h} écrit $H = \sum_i h^i H_i$,

$$\text{ad}(H)E_\alpha = \alpha(H)E_\alpha, \quad (1.6)$$

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} . On peut donc considérer la *racine* α de composantes $\alpha_{(i)}$ comme un vecteur de l'espace dual de \mathfrak{h} , ce qu'on note $\alpha \in \mathfrak{h}^*$. 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).

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. 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

$$(H_i, E_\alpha) = 0 \quad (E_\alpha, E_\beta) = 0 \quad \text{sauf si } \alpha + \beta = 0. \quad (1.7)$$

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

$$(H, [H', E_\alpha]) = \text{tr}(\text{ad } H[\text{ad } H', \text{ad } E_\alpha]) = \text{tr}([\text{ad } H, \text{ad } H'] \text{ad } E_\alpha) = 0 \quad (1.8)$$

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

$$([H, E_\alpha], E_\beta) = \alpha(H)(E_\alpha, E_\beta) = -(E_\alpha, [H, E_\beta]) = -\beta(H)(E_\alpha, E_\beta) \quad (1.9)$$

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 (1.7) : 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

$$\forall H \in \mathfrak{h} \quad \alpha(H) = (H_\alpha, H) . \tag{1.10}$$

(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

$$\langle \alpha, \beta \rangle = (H_\alpha, H_\beta) , \tag{1.11}$$

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

$$\text{ad } H_i [E_\alpha, E_\beta] = [H_i, [E_\alpha, E_\beta]] = [E_\alpha, [H_i, E_\beta]] - [E_\beta, [H_i, E_\alpha]] = (\alpha + \beta)_{(i)} [E_\alpha, E_\beta] . \tag{1.12}$$

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 (1.9)

$$\begin{aligned} (H_i, [E_\alpha, E_{-\alpha}]) &= \text{tr}(\text{ad } H_i [\text{ad } E_\alpha, \text{ad } E_{-\alpha}]) = \text{tr}([\text{ad } H_i, \text{ad } E_\alpha] \text{ad } E_{-\alpha}) \\ &= \alpha_{(i)}(E_\alpha, E_{-\alpha}) = (H_i, H_\alpha)(E_\alpha, E_{-\alpha}) \end{aligned} \tag{1.13}$$

donc

$$[E_\alpha, E_{-\alpha}] = (E_\alpha, E_{-\alpha})H_\alpha . \tag{1.14}$$

Pour récapituler, nous avons construit une base canonique de l'algèbre \mathfrak{g}

$$\begin{aligned} [H_i, H_j] &= 0 \\ [H_i, E_\alpha] &= \alpha_{(i)} E_\alpha \\ [E_\alpha, E_\beta] &= \begin{cases} N_{\alpha\beta} E_{\alpha+\beta} & \text{si } \alpha + \beta \text{ est une racine} \\ (E_\alpha, E_{-\alpha}) H_\alpha & \text{si } \alpha + \beta = 0 \\ 0 & \text{sinon} \end{cases} \end{aligned}$$

(1.15)

va montrer que la géométrie –longueurs et angles– des racines est fortement contrainte. Il suffit de se rappeler les leçons de l’algèbre $\mathfrak{su}(2)$: dans une représentation de dimension finie, J_z a des valeurs propres entières ou demi-entières. Donc ici, où chaque $\frac{H_\alpha}{\langle \alpha, \alpha \rangle}$ joue le rôle d’un J_z et a les E_β pour vecteurs propres, $\text{ad } H_\alpha E_\beta = \langle \alpha, \beta \rangle E_\beta$, c’est-à-dire

$$[H_\alpha, E_\beta] = \langle \alpha, \beta \rangle E_\beta \quad (2.2)$$

nous concluons que

$$2 \frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} = m \in \mathbb{Z} . \quad (2.3)$$

o *Chaînes de racines*

Il est en fait utile de reprendre la discussion précédente et de l’affiner. L’idée est comme dans $\mathfrak{su}(2)$ d’appliquer de façon répétée les opérateurs “montant” E_α et “descendant” $E_{-\alpha}$ (*ladder 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 $\mathfrak{su}(2)$ engendrée par H_α et $E_{\pm\alpha}$. D’après ce que nous savons de ces représentations de $\mathfrak{su}(2)$, 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 (2.3)

$$m = -p - q . \quad (2.4)$$

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 § 1 a laissé les coefficients $N_{\alpha\beta}$ indéterminés. On démontre (cf Exercice C) 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

$$|N_{\alpha\beta}| = \sqrt{\frac{1}{2}(1-p)q\langle \alpha, \alpha \rangle} . \quad (2.5)$$

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

$$w_\alpha(x) = x - 2 \frac{\langle \alpha, x \rangle}{\langle \alpha, \alpha \rangle} \alpha . \quad (2.6)$$

C'est une réflexion dans l'hyperplan orthogonal à α : $(w_\alpha)^2 = I$, $w_\alpha(\alpha) = -\alpha$, 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 (2.4), 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.

Exemple : pour l'algèbre $\mathfrak{su}(n)$, on montre que $W = S_n$, le groupe de permutations de n objets.

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 .

◦ *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 α_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.

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.

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. (P)

On définit alors la matrice de Cartan comme

$$C_{ij} = 2 \frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle}. \quad (2.7)$$

Attention, cette matrice n'est pas *a priori* symétrique. 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 (2.7) est défini positif. Selon l'inégalité de Schwartz, $\langle \alpha, \beta \rangle^2 \leq \langle \alpha, \alpha \rangle \langle \beta, \beta \rangle$ 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.

Notons d'abord qu'en multipliant ou en divisant les deux équations (2.3) pour la paire $\alpha_i, \alpha_j, i \neq j$, c'est-à-dire $C_{ij} = m_1 \leq 0$ et $C_{ji} = m_2 \leq 0$, où on a pris en compte la propriété (P) ci-dessus, on obtient que si $i \neq j$,

$$\left. \begin{aligned} \cos \widehat{\alpha_i, \alpha_j} &= -\frac{\sqrt{m_1 m_2}}{2} \\ \frac{\|\alpha_i\|}{\|\alpha_j\|} &= \sqrt{\frac{m_1}{m_2}} \end{aligned} \right\} \text{ avec } m_1, m_2 \in \mathbb{N}, \quad (2.8)$$

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 ?(indéterminé), $1, \sqrt{2}, \sqrt{3}$.

Il n'existe bien sûr qu'une seule algèbre de rang 1, c'est l'algèbre $\mathfrak{su}(2)$ ((1.1) ou (1.18)) complexifiée. 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 figure 1, avec leurs matrices de Cartan s'écrivant

$$A_2 : \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad B_2 : \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix} \quad G_2 : \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \quad D_2 : \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}. \quad (2.9)$$

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

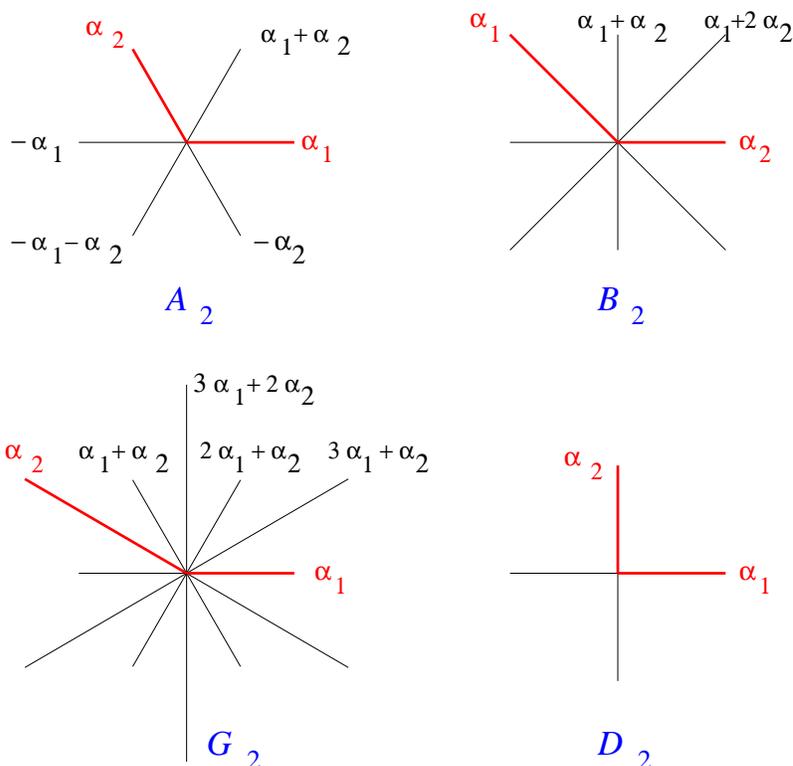


Fig. 1: Systèmes de racines de rang 2. Les deux racines simples positives sont en trait gras. Pour les algèbres B_2 , G_2 et D_2 , on n'a noté que les racines positives.

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.

o 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 est associé un vertex du diagramme. Deux vertex sont unis par une ligne sissi $\langle \alpha_i, \alpha_j \rangle \neq 0$; 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

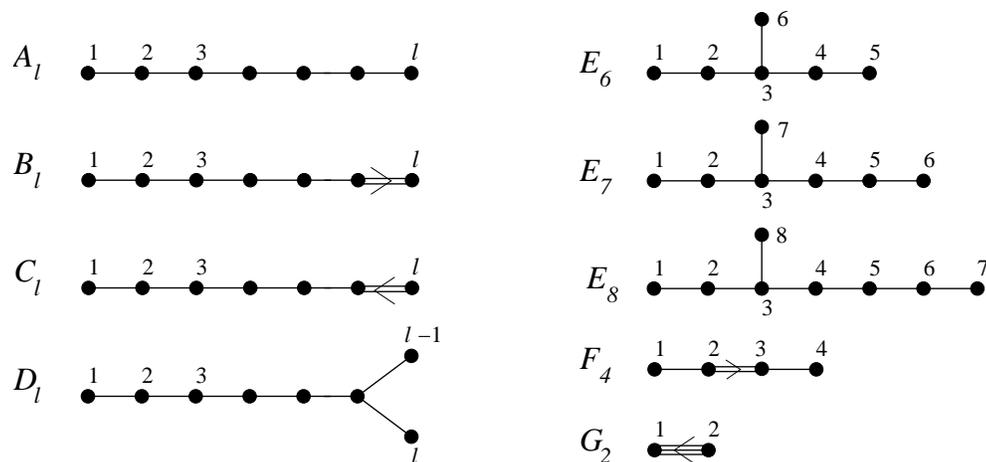


Fig. 2: Diagrammes de Dynkin

un rapport des longueurs de $\sqrt{2}$, resp. $\sqrt{3}$) et porte alors une flèche (ou plutôt un signe $>$) indiquant quelle racine est la plus longue.

2.2. Systèmes de racines des algèbres simples. La classification de Cartan

Une analyse de tous les cas possibles a mené Cartan ³ 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

$$A_\ell, B_\ell, C_\ell, D_\ell, E_6, E_7, E_8, F_4, G_2. \quad (2.10)$$

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 figure 2.

Les quatre familles infinies sont identifiées aux algèbres de Lie (complexifiées) de groupes classiques

$$A_\ell = \mathfrak{sl}(\ell + 1, \mathbb{C}), \quad B_\ell = \mathfrak{so}(2\ell + 1, \mathbb{C}), \quad C_\ell = \mathfrak{sp}(2\ell, \mathbb{C}), \quad D_\ell = \mathfrak{so}(2\ell, \mathbb{C}). \quad (2.11)$$

ou à leur unique forme réelle compacte, respectivement $\mathfrak{su}(\ell + 1)$, $B_\ell = \mathfrak{so}(2\ell + 1)$, $C_\ell = \mathfrak{usp}(\ell)$, $D_\ell = \mathfrak{so}(2\ell)$.

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.

³ Ce travail de classification, entrepris pas Killing a été complété et corrigé par É. Cartan, puis simplifié par van der Waerden, Dynkin, ...

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 $\mathrm{su}(2)$, singularités “simples”, etc, sont classifiés selon ce schéma ADE . . . mais cela est une autre histoire !

Les formes réelles de ces algèbres simples complexes ont aussi été classifiées par Cartan. On trouve 12 séries infinies et 23 cas exceptionnels !

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

$$e_i = \left(\frac{2}{\langle \alpha_i, \alpha_i \rangle} \right)^{\frac{1}{2}} E_{\alpha_i} , \quad f_i = \left(\frac{2}{\langle \alpha_i, \alpha_i \rangle} \right)^{\frac{1}{2}} E_{-\alpha_i} , \quad h_i = \frac{2\alpha_i \cdot H}{\langle \alpha_i, \alpha_i \rangle} . \quad (2.12)$$

Leurs relations de commutation s'écrivent

$$\begin{aligned} [h_i, h_j] &= 0 \\ [h_i, e_j] &= C_{ji} e_j \\ [h_i, f_j] &= -C_{ji} f_j \\ [e_i, f_j] &= \delta_{ij} h_j \end{aligned} \quad (2.13)$$

(le vérifier). L'algèbre est engendrée par les e_i, f_i, h_i et tous leurs commutateurs contraints par (2.13) et par les “relations de Serre”

$$\begin{aligned} \mathrm{ad}(e_i)^{1-C_{ji}} e_j &= 0 \\ \mathrm{ad}(f_i)^{1-C_{ji}} f_j &= 0 . \end{aligned} \quad (2.14)$$

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*.

2.4. Coracines. Plus haute racine. Nombres et 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

$$\alpha_i^\vee := \frac{2}{\langle \alpha_i, \alpha_i \rangle} \alpha_i , \quad (2.15)$$

pour α_i une racine simple, apparaît fréquemment, on lui donne le nom de *coracine*. La matrice de Cartan se réécrit donc

$$C_{ij} = \langle \alpha_i, \alpha_j^\vee \rangle . \quad (2.16)$$

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

$$\theta = \sum_i a_i \alpha_i \quad , \quad \frac{2}{\langle \theta, \theta \rangle} \theta = \sum_i a_i^\vee \alpha_i^\vee , \quad (2.17)$$

appelées aussi *indices de Kac* ou *indices de Kac duaux*, jouent aussi un rôle, en particulier par leurs sommes,

$$h = 1 + \sum_i a_i \quad , \quad h^\vee = 1 + \sum_i a_i^\vee . \quad (2.18)$$

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

$$\widehat{C}_{ij} := 2 \frac{\langle \alpha_i, \alpha_j \rangle}{\sqrt{\langle \alpha_i, \alpha_i \rangle \langle \alpha_j, \alpha_j \rangle}} \quad (2.19)$$

donne un spectre de valeurs propres

$$\text{valeurs propres de } \widehat{C} = \left\{ 4 \sin^2 \left(\frac{\pi}{2h} m_i \right) \right\}, \quad i = 1, \dots, \ell, \quad (2.20)$$

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 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.

3. Représentations des algèbres semi-simples

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

$$H|\lambda_a\rangle = \lambda(H)|\lambda_a\rangle \quad (3.1)$$

ou de façon équivalente

$$H_i|\lambda_a\rangle = \lambda_{(i)}|\lambda_a\rangle, \quad (3.2)$$

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 figure 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 sont les racines. Les racines étudiées aux paragraphes précédents appartiennent donc à l'ensemble des poids dans \mathfrak{h}^* .

Les vecteurs $|\lambda_\alpha\rangle$ 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 § 2, engendrées par $\{H_\alpha, E_\alpha, E_{-\alpha}\}$. Par le même argument qu'au § 2, on va montrer que tout poids λ satisfait

$$\forall \alpha, \quad 2 \frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} = m' \in \mathbb{Z}, \quad (3.3)$$

et réciproquement, on montre que tout $\lambda \in \mathfrak{h}^*$ satisfaisant (3.3) est le poids d'une représentation de dimension finie. On peut donc utiliser (3.3) comme autre définition d'un poids. Pour se convaincre que le poids de toute représentation satisfait (3.3), on peut, comme au § 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 $(E_{-\alpha})^{|p'|} |\lambda_\alpha\rangle \neq 0$, et q' le plus grand entier ≥ 0 tel que $(E_\alpha)^{q'} |\lambda_\alpha\rangle \neq 0$, 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

$$2q' + 2 \frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} = 2j \quad 2p' + 2 \frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} = -2j.$$

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

$$2 \frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} = -(q' + p') =: m', \quad \text{comme annoncé en (3.3).}$$

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

$$\mathcal{C}_w = \{ \lambda | \langle w\lambda, \alpha_i \rangle \geq 0, \forall i = 1, \dots, \ell \}, \quad (3.4)$$

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 la discussion ci-dessus, et $\langle \Lambda, \alpha \rangle = j > 0$, 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

$$\lambda_i = 2 \frac{\langle \lambda, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} \in \mathbb{Z} \quad (3.5)$$

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

$$2 \frac{\langle \Lambda_j, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = \delta_{ij} . \quad (3.6)$$

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 (3.5),

$$\Lambda = \sum_{j=1}^{\ell} \lambda_j \Lambda_j \quad , \quad \lambda_i \in \mathbb{N} . \quad (3.7)$$

et tout Λ de la forme (3.7) 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.

La représentation adjointe doit donc avoir un plus haut poids, dominant. Montrer en utilisant les propriétés énoncées ci-dessus que ce ne peut être que θ , défini au § 2.4.

◦ 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_i . 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.

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

$$\begin{aligned} \rho &= \frac{1}{2} \sum_{\alpha > 0} \alpha \\ &= \sum_j \Lambda_j . \end{aligned} \quad (3.8)$$

Une formule remarquable, due à Weyl, exprime la dimension de la représentation de plus haut poids Λ comme un produit sur les racines positives

$$\dim(\Lambda) = \prod_{\alpha > 0} \frac{\langle \Lambda + \rho, \alpha \rangle}{\langle \rho, \alpha \rangle} \quad (3.9)$$

tandis que la valeur propre du Casimir quadratique est

$$C_2(\Lambda) = \frac{1}{2} \langle \Lambda, \Lambda + 2\rho \rangle . \quad (3.10)$$

Une question liée est celle de la trace des générateurs de \mathfrak{g} dans la représentation (Λ) . Soit t_a la base de \mathfrak{g} telle que $\text{tr } t_a t_b = T_A \delta_{ab}$, avec un coefficient T_A ("A" comme adjointe) dont le signe dépend des conventions (t hermitiques ou antihermitiques, cf Chapitre 1). Dans la représentation de plus haut poids Λ , on doit avoir

$$\text{tr } d_\Lambda(t_a) d_\Lambda(t_b) = T_\Lambda \delta_{ab} . \quad (3.11)$$

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,

$$\begin{aligned} \text{tr } C_2 &= \sum_a \text{tr } (d_\Lambda(t_a))^2 = T_\Lambda \sum_a 1 = T_\Lambda \dim \mathfrak{g} \\ &= C_2(\Lambda) \text{tr } I_\Lambda = C_2(\Lambda) \dim(\Lambda) \end{aligned}$$

d'où

$$T_\Lambda = C_2(\Lambda) \frac{\dim(\Lambda)}{\dim \mathfrak{g}} , \quad (3.12)$$

une formule souvent utile dans les calculs (théories de jauge...). Dans la représentation adjointe, $\dim(A) = \dim \mathfrak{g}$, donc $T_\Lambda = T_A = C_2(A)$.

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

$$\begin{aligned} A_\ell = \mathfrak{su}(\ell + 1) & \quad \lambda_i \leftrightarrow \lambda_{\ell+1-i} & \ell > 1 \\ D_{2r+1} = \mathfrak{so}(4r + 2) & \quad \lambda_\ell \leftrightarrow \lambda_{\ell-1}, & \ell = 2r + 1 \\ E_6 & \quad \lambda_i \leftrightarrow \lambda_{6-i}, & i = 1, 2. \end{aligned} \quad (3.13)$$

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 l’hyperplan orthogonal à $\hat{\rho} := \sum_{i=1}^n \hat{e}_i$, donc $e_i = \hat{e}_i - \frac{1}{n}\hat{\rho}$. (En d’autres termes, on s’est placé dans l’hyperplan $\sum_i x^i = 1$ de l’espace \mathbb{R}^n). Ces vecteurs ont pour produits scalaires

$$(e_i, e_j) = \delta_{ij} - \frac{1}{n}. \quad (3.14)$$

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

$$\alpha_{ij} = e_i - e_j, \quad 1 \leq i < j \leq n, \quad (3.15)$$

et les $\ell = n - 1$ racines simples sont

$$\alpha_i = \alpha_{i, i+1} = e_i - e_{i+1}, \quad 1 \leq i \leq n - 1. \quad (3.16)$$

Ces racines ont été normalisées à $\langle \alpha, \alpha \rangle = 2$. La somme des racines positives se calcule aisément

$$\begin{aligned} 2\rho &= (n-1)e_1 + (n-3)e_2 + \dots + (n-2i+1)e_i + \dots - (n-1)e_n \\ &= (n-1)\alpha_1 + 2(n-2)\alpha_2 + \dots + i(n-i)\alpha_i + \dots + (n-1)\alpha_{n-1}. \end{aligned} \quad (3.17)$$

On vérifie que la matrice de Cartan est bien $\langle \alpha_i, \alpha_j \rangle = \begin{cases} 2 & \text{if } i = j \\ -1 & \text{if } i = j \pm 1 \end{cases}$, 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

$$\Lambda_i = \sum_{j=1}^i e_j, \quad (3.18)$$

$$e_1 = \Lambda_1, \quad e_i = \Lambda_i - \Lambda_{i-1} \text{ pour } i = 2, \dots, n-1, \quad e_n = -\Lambda_{n-1}$$

avec des produits scalaires

$$\langle \Lambda_i, \Lambda_j \rangle = \frac{i(n-j)}{n}, \quad i \leq j. \quad (3.19)$$

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 (3.9) et (3.15), démontrer l'expression suivante

$$\dim(\Lambda) = \prod_{1 \leq i < j \leq n} \frac{f_i - f_j + j - i}{j - i} \quad \text{où} \quad f_i := \sum_{k=i}^{n-1} \lambda_k, \quad f_n = 0. \quad (3.20)$$

◦ *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

$$\nu(\lambda) := \lambda_1 + 2\lambda_2 + \dots + (n-1)\lambda_{n-1} \pmod{n}, \quad (3.21)$$

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 $\nu(\lambda) = 0$.

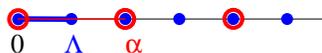


Fig. 3: Poids de $\mathfrak{su}(2)$. Les parties positives des réseaux des poids (petits points) et des racines (gros points).

◦ Exemples de $su(2)$ et $su(3)$.

Dans le cas de $su(2)$, 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 figure 3. L'indice de Dynkin λ_1 s'identifie à l'entier $2j$, les deux classes de congruence 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 l'expression bien connue.

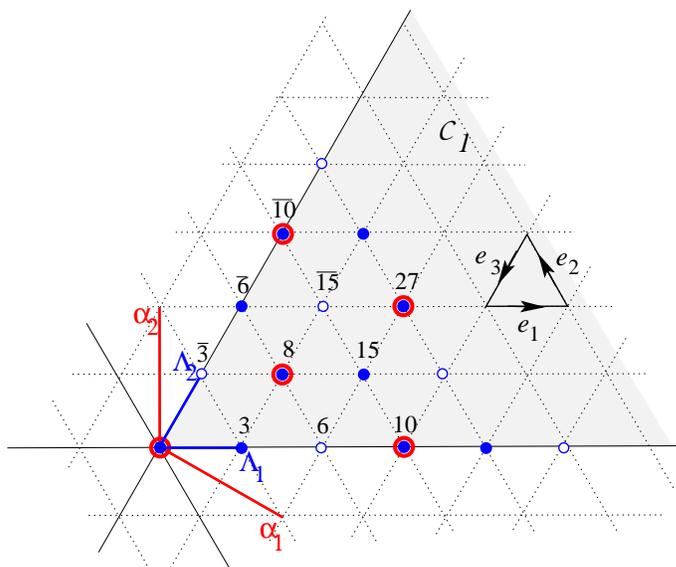


Fig. 4: Poids de $su(3)$. Seule la première chambre de Weyl C_1 a été détaillée, avec quelques plus hauts poids. Les poids de trialité 0 (réseau des racines) sont représentés par un gros disque, ceux de trialité 1, resp. 2, par un disque plein, resp. vide.

Pour $su(3)$, le réseau des poids est triangulaire, voir figure 4, sur laquelle on a indiqué la trialité $\tau(\lambda) := \lambda_1 + 2\lambda_2 \pmod{3}$ et porté les poids fondamentaux des représentations les plus basses. Selon l'usage, les représentations sont repérées par leur dimension⁴

$$\dim(\Lambda) = \frac{1}{2}(\lambda_1 + 1)(\lambda_2 + 1)(\lambda_1 + \lambda_2 + 2), \quad (3.22)$$

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.

⁴ ce qui est parfois ambigu ; par exemple, identifier sur la figure 4 le poids d'une autre représentation de dimension 15.

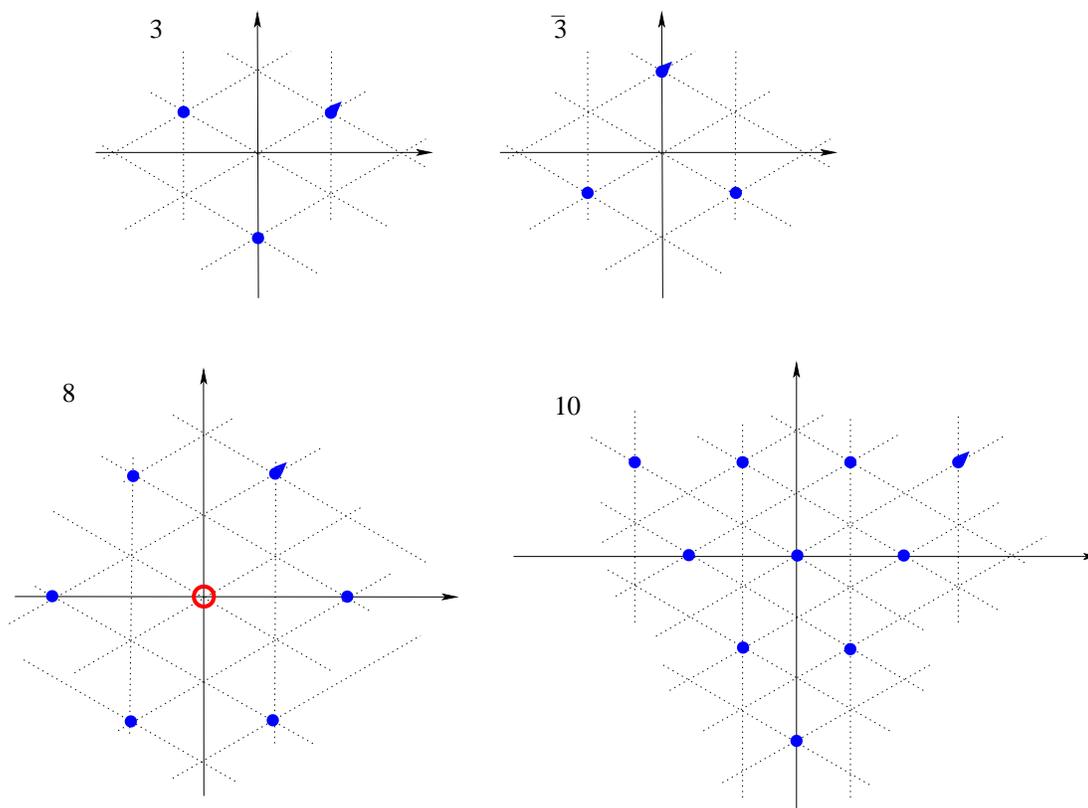


Fig. 5: Les diagrammes de poids des représentations les plus basses de $\mathfrak{su}(3)$, désignées par leur dimension. Noter qu’une rotation de 30° du réseau des poids a été effectuée par rapport aux figures précédentes. Dans chaque représentation, le poids le plus haut est indiqué par une petite indentation. Les points ont une multiplicité 1, les points ouverts ont la multiplicité 2.

Exercice. Calculer la valeur propre de l’opérateur de Casimir quadratique en termes des indices de Dynkin λ_1, λ_2 grâce aux formules (3.10) et (3.17).

L’ensemble des poids des représentations les plus basses est représenté sur la figure 5, après rotation des axes des figures précédentes. L’axe horizontal, porté par α_1 , et l’axe vertical, porté 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.

4. Produit tensoriel des représentations de $\mathfrak{su}(n)$

4.1. Règles de Littlewood-Richardson

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 chapitre 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 poids dominant 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 .

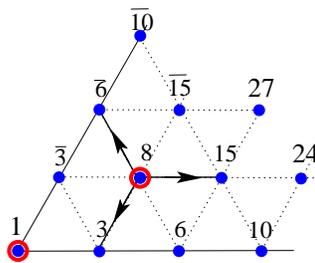


Fig. 6: Opération de produit tensoriel par la représentation 3 figurée sur le diagramme des poids de $\mathfrak{su}(3)$.

Illustrons ceci sur le cas de $\text{su}(3)$. 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

$$\begin{aligned}
 3 \otimes 3 &= \bar{3} + 6 & 3 \otimes \bar{3} &= 1 + 8 \\
 3 \otimes 6 &= 8 + 10 & 3 \otimes \bar{6} &= \bar{3} + \bar{15} \\
 3 \otimes 8 &= 3 + \bar{6} + 15 \\
 3 \otimes 15 &= 6 + \bar{15} + 24
 \end{aligned} \tag{4.1}$$

etc, et leurs conjuguées, cf figure 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 dominants 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

$$8 \otimes (1 + 8) = 8 \otimes 3 \otimes \bar{3} = (3 + \bar{6} + 15) \otimes \bar{3} = 1 + 8 + 8 + 8 + 10 + \bar{10} + 27$$

d'où on tire une expression qui nous sera utile par la suite

$$8 \otimes 8 = 1 + 8_s + 8_a + 10 + \bar{10} + 27 . \tag{4.2}$$

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 $\text{SU}(3)$ de saveur.

Bien que fastidieux, ce processus est simple et systématique. (Il existe aussi des codes qui calculent ces décompositions...)

4.2. Construction tensorielle explicite des représentations de $SU(2)$ et $SU(3)$

Considérons un vecteur $V \in \mathbb{C}^n$ de la représentation de définition de $SU(n)$. Sous l'action de $U \in SU(n)$, $V \mapsto V' = UV$, 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 $V.W := v^i w_i$ est invariant, en vertu de $U^\dagger.U = I$. Autrement dit le tenseur mixte δ_j^i est invariant

$$\delta'^i_j = U^i_{i'} U^{\dagger j'}_j \delta^{i'}_{j'} = (U.U^\dagger)^i_j = \delta^i_j .$$

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

$$t^{i_1, \dots, i_p}_{k_1, \dots, k_m} = U^{i_1}_{j_1} \dots U^{i_p}_{j_p} U^{\dagger l_1}_{k_1} \dots U^{\dagger l_m}_{k_m} t^{j_1, \dots, j_p}_{l_1, \dots, l_m} . \quad (4.3)$$

• Dans le cas de $SU(2)$, on sait que les représentations U et U^* sont équivalentes. Cela résulte de l'existence d'une matrice $C = i\sigma_2$, telle que $CU C^{-1} = U^*$, 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 ($\epsilon^{ij} = -\epsilon_{ij}$, $\epsilon_{ij} \epsilon^{jk} = \delta_i^k$), 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 rang ⁵. 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 $SU(2)$ par produits tensoriels symétrisés de la représentation de dimension 2, cf Chap 00, et le rang p s'identifie à $2j$. On vérifie en particulier que le nombre de composantes indépendantes d'un tenseur de rang p complètement symétrique dans l'espace \mathbb{C}^2

⁵ Il est bon de se rappeler les identités $\epsilon_{ab} \epsilon_{cd} = \delta_{ac} \delta_{bd} - \delta_{ad} \delta_{bc}$ et donc $\epsilon_{ab} \epsilon_{bc} = -\delta_{ac}$.

est $p + 1$, puisque ces composantes possèdent $0, 1, \dots, p$ indices égaux à 1, les autres étant égaux à 2.

Un tenseur de rang p complètement symétrique sera représenté par un “diagramme d’Young” avec p boîtes $\underbrace{\square\square\square}_p$. 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 dominant $p\Lambda_1 + m\Lambda_2$ 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 (3.22), cf Exercice D. 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 § 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 § 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 $t_j^i \mapsto t'^i_j = U^i_{i'} U^{*j'}_j t'^i_{j'}$, soit sous forme matricielle

$$t' = UtU^\dagger, \tag{4.4}$$

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 ?

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 $GL(n)$ est naturellement représenté dans E

$$g \in GL(E) \sim GL(n), \quad x \in E \mapsto x' = g.x. \tag{5.1}$$

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 (5.1)

$$g \in GL(n), \quad D(g)x^{(1)} \otimes \dots \otimes x^{(m)} = g.x^{(1)} \otimes \dots \otimes g.x^{(m)} \tag{5.2}$$

qui est en général réductible. Mais dans F , agit aussi le groupe symétrique S_m selon

$$\sigma \in S_m, \quad \widehat{D}(\sigma)x^{(1)} \otimes \dots \otimes x^{(m)} = x^{(\sigma^{-1}1)} \otimes \dots \otimes x^{(\sigma^{-1}m)}. \tag{5.3}$$

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

$$\mathcal{D}(g)_{\{i_1, \dots, i_m\}, \{j_1, \dots, j_m\}} = \prod_{k=1}^m g_{i_k, j_k} \tag{5.4}$$

et celle de S_m

$$\widehat{\mathcal{D}}(\sigma)_{\{i_1, \dots, i_m\}, \{j_1, \dots, j_m\}} = \prod_{k=1}^m \delta_{i_{\sigma k}, j_k} \tag{5.5}$$

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 $t'_i = \mathcal{D}_{i, \cdot, j} t_j$, resp. $\widehat{\mathcal{D}}_{i, \cdot, j} t_j$. Ces deux ensembles de matrices commutent

$$\begin{aligned} \sum_{\{j\}} \mathcal{D}(g)_{\{i\}, \{j\}} \widehat{\mathcal{D}}(\sigma)_{\{j\}, \{k\}} &= \prod_l g_{i_l j_l} \delta_{j_l, k_{\sigma^{-1} l}} = \prod_l g_{i_{\sigma l} k_l} \\ &= \sum_{\{j\}} \widehat{\mathcal{D}}(\sigma)_{\{i\}, \{j\}} \mathcal{D}(g)_{\{j\}, \{k\}} \end{aligned} \tag{5.6}$$

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 : $\ell_1 > \ell_2 > \dots > \ell_k$. Alors

$$n_Y = \frac{n!}{\prod_i \ell_i!} \prod_{i < j} (\ell_i - \ell_j) \tag{5.7}$$

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 représentations irréductibles et les diagrammes d'Young à m boîtes. La dimension de cette représentation est donnée par le nombre de tableaux standards (5.7).

On appelle tenseur de type (de symétrie) Y un tenseur se transformant par S_m selon cette représentation. 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 $GL(n)$

$$t_{i_1, i_2} = \frac{1}{2} (t_{i_1, i_2} + t_{i_2, i_1}) + \frac{1}{2} (t_{i_1, i_2} - t_{i_2, i_1}) \tag{5.8}$$

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

$$\begin{aligned} \begin{array}{|c|} \hline \square \square \square \\ \hline \end{array} A &= t_{i_1, i_2, i_3} + t_{i_2, i_3, i_1} + t_{i_3, i_1, i_2} + t_{i_2, i_1, i_3} + t_{i_3, i_2, i_1} + t_{i_1, i_3, i_2} \\ \begin{array}{|c|} \hline \square \\ \hline \square \square \\ \hline \end{array} B &= t_{i_1, i_2, i_3} + t_{i_2, i_3, i_1} + t_{i_3, i_1, i_2} - t_{i_2, i_1, i_3} - t_{i_3, i_2, i_1} - t_{i_1, i_3, i_2} \\ \begin{array}{|c|} \hline \square \square \\ \hline \square \\ \hline \end{array} C_1 &= t_{i_1, i_2, i_3} - t_{i_2, i_3, i_1} + t_{i_2, i_1, i_3} - t_{i_3, i_2, i_1} \\ \begin{array}{|c|} \hline \square \square \\ \hline \square \\ \hline \end{array} D_1 &= t_{i_1, i_2, i_3} - t_{i_2, i_1, i_3} + t_{i_3, i_2, i_1} - t_{i_3, i_1, i_2} \end{aligned}$$

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 $t_{i, j, k}$ 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 pour n donné. Il est clair qu'un tableau à $k > n$ lignes implique une antisymétrisation de k indices prenant leurs valeurs dans $\{1, \dots, n\}$ et donne donc un résultat nul. Par contre 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

$$\dim_Y^{(n)} = \frac{\Delta(f_1 + n - 1, f_2 + n - 2, \dots, f_n)}{\Delta(n - 1, n - 2, \dots, 0)} \tag{5.8}$$

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 (3.20).

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 $1 \leq i_1 \leq i_2 \leq \dots \leq i_m \leq n$. 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, ..., n . La dimension cherchée est donc le coefficient binomial C_{n+m-1}^m , en accord dans ce cas particulier avec (5.8).

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

$$\square\square + \begin{array}{|c|} \hline \square \\ \hline \end{array} + 2 \begin{array}{|c|c|} \hline \square & \square \\ \hline \end{array}$$

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 H. 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.

o *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 $\Lambda_3 = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}$. 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

$$\Lambda = (\lambda_1, \dots, \lambda_{n-1}) \leftrightarrow Y = (f_i = \sum_{j=i}^{n-1} \lambda_j) . \tag{5.9}$$

Autrement dit, λ_k est le nombre de colonnes de Y de hauteur k .

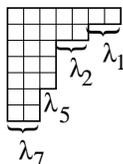


Fig. 7: Correspondance entre diagramme d'Young et indices de Dynkin. Ici $Y \leftrightarrow \Lambda = (2, 2, 0, 0, 1, 0, 2)$

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 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 A. The classical algebras of type A, B, C, D

 1. $sl(N) = A_{N-1}$

Rank = $r = N - 1$, dimension $N^2 - 1$, Coxeter number $h = N$, dual Coxeter number $h^\vee = N$.

$e_i, i = 1, \dots, N$ a set of vectors in \mathbb{R}^N such that $\sum_1^N e_i = 0, \langle e_i, e_j \rangle = \delta_{ij} - \frac{1}{N}$.

Roots $\alpha_{ij} = e_i - e_j, i \neq j = 1, \dots, N$; positive roots $\alpha_{ij} = e_i - e_j, i < j$; their number $|\Delta_+| = N(N-1)/2$; simple roots $\alpha_i := \alpha_{i, i+1} = e_i - e_{i+1} i = 1, \dots, N-1$.

Highest root $\theta = \alpha_1 + \dots + \alpha_{N-1} = 2e_1 + e_2 + \dots + e_{N-1} = \Lambda_1 + \Lambda_{N-1} = (1, 0, \dots, 0, 1)$.

Sum of positive roots

$$\begin{aligned} 2\rho &= (N-1)e_1 + (N-3)e_2 + \dots + (N-2i+1)e_i + \dots - (N-1)e_N \\ &= (N-1)\alpha_1 + 2(N-2)\alpha_2 + \dots + i(N-i)\alpha_i + \dots + (N-1)\alpha_{N-1}. \end{aligned}$$

$$\text{Cartan matrix } \langle \alpha_i, \alpha_j \rangle = \begin{cases} 2 & \text{if } i = j \\ -1 & \text{if } i = j \pm 1 \\ 0 & \text{otherwise} \end{cases}$$

Fundamental weights $\Lambda_i i = 1, \dots, N-1, \Lambda_i = \sum_{j=1}^i e_j, e_1 = \Lambda_1, e_i = \Lambda_i - \Lambda_{i-1}$ for $i = 2, \dots, N-1, e_N = -\Lambda_{N-1}$.

$\langle \Lambda_i, \Lambda_j \rangle = \frac{i(N-j)}{N}$ for $i \leq j$.

Weyl group: $W \equiv \mathcal{S}_N$ acts on the weights by permuting the e_i : $w \in W \leftrightarrow \bar{w} \in \mathcal{S}_N: w(e_i) = e_{\bar{w}(i)}$

Coxeter exponents $\{1, 2, \dots, N-1\}$.

 2. $so(2l+1) = B_l, l \geq 2$

Rank = l , dimension $l(2l+1)$, Coxeter number $h = 2l$, dual Coxeter number $h^\vee = 2l-1$

$e_i, i = 1, \dots, l, \langle e_i, e_j \rangle = \delta_{ij}$ a basis of \mathbb{R}^l .

Roots $\pm e_i, 1 \leq i \leq l$ and $\pm e_i \pm e_j, 1 \leq i < j \leq l$. Basis of simple roots $\alpha_i = e_i - e_{i+1}, i = 1, \dots, l-1$, and $\alpha_l = e_l$.

Positive roots

$$\begin{aligned} e_i &= \sum_{i \leq k \leq l} \alpha_k, \quad 1 \leq i \leq l, \\ e_i - e_j &= \sum_{i \leq k < j} \alpha_k, \quad 1 \leq i < j \leq l, \\ e_i + e_j &= \sum_{i \leq k < j} \alpha_k + 2 \sum_{j \leq k \leq l} \alpha_k, \quad 1 \leq i < j \leq l, \end{aligned}$$

their number is $|\Delta_+| = l^2$.

Highest root $\theta = e_1 + e_2 = \alpha_1 + 2\alpha_2 + \dots + 2\alpha_l$.

Sum of positive roots

$$\begin{aligned} 2\rho &= (2l-1)e_1 + (2l-3)e_2 + \dots + (2l-2i+1)e_i + \dots + 3e_{l-1} + e_l \\ &= (2l-1)\alpha_1 + 2(2l-2)\alpha_2 + \dots + i(2l-i)\alpha_i + \dots + l^2\alpha_l. \end{aligned}$$

$$\text{Cartan matrix } \langle \alpha_i, \alpha_j^\vee \rangle = \begin{cases} 2 & \text{if } 1 \leq i = j \leq l \\ -1 & \text{if } 1 \leq i = (j \pm 1) \leq l-1 \\ -2 & \text{if } i = l-1, j = l \\ -1 & \text{if } i = l, j = l-1 \\ 0 & \text{otherwise} \end{cases}$$

Fundamental weights $\Lambda_i = \sum_{j=1}^i e_j, i = 1, \dots, l-1, \Lambda_l = \frac{1}{2} \sum_{j=1}^l e_j$; hence $e_1 = \Lambda_1 = (1, 0, \dots, 0)$,

$e_i = \Lambda_i - \Lambda_{i-1} = (0, \dots, -1, 1, 0, \dots), i = 2, \dots, l-1, e_l = 2\Lambda_l - \Lambda_{l-1} = (0, \dots, 0, -1, 2)$.

Dynkin indices of the roots $\alpha_1 = (2, -1, 0, \dots), \alpha_i = (0, \dots, -1, 2, -1, 0, \dots), i = 2, \dots, l-2; \alpha_{l-1} = (0, \dots, 0, -1, 2, -2); \alpha_l = (0, \dots, 0, -1, 2)$ and $\theta = (0, 1, 0, \dots, 0)$

Weyl group: $W \equiv \mathcal{S}_N \times (\mathbb{Z}_2)^l$, of order $2^l \cdot l!$, acts on the weights by permuting the e_i and $e_i \mapsto (\pm 1)_i e_i$.
Coxeter exponents $\{1, 3, 5, \dots, 2l - 1\}$.

3. $sp(2l) = C_l, l \geq 2$

Rank = l , dimension $l(2l + 1)$, Coxeter number $h = 2l$, dual Coxeter number $h^\vee = l + 1$
 $e_i, i = 1, \dots, l$, $\langle e_i, e_j \rangle = \frac{1}{2} \delta_{ij}$ a basis of \mathbb{R}^l (**Beware !** factor 2 to enforce the normalisation $\theta^2 = 2$).
Basis of simple roots $\alpha_i = e_i - e_{i+1}, i = 1, \dots, l - 1$, and $\alpha_l = 2e_l$.
Roots $\pm 2e_i, 1 \leq i \leq l$ and $\pm e_i \pm e_j, 1 \leq i < j \leq l$.
Positive roots

$$\begin{aligned} e_i - e_j &= \sum_{i \leq k < j} \alpha_k, \quad 1 \leq i < j \leq l, \\ e_i + e_j &= \sum_{i \leq k < j} \alpha_k + 2 \sum_{j \leq k < l} \alpha_k + \alpha_l, \quad 1 \leq i < j \leq l, \\ 2e_i &= 2 \sum_{i \leq k < l} \alpha_k + \alpha_l, \quad 1 \leq i \leq l, \end{aligned}$$

their number is $|\Delta_+| = l^2$.
Highest root $\theta = 2e_1 = 2\alpha_1 + 2\alpha_2 + \dots + 2\alpha_{l-1} + \alpha_l$.
Sum of positive roots

$$\begin{aligned} 2\rho &= 2le_1 + (2l - 2)e_2 + \dots + (2l - 2i + 2)e_i + \dots + 4e_{l-1} + 2e_l \\ &= 2l\alpha_1 + 2(2l - 1)\alpha_2 + \dots + i(2l - i + 1)\alpha_i + \dots + (l - 1)(l - 2)\alpha_{l-1} + \frac{1}{2}l(l + 1)\alpha_l. \end{aligned}$$

$$\text{Cartan matrix } \langle \alpha_i, \alpha_j^\vee \rangle = \begin{cases} 2 & \text{if } 1 \leq i = j \leq l \\ -1 & \text{if } 1 \leq i = (j \pm 1) \leq l - 1 \\ -1 & \text{if } i = l - 1, j = l \\ -2 & \text{if } i = l, j = l - 1 \\ 0 & \text{otherwise} \end{cases}$$

Fundamental weights

$\Lambda_i = \sum_{j=1}^i e_j, i = 1, \dots, l$, hence $e_1 = \Lambda_1 = (1, 0, \dots, 0)$, $e_i = \Lambda_i - \Lambda_{i-1} = (0, \dots, -1, 1, 0 \dots)$, $i = 2, \dots, l$.

Dynkin indices of the roots

$\alpha_1 = (2, -1, 0, \dots)$, $\alpha_i = (0, \dots, -1, 2, -1, 0 \dots)$, $i = 2, \dots, l - 1$; $\alpha_l = (0, \dots, 0, -2, 2)$ and $\theta = (2, 0, \dots, 0)$

Weyl group: $W \equiv \mathcal{S}_N \times (\mathbb{Z}_2)^l$, of order $2^l \cdot l!$, acts on the weights by permuting the e_i and $e_i \mapsto (\pm 1)_i e_i$.

4. $so(2l) = D_l, l \geq 3$

Rank = l , dimension $l(2l - 1)$, Coxeter number = dual Coxeter number $h = 2l - 2 = h^\vee$
 $e_i, i = 1, \dots, l$, $\langle e_i, e_j \rangle = \delta_{ij}$ a basis of \mathbb{R}^l .
Basis of simple roots $\alpha_i = e_i - e_{i+1}, i = 1, \dots, l - 1$, and $\alpha_l = e_{l-1} + e_l$.
Positive roots

$$\begin{aligned} e_i - e_j &= \sum_{i \leq k < j} \alpha_k, \quad 1 \leq i < j \leq l, \\ e_i + e_j &= \sum_{i \leq k < j} \alpha_k + 2 \sum_{j \leq k < l-1} \alpha_k + \alpha_{l-1} + \alpha_l, \quad 1 \leq i < j \leq l - 1, \\ e_i + e_l &= \sum_{i \leq k \leq l-2} \alpha_k + \alpha_l, \quad 1 \leq i \leq l - 1, \end{aligned}$$

their number is $|\Delta_+| = l(l - 1)$.
Highest root $\theta = e_1 + e_2 = \alpha_1 + 2\alpha_2 + \dots + 2\alpha_{l-2} + \alpha_{l-1} + \alpha_l$.

Sum of positive roots

$$\begin{aligned} 2\rho &= 2(l-1)e_1 + 2(l-2)e_2 + \cdots + 2e_{l-1} \\ &= 2(l-1)\alpha_1 + 2(2l-3)\alpha_2 + \cdots + i(2l-i-1)\alpha_i + \cdots + \frac{l(l-1)}{2}(\alpha_{l-1} + \alpha_l). \end{aligned}$$

Weyl group: $W \equiv \mathcal{S}_N \ltimes (\mathbb{Z}_2)^{l-1}$, of order $2^{l-1}l!$, acts on the weights by permuting the e_i and $e_i \mapsto (\pm 1)_i e_i$, with $\prod_i (\pm 1)_i = 1$.

Coxeter exponents $\{1, 3, 5, \dots, 2l-3, l-1\}$, with $l-1$ appearing twice if l is even.

$$\text{Cartan matrix } \langle \alpha_i, \alpha_j \rangle = \begin{cases} 2 & \text{if } 1 \leq i = j \leq l \\ -1 & \text{if } 1 \leq i = (j \pm 1) \leq l-2 \\ -1 & \text{if } (i, j) = (l-2, l) \text{ or } (l, l-2) \\ 0 & \text{otherwise} \end{cases}$$

Fundamental weights

$$\begin{aligned} \Lambda_i &= \sum_{j=1}^i e_j = \alpha_1 + 2\alpha_2 + \cdots + (i-1)\alpha_{i-1} + i(\alpha_i + \cdots + \alpha_{l-2}) + \frac{i}{2}(\alpha_{l-1} + \alpha_l) \text{ for } i = 1, \dots, l-2; \\ \Lambda_{l-1} &= \frac{1}{2}(e_1 + \cdots + e_{l-1} - e_l) = \frac{1}{2}(\alpha_1 + 2\alpha_2 + \cdots + (l-2)\alpha_{l-2}) + \frac{l}{2}\alpha_{l-1} + \frac{l-2}{2}\alpha_l; \\ \Lambda_l &= \frac{1}{2}(e_1 + \cdots + e_{l-1} + e_l) = \frac{1}{2}(\alpha_1 + 2\alpha_2 + \cdots + (l-2)\alpha_{l-2}) + \frac{l-2}{2}\alpha_{l-1} + \frac{l}{2}\alpha_l. \end{aligned}$$

For the exceptional algebras of types E, F, G , see Bourbaki.

Exercices

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 § 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 § 1.2), la forme de Killing s'écrit

$$(H, H') = \text{tr}(\text{ad}(H)\text{ad}(H')) = \sum_{\alpha \in \Delta} \alpha(H)\alpha(H'). \quad (\text{A.1})$$

2. On veut montrer que les racines α définies par (1.5) ou (1.6) 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

$$\forall \alpha \in \Delta \quad \alpha(H) = 0. \quad (\text{A.2})$$

Montrer en utilisant (A.1) 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 (1.10)).

3. Variante de l'argument précédent : sous l'hypothèse du 2. et donc de (A.2), 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

$$N_{\alpha\beta} = -N_{-\alpha, -\beta} . \quad (\text{A.3})$$

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

$$N_{\alpha\beta} = N_{\beta, -\alpha-\beta} = N_{-\alpha-\beta, \alpha} . \quad (\text{A.4})$$

3. Considérant la α -chaîne passant par β et les deux entiers p et q définis au § 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 (A.4) que l'équation précédente se réécrit

$$\langle \alpha, \beta + k\alpha \rangle = f(k) - f(k-1) . \quad (\text{A.5})$$

4. Que valent $f(q)$ et $f(q-1)$? Montrer que la récurrence (A.5) se résout en

$$f(k) = -(N_{\alpha, \beta+k\alpha})^2 = (k-q)\langle \alpha, \beta + \frac{1}{2}(k+q+1)\alpha \rangle . \quad (\text{A.6})$$

Que vaut $f(p-1)$? Montrer que l'on retrouve l'expression (2.3). Montrer que la formule (A.6) est en accord avec (2.5). 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 cours, §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 \mathbb{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 $so(2l+1)$.

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

$$\begin{aligned} e_i &= \sum_{i \leq k \leq l} \alpha_k, & 1 \leq i \leq l, \\ e_i - e_j &= \sum_{i \leq k < j} \alpha_k, & 1 \leq i < j \leq l, \\ e_i + e_j &= \sum_{i \leq k < j} \alpha_k + 2 \sum_{j \leq k \leq l} \alpha_k, & 1 \leq i < j \leq l, \end{aligned}$$

se qualifient comme racines positives. Vérifier cette assertion sur le cas de $B_2 = so(5)$.

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") $W \equiv \mathcal{S}_l \ltimes (\mathbb{Z}_2)^l$, 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. Vérifier cette assertion sur le cas de $B_2 = so(5)$.

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 $SO(5)$?

k. Dessiner sur un même dessin les racines et les premiers poids de $so(5)$.

2. G_2

Dans l'espace \mathbb{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.

- Que peut-on dire alors de la dimension de l'algèbre G_2 ?
- 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.
- Que sont les racines positives? Calculer le vecteur ρ , demi-somme des racines positives.
- 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.
- Vérifier que les poids fondamentaux sont

$$\Lambda_1 = 2\alpha_1 + \alpha_2 \quad \Lambda_2 = 3\alpha_1 + 2\alpha_2$$

- Que sont les dimensions de représentations fondamentales?
- 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

Pourquoi les groupes $SO(5)$ ou G_2 ne convenaient-ils pas comme groupe de symétrie étendant le $SU(2)$ d'isospin, sachant qu'on avait observé plusieurs "octets" de particules?

D. Dimensions des représentations de $SU(3)$

Admettant que la construction du § 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 dominant (p, m) de $SU(3)$, on se propose de calculer la dimension $d(p, m)$ de l'espace de ces tenseurs.

- 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 $d(p, m) = d(p, 0)d(0, m) - d(p-1)d(0, m-1)$.

- Montrer par un calcul semblable à celui de $SU(2)$ que

$$d(p, 0) = d(0, p) = \frac{1}{2}(p+1)(p+2) .$$

- En déduire l'expression de $d(p, m)$ et comparer à (3.22).

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 $SU(3)$.

1. Symétries globales exactes ou brisées. Brisure spontanée

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 $\mathcal{D}(g)$, indépendamment du point x

$$\phi(x) \mapsto \mathcal{D}(g)\phi(x) . \tag{1.1}$$

Selon le théorème de Wigner, on suppose que cette transformation est aussi réalisée sur les états de l’espace de Hilbert de la théorie par un opérateur unitaire $U(g)$.

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_i^μ de divergence nulle (cf chap.0), à des charges conservées Q_i , $i = 1, \dots, \dim G$. Ces charges agissent sur les champs comme générateurs infinitésimaux, classiquement au sens du crochet de Poisson, $\{Q_i, \phi(x)\}\delta\alpha^i = \delta\phi(x)$, 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 $[Q_i, \phi(x)]\delta\alpha^i = -i\hbar\delta\phi(x)$ et entre eux $[Q_i, Q_j] = iC_{ij}^k Q_k$. 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

$$\phi(x) \mapsto e^{iq\alpha} \phi(x) ; \quad \phi^\dagger(x) \mapsto e^{-iq\alpha} \phi^\dagger(x),$$

et il y a invariance (du lagrangien) si tous les champs se transforment ainsi, avec 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 $Q = \sum q_i$ est conservée. La théorie quantique est l'électrodynamique quantique, et on y démontre que la symétrie classique par le groupe $U(1)$, 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

$$\partial_\mu j_i^\mu(x) = \frac{\partial \mathcal{L}(x)}{\partial \alpha^i}, \quad (1.2)$$

cf. Chapitre 0, § 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), 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 figure 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 Higgs– et à la cosmologie.

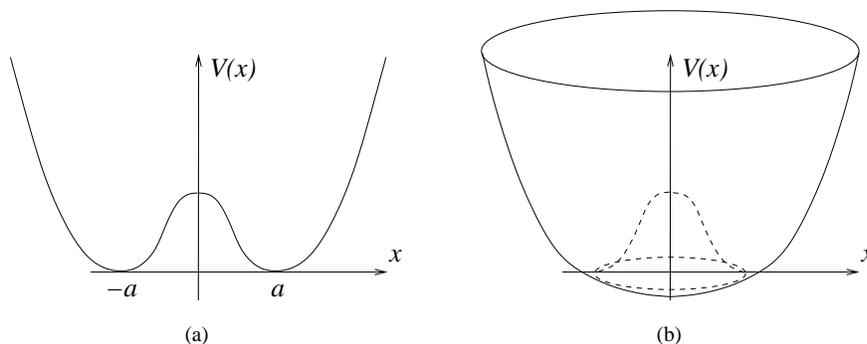


Fig. 1: Potentiels (a) à “double puits” ; (b) en “chapeau mexicain”

▷ *Exemple. Brisure spontanée du modèle $O(n)$*

Le lagrangien du “modèle $O(n)$ ” bosonique (minkovskien, ici), pour un champ $\phi = \{\phi^i\}$ réel à n composantes,

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4}(\phi^2)^2 \quad (1.3)$$

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

$$\langle 0|\phi|0\rangle = v\hat{n} . \quad (1.4)$$

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

$$\left. \frac{\partial^2 V}{\partial\phi_i\partial\phi_j} \right|_{\phi=v} \delta^a v_i = 0 , \tag{1.5}$$

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 (1.5) 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.

1.2. Brisure de la symétrie chirale

Considérons un lagrangien impliquant des fermions de masse nulle,

$$\mathcal{L} = \bar{\psi}i\not{\partial}\psi + g(\bar{\psi}\gamma^\mu\psi)(\bar{\psi}\gamma_\mu\psi) , \tag{1.6}$$

où $\psi = \{\psi_\alpha\}_{\alpha=1,\dots,N}$ est un vecteur à N composantes. Ce lagrangien est invariant sous l’action des deux types de transformations infinitésimales

$$\begin{aligned} \delta\psi(x) &= \delta A\psi(x) \\ \delta\psi(x) &= \delta B\gamma_5\psi(x) , \end{aligned} \tag{1.7}$$

où les matrices A et B sont antihermitiennes infinitésimales, de taille $N \times N$. Les courants de Noether conservés sont respectivement

$$J_\mu^a = \bar{\psi} T^a \gamma_\mu \psi \quad J_\mu^{a(5)} = \bar{\psi} T^a \gamma_5 \gamma_\mu \psi , \quad (1.8)$$

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 $(\gamma_5)^2 = I$, que $\frac{1}{2}(I \pm \gamma_5)$ sont donc des projecteurs et que $\gamma_5 = \gamma_5^\dagger$ anticommute avec les γ_μ ; 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

$$\partial^\mu J_\mu^{a(5)}(x) \propto m \bar{\psi} T^a \gamma_5 \psi . \quad (1.9)$$

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 (1.6) avec $N = 2$ et (1.3) pour $n = 4$, avec un terme couplant les fermions à quatre champs de bosons, nommés traditionnellement σ et $\boldsymbol{\pi}$

$$\mathcal{L} = \bar{\psi}(i \not{\partial} + g(\sigma + i\boldsymbol{\pi} \cdot \boldsymbol{\tau} \gamma_5))\psi + \frac{1}{2}((\partial\boldsymbol{\pi})^2 + (\partial\sigma)^2) - \frac{1}{2}m^2(\sigma^2 + \boldsymbol{\pi}^2) - \frac{\lambda}{4}(\sigma^2 + \boldsymbol{\pi}^2)^2 , \quad (1.10)$$

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 B). 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 $v = -m^2/\lambda$, 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 π 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. Nous en retrouverons des éléments dans le modèle standard.

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 celui 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.

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.

2.1. Pourquoi le groupe SU(3) ?

On a vu (Chapitre 00) que 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 lagrangien) des interactions fortes est invariant sous l'action de ce groupe SU(2) et en conséquence, 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 liée à la valeur propre de la troisième composante I_z de l'isospin par la relation

$$Q = \frac{1}{2}\mathcal{B} + I_z \quad [\text{pour SU}(2)] \quad (2.1)$$

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.

Cette relation entre Q et I_z doit être amendée pour une nouvelle famille de mésons ($K^\pm, K^0, \overline{K^0} \dots$) ou de baryons $\Lambda^0, \Sigma, \Xi, \dots$ découverts à la fin des années 50. On leur attribue un nouveau nombre quantique, l'*étrangeté* S . Cette étrangeté 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'étrangeté, tandis que la désintégration observée $\Lambda^0 \rightarrow p + \pi^-$ viole cette loi de conservation, car elle procède par les interactions faibles. La relation (2.1) doit être modifiée en la *relation de Gell-Mann–Nishima*

$$Q = \frac{1}{2}\mathcal{B} + \frac{1}{2}S + I_z = \frac{1}{2}Y + I_z, \quad (2.2)$$

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 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 $\mathfrak{su}(3)$, faite de 8 matrices hermitiennes λ_a qui jouent le rôle des matrices de Pauli σ_i pour $\mathfrak{su}(2)$. Ces matrices sont normalisées par

$$\text{tr} \lambda_a \lambda_b = 2\delta_{ab}. \quad (2.3)$$

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. Les deux générateurs de l'algèbre de Cartan sont

$$\lambda_3 = \begin{pmatrix} 1 & \cdot & \cdot \\ \cdot & -1 & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & \cdot & \cdot \\ \cdot & 1 & \cdot \\ \cdot & \cdot & -2 \end{pmatrix}. \quad (2.4)$$

Les charges I_z et Y sont alors les représentants dans la représentation considérée de $\frac{1}{2}\lambda_3$ et $\frac{1}{\sqrt{3}}\lambda_8$.

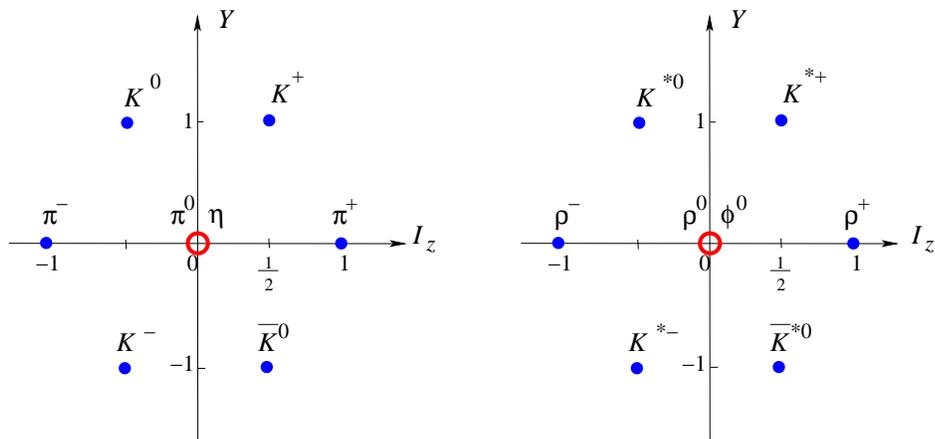


Fig. 2: Les octets de mésons pseudoscalaires ($J^P = 0^-$) et vecteurs ($J^P = 1^-$).

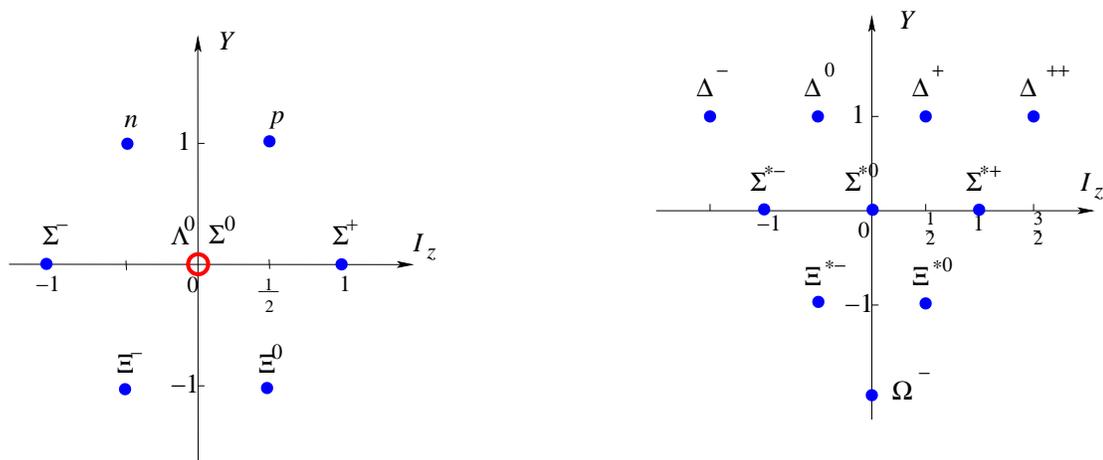


Fig. 3: L'octet ($J^P = \frac{1}{2}^+$) et le décuplet ($J^P = \frac{3}{2}^+$) de baryons.

Les matrices λ_a satisfont des relations de commutation

$$[\lambda_a, \lambda_b] = 2if_{abc}\lambda_c \quad (2.5)$$

avec les constantes de structure (réelles, complètement antisymétriques) f_{abc} de l'algèbre $\text{su}(3)$. Il est utile de considérer aussi les anticommutateurs

$$\{\lambda_a, \lambda_b\} = \frac{4}{3}\delta_{ab} + 2d_{abc}\lambda_c. \quad (2.6)$$

Grâce à (2.3), (2.5) et (2.6) 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 de (2.5), la relation (2.6) et les constantes d_{abc} (réelles, complètement symétriques) sont propres à la représentation de dimension 3.

Les hadrons s'organisent en représentations de $\text{SU}(3)$. 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 trialité nulle, telle la représentation 27, ou de celles de trialité non nulle, comme la 3 et la $\bar{3}$. On y reviendra au § 2.5.

2.2. Conséquences de la symétrie $SU(3)$

o 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, § 4.2, il est naturel de regrouper les 8 champs associés à ces particules sous forme d'une matrice de trace nulle.

$$\Phi = \begin{pmatrix} \frac{1}{\sqrt{2}}\pi^0 - \frac{1}{\sqrt{6}}\eta & \pi^+ & K^+ \\ \pi^- & -\frac{1}{\sqrt{2}}\pi^0 - \frac{1}{\sqrt{6}}\eta & K^0 \\ K^- & \bar{K}^0 & \sqrt{\frac{2}{3}}\eta \end{pmatrix}, \quad (2.7)$$

et

$$\Psi = \begin{pmatrix} \frac{1}{\sqrt{2}}\Sigma^0 - \frac{1}{\sqrt{6}}\Lambda & \Sigma^+ & p \\ \Sigma^- & -\frac{1}{\sqrt{2}}\Sigma^0 - \frac{1}{\sqrt{6}}\Lambda & n \\ \Xi^- & \Xi^0 & \sqrt{\frac{2}{3}}\Lambda \end{pmatrix}. \quad (2.8)$$

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

$$Q = I_z + \frac{1}{2}Y = \frac{1}{3} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad Y = \frac{1}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad (2.9)$$

agissent dans la représentation adjointe par commutation et on a bien

$$[Q, \Phi] = \begin{pmatrix} 0 & \pi_+ & K_+ \\ -\pi_- & 0 & 0 \\ -K_- & 0 & 0 \end{pmatrix} \quad [Y, \Phi] = \begin{pmatrix} 0 & 0 & K_+ \\ 0 & 0 & K_0 \\ -K_- & -\bar{K}_0 & 0 \end{pmatrix}.$$

¹ dite "de saveur", selon la terminologie moderne, mais appelée "symétrie unitaire" ou "voie octuple" à l'époque de Gell-Mann et Ne'eman...

Exercice : sans aucun calcul, que doit valoir $[I_z, \Phi]$? Vérifier.

o *Produits tensoriels dans $SU(3)$ et couplages invariants*

On rappelle que pour $SU(3)$, avec les notations du chapitre 3,

$$8 \otimes 8 = 1 \oplus 8 \oplus 8 \oplus 10 \oplus \overline{10} \oplus 27 . \quad (2.10)$$

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 $\overline{\Psi}\Phi\Psi$, 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 § 3.2), ce nombre est égal au nombre de fois où la représentation 8 apparaît dans $8 \otimes 8$, soit selon (2.10), 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

$$\text{tr}\overline{\Psi}\Psi\Phi = \overline{\psi}_j^i \psi_i^k \phi_k^j \quad \text{et} \quad \text{tr}\overline{\Psi}\Phi\Psi = \overline{\psi}_j^i \psi_i^k \phi_k^j \quad (2.11)$$

(cette écriture omet les indices des spineurs de Dirac, l'éventuelle matrice γ_5 etc). On préfère souvent récrire ces deux termes en termes de leurs somme et différence, donc de $\text{tr}\overline{\Psi}[\Phi, \Psi]$ et $\text{tr}\overline{\Psi}\{\Phi, \Psi\}$, appelés terme f et terme d, par référence à (2.5) et (2.6).

• Autre question de même nature : quel est *a priori* le nombre d'amplitudes invariante 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 (2.10), ce nombre est $\sum_i m_i^2 = 8$. Il y a donc huit amplitudes invariante. Autrement dit on peut écrire *a priori* l'amplitude de diffusion sous la forme

$$\langle \mathcal{N}_f \mathcal{P}_f | \mathcal{T} | \mathcal{N}_i \mathcal{P}_i \rangle = \sum_{r=1}^8 A_r(s, t) \langle (I_z, Y)_{(\mathcal{N}_f)}, (I_z, Y)_{(\mathcal{P}_f)} | r, (I_z, Y)_{(r)} \rangle \langle r, (I_z, Y)_{(r)} | (I_z, Y)_{(\mathcal{N}_i)}, (I_z, Y)_{(\mathcal{P}_i)} \rangle$$

(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 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 pour toutes les permutations P . 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 en fin de chapitre...

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

$$\mathcal{L}_{em} = -qj^\mu A_\mu \quad (2.12)$$

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 (2.2), 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).

o Moments magnétiques

Les facteurs de forme électromagnétiques de l'octet de baryons sont définis par

$$\langle B | j_\mu(x) | B' \rangle = e^{ikx} \bar{u}(F_e^{BB'}(k^2)\gamma_\mu + F_m^{BB'}(k^2)\sigma_{\mu\nu}k^\nu)u' \quad (2.13)$$

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 $B = B'$, $F_e(0) = q_B$, 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 (2.11), 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 (2.9)

$$Q = \begin{pmatrix} \frac{2}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & -\frac{1}{3} \end{pmatrix},$$

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

$$\begin{aligned} \mu(n) = \mu(\Xi^0) = 2\mu(\Lambda) = -2\mu(\Sigma^0) & \quad \mu(\Sigma^+) = \mu(p) \\ \mu(\Xi^-) = \mu(\Sigma^-) = -(\mu(p) + \mu(n)) & \quad \mu(\Sigma^0 \rightarrow \Lambda) = \frac{\sqrt{3}}{2}\mu(n), \end{aligned} \quad (2.14)$$

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" sont mesurés par leur précession de spin dans un champ magnétique ou dans des transitions dans des atomes exotiques. 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

$$\begin{aligned} \mu(p) &= 2.792847351 \pm 0.000000028\mu_N & \mu(n) &= -1.9130427 \pm 0.00000005\mu_N \\ \mu(\Lambda) &= -0.613 \pm 0.004\mu_N & |\mu(\Sigma^0 \rightarrow \Lambda)| &= 1.61 \pm 0.08\mu_N \\ \mu(\Sigma^+) &= 2.458 \pm 0.010\mu_N & \mu(\Sigma^-) &= -1.160 \pm 0.025\mu_N \\ \mu(\Xi^0) &= -1.250 \pm 0.014\mu_N & \mu(\Xi^-) &= -0.6507 \pm 0.0025\mu_N \end{aligned} \quad (2.15)$$

où μ_N est le magnéton nucléaire, $\mu_N = \frac{e\hbar}{2m_p} = 3.152 \cdot 10^{-14} \text{ MeV T}^{-1}$.

2.4. Écart 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 $G_{11} = \frac{1}{2}I_z + Y = Q$, 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 , $\text{cof}G_{33} = \frac{1}{4}Y^2 - I_z^2 - I_+I_- = \frac{1}{4}Y^2 - \tilde{I}^2$.

On obtient ainsi, pour toute représentation (tout multiplet), une formule de masse

$$M = m_1 + m_2Y + m_3(I(I+1) - \frac{1}{4}Y^2) \quad (2.16)$$

ce qui laisse trois constantes indéterminées (dépendant du multiplet). Par exemple pour l’octet de baryons, on a quatre masses expérimentales, ce qui conduit à la règle de somme

$$\frac{M_\Xi + M_N}{2} = \frac{3M_\Lambda + M_\Sigma}{4} \quad (2.17)$$

bien vérifiée expérimentalement : on trouve 1128,5 MeV/ c^2 au membre de gauche, 1136 MeV/ c^2 à celui de droite². Pour le décuplet, vérifier que cette même formule donne

² Les masses observées de ces hadrons sont $M_N \approx 939$ MeV/ c^2 , $M_\Lambda = 1116$ MeV/ c^2 , $M_\Sigma \approx 1195$ MeV/ c^2 , $M_\Xi \approx 1318$ MeV/ c^2 ; celles des mésons pseudoscalaires $m_\pi \approx 137$ MeV/ c^2 , $m_K \approx 496$ MeV/ c^2 et $m_\eta = 548$ MeV/ c^2 .

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

$$m_K^2 = \frac{3m_\eta^2 + m_\pi^2}{4}.$$

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 constituants élémentaires de tous les hadrons (connus à l'époque). Leurs charges et hypercharges sont respectivement

<i>Quarks</i>	:	u	d	s	\bar{u}	\bar{d}	\bar{s}
<i>Isospin I_z</i>	:	$\frac{1}{2}$	$-\frac{1}{2}$	0	$-\frac{1}{2}$	$\frac{1}{2}$	0
<i>Charge baryonique \mathcal{B}</i>	:	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$
<i>Étrangeté S</i>	:	0	0	-1	0	0	1
<i>Hypercharge Y</i>	:	$\frac{1}{3}$	$\frac{1}{3}$	$-\frac{2}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$
<i>Charge électrique Q</i>	:	$\frac{2}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

Table 1. Nombres quantiques des quarks u, d, s

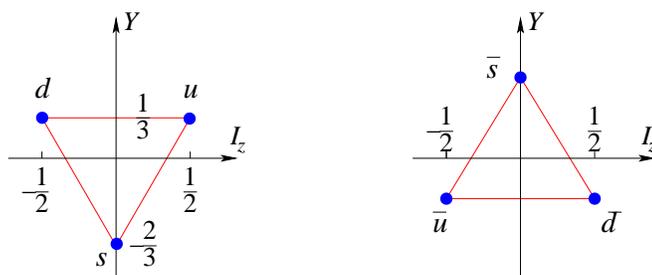


Fig. 4: Les triplets de quarks et antiquarks.

On se rappelle (chapitre 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 trialité nulle peuvent donner lieu à des particules observables. Ainsi,

$$p = uud, \quad n = udd, \quad \Omega^- = sss, \quad \Delta^{++} = uuu, \quad \dots, \quad \Delta^- = ddd,$$

$$\pi^+ = u\bar{d}, \quad \pi^0 = \frac{(u\bar{u} - d\bar{d})}{\sqrt{2}}, \quad \pi^- = d\bar{u}, \quad \eta_8 = \frac{(u\bar{u} + d\bar{d} - 2s\bar{s})}{\sqrt{6}}, \quad K^+ = u\bar{s}, \quad K^0 = d\bar{s} \quad \text{etc.}$$

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) dû aux interactions brisant $SU(3)$ de ces η_1 et η_8 . Exercice : compléter sur la figure 3 les interprétations des baryons comme états liés des quarks en s'aidant des charges et nombres quantiques.

2.6. Courants hadroniques et interactions faibles

Les interactions faibles sont phénoménologiquement bien décrites par un lagrangien effectif de la forme "courant-courant" (Fermi)

$$\mathcal{L}_{\text{Fermi}} = -\frac{G}{\sqrt{2}} J^\rho(x) J_\rho^\dagger(x) \quad (2.18)$$

où G est la constante de Fermi qui vaut (dans les unités où $\hbar = c = 1$)

$$G = (1,026 \pm 0,001) \times 10^{-5} m_p^{-2}. \quad (2.19)$$

(Mais 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

$$J_\rho(x) = l_\rho(x) + h_\rho(x) \quad (2.20)$$

Le courant leptonique

$$l_\rho(x) = \bar{\psi}_e(x) \gamma_\rho (1 - \gamma_5) \psi_{\nu_e} + \bar{\psi}_\mu(x) \gamma_\rho (1 - \gamma_5) \psi_{\nu_\mu} + \bar{\psi}_\tau(x) \gamma_\rho (1 - \gamma_5) \psi_{\nu_\tau}$$

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

$$h_\rho = \cos \theta_C h_\rho^{(\Delta S=0)} + \sin \theta_C h_\rho^{(\Delta S=1)} \quad (2.21)$$

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 suivant.) 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,

$$\begin{aligned} h_\rho^{(\Delta S=0)} &= (V_\rho^1 - iV_\rho^2) - (A_\rho^1 - iA_\rho^2) \\ h_\rho^{(\Delta S=1)} &= (V_\rho^4 - iV_\rho^5) - (A_\rho^4 - iA_\rho^5). \end{aligned} \quad (2.22)$$

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

$$V_\rho^a(x) = \bar{q}(x)\frac{\lambda^a}{2}\gamma_\rho q(x) \quad A_\rho^a(x) = \bar{q}(x)\frac{\lambda^a}{2}\gamma_\rho\gamma_5 q(x). \quad (2.23)$$

Nous les retrouverons dans le Modèle Standard.

3. De SU(3) à SU(4) et aux six saveurs

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.

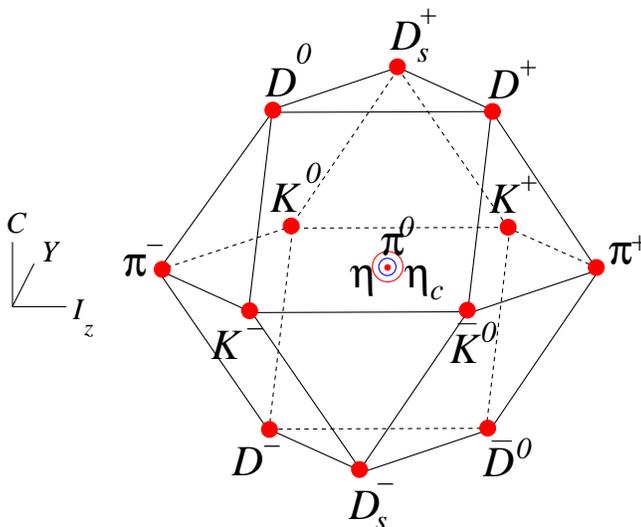


Fig. 5: Les mésons de spin $J^P = 0^-$ de la représentation 15 de $SU(4)$.

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 BaBar au SLAC, 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 quarks³

$$m_u \approx 1.5 - 4 \text{ MeV}, \quad m_d \approx 4 - 8 \text{ MeV}, \quad m_s \approx 80 - 130 \text{ MeV}$$

$$m_c \approx 1.15 - 1.35 \text{ GeV}, \quad m_b \approx 4 - 5 \text{ GeV}, \quad m_t \approx 175 \text{ GeV}$$

ce qui réduit son utilité. On peut toutefois récrire (2.2) sous la forme

$$Q = \frac{1}{2}Y + I_z \quad Y = \mathcal{B} + \mathcal{S} + \mathcal{C} + \mathcal{B} + \mathcal{T}$$

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

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.

<i>Quarks</i>	:	<i>u</i>	<i>d</i>	<i>s</i>	<i>c</i>	<i>b</i>	<i>t</i>
<i>Isospin</i> I_z	:	$\frac{1}{2}$	$-\frac{1}{2}$	0	0	0	0
<i>Charge baryonique</i> \mathcal{B}	:	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
<i>Étrangeté</i> S	:	0	0	-1	0	0	0
<i>Charme</i> C	:	0	0	0	1	0	0
<i>Beauté</i> B	:	0	0	0	0	-1	0
<i>Vérité</i> T	:	0	0	0	0	0	1
<i>Hypercharge</i> Y	:	$\frac{1}{3}$	$\frac{1}{3}$	$-\frac{2}{3}$	$\frac{4}{3}$	$-\frac{2}{3}$	$\frac{4}{3}$
<i>Charge électrique</i> 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

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 constituants 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 faisant partie 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

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/2010/tables/contents_tables.html

★

Exercices pour le chapitre 4

A. Modèle sigma et brisure de la symétrie chirale

On considère le lagrangien (1.10) 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 § 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. 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 figures 2, 3 et 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. Compléter et justifier tous les arguments esquissés aux § 2.2, 2.3 et 2.4. En particulier vérifier que la formule (2.16) conduit bien pour l'octet de baryon à la règle (2.17), et pour le décuplet, à des écarts de masse constants.

D. 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)

$$\sum_P \text{tr}(A_{P_1} A_{P_2} A_{P_3} A_{P_4}) = a \sum_P \text{tr}(A_{P_1} A_{P_2}) \text{tr}(A_{P_3} A_{P_4}) \quad (3.1)$$

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 quadrilinéaires en A_1, \dots, A_4 , invariants par l'action de $SU(3)$ $A_i \rightarrow U A_i U^\dagger$ et linéairement indépendants peut-on écrire ? Pourquoi l'identité (3.1) 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 (2.7).

- a. Pourquoi est-il naturel d'imposer que ce lagrangien soit pair dans le champ Φ ?
- b. En utilisant les résultats du 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 SU(3) ?
- d. En déduire que tout lagrangien de degré 4 en Φ invariant par SU(3) est en fait invariant par ce groupe G .

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.

1. Invariance de jauge. Couplage minimal. Lagrangien de Yang–Mills

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

$$\mathcal{L} = \bar{\psi}(i\cancel{\partial} - e\cancel{A} - m)\psi - \frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) \quad (1.1)$$

est invariant sous l'effet de transformations de jauge infinitésimale

$$\begin{aligned} \delta A_\mu(x) &= -\partial_\mu \delta\alpha(x) \\ \delta\psi(x) &= ie\delta\alpha(x)\psi(x) , \end{aligned} \quad (1.2)$$

puisque le tenseur de champ

$$F_{\mu\nu} = (\partial_\mu A_\nu - \partial_\nu A_\mu)$$

est invariant, et que la combinaison

$$i\cancel{D}\psi(x) := (i\cancel{\partial} - e\cancel{A})\psi(x)$$

se transforme comme ψ . La forme finie de ces transformations est aussi aisée à écrire

$$\begin{aligned} A_\mu(x) &\mapsto A_\mu(x) - \partial_\mu \alpha(x) \\ \psi(x) &\mapsto e^{ie\alpha(x)}\psi(x) , \end{aligned} \quad (1.3)$$

ce qui montre bien que les transformations sont celles d'une version locale (dépendant de x) du groupe $U(1)$ ou \mathbb{R} (selon qu'on identifie ou non α et $\alpha + 2\pi$). 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 électromagnétique. 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

$$\delta\mathcal{L} = (\partial_\mu - iqA_\mu)\phi^*(\partial^\mu + iqA_\mu)\phi - V(\phi^*\phi) \quad (1.4)$$

qui est bien invariant sous $\phi(x) \mapsto e^{iq\alpha(x)}\phi(x)$, $A_\mu(x) \mapsto A_\mu(x) - \partial_\mu\alpha(x)$.

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

$$\delta\psi(x) = T_a\delta\alpha^a\psi(x) . \quad (1.5)$$

(Dans ce chapitre, 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 , et porte donc des indices de la représentation adjointe

$$A_\mu(x) = \{A_\mu^a(x)\} \quad (1.6)$$

ou encore est représenté dans toute représentation par la matrice antihermitique¹

$$A_\mu(x) = T_a A_\mu^a(x) . \quad (1.7)$$

¹ Attention, cette convention implique que certaines expressions diffèrent par un facteur i du cas abélien

La dérivée covariante s'écrit

$$D_\mu \psi(x) := (\partial_\mu - A_\mu(x))\psi(x) , \quad (1.8)$$

ou encore, en composantes

$$D_\mu \psi_A(x) := (\partial_\mu \delta_{AB} - A_\mu^a(x) (T_a)_A^B) \psi_B(x) . \quad (1.9)$$

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

$$\delta A_\mu^a(x) = \partial_\mu \delta \alpha^a(x) + C_{bc}^a \delta \alpha^b(x) A_\mu^c(x) . \quad (1.10)$$

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 $(t_a)_b^c = -C_{ab}^c$). 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

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - [A_\mu, A_\nu] \quad (1.11)$$

ou en composantes

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - C_{bc}^a A_\mu^b A_\nu^c . \quad (1.12)$$

On démontre, au prix d'un peu d'algèbre, et en utilisant l'identité de Jacobi, que

$$\delta F_{\mu\nu}^a(x) = C_{bc}^a \delta \alpha^b(x) F_{\mu\nu}^c(x) , \quad (1.13)$$

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 ,

$$\begin{aligned} \psi(x) &\mapsto \mathcal{D}(g(x))\psi(x) \\ A_\mu = A_\mu^a T_a &\mapsto \mathcal{D}(g(x))(-\partial_\mu + A_\mu(x))\mathcal{D}(g^{-1}(x)) , \end{aligned} \quad (1.14)$$

et pour la dérivée covariante agissant sur ψ ,

$$D_\mu \psi(x) \mapsto \mathcal{D}(g(x))D_\mu \psi(x) \quad (1.15)$$

ou encore²

$$D_\mu \mapsto \mathcal{D}(g(x))D_\mu\mathcal{D}(g^{-1}(x)) . \quad (1.16)$$

Or on vérifie aisément que dans une représentation donnée

$$[D_\mu, D_\nu] = -F_{\mu\nu} := -F_{\mu\nu}^a T_a \quad (1.17)$$

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

$$F_{\mu\nu}(x) \mapsto g(x)F_{\mu\nu}(x)g^{-1}(x) , \quad (1.18)$$

dont (1.13) est la version infinitésimale.

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

$$F_{\mu\nu} = 0 \Leftrightarrow A_\mu(x) = (\partial_\mu g(x))g^{-1}(x) . \quad (1.19)$$

(Le \Leftarrow se démontre au prix d'une ligne de calcul, pour le \Rightarrow , voir huit lignes plus bas...)

Insistons sur le caractère local de cette propriété.

Un autre objet intéressant est l'élément du groupe attaché à une courbe C allant de x_0 à x

$$g(C) := P \exp \left(\int_C dx^\mu A_\mu(x) \right) \quad (1.20)$$

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 (1.14)

$$g(C) \mapsto g(x)g(C)g^{-1}(x_0) . \quad (1.21)$$

On montre aussi que si $F = 0$, $g(C)$ ne dépend que des extrémités x_0 et x de C . Le $g(x, x_0) := g(C)$ qui en résulte satisfait $(\partial_\mu - A_\mu)g(x, x_0) = 0$, soit (1.19).

Boucle de Wilson. Considérons le cas où $x = x_0$. La boucle C est fermée et $g(C)$ se transforme de façon covariante, $g(C) \mapsto g(x_0)g(C)g^{-1}(x_0)$. Sa trace

$$W(C) = \text{tr} g(C) = \text{tr} P \exp \oint dx^\mu A_\mu(x) \quad (1.22)$$

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 $\text{tr} F_{\mu\nu} F^{\mu\nu}$, $\bar{\psi} i \not{D} \psi$ 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 § 3.1 et voir le Problème à la fin de ce chapitre pour une version discrétisée de cette quantité.

² 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 (1.14), elle n'agit que sur $\mathcal{D}(g^{-1}(x))$.

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 champ 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 (1.17). 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 (1.14). 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.

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

$$\mathcal{L} = \frac{1}{2g^2} \text{tr}(F_{\mu\nu}F^{\mu\nu}) + \bar{\psi}(i(\not{\partial} - A) - m)\psi, \quad (1.23)$$

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 1 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 et 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, \mathcal{L} de (1.23) ressemble beaucoup au lagrangien du cas abélien (1.1), 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 de type universel, via la dérivée covariante ;
- 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 à introduire des mécanismes “doux” de brisure de l’invariance de jauge (brisure spontanée) pour y remédier ;

- contrairement au cas abélien, le champ de jauge porte une charge du groupe : on a vu que pour les transformations globales (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.

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 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 *systèmes contraints*, (voir par exemple [IZ, chap. 12]), conduit à l’action suivante

$$S_{FP} = \int d^4x \mathcal{L}_{FP}, \quad \mathcal{L}_{FP} = -\frac{1}{4g^2} F_{\mu\nu a} F^{\mu\nu a} - \frac{1}{2\alpha} (\partial_\mu A_a^\mu)^2 + \partial \bar{\eta}_a (D^\mu \eta)_a + \mathcal{L}_{\text{matter}} \quad (1.24)$$

Les champs auxiliaires η et $\bar{\eta}$, dits “fantômes de Faddeev–Popov”, dotés d’un spin nul mais anticommuteurs (théorème spin–statistique !?), ont pour seul but d’exprimer le “déterminant de Faddeev–Popov” $\det(-\partial_\mu D^\mu)$, où D est la dérivée covariante dans la représentation adjointe. Le “terme de jauge” $-\frac{1}{2\alpha} (\partial_\mu A_a^\mu)^2$ (avec α un paramètre arbitraire) a pour effet de briser l’invariance de jauge de l’action et de rendre la forme quadratique en A inversible. La partie du lagrangien impliquant les champs de matière, tel celui du champ de fermion $\bar{\psi}(i(\not{\partial} - \not{A}) - m)\psi$, n’a pas été explicitée. Il est commode d’effectuer alors un changement de variable $A_\mu \rightarrow gA_\mu$ qui fait apparaître plus explicitement la constante de couplage devant les termes de degré ≥ 3 dans les champs.

Cette méthode suppose implicitement qu’une condition telle $\partial_\mu A^\mu = 0$ “fixe complètement la jauge”, c’est-à-dire détermine de façon unique le g tel que A_μ^g soit solution. Si cela est justifié pour des A au voisinage de 0, donc dans le cadre de la théorie des perturbations, cela ne l’est pas de façon globale. C’est le *phénomène de Gribov*, qui reste toujours l’objet de travaux.

Au final, on lit sur cette action les règles de Feynman suivantes

$$\begin{array}{c} a \\ \lambda \end{array} \text{---} \begin{array}{c} q \\ \mu \end{array} \text{---} \begin{array}{c} b \\ \mu \end{array} = -i\delta_{ab} \left(\frac{g_{\mu\nu}}{q^2+i\epsilon} - (1-\alpha) \frac{q_\mu q_\nu}{(q^2+i\epsilon)^2} \right) \quad \begin{array}{c} A \\ \alpha \end{array} \xrightarrow{p} \begin{array}{c} B \\ \beta \end{array} = \delta_{AB} \left(\frac{i}{p-m} \right)$$

$$\begin{array}{c} c \\ \nu \end{array} \begin{array}{c} a \\ \lambda \end{array} \text{---} \begin{array}{c} p \\ \mu \end{array} \text{---} \begin{array}{c} q \\ \mu \end{array} \text{---} \begin{array}{c} b \\ \mu \end{array} = g C_{abc} (2\pi)^d \delta^d(p+q+r) (g_{\lambda\mu}(p-q)_\nu + g_{\mu\nu}(q-r)_\lambda + g_{\nu\lambda}(r-p)_\nu)$$

$$\begin{array}{c} p \\ a \end{array} \begin{array}{c} \lambda \end{array} \text{---} \begin{array}{c} p \\ \mu \end{array} \text{---} \begin{array}{c} p \\ \nu \end{array} \text{---} \begin{array}{c} p \\ d \end{array} \begin{array}{c} s \end{array} \\ \begin{array}{c} b \\ q \end{array} \begin{array}{c} \mu \end{array} \text{---} \begin{array}{c} c \\ v \end{array} \begin{array}{c} r \end{array} \end{array} = -ig^2 (2\pi)^d \delta^d(p+q+r+s)$$

$$(C_{abp}C_{cdp}(g_{\lambda\nu}g_{\rho\mu} - g_{\lambda\rho}g_{\nu\mu}) + C_{acp}C_{bdp}(g_{\lambda\rho}g_{\nu\mu} - g_{\lambda\mu}g_{\nu\rho}) + C_{adp}C_{bcp}(g_{\lambda\mu}g_{\rho\nu} - g_{\lambda\nu}g_{\rho\mu}))$$

$$\begin{array}{c} B \\ \beta \end{array} \begin{array}{c} r \end{array} \\ \begin{array}{c} a \\ \mu \end{array} \text{---} \begin{array}{c} p \\ \mu \end{array} \text{---} \begin{array}{c} q \\ \mu \end{array} \text{---} \begin{array}{c} A \\ \alpha \end{array} \end{array} = g(2\pi)^d \delta^d(p+q-r) (\gamma_\mu)_{\alpha\beta} T_{AB}^a$$

et des règles analogues pour les champs de fantômes.

Illustrons cela sur quelques diagrammes à une boucle

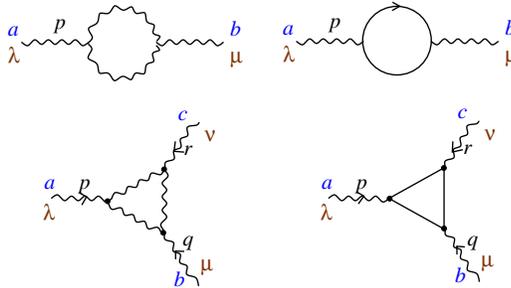


Fig. 1: Quelques diagrammes à une boucle dans une théorie de jauge

La contribution du premier diagramme à la fonction à deux points du champ de jauge est

$$\frac{g^2}{2} \int \frac{d^d k}{(2\pi)^d} \frac{1}{k^2(q-k)^2} C_{acd}C_{bcd}(\dots)$$

où \dots représente un polynôme de degré 2 dans les impulsions p, q, k . Le “coefficient groupiste”, qui ne dépend que du groupe de jauge et de sa représentation adjointe, est égal à

$$C_{acd}C_{bcd} = C_2 \delta_{ab}$$

($C_2 = c_2(\text{adj})$ avec les notations de l’exercice A.1, et $C_2 = N$ pour $SU(N)$, cf exercice A.2). La contribution du second diagramme est, en ce qui concerne le coefficient groupiste,

proportionnelle à $\text{tr}T_a T_b$. On montre (voir Exercice 1 pour le calcul de ces coefficients) que $\text{tr}T_a T_b = T_f \delta_{ab}$ où T_f est proportionnel à la valeur du Casimir quadratique dans la représentation portée par les fermions.

1.6. Renormalisabilité

On démontre, et cela a été une étape décisive dans la construction du Modèle Standard³, 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. Ainsi à l’ordre à une boucle, les diagrammes de la figure 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

$$g \mapsto g_0 = \left(1 - \frac{g^2}{(4\pi)^2} \left(\frac{11}{3} C_2 - \frac{4}{3} T_f \right) \log \frac{\Lambda}{\mu} \right) g, \quad (1.25)$$

où Λ est une échelle de coupure ultraviolette (“cutoff”) et μ une échelle de masse indispensable à la définition de la procédure de renormalisation.

2. Champs de jauge massifs

2.1. Interactions faibles et bosons intermédiaires

On a vu au chapitre 3 (équ. (2.17)) que le lagrangien de Fermi

$$\mathcal{L}_{\text{Fermi}} = -\frac{G}{\sqrt{2}} J^\rho(x) J_\rho^\dagger(x) \quad (2.1)$$

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.

³ G. ’t Hooft et M. Veltman, prix Nobel 1999

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 $\sqrt{s} \sim G^{-\frac{1}{2}} \sim 300 \text{ GeV}$. Et la non-renormalisabilité de la théorie empêche d'améliorer ce terme de Born.

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 W de masse M , dans la limite de grande masse $M^\#$. Considérons le nouveau lagrangien

$$\mathcal{L}_{\text{boson int.}} = gJ^\rho(x)W_\rho^\dagger(x) + \text{h.c.} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + M^2W_\rho^\dagger W^\rho. \quad (2.2)$$

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

$$\frac{G}{\sqrt{2}} = \frac{g^2}{M^2} \quad (2.3)$$

reliant le nouveau couplage g au couplage de Fermi. La théorie (2.2) avec son "boson intermédiaire" W , vecteur des interactions faibles, est-elle la bonne théorie des interactions faibles ? En fait le propagateur du champ W est

$$-i \frac{g_{\mu\nu} - \frac{k_\mu k_\nu}{M^2}}{k^2 - M^2} \quad (2.4)$$

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.

2.2. Brisure spontanée de symétrie de jauge. Mécanisme de Brout–Englert–Higgs

Revenons au cas abélien décrit par (1.1)-(1.4) et supposons maintenant que le potentiel V soit tel que son minimum soit 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

$$\phi(x) = e^{iq\theta(x)/v} \frac{v + \varphi(x)}{\sqrt{2}} \quad (2.5)$$

[#] 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).

avec v réel et φ hermitien, et accompagnant cela d'une transformation de jauge

$$\begin{aligned}\phi(x) &\mapsto \phi'(x) = e^{-iq\theta(x)/v} \phi(x) = \frac{v + \varphi(x)}{\sqrt{2}} \\ A_\mu(x) &\mapsto A'_\mu(x) = A_\mu(x) + \frac{1}{v} \partial_\mu \theta(x)\end{aligned}$$

et de la transformation correspondante pour les éventuels autres champs chargés ($\psi \dots$), on voit que le lagrangien $\delta\mathcal{L}$ de (1.4) s'écrit

$$\begin{aligned}\delta\mathcal{L} &= (\partial_\mu - iqA'_\mu)\phi'(\partial^\mu + iqA'_\mu)\phi' - V(\phi'^2) \\ &= \frac{1}{2}|(\partial_\mu - iqA'_\mu)\phi|^2 + \frac{1}{2}q^2v^2 A'_\mu A'^\mu - V\left(\frac{1}{2}(v + \varphi)^2\right).\end{aligned}$$

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–Higgs, 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 .

Ce mécanisme de Brout–Englert–Higgs 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 § 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 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.

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 : $U(1) \times SU(2) \times SU(3)$, 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 !

3.1. Le secteur fort

Le groupe $SU(3)$ est celui de la couleur (cf chap. 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 (1.23), avec des masses fermioniques dépendant de la saveur, engendrées par le secteur faible.

o Liberté asymptotique

Connaissant les règles de Feynman et la renormalisabilité de la théorie, cf § 1.6, on peut calculer la renormalisation de la constante de couplage g , (1.25), et la fonction beta correspondante. On trouve⁴

$$\beta(g) = -\Lambda \frac{\partial}{\partial \Lambda} g(\Lambda)|_{g_0} = -\frac{g^3}{(4\pi)^2} \left(\frac{11}{3} C_2 - \frac{4}{3} T_f \right) + O(g^5) \quad (3.1)$$

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 : $\frac{dg^2(\lambda)}{d \log \lambda} < 0 \Rightarrow g^2(\lambda) \sim (2b \log \lambda)^{-1} \rightarrow 0$ quand $\lambda \rightarrow \infty$, avec $b =$ coefficient du terme $-g^3$ dans (3.1). 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 ?

⁴ David J. Gross, H. David Politzer, Frank Wilczek, prix Nobel 2004

Cette théorie de jauge non abélienne est la seule théorie des champs locale et renormalisable à 4 dimensions 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 la structure interne de ces derniers comme faite de constituants quasi-libres à très courte distance.

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, § 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, 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 (§ 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 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 omni-présentes et l'observation de toute autre interaction, de tout autre effet, présuppose une connaissance aussi précise que possible de la contribution forte. Alors que le LHC commence à fournir des données, les calculs des contributions de QCD gardent une importance fondamentale.

3.2. Le secteur électro-faible, une esquisse.

Le groupe $U(1) \times SU(2)$ est celui qui décrit les interactions électro-faibles (modèle de Glashow–Salam–Weinberg⁵). 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

⁵ S. Glashow, A. Salam, S. Weinberg, prix Nobel 1979

les 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

$$D_\mu\psi = (\partial_\mu - g_3 A_\mu^a T_a - g_2 W_\mu^j t_j - i \frac{g_1}{2} y B_\mu) \psi \quad (3.2)$$

où 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 :

<i>Quarks & Leptons</i>	:	(ν_L^e, e_L)	ν_R^e	e_R	(u_L, d_L)	u_R	d_R
<i>Isospin faible t_z</i>	:	$(\frac{1}{2}, -\frac{1}{2})$	0	0	$(\frac{1}{2}, -\frac{1}{2})$	0	0
<i>Hypercharge faible y</i>	:	$(-1, -1)$	0	-2	$(\frac{1}{3}, \frac{1}{3})$	$\frac{4}{3}$	$-\frac{2}{3}$
<i>Charge électrique $Q = \frac{1}{2}y + t_z$</i>	:	$(0, -1)$	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.

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_3^μ , 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)⁶ a été la première confirmation de la validité du Modèle Standard.

Le groupe U(1)_{em} de l'électromagnétisme est 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 $A_\mu^{em} = \cos \theta_W B_\mu + \sin \theta_W W_\mu^0$, 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^0 et B . On les lit sur les dérivées covariantes (3.2) avec les nombres quantiques de la Table 1

$$\frac{1}{2} i \left[\bar{e}_L (-g_2 W^0 - g_1 B) e_L + \bar{\nu}_R (-2g_1 B) e_R + \bar{\nu}_e (g_2 W^0 - g_1 B) \nu_e \right]$$

⁶ En lire l'histoire dans <http://cerncourier.com/cws/article/cern/29168>

La rotation $W^0 = \cos \theta_W Z^0 + \sin \theta_W A$, $B = -\sin \theta_W Z^0 + \cos \theta_W A$ 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

$$2e = g_2 \sin \theta_W + g_1 \cos \theta_W = 2g_1 \cos \theta_W \quad \text{et} \quad g_2 \sin \theta_W - g_1 \cos \theta_W = 0$$

qui sont bien compatibles et fournissent

$$\tan \theta_W = \frac{g_1}{g_2} \quad e = g_1 \cos \theta_W = g_2 \sin \theta_W . \quad (3.3)$$

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 $y = +1$), $|D\Phi|^2$ où $D_\mu \Phi = (\partial_\mu - g_2 W_\mu^i \frac{\tau_i}{2} - \frac{1}{2} g_1 B_\mu) \Phi$. Le champ Φ est doté d'un potentiel $V(\Phi)$ en "chapeau mexicain", responsable de la brisure spontanée de $U(1) \times SU(2)$ en $U(1)_{em}$, et donc de l'apparition des masses des champs vectoriels selon le mécanisme décrit au § 2, et même de celle des fermions. Ce champ dit de Higgs (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 cette composante φ (et la "particule de Higgs" qu'elle crée) que l'on espère découvrir prochainement dans les expériences 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 $U(1) \times SU(2)$ 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 (2.5) au groupe $SU(2)$ de générateurs $i\frac{\tau^j}{2}$ ($\tau^j =$ matrices de Pauli)

$$\Phi(x) = e^{i\xi_j(x) \frac{\tau^j}{2}} \begin{pmatrix} 0 \\ \frac{v + \varphi(x)}{\sqrt{2}} \end{pmatrix} ,$$

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

$$\boxed{M_{W^\pm} = \frac{1}{2} v g_2 \quad M_{Z^0} = \frac{1}{2} v \sqrt{g_1^2 + g_2^2}} \quad (3.4)$$

et en utilisant (3.3), la relation $G = \frac{g_2^2}{8M_W^2}$ qu'on lit sur le lagrangien et la valeur expérimentale de e et de $G = 10^{-5}m_p^2$

$$M_{W^\pm} = \frac{38}{\sin \theta_W} \text{GeV} \quad M_{Z^0} = \frac{M_W}{\cos \theta_W} = \frac{38}{\sin \theta_W \cos \theta_W} \text{GeV} .$$

Ces expressions subissent ensuite de petites corrections perturbatives. Enfin la masse du fameux et tant attendu boson de Higgs φ n'est pas prédite par la théorie mais devrait être dans l'intervalle 100–200 GeV.

Les “bosons intermédiaires” associés aux champs vectoriels massifs W^\pm et Z^0 ont été découverts expérimentalement dès la fin des années 70⁷; ils ont des masses $M_{W^\pm} = 80.4 \text{ GeV}$ et $M_{Z^0} = 91.2 \text{ GeV}$ compatibles avec une valeur de l'angle de Weinberg donnée par

$$\sin^2 \theta_W \approx 0.23 , \quad (3.5)$$

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

$$\mathcal{L} = -\frac{1}{4} \mathbf{F}_{\mu\nu} \mathbf{F}^{\mu\nu} + \sum_{\substack{\text{quarks \& leptons} \\ \text{gauches et droits}}} \bar{\psi} \gamma^\mu D_\mu \psi + |D\Phi|^2 - V(\Phi) + \text{couplages fermions - Higgs} , \quad (3.6)$$

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),

$$\mathcal{L}_Y = -Y_{ij}^d \bar{\psi}_{Li} \cdot \Phi d_{Rj} - Y_{ij}^u \bar{\psi}_{Li} \cdot \tilde{\Phi}^\dagger u_{Rj} + \text{h.c.} , \quad (3.7)$$

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

$$\psi_{Li} = \begin{pmatrix} u_i \\ d_i \end{pmatrix}_L = \left(\begin{pmatrix} u \\ d \end{pmatrix}_L , \begin{pmatrix} c \\ s \end{pmatrix}_L , \begin{pmatrix} t \\ b \end{pmatrix}_L \right) .$$

⁷ Carlo Rubbia et Simon van der Meer, prix Nobel 1984

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 (3.6) 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 (3.6) : si (u_L, c_L, t_L) et (d_L, s_L, b_L) désignent les états propres de masse, le courant hadronique chargé couplé au champ W^+ est

$$J_\mu = (\overline{u\bar{c}t})_L \gamma_\mu M \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L \quad (3.8)$$

avec M la *matrice unitaire de Cabibbo-Kobayashi-Maskawa*⁸. 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

$$M = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}$$

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. La mesure de ses éléments de matrice 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 $e^{i\delta}$).

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...

4. Compléments

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 (3.6), 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

⁸ M. Kobayashi, T. Maskawa, prix Nobel 2008, avec Y. Nambu

certaines termes doivent être ajustés de façon extrêmement fine ; la question du boson de Higgs dont la découverte est espérée au LHC, 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, et prédisent aussi l'existence de partenaires supersymétriques pour toutes les particules connues. Pour confirmation, rendez-vous peut-être dans quelques mois au LHC...

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 10^{16} GeV a été une forte incitation dans le sens d'une grande unification, voir fig. 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.

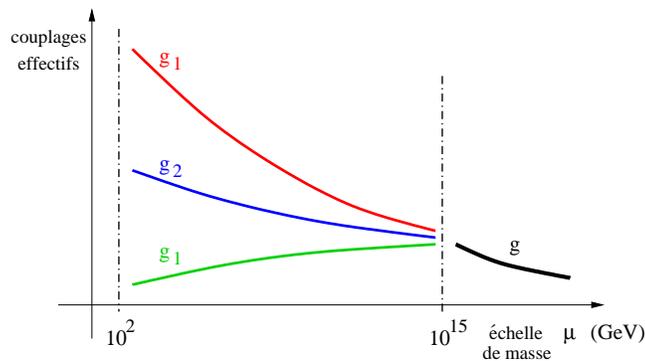


Fig. 2: Évolutions schématisées des 3 couplages effectifs du modèle standard et de celui de la théorie grand-unifiée

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 : 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 $\bar{5}$ (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 électro-faible, 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. 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 ;
- elle place dans une même représentation leptons et quarks et explique donc la commensurabilité de leurs charges électriques et la compensation des anomalies (voir § suivant) ;
- elle réduit le nombre de paramètres du modèle standard et prédit la valeur de l'angle de Weinberg (à l'échelle d'unification) ;
- à l'inverse elle n'explique pas le pourquoi des trois générations observées ;
- elle n'élucide 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.

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

$$\mathcal{L} = i\bar{\psi}(\not{\partial} - \not{A}) \frac{(1 - \gamma_5)}{2} \psi$$

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 A.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, $T_a^* = C T_a C^{-1}$. 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 A.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

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.

Pour une revue historique de ces développements, voir *Broken Symmetries*, note du Comité Nobel, 2008,

http://nobelprize.org/nobel_prizes/physics/laureates/2008/phyadv08.pdf

Pour une revue détaillée du Modèle Standard et une compilations 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.**A. Facteurs groupistes...**

1. *Opérateurs de Casimir* (cet exercice aurait dû se trouver au chapitre 1 !)

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

$$C_2^{(R)} = c_2(R) \mathbb{I} .$$

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 (3.10) du chap. 3 (modifiée par un facteur 2 !) $c_2(\Lambda) = \frac{1}{2} \langle \Lambda, \Lambda + 2\rho \rangle$ 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 (3.8) et (3.19) 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 A 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

$$(T_\alpha)^* = UT_\alpha U^{-1} .$$

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 ?

B. 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 $\vec{\Phi}$ 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. 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 :

$$\int_G \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g.g_1.g^{-1}.g_2) = \frac{1}{n_\rho} \chi^{(\rho)}(g_1) \chi^{(\rho)}(g_2) , \quad (4.1)$$

et

$$\int_G \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g.g_1) \chi^{(\sigma)}(g^{-1}.g_2) = \frac{\delta_{\rho,\sigma}}{n_\rho} \chi^{(\rho)}(g_1.g_2) . \quad (4.2)$$

Rappeler pourquoi une représentation de G peut toujours être considérée comme unitaire et montrer qu'alors

$$\chi^{(\rho)}(g^{-1}) = \chi^{(\bar{\rho})}(g) = (\chi^{(\rho)}(g))^* , \quad (4.3)$$

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 (4.3), montrer que les fonctions $b_\rho(\beta)$ sont réelles, $b_\rho(\beta) = (b_\rho(\beta))^* = b_{\bar{\rho}}(\beta)$.

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 $\text{SU}(3)$? 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 $-\ell = \vec{ji}$, 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

$$E = - \sum_{\text{plaquettes } p} \chi(g_p) \quad (4.4)$$

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

$$Z = \prod_{\text{liens } \ell} \int_G \frac{d\mu(g_\ell)}{v(G)} \prod_{\text{plaquettes}} e^{\beta \chi(g_p)} . \quad (4.5)$$

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 § 1. Les degrés de liberté g_{ij} représentent les variables de chemin définies en (1.20), $g_{ij} = g(j, i)$ 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 (4.4) 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 de la question 3 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 figure 1), 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

$$W^{(\sigma)}(C) := \langle \chi^{(\sigma)}(g_C) \rangle = \frac{1}{Z} \prod_{\text{liens } \ell} \int_G \frac{d\mu(g_\ell)}{v(G)} \chi^{(\sigma)} \left(\prod_{\ell \in C} g_\ell \right) \prod_p e^{\beta \chi(g_p)}. \quad (4.6)$$

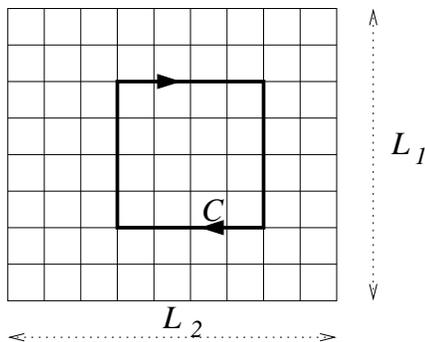


Figure 1.

d) Montrer en utilisant les résultats de la question 2 qu'on peut développer chaque

$\exp \beta \chi(g_p)$ sur les caractères des représentations irréductibles de G selon

$$e^{\beta \chi(g_p)} = \sum_{\rho} n_{\rho} b_{\rho} \chi^{(\rho)}(g_p) . \tag{4.7}$$

e) On insère dans (4.5) ou (4.6) le développement (4.7) 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

$$Z = b_1^{\mathcal{N}} \quad W^{(\sigma)}(C) = n_{\sigma} \left(\frac{b_{\sigma}}{b_1} \right)^A \tag{4.8}$$

où A désigne l'aire de la courbe C , c'est-à-dire le nombre de plaquettes qu'elle enferme, et l'indice 1 se rapporte à la représentation identité.

f) 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 (4.4), 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 .

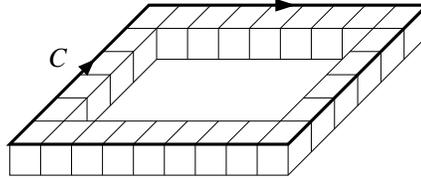


Figure 2.

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 (figure 2).

– Montrer en effet que pour une telle configuration l'application répétée des formules (4.1) et (4.2) sur toutes les variables g_{ℓ} conduit à l'expression suivante

$$W^{(\sigma)}(C) \Big|_{\text{tube}} = \sum_{\rho} \left(\frac{b_{\rho}}{b_1} \right)^P \int_G \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g) \chi^{(\rho)}(g^{-1}) \chi^{(\sigma)}(g) \tag{4.9}$$

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 (4.9) 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 $SU(2)$) 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 (4.9) si la condition \mathcal{C} peut être satisfaite, et du type (4.8) 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 (4.8), soit de la contribution à (4.9) 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 dimensions, groupe fini) 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 . . .