

Des Groupes de Lie et de leurs applications en Physique Fondamentale

Mémoire M1 - Université Paul Sabatier

Marc Briant, Lucas Chesnel, Baptiste Schubnel

Juin 2008

Table des matières

1	Introduction	4
2	Définitions préliminaires	5
2.1	Représentation des groupes compacts	5
2.2	Caractères et quasi-caractères	7
2.3	Coefficients matriciels d'un groupe	9
3	Une notion importante : la mesure de Haar	11
3.1	Un détour par la topologie	12
3.1.1	La topologie faible	13
3.1.2	La topologie faible-*	13
3.2	Un nouveau théorème du point fixe	15
3.3	L'existence de mesures de Haar	16
3.3.1	Vers la mesure fixe	16
3.3.2	Vers la mesure de Haar	18
3.3.3	Remarques importantes	20
3.4	Application aux caractères	22
4	Vers les groupes de Lie	24
4.1	Incursion dans la topologie différentielle	24
4.2	Groupes de Lie	26
4.2.1	Définition et premières propriétés	26
4.2.2	L'espace tangent et mesure de Haar	27
4.2.3	Les sous-groupes de Lie de $GL_n(\mathbb{C})$	28
5	Vers les algèbres de Lie	32
5.1	Les algèbres de Lie	32
5.2	Du groupe de Lie à l'algèbre de Lie	33
5.2.1	A partir des sous-groupes de Lie de $GL_n(\mathbb{C})$	34
5.2.2	Le cas général	36
6	Etude des algèbres de Lie semi-simples	37
6.1	Tore maximal et décomposition en système de racines	37
6.2	Propriétés de la décomposition en système de racines	40
7	Etude des systèmes de racines	42
7.1	Définition d'un système de racines	42
7.2	Exemples	43

TABLE DES MATIÈRES

7.3	Paires de racines	44
7.4	Racines simples et groupe de Weyl	45
7.5	Bases et chambres de Weyl	45
7.6	Quelques résultats sur les racines simples	50
7.7	Le groupe de Weyl	51
7.8	Systèmes de racines irréductibles	52
8	Classification des systèmes de racines	55
8.1	Matrice de Cartan de Φ	55
8.2	Graphes de Coxeter et diagrammes de Dynkin	56
8.3	Composantes irréductibles	58
8.4	Théorème de classification	58
9	Des symétries et des groupes : pourquoi et comment utiliser des groupes de Lie en physique ?	61
9.1	Comment décrire une transformation continue en physique ?	61
9.2	Etude de $SU(2)$	64
9.2.1	Représentations irréductibles de $\mathfrak{sl}(2, \mathbb{C})$	65
9.2.2	Représentations irréductibles de $SU(2)$	67
9.3	$SU(2)$ et le groupe d'isospin : une immersion dans les propriétés de l'interaction nucléaire forte	71
9.3.1	$SU(2)$ et l'interaction forte	71
9.3.2	Décomposition de Clebsch-Gordan	75
9.3.3	Hypothèse d'invariance de l'interaction forte par le groupe d'isospin	76
9.3.4	Description des états à plusieurs particule : prévoir la probabilité de certaines réactions nucléaires	78
9.4	Représentations de $SU(3)$	83
9.4.1	Quelques notions mathématiques supplémentaires	83
9.4.2	Etude des racines de $\mathfrak{sl}(3, \mathbb{C})$	84
9.4.3	Représentation 8	85
9.4.4	Représentations 3 et $\bar{\mathbf{3}}$	86
9.4.5	Représentation 10	88
9.5	$SU(3)$ et les quarks	90
9.6	Pour aller plus loin	97
	Conclusion	99
	Bibliographie	100

1 Introduction

La théorie de la physique des particules fondamentales, comme les neutrons, protons et quarks, se base désormais sur l'étude des symétries générales que présentent ces éléments à la base de la matière. Cette recherche de propriétés de symétrie a entraîné depuis quelques temps le concept de spin des particules. Cependant, il semblerait que des particules élémentaires de même famille soient bien plus liées entre elles que par un simple spin de même valeur. Elles seraient au contraire issues d'un même groupe qu'elles représenteraient chacune à leur manière. Ces structures qu'elles décrivent sont les groupes et algèbres de Lie.

Nous allons donc, afin de pouvoir approcher ces nouvelles théories de physique fondamentale, faire une étude des groupes de Lie de manière purement mathématique.

C'est ainsi que dans un premier temps nous allons développer les notions et théorèmes généraux qui seront à la base de la théorie des groupes de Lie. Puis nous irons étudier ces groupes si spécifiques et comprendre ce qui en fait des structures si fortes. Une telle étude nous amènera naturellement à l'étude des algèbres de Lie puis aux algèbres de Lie semi-simples, qui nous seront si utiles en physique. Nous verrons alors que nous pouvons classifier ces algèbres grâce à des représentations, des systèmes, de racines. Leur étude mathématique est importante car physiquement nous apprendrons que chaque particule est une racine d'une algèbre. Enfin, nous pourrons alors nous plonger dans le monde des particules élémentaires et tisser les liens forts qui existent entre celles-ci et les structures développées en amont.

Il est important de noter que la théorie des groupes de Lie est un domaine très vaste qui recouvre de l'algèbre, aussi bien que du calcul différentiel ou de la topologie. Cependant, notre but est réellement de comprendre ce lien entre physique et mathématiques. Dès lors, nos chapitres s'articuleront autour de cette idée et nous ne démontrerons et ne parlerons que de ce qui nous servira en physique. Ainsi, un seul type d'algèbre de Lie sera vue en détail et nos exemples développeront les groupes nécessaires à la compréhension de notre dernière partie.

2 Définitions préliminaires

Toute l'étude des Groupes de Lie se fonde sur des notions générales de groupe. Commençons donc par apporter quelques définitions et comprendre leurs fonctionnalités algébriques.

2.1 Représentation des groupes compacts

Nous nous placerons, sauf mention explicite du contraire, dans des groupes compacts. Le but est d'arriver à les classer, à leur donner une représentation qui permet de les comprendre au travers de leur classification. Les groupes et algèbres de Lie sont des structures qui permettent, comme nous le verrons, une classification aisée grâce au système de racines. Avant tout il est important de s'extraire des groupes pour travailler sur des structures algébriques plus générales.

Définition 1 *On appelle (π, V) une représentation de G , groupe compact, lorsque V est un espace de Banach sur un corps K et $\pi : G \rightarrow GL(V)$ un homéomorphisme continu.*

Dans cette définition nous voyons que l'homéomorphisme π joue un rôle important, ce que nous étudierons dans la partie suivante, mais que les propriétés algébriques de V vont permettre de définir différentes représentations d'un même groupe. Parmi ces représentations nous nous intéresserons par la suite à celle dont l'espace de Banach est de dimension finie : $\dim(V) < +\infty$.

La dimension de V s'appelle alors le *degré* de la représentation.

Définition 2 *Une représentation (π, V) d'un groupe compact G est dite irréductible si les seuls sous-espaces stables sont V et le vecteur nul.*

Les représentations irréductibles d'un groupe possèdent de nombreuses propriétés qui permettent de manipuler aisément des objets plus généraux. Elles se rapprochent de l'arithmétique en ce sens que nous avons une décomposition en représentations irréductibles.

Remarque 1 *Une représentation de dimension un est nécessairement irréductible.*

Proposition 1 *Une représentation de degré fini sur un espace de Hilbert d'un groupe compact est la somme directe de représentations irréductibles.*

Preuve Donnons-nous une représentation (π, V) de G , compact, de degré fini n . Si V ne possède aucun sous-espace stable autre que lui-même et le vecteur nul

2.1 Représentation des groupes compacts

alors on a la proposition.

Dans le cas contraire supposons la proposition vraie jusqu'au rang n , avec $\dim(V) = n + 1$.

On prend alors V_1 un sous-espace invariant de V non nul de dimension minimale (il existe bien car sinon on aurait V irréductible). Ce sous-espace est clairement irréductible (sinon il ne serait pas de dimension minimale). En considérant son supplémentaire orthogonal, respectant un produit interne G -invariant, nous obtenons un sous-espace stable de dimension inférieure ou égale à n . En appliquant notre hypothèse de récurrence à ce dernier et en remarquant que V_1 et son orthogonal sont en somme directe nous obtenons la proposition au rang $n + 1$. ■

Une représentation décomposable en somme directe hilbertienne de représentations irréductibles est appelée une représentation semi-simple. Ce concept est à la base de la théorie physique que nous étudierons ensuite.

Remarque 2 On appelle représentation triviale de G sur un espace vectoriel V de dimension 1 la représentation (π, V) telle que :

$$\forall g \in G, \forall v \in V, \pi(g)v = v$$

Regardons quelques exemples classiques pour se familiariser avec cette notion.

Exemple 1 En tant que groupe compact, G agit sur lui-même par multiplication à droite. Cette action définit une représentation sur $K[G]$ et on la nomme représentation régulière.

Exemple 2 Si le degré de la représentation est 1 et que G est de cardinal fini n alors pour tout élément g de G on a $\pi(g)^n = 1$ et ainsi la représentation prend ses valeurs dans le groupe des racines de l'unité de K . La représentation est dite représentation unité et nous verrons plus loin une idée plus générale avec G de cardinal infini.

Notre étude des groupes de Lie en physique fondamentale nécessitera la connaissance d'une représentation irréductible du groupe spécial unitaire de $M_n(\mathbb{C})$:

$$SU_n(\mathbb{C}) = \{A \in GL_n(\mathbb{C}) \mid AA^* = I \text{ et } \det(A) = 1\}$$

Cependant, cela nécessite des outils plus puissants et donc une étude approfondie des représentations. Ce que nous allons développer par la suite.

Il existe des ponts mathématiques entre diverses représentations d'un même groupe.

2.2 Caractères et quasi-caractères

Définition 3 *Etant données (π_1, V_1) et (π_2, V_2) deux représentations d'un même groupe G , on appelle morphisme de représentation une application linéaire T de V_1 dans V_2 qui vérifie la propriété suivante :*

$$\forall g \in G, T \circ \pi_1(g) = \pi_2(g) \circ T$$

Nous noterons $\text{Hom}_G(V_1, V_2)$ l'ensemble des morphismes de représentations de V_1 dans V_2 .

Comme dans le cas de morphismes entre espaces vectoriels, nous dirons qu'un isomorphisme de représentation est une morphisme de représentation bijectif et que deux représentations sont isomorphes s'il existe entre elles un isomorphisme de représentation. Cependant, les représentations irréductibles sont liées entre elles de manière très forte.

Proposition 2 *(Lemme de Schur) Si (π_1, V_1) et (π_2, V_2) sont deux représentations irréductibles de G alors soit elles sont isomorphes, soit il n'existe aucun isomorphisme de représentation entre V_1 et V_2 .*

Preuve Soit T un morphisme de représentation de (π_1, V_1) dans (π_2, V_2) . Le noyau de T est un sous-espace invariant de V_1 et comme la représentation est irréductible on a $\ker(T)$ qui est nul ou vaut V_1 . Soit T est nulle, soit il est injectif.

Si T n'est pas nul alors l'image de T est un sous-espace invariant de V_2 et donc est soit le vecteur nul soit V_2 . T n'étant pas nul on a donc T qui est surjectif donc bijectif.

Au final, T est soit nul soit bijectif. ■

Maintenant que nous avons pris connaissance avec les espaces de Banach associé à une représentation d'un groupe compact, nous allons étudier les premières propriétés des morphismes π d'une représentation (π, V) .

2.2 Caractères et quasi-caractères

Nous avons un groupe compact G auquel nous avons associé une représentation (π, V) . Nous nous plaçons toujours dans le cas de représentations de degré fini donc $\dim(V) < +\infty$.

Définition 4 *On appelle caractère de π la fonction :*

$$\forall g \in G, \chi_\pi(g) = \text{tr}(\pi(g))$$

2.2 Caractères et quasi-caractères

L'analyse de Fourier sur les groupes abéliens localement compacts, théorie que nous ne reprendrons pas ici, rend les caractères d'un groupe compacts intéressants en ceci qu'ils sont orthogonaux dans $L^2(G)$. Nous avons notamment les résultats classiques tels que l'existence de bases orthonormales de caractères de G ainsi que la formule de Plancherel. Si nous notons $\chi(G)$ l'ensemble des caractères de G , nous avons pour toute fonction f dans $L^2(G)$:

$$\forall g \in G, f(g) = \sum_{\chi \in \chi(G)} a_\chi(f) \chi(g)$$

$$\int_G |f(g)|^2 dg = \sum_{\chi \in \chi(G)} |a_\chi(f)|^2$$

Où l'on a $a_\chi(f) = \int_G f(g) \overline{\chi(g)} dg$

Lorsque nous avons V de dimension un alors il peut être assimilé à son corps K et le morphisme π est continu de G dans le groupe multiplicatif $(\mathbb{C}, \times)$.

Exemple 3 *Considérons le cercle unité de $\mathbb{C}$, U , qui est un groupe compact commutatif. Nous avons les caractères*

$$\chi_n : U \rightarrow (\mathbb{C}, \times)$$

$$z \rightarrow z^n$$

Les caractères χ_n forment un système orthonormal de $L^2(U)$.

Nous verrons plus loin que ces propriétés d'orthogonalités sont encore plus fortes mais pour cela il nous faut développer un concept plus faible, celui de quasi-caractère.

Définition 5 *On appelle quasi-caractère d'un groupe topologique G , un homéomorphisme continu χ de G dans $(\mathbb{C}, \times)$.*

On dit qu'il est unitaire si : $\forall g \in G, |\chi(g)| = 1$.

Cette notion de quasi-caractère incarne finalement les représentations de degré un d'un groupe. Par là, nous pouvons assimiler un quasi-caractère comme un caractère en considérant la représentation (π, V) avec $V = \mathbb{C}$ et pour tout g de G :

$$\pi(g)v = \chi(g)v, \forall v \in V$$

2.3 Coefficients matriciels d'un groupe

Cette partie va nous permettre de généraliser le concept de quasi-caractère, qui n'est que la trace de l'homéomorphisme associé à une représentation. Nous considérerons donc ici un groupe G compact et une représentation de G (π, V) , de degré fini avec V hilbertien donc muni d'un produit scalaire $\langle \cdot, \cdot \rangle$.

En considérant une base orthonormale $(e_1, \dots, e_n)$ de V , nous avons, pour tout g élément de G , la matrice $\pi(g)$ qui s'écrit $(\pi_{ij}(g))_{1 \leq i, j \leq n}$.

$$\pi_{ij}(g) = \langle \pi(g)e_j, e_i \rangle = L_i(\pi(g)e_j)$$

avec $L_i(\sum_{j=1}^n v_j e_j) = v_i$ qui est une application linéaire continue (car dimension finie).

Nous pouvons alors généraliser ce concept de coefficients de matrice.

Définition 6 Soient G un groupe compact et (π, V) une représentation de degré fini de G .

On appelle coefficient matriciel sur G une application de la forme $\phi(g) = L(\pi(g)v)$ où L est une application linéaire : $V \rightarrow \mathbb{C}$.

Nous allons voir dans ce chapitre que le concept d'orthogonalité des caractères va se généraliser via ces applications coefficients matriciels. Il est important de souligner que dans notre étude des groupes de Lie, la théorie des coefficients matriciels n'est qu'un moyen d'arriver aux théorèmes sur les représentations. Nous n'irons pas plus loin dans cette étude. Nous pouvons dès à présent remarquer que des coefficients matriciels sont continus puisque nous sommes en présence d'applications linéaires en dimension finie.

A partir de deux représentations (π_1, V_1) et (π_2, V_2) d'un même groupe nous pouvons créer les représentations somme directe et produit tensoriel : $(\pi_1 \oplus \pi_2, V_1 \oplus V_2)$ et $(\pi_1 \otimes \pi_2, V_1 \otimes V_2)$.

Proposition 3 Muni des lois somme directe et produit tensoriel, l'ensemble des coefficients matriciels de G forme un anneau.

Preuve Tout d'abord, les espaces somme directe et produit tensoriel sont toujours de dimension finie.

En prenant v_i dans V_i et les applications L_i dans V_i^* nous créons $L : V_1 \oplus V_2 \rightarrow \mathbb{C}$ telle que : $L(x_1, x_2) = L_1(x_1) \pm L_2(x_2)$. Grâce à cette application nous avons un coefficient matriciel $L((\pi_1 \oplus \pi_2)(g)(v_1, v_2))$. De même pour le produit avec $L = L_1 \otimes L_2$ qui devient bien un coefficient de matrice associé à la représentation produit tensoriel. ■

2.3 Coefficients matriciels d'un groupe

Désormais nous pouvons étudier la généralisation de l'orthogonalité des caractères dans le cas où l'on a des représentations irréductibles dans des espaces hilbertiens.

Théorème 1 (*Théorème de Schur*) Soient (π_1, V_1) et (π_2, V_2) des représentations irréductibles, où les V_i sont hilbertiens, d'un même groupe compact G .

Alors soit les deux représentations sont isomorphes, soit les coefficients matriciels de π_1 sont orthogonaux aux coefficients matriciels de π_2 dans $L^2(G)$.

Preuve Supposons qu'ils existent $\phi_1(g) = L(\pi_1(g)v_1)$ et $\phi_2(g) = L(\pi_2(g)v_2)$ des coefficients matriciels respectivement de V_1 et V_2 qui ne soient pas orthogonaux dans $L^2(G)$. ϕ_i est une forme linéaire continue non nulle sur un espace de Hilbert donc il existe w_i dans V_i tel que $\forall v_i \in V_i, \phi_i(v_i) = \langle \pi_i(g)v_i, w_i \rangle$

Pour v_2 dans V_2 on peut définir l'opérateur :

$$\begin{aligned} T(v_1) &= \int_G \phi_1(v_1) \pi_2(g^{-1}) v_2 dg \\ &= \int_G \langle \pi_1(g)v_1, w_1 \rangle \pi_2(g^{-1}) v_2 dg \end{aligned}$$

T est un morphisme de représentation entre V_1 et V_2 (simple calcul) et d'après le Lemme de Schur il est soit nul, soit bijectif. Or nous avons dans $L^2(G)$:

$$\begin{aligned} \langle v_2, T(v_1) \rangle &= \int_G \langle \pi_1(g)v_1, w_1 \rangle \langle \pi_2(g^{-1})v_2, w_2 \rangle dg \\ &= \int_G \phi_1(v_1) \overline{\phi_2(v_2)} dg \neq 0 \end{aligned}$$

Dès lors T est bijectif car non nul et donc (π_1, V_1) et (π_2, V_2) sont isomorphes. ■

Ce résultat est très important car il est à la base du théorème de Peter et Weil, que nous ne démontrerons pas ici car il nécessite l'étude préalable de l'anneau de convolution des fonctions continues sur G , de théorèmes d'équicontinuité et de la théorie des opérateurs compacts sur un espace de Hilbert. Ce théorème est très puissant.

Théorème 2 (*Peter et Weil*) Pour G un groupe compact, les coefficients matriciels de G sont denses dans $L^2(G)$.

Nous avons réuni dans cette première partie les notions clés au sein des groupes. Il est à noter que celles que nous avons mis en avant sont les caractères car ils vont permettre de classer au sein des groupes et algèbres de Lie. Les autres notions élargissant cette notion nous ont paru importante car en dimension 1 elles sont équivalentes et que nous allons voir certaines qui sont essentielles.

3 Une notion importante : la mesure de Haar

Depuis le début de notre étude des groupes, nous nous plaçons dans des groupes compacts et donc a fortiori topologique, ce qui nous a permis d'utiliser une mesure lors de nos différentes approches. Cependant, contrairement aux aspects généraux d'ensemble topologique, un groupe topologique possède des invariances multiplicatives, des inverses...Il paraît alors fondamental de se préoccuper des mesures régulières, c'est-à-dire telles que $\mu(K) < +\infty$ pour tout compact K de G , qui prennent en compte ces invariances, allant jusqu'à en créer des nouvelles. On appelle alors mesure invariante par translation à gauche une mesure μ_L une mesure régulière qui vérifie, pour tout h de G :

$$\mu_L(hB) = \mu_L(B)$$

Une autre façon d'appréhender ce problème est de se ramener au calcul d'intégrale suivant, pour tout h de G :

$$\forall f \in L^1(G), \int_G f(hg) d\mu_L(g) = \int_G f(g) d\mu_L(g)$$

Nous avons la définition similaire pour une mesure μ_R invariante par translation à droite.

Définition 7 *On appelle mesure de Haar à gauche sur G une mesure régulière invariante par translation à gauche. Par extension on nomme mesure de Haar à droite toute mesure régulière invariante par translation à droite.*

Dans le cas où de telles mesures existent alors elles ne sont pas nécessairement égales. Si une mesure de Haar à gauche est également invariante par translation à droite alors on dit que le groupe G est *unimodulaire*.

De tels objets sont très importants car ils permettent de simplifier les calculs en ramenant, via les translations, à des problèmes mieux connus. Le but de cette partie est de démontrer ce théorème :

Théorème 3 *Si G est un groupe compact séparable alors il existe une unique mesure de Haar à gauche normalisée (telle que $\mu(G) = 1$) sur G . De plus, cette mesure est également invariante par translation à droite.*

En conséquence de quoi, tout groupe compact est unimodulaire.

Le fait qu'on la veuille normalisée permet l'unicité et, surtout, de travailler avec des mesures de probabilité. Nous noterons par la suite :

$$Prob(G) = \{\mu, \text{ mesure}, \mu(G) = 1\}$$

3.1 Un détour par la topologie

La notion de mesure de Haar fait référence à une invariance de la mesure par un opérateur de translation sur tous les éléments de G . En effet, si nous prenons $E = C(G)$ l'espace vectoriel des fonctions continues de G dans (R) nous pouvons définir, pour μ dans $Prob(G)$ une forme linéaire sur E :

$$\forall f \in E, \mu(f) = \int_G f(g) d\mu(g)$$

Notre groupe G étant compact et séparable il vient que nous avons une suite d'éléments de G dense dans G que nous notons $(g_n)_{n \in \mathbb{N}}$. Alors par continuité de la mesure, il suffit que cette dernière soit invariante par translation à gauche sur tous les g_n . On construit alors un opérateur qui nous donne cette propriété d'invariance :

$$\begin{aligned} T : Prob(G) &\rightarrow Prob(G) \\ \mu &\rightarrow \sum_{n \geq 1} \frac{g_n \mu}{2^n} \end{aligned}$$

Où l'on note :

$$\begin{aligned} g\mu : Prob(G) &\rightarrow Prob(G) \\ \mu &\rightarrow g\mu / \forall f \in E, g\mu(f) = \mu(f_g) \end{aligned}$$

Avec $\forall x \in G, f_g(x) = f(gx)$.

Il vient donc que notre mesure de Haar, si elle existe, doit être un point fixe de T (ce que nous démontrerons dans le paragraphe 2.3).

Or, dans le cas de fonctions continues sur un compact métrique et contractantes ($x \neq y \Rightarrow d(f(x), f(y)) < d(x, y)$) nous avons le théorème dit du point fixe : il existe un point invariant sous l'action de cette fonction. Ce résultat vient du fait que la distance est continue donc en considérant la fonction $x \rightarrow d(f(x), x)$ nous savons qu'elle admet un minimum sur le compact et ce minimum vaut 0 (sinon prendre $f(x_{min})$ et on obtient une absurdité).

Or, ce théorème nécessite de fortes hypothèses topologiques ce qui nous pousse à réduire notre topologie pour la rendre plus faible et ainsi pouvoir gérer des nouveaux problèmes qui n'auraient pas de solution dans la topologie forte. Cependant, nous allons voir que nous pouvons dégager deux sortes de « topologie faible » qui, suivant les espaces, peuvent être égales.

Depuis le début nous travaillons avec des espaces de Banach donc nous allons définir ces différentes topologies dans de tels espaces.

3.1.1 La topologie faible

Nous prenons un espace de Banach E et la première idée est d'étudier son dual E' des formes linéaires continues sur E . Les formes continues sont telles que si on a une suite $(u_n)_{n \in \mathbb{N}}$ qui converge vers u dans E alors pour tout ϕ de E' on a $(\phi(u_n))_{n \in \mathbb{N}}$ qui converge vers $\phi(u)$. Ceci est à la base de la topologie faible.

Définition 8 *Dans un espace de Banach E , on dit que la suite $(u_n)_{n \in \mathbb{N}}$ converge faiblement vers u dans E si et seulement si :*

$$\forall \phi \in E', \lim_{n \rightarrow +\infty} \phi(u_n) = \phi(u)$$

Cette nouvelle notion de limite définit sur E une topologie que l'on nomme *topologie faible*.

On parle de topologie faible car elle est moins précise que la topologie initiale, que l'on qualifie alors de forte. En effet, on a toujours :

$$\forall \phi \in E', |\phi(u_n) - \phi(u)| \leq \|\phi\|_{E'} \|u_n - u\|_E$$

Donc toute suite convergent fortement converge faiblement.

En dimension finie il est clair que ces deux topologies sont égales (il suffit de considérer les applications projections sur les vecteurs d'une base de E) mais en dimension infinie elles ne coïncident jamais (résultat admis).

3.1.2 La topologie faible-*

Nous nous sommes attachés aux limites faibles d'éléments de E mais nous voyons que les rôles peuvent être aisément permutés. En effet, un espace de Banach s'identifie à un sous-espace vectoriel de E'' , l'espace dual de son dual. Nous obtenons donc une notion de limite encore plus faible.

Définition 9 *Dans un espace de Banach E , on dit que la suite $(\phi_n)_{n \in \mathbb{N}}$ de E' converge faiblement-* vers ϕ dans E' si et seulement si :*

$$\forall u \in E, \lim_{n \rightarrow +\infty} \phi_n(u) = \phi(u)$$

Encore une fois, cette notion de limite va définir une topologie. Pour s'en persuader, nous pouvons considérer l'application :

$$\forall (\phi, \psi) \in E'^2, d(\phi, \psi) = \sum_{n \geq 0} \frac{|\phi(u_n) - \psi(u_n)|}{2^n \|u_n\|_E}$$

3.1 Un détour par la topologie

Avec $(u_n)_{n \in \mathbb{N}}$ une suite dense d'éléments de E (existe si E de Banach séparable). Cette application est une distance pour la topologie faible-* sur la boule unité fermée de E' .

Grâce à cette métrique il est aisé de démontrer que la boule unité fermée de E' est compacte pour la topologie faible-*. Cependant, cette nouvelle topologie est bien plus puissante grâce au théorème suivant que nous admettrons.

Théorème 4 (*Théorème de d'Alaoglu-Banach-Bourbaki*) *Les parties convexes, bornées et fermées de E^* sont compactes pour la topologie faible-**.

Remarquons que pour un espace de Hilbert nous avons le théorème suivant :

Théorème 5 (*Théorème de Riez*)

$$\forall \phi \in E', \exists ! a \in E / \forall u \in E, \phi(u) = \langle a, u \rangle$$

Preuve Pour ϕ dans E' non nulle (sinon prendre $a = 0$) on a $\text{Ker}(\phi)$ fermé dans E , car ϕ est continue, et donc on a

$$E = \text{Ker}(\phi) \oplus \text{Ker}(\phi)^\perp$$

Dès lors, comme $\phi \neq 0$ il vient que $\text{Ker}(\phi)^\perp$ n'est pas nul donc on peut considérer z dans $\text{Ker}(\phi)^\perp$ non nul et comme ϕ est une forme linéaire il vient :

$$\text{Ker}(\phi)^\perp = \text{Vect}(z)$$

Donc pour u dans E on a $u = u_{\text{ker}} + \lambda.z$ et donc $\phi(u) = \lambda.\phi(z)$.

En posant $a = \frac{\phi(z)}{\|z\|^2}.z$ on a le résultat.

L'unicité venant du fait que le seul vecteur orthogonal à tous les vecteurs de E est le vecteur nul. ■

Ce théorème montre donc que l'on peut assimiler E avec son dual E' et vice-versa. Ainsi nous pouvons, dans un espace hilbertien, parler de la topologie faible-* sur E .

Nous avons donc défini l'une après l'autre, des topologies qui nous offrent de la compacité plus facilement que la topologie forte initiale. Le passage par la topologie faible-* est indispensable puisque nous devons, rappelons-le, démontrer l'existence d'un point fixe parmi les mesures pour l'application translation à droite sur celles-ci, et donc travailler dans un espace dual.

3.2 Un nouveau théorème du point fixe

Nous avons étudié des notions topologiques sur les espaces dual d'un espace de Banach et avons remarqué qu'elle nous donne aisément des compacts. Prenons donc un espace de Banach.

Rappelons d'abord que si K est une partie convexe de E' alors une application T de $K \rightarrow K$ est affine si :

$$\forall (x, y) \in E'^2, \forall t \in [0, 1], T(tx + (1 - t)y) = tT(x) + (1 - t)T(y)$$

Grâce à la topologie faible-* nous pouvons obtenir un théorème de point fixe dans E' .

Théorème 6 (*Théorème de Markov-Kakutani*) Soient E un espace vectoriel normé et K une partie convexe, bornée, fermée et non vide de E' . Soit $T : K \rightarrow K$ une application affine et continue pour la topologie faible-*. Alors T possède un point fixe dans K .

Preuve Tout d'abord, remarquons que, en vertu du théorème d'Alaoglu-Banach-Bourbaki, K est compact pour la topologie faible-*. Dans la suite on munit K de cette topologie uniquement.

Fixons x_0 dans K et on construit la suite de points :

$$\forall n \in \mathbb{N}, x_n = \frac{1}{n+1} \sum_{k=0}^n T^k(x_0)$$

Remarque 3 Pour le théorème du point fixe dans un espace vectoriel compact, la suite $(x_n)_{n \in \mathbb{N}}$ vaut seulement $x_n = T^n(x_0)$. Mais ici il faut se ramener à des points dans le convexe K ce qui est le cas ici.

Pour étudier l'hypothétique convergence de $(x_n)_{n \in \mathbb{N}}$ considérons la suite décroissante d'ensembles fermés suivante :

$$X_n = \overline{\{x_m, m \geq n\}}$$

Ainsi que leur intersection :

$$X = \bigcap_{n \in \mathbb{N}} X_n$$

K étant compact il vient que X est non vide. Considérons alors x dans X , remarquons que x est dans K puisque K est fermé dans E .

3.3 L'existence de mesures de Haar

Par définition de X on a alors une suite extraite $(x_{\phi(n)})_{n \in \mathbb{N}}$ qui converge vers x dans K pour la topologie faible-*, c'est-à-dire :

$$\forall u \in E, \lim_{n \rightarrow +\infty} x_{\phi(n)}(u) = x(u)$$

Soit u dans E . On a alors :

$$\begin{aligned} |Tx_{\phi(n)}(u) - x_{\phi(n)}| &= \frac{1}{1 + \phi(n)} \left| \sum_{k=0}^{\phi(n)} T^{k+1}(x_0) - \sum_{k=0}^{\phi(n)} T^k(x_0) \right| \\ &= \frac{1}{1 + \phi(n)} |T^{\phi(n)+1}(x_0) - T(x_0)| \end{aligned}$$

Nous savons que K est borné dans E par M . Dès lors on a :

$$|Tx_{\phi(n)}(u) - x_{\phi(n)}| \leq \frac{2M}{1 + \phi(n)}$$

Par passage à la limite quand n tend vers $+\infty$, et comme T est continue nous obtenons :

$$\forall u \in E, Tx(u) = x(u)$$

Ce qui est la définition de $Tx = x$. ■

Nous avons désormais en main les outils pour démontrer notre théorème initial.

3.3 L'existence de mesures de Haar

Reprenons notre cheminement initial en reliant nos notations aux notations précédentes pour voir clairement ce que l'on veut.

Notre espace de banach est $E = C(G)$ l'espace des fonctions continues de $G \rightarrow \mathbb{R}$. Et nous travaillons avec $K = \text{Prob}(G)$ l'ensemble des mesures de probabilité de G . Nous avons vu que $\text{Prob}(G)$ peut être assimilé à une partie de E' .

3.3.1 Vers la mesure fixe

Proposition 4 *Si G est un groupe compact alors $\text{Prob}(G)$ est une partie convexe, fermée et bornée de $C(G)$.*

Preuve L'ensemble des mesures ramenées à E' est un sous-espace vectoriel donc il est convexe. De plus, il est fermé car la limite d'une suite de mesures de norme

3.3 L'existence de mesures de Haar

1 est une mesure de norme 1.

Enfin :

$$\forall f \in C(G), \mu(f) \leq \|f\|_\infty \mu(G) = \|f\|_\infty$$

Remarquons que la norme $\|f\|_\infty = \sup_G |f|$ existe bien car G est compact et f continue sur G . Dès lors on a, avec les normes subordonnées :

$$\forall \mu \in \text{Prob}(G), \|\mu\| \leq 1$$

Donc $\text{Prob}(G)$ est bornée. ■

G est séparable donc admet $(g_n)_{n \in \mathbb{N}}$ une suite dense dans G . A partir de celle-ci nous sommes en mesure de définir l'application T suivante (nous rappelons ici les formules établies au début du chapitre) :

$$\begin{aligned} T : \text{Prob}(G) &\rightarrow \text{Prob}(G) \\ \mu &\rightarrow \sum_{n \geq 1} \frac{g_n \mu}{2^n} \end{aligned}$$

Où l'on note :

$$\begin{aligned} g\mu : \text{Prob}(G) &\rightarrow \text{Prob}(G) \\ \mu &\rightarrow g\mu / \forall f \in E, g\mu(f) = \mu(f_g) \end{aligned}$$

Avec $\forall x \in G, f_g(x) = f(gx)$.

Proposition 5 *T est continue, pour la topologie faible-* de $\text{Prob}(G) \rightarrow \text{Prob}(G)$ et affine.*

Preuve Le caractère affine ne pose aucun problème vu que une somme est linéaire et on a bien T à valeurs dans $\text{Prob}(G)$ puisque $\sum_{i \geq 1} \frac{1}{2^i} = 1$.

Soit $(\mu_n)_{n \in \mathbb{N}}$ une suite de $\text{Prob}(G)$ qui converge vers μ dans $\text{Prob}(G)$ pour la topologie faible-*. Pour tout f de $C(G)$:

$$|T\mu_n(f) - T\mu(f)| \leq \sum_{i \geq 1} \frac{1}{2^i} |\mu_n(f_{g_i}) - \mu(f_{g_i})|$$

Par convergence en topologie faible-* on a :

$$\forall f \in C(G), \exists N_f \in \mathbb{N}, \forall n \geq N_f, |\mu_n(f) - \mu(f)| \leq \epsilon$$

3.3 L'existence de mesures de Haar

or, par construction des f_{g_i} nous avons que $N_{f_{g_i}} = N_f$. Et donc, pour tout $n \geq N_f$:

$$\begin{aligned} |T\mu_n(f) - T\mu(f)| &\leq \epsilon \sum_{i \geq 1} \frac{1}{2^i} \\ &\leq \epsilon \end{aligned}$$

Dès lors : $\forall f \in C(G), \lim_{n \rightarrow +\infty} T\mu_n(f) = T\mu$.

Ceci prouve la faible-* continuité de T . ■

3.3.2 Vers la mesure de Haar

Nous avons donc désormais les hypothèses nécessaires à l'application de théorème de Markov-Kakutani. Nous en déduisons donc que T admet une mesure fixe μ dans $Prob(G)$.

Proposition 6 μ , mesure fixe de T est une mesure de Haar à gauche de norme 1.

Preuve Soient f dans $C(G)$ et g dans G .

On a déjà :

$$T\mu(f) = \mu(f) \tag{1}$$

$$\sum_{n \geq 1} \frac{1}{2^n} \int_G f(g_n u) d\mu(u) = \int_G f(u) d\mu(u) \tag{2}$$

ceci nous donne donc :

$$\sum_{n \geq 1} \frac{1}{2^n} \int_G f(g g_n u) d\mu(u) = \int_G f(g u) d\mu(u)$$

En posant :

$$\begin{aligned} F : G &\rightarrow \mathbb{R} \\ g &\rightarrow \int_G f(g u) d\mu(u) \end{aligned}$$

Nous obtenons d'après (2) que :

$$\forall g \in G, \sum_{n \geq 1} \frac{1}{2^n} F(g g_n) = F(g)$$

3.3 L'existence de mesures de Haar

f étant continue sur le compact G elle y est uniformément continue et donc, comme nous intégrons sur un compact nous avons que F est continue sur G . F est donc bornée et atteint ses bornes. Notons g_{max} un élément de G tel que $F(g_{max}) = \max_G F$. Dès lors :

$$\sum_{n \geq 1} \frac{1}{2^n} F(g_{max} g_n) = F(g_{max})$$

Comme nous avons que $\sum_{n \geq 1} \frac{1}{2^n} = 1$ l'égalité ci-dessus n'est possible que si :

$$\forall n \in \mathbb{N}, F(g_{max} g_n) = F(g_{max})$$

Or, $(g_n)_{n \in \mathbb{N}}$ est dense dans G et F est continue ce qui nous donne finalement que :

$$\forall g \in G, F(g_{max} g) = F(g_{max})$$

On vient donc de prouver que F est constante sur G . Dès lors on a $F(g) = F(1)$ pour tout g de G .

$$\forall g \in G, \int_G f(gu) d\mu(u) = \int_G f(u) d\mu(u)$$

μ est donc bien invariante par translation à gauche. ■

Nous pouvons désormais parler du sous-ensemble de $Prob(G)$ constitué des mesures de probabilité sur G qui sont invariantes par translation à gauche. Cet ensemble est convexe mais également compact pour la topologie faible-* de $Prob(G)$ puisque l'intégrale sur un compact de fonctions continues sur ce compact est uniformément convergente. Dès lors, nous reprenons les axes des démonstrations ci-dessus appliqués sur ce nouvel ensemble et nous voyons qu'il existe au sein de ces mesures une mesure invariante par translation à gauche. Prenons ν une de ces mesures.

Nous avons donc l'existence de ν mesure de norme 1 invariante par translation à droite et à gauche. Il ne nous reste plus maintenant qu'à nous occuper de l'unicité

de la mesure de Haar, ce qui est donné par la proposition suivante.

Proposition 7 *La mesure ν est l'unique mesure de Haar à gauche de norme 1.*

Preuve Soit m une mesure de $Prob(G)$ invariante par translation à gauche. Nous sommes sur un compact donc nous pouvons appliquer Fubini à des fonctions

3.3 L'existence de mesures de Haar

continues sur G , ce qui nous donne, puisque m est invariante par translation à droite et ν par translation à droite et à gauche :

$$\begin{aligned}
 \int_G \int_G f(ug) dm(g) d\nu(u) &= \int_G \int_G f(g) dm(g) d\nu(u) \\
 &= \nu(G) \int_G f(g) dm(g) \\
 &= \int_G \int_G f(u) d\nu(u) dm(g) \\
 &= m(G) \int_G f(u) d\nu(u)
 \end{aligned}$$

Comme nous sommes avec des mesures de probabilité il vient au final :

$$\int_G f(g) dm(g) = \int_G f(u) d\nu(u)$$

Ce qui correspond, puisque cela est vrai pour tout f de $C(G)$, à : $m = \nu$. ■

3.3.3 Remarques importantes

Nous avons donc montré que pour un groupe compact séparable il existe une unique mesure de probabilité qui soit mesure de Haar et elle est alors également invariante par translation à droite. Cependant, nous avons vu dans la preuve d'unicité précédente que si nous avons deux mesures quelconques invariantes par translation à gauche alors elles sont proportionnelles. Ainsi, une mesure de Haar à gauche, ou à droite, est unique à une constante multiplicative près. En fait, il est possible de se soustraire aux hypothèses de compacité et de séparabilité ; le théorème suivant se veut plus général.

Théorème 7 *Dans un groupe G localement compact, il existe une unique mesure de Haar à gauche, à une constante multiplicative près. Il existe également une unique mesure de Haar à droite, à une constante multiplicative près.*

Il est à noter que ce théorème n'assure nullement l'unimodularité des groupes seulement localement compacts. En général, les mesures de Haar à droite et à gauche ne coïncident pas. Remarquons que l'on a cherché les mesures régulières, c'est-à-dire de mesure finie sur les compacts inclus dans G . donc si G est compact alors on a $\mu_L(G) < +\infty$ et on peut se ramener, comme nous l'avons vu, à une mesure normalisée mais dans le cas de groupes seulement localement compacts il est possible d'avoir $\mu_L(G) = +\infty$.

3.3 L'existence de mesures de Haar

Exemple 4 *Il faut un groupe non commutatif pour obtenir des exemples de groupes non unimodulaire. Par exemple prenons le groupe suivant :*

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

Dans ce groupe nous avons la multiplication matricielle qui donne :

$$\begin{pmatrix} x' & y' \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} x'x & x'y + y' \\ 0 & 1 \end{pmatrix}$$

Dès lors, en faisant le changement de variable $u = xx'$ et $v = xy' + y$ nous aurons une invariance par multiplication à gauche en rendant invariant le facteur en x' (y' étant un terme linéaire, il ne joue pas dans la jacobienne du changement de variable et quand on intègre sur $\mathbb{R} \times \mathbb{R}^{+}$ les changements affines ne jouent pas sur les bornes puisque x' est positif) qui est de degré 2 dans la jacobienne. Donc nous obtenons de là :*

$$d\mu_L = \frac{dx dy}{x^2}$$

De même, pour avoir une invariance par translation à droite il faut faire le même changement de variable mais cette fois-ci pour il faut rendre invariant le facteur en x et par le même raisonnement on obtient :

$$d\mu_R = \frac{dx dy}{x}$$

Ces deux mesures ne coïncident clairement pas.

Nous verrons dans les prochaines sections que chez certains groupes il la mesure de Haar s'obtient facilement (dans les groupes de Lie par exemple). C'est donc souvent l'unicité du théorème qui est importante.

Exemple 5 *Sur $\mathbb{R}^{+*}$ on a une mesure de Haar qui est $\frac{dx}{x}$. Cela se voit facilement puisque, pour f telle que $\frac{f(x)}{x}$ intégrable sur $\mathbb{R}^{+*}$:*

$$\int_{\mathbb{R}^{+*}} \frac{f(ax)}{x} dx = \int_{\mathbb{R}^{+*}} \frac{f(u)}{u} du$$

De même que la mesure de Lebesgue sur $\mathbb{R}$ est une mesure de Haar mais nous voyons qu'elle est bien régulière mais n'est pas finie sur $\mathbb{R}$.

De manière général, sur tout espace vectoriel de dimension finie, la mesure de Lebesgue est une mesure de Haar.

3.4 Application aux caractères

Faisons un petit clin d'oeil à notre première partie afin de montrer que la mesure de Haar, au delà du simple concept de mesure, est très attaché à son groupe. En effet, à partir de la mesure de Haar à gauche nous sommes en mesure d'extraire un quasi-caractère.

Pour tout g de G , l'application de $G \rightarrow G$ qui à h associe $g^{-1}hg$ est un automorphisme de G . Dès lors, elle envoie une mesure de Haar à gauche sur une mesure de Haar à gauche. Or elles ne diffèrent que par une constante multiplicative. D'où :

$$\forall g \in G, \exists \delta(g) > 0, \int_G f(g^{-1}hg) d\mu_L(h) = \delta(g) \int_G f(h) d\mu_L(h)$$

Proposition 8 $\delta : G \rightarrow \mathbb{R}^{+*}$ est un quasi-caractère.

Preuve

$$\begin{aligned} \delta(gt) \int_G f(h) d\mu_L(h) &= \int_G f((gt)^{-1}h(gt)) d\mu_L(h) \\ &= \int_G f(t^{-1}(g^{-1}hg)t) d\mu_L(h) \\ &= \delta(t) \int_G f(g^{-1}hg) d\mu_L(h) \\ &= \delta(t)\delta(g) \int_G f(h) d\mu_L(h) \end{aligned}$$

On a donc δ qui est un homéomorphisme de $G \rightarrow \mathbb{R}^{+*}$. De plus, il est continu car si $(g_n)_{n \in \mathbb{N}}$ tend vers g dans G on a, pour f dans $C(G) : (f(g_n))_{n \in \mathbb{N}}$ qui converge vers $f(g)$. Or on a :

$$\forall n \in \mathbb{N}, |f(g_n)| \leq \sup_G |f|$$

D'après le principe de convergence dominée (G est compact donc une constante y est intégrable) on a le résultat. ■

L'existence de la mesure de Haar permet d'obtenir des résultats forts comme celui-ci :

Proposition 9 Soit (π, V) une représentation irréductible de G compact. On note χ son caractère. On a alors :

$$\int_G \chi(g) dg = \begin{cases} 1 & \text{si } (\pi, V) \text{ est triviale} \\ 0 & \text{sinon} \end{cases}$$

Preuve La représentation triviale a pour caractère la fonction constante égale à 1 (trace de l'identité). Dès lors, en normalisant la mesure de Haar de G nous obtenons le résultat lorsque π est triviale.

Si π n'est pas triviale alors on a :

$$\int_G \chi(g) dg = \int_G \chi(g) \chi_{trivial}(g) dg$$

On reconnaît ici le produit scalaire dans $L^2(G)$ de deux coefficients matriciels de deux représentations irréductibles. D'après le théorème d'orthogonalité de Schur il vient que les deux représentations sont isomorphes, ce qui n'est pas possible puisque π n'est pas triviale, ou alors leurs coefficients matriciels sont orthogonaux. Donc on obtient :

$$\int_G \chi(g) dg = 0$$

■ Cette proposition permet d'obtenir la suivante.

Prenons (π, V) une représentation de G , de caractère χ , de degré fini. On note V^G le sous-espace de V défini par :

$$V^G = \{v \in V / \forall g \in G, \pi(g)v = v\}$$

Proposition 10

$$\int_G \chi(g) dg = \dim(V^G)$$

Preuve (π, V) étant de degré fini on a, d'après la proposition 1, $V = \oplus_{i \in I} V_i$ avec les V_i irréductibles. Si nous appelons χ_i la restriction de χ à V_i la proposition précédente nous donne :

$$\int_G \chi_i(g) dg = 1 \iff \pi \text{ restreint à } V_i \text{ est triviale}$$

Dès lors on a $\int_G \chi_i(g) dg$ qui est égal au nombre de π_i triviaux. Le nombre de π_W triviaux (π_W restriction à W sous-espace vectoriel de V) est la dimension du sous-espace vectoriel V^G . les V_i étant en somme directe nous avons que les π_i triviaux décrivent tout l'ensemble π_W triviaux, π_W restriction à W sous-espace vectoriel de V et donc nous avons le résultat. ■

Nous n'irons pas plus loin dans cette étude de la mesure de Haar. Il nous a semblé important d'en étudier les premières caractéristiques et son existence car cela livre, dans la théorie des groupes, des mesures invariantes. De plus, elles sont descriptibles efficacement dans les groupes de Lie.

4 Vers les groupes de Lie

Nous nous attaquons désormais au problème des groupes de Lie. Nous allons pour cela faire un détour par la topologie différentielle car les groupes de Lie sont des groupes différentiels. Par cette occasion, ils tissent un lien entre algèbre et analyse. Nous verrons par la suite que les algèbres de Lie, elles, ne nécessitent que des définitions algébriques ce qui les rendent différentes des groupes de Lie.

4.1 Incursion dans la topologie différentielle

Nous n'étudierons ici que très superficiellement les notions. Il nous est juste nécessaire d'obtenir des définitions et quelques résultats pour nos démonstrations ultérieures. L'aspect le plus important va être de bien comprendre les notions de variétés différentielles, que nous appellerons seulement variétés (aucune ambiguïté possible avec les variétés algébriques dont nous ne parlons pas ici).

Définition 10 *Une variété M de dimension d est un espace topologique séparé qui est localement difféomorphe à un ouvert de $\mathbb{R}^d$. En d'autres termes pour tout x de M :*

$$\exists U_x \text{ voisinage de } x \text{ dans } M, \exists V_x \text{ ouvert de } \mathbb{R}^d, \exists \phi_x \in C^k(U_x, V_x) \text{ difféomorphisme}$$

La donnée de (ϕ_x, U_x) est appelé une carte locale de M .

Remarque 4 *Par difféomorphisme de $U \rightarrow V$ nous entendons un isomorphisme de classe C^∞ sur U dont l'application réciproque est de classe C^∞ sur V .*

De telles structures sont suffisamment lisses pour développer du calcul différentiel sur elle. Nous allons alors pouvoir définir chez ces variétés la notion d'espace tangent, qui nous permettra, dans les groupes de Lie, de trouver la mesure de Haar.

Remarquons tout d'abord que, pour x dans la variété M , il existe une relation d'équivalence sur l'ensemble des couples (f, U_x) tels que f soit une fonction de classe C^∞ de $U_x \rightarrow \mathbb{R}$. On dit que (f, U_x) et (g, W_x) sont équivalentes si $f = g$ sur un ouvert contenu dans $U_x \cap W_x$.

Pour un point x de M on appelle Eqv_x l'ensemble des classes d'équivalence de tels couples.

Définition 11 *Pour x dans une variété M on appelle dérivée locale de Eqv_x une application linéaire D de $Eqv_x \rightarrow \mathbb{R}$ vérifiant la propriété :*

$$\forall (f, g) \in Eqv_x, D(fg) = f(x)D(g) + g(x)D(f)$$

4.1 Incursion dans la topologie différentielle

Nous reconnâtrons ici l'opérateur de dérivée partielle en x si la variété est plongée dans $\mathbb{R}^n$. Une dérivée locale peut donc être comprise comme la généralisation de notre dérivée partielle en x classique. L'espace tangent à une variété M , plongée dans $\mathbb{R}^n$, au point x est l'ensemble des vecteurs tangents en x aux courbes tracées sur M . Par analogie nous allons définir un espace tangent sur une variété quelconque en prenant en compte toutes les dérivées locales de Eqv_x .

Pour x de M , nous noterons $T_x(M)$ l'espace tangent à M au point x . Un élément de $T_x(M)$ sera appelé un vecteur tangent à M en x . La notion d'espace est bien établit par la proposition ci-dessous.

Proposition 11 *Si x est un point d'une variété M de dimension d alors l'ensemble des dérivées locales de Eqv_x est un espace vectoriel de dimension d .*

Preuve Prenons un point x dans une variété M de dimension d .

Cette proposition vient du fait qu'en dimension finie nous avons des applications coordonnées et pour des fonctions C^∞ nous avons des dérivées partielles qui forment une base des dérivées locales. Sur $\mathbb{R}^d$ on a $df_a(h) = \sum_{i=1}^d h_i \partial_{x_i} f(a)$. Il faut se rendre compte que nos dérivées locales sont dans des voisinages de x dans M et qu'elles s'effectuent donc en x .

En prenant les notations de la définition de variété : sur le voisinage $\phi_x(U_x)$ de $\mathbb{R}^d$ on considère un système de coordonnées locales $y = (y_1, \dots, y_d)$.

On prend x dans U_x , il existe un unique x' dans $\mathbb{R}^d$ tel que $(\phi_x(x) = x')$. On a donc $D(f) = D(f(\phi_x^{-1}(x')))$ qui est donc une dérivée partielle en x' de $\mathbb{R}^d$ de $f \circ \phi_x^{-1}$. Supposons que ce soit la dérivée partielle par rapport à y_i

$$\begin{aligned} D(f) &= \partial_{y_i}(f(\phi_x^{-1}(x'))) \\ &= \sum_{k=1}^d \partial_{y_i}[(\phi_x^{-1})_k](x') \partial_{(\phi_x^{-1})_k} f(\phi_x^{-1}(x')) \\ &= \sum_{k=1}^d a_k(D) \partial_{x_k} f(x) \end{aligned}$$

On obtient donc que toutes les dérivées locales de Eqv_x s'écrivent sur la famille $(\partial_{x_1}, \dots, \partial_{x_d})$ qui est libre, on peut même alors définir la notion de coordonnées locales sur M avec les applications coordonnées $(x_1, \dots, x_d)$, et donc forme une base de $T_x(M)$. ■

Nous avons donc maintenant défini le cadre des variétés et leur avons assigné un espace tangent dont le formalisme va nous être nécessaire pour les groupes de Lie. Il ne nous reste plus qu'à énoncer une nouvelle définition qui va nous permettre de lier points de la variété à vecteur tangent.

Définition 12 *On appelle champ vectoriel sur M , variété, une application de classe C^∞ sur M , qui à tout point m de M associe un vecteur tangent D_m de $T_m(M)$*

4.2 Groupes de Lie

Nous allons ici définir ce qu'est un groupe de Lie. Une telle définition nécessite les notions de groupe topologique, de séparabilité, de variétés différentielles et de tout ce qu'elles impliquent. C'est pour cela qu'il nous a fallu développer les divers aspects dans les parties précédentes. Puis nous étudierons deux exemples importants pour nous puisque nous en aurons besoin lors de l'application physique des groupes de Lie.

4.2.1 Définition et premières propriétés

Commençons par définir un groupe de Lie. L'important dans ces groupes est notre volonté à travailler dans des structures lisses où il est possible d'avoir une mesure et d'effectuer du calcul différentiel. D'ailleurs, la théorie des groupes et algèbres de Lie est présente dans l'étude actuelle des équations différentielles mais nous n'aborderons pas cet aspect très éloigné de nos connaissances actuelles.

Définition 13 *Un groupe de Lie G est un groupe topologique, localement compact, qui est également une variété différentielle telle que les applications de multiplication de $G \times G \rightarrow G$ et passage à l'inverse sont de classe C^∞ .*

A partir de cette définition on peut parler de sous-groupe de Lie d'un groupe de Lie G .

Définition 14 *G' est un sous-groupe de Lie de G , un groupe de Lie, si G' est un sous-groupe fermé de G qui vérifie les propriétés d'un groupe de Lie.*

Remarque 5 *Un sous-groupe de Lie est lui-même une variété différentielle, c'est même une sous-variété de G , mais sa dimension n'est pas nécessairement la même que celle de G . En revanche, elle lui est nécessairement inférieure ou égale. La définition veut qu'un sous-groupe de Lie soit fermé. Une définition plus générale peut être donnée mais elle permet des groupes peu métrisables donc d'une manière générale on ne considère que des sous-groupes fermés.*

Nous travaillons donc sur un type de groupes très réguliers ce qui fait que pour les lier entre eux il est nécessaire de fortifier nos outils. C'est ainsi qu'un homéomorphisme de groupes de Lie est un homéomorphisme de groupes qui est de classe C^∞ .

4.2.2 L'espace tangent et mesure de Haar

Un groupe de Lie est également une variété, structure sur laquelle nous avons défini un espace tangent sur lequel nous pouvons créer des champs vectoriels. Or G possède une multiplication interne de classe C^∞ et donc nous allons combiner nos deux aspects pour en déduire la mesure de Haar sur G . Nous apprendrons au chapitre suivant que cet espace tangent est très important dans la théorie des groupes et algèbres de Lie. Considérons, pour g dans G l'application :

$$\begin{aligned} Mult_g : G &\rightarrow G \\ h &\rightarrow gh \end{aligned}$$

Par définition même, cette application est un difféomorphisme et donc elle induit un difféomorphisme sur chaque espace tangent $T_h(G)$ dans l'espace tangent $T_{gh}(G)$. Nous noterons cet application L_g .

On dira alors qu'un champ vectoriel D sur G est *invariant à gauche* si :

$$\forall g \in G, \forall h \in G, L_g(D(h)) = D(gh)$$

La proposition suivante nous donne l'existence de tels champs vectoriels.

Proposition 12 *L'espace vectoriel des champs vectoriels invariants à gauche est de dimension la dimension de G en tant que variété.*

Preuve Soit X un champ vectoriel invariant à gauche. On a alors, en appelant 1 le neutre du groupe G :

$$\forall g \in G, D(g) = D(g \times 1) = L_g(D(1))$$

On a donc qu'un champ vectoriel invariant à gauche est parfaitement défini par la donnée du vecteur tangent qu'il associe à l'élément neutre de G .

Réciproquement, en prenant D_1 dans T_1 nous pouvons créer un champ vectoriel en posant pour tout g de G : $D(g) = L_g(D_1)$. Or on a, par construction : $L_g L_h = L_{gh}$ ce qui nous donne finalement un champ invariant à gauche :

$$\begin{aligned} L_g(D(h)) &= L_g(L_h(D_1)) \\ &= L_{gh}(D_1) \\ &= D(gh) \end{aligned}$$

Finalement on obtient que l'ensemble des champs vectoriels invariants à gauche est isomorphe à l'espace tangent en l'élément neutre de G . D'après la proposition 11 il est de dimension celle de G . ■

La démonstration ci-dessus nous a également permis d'apprendre qu'à un vecteur tangent D_1 de T_1 il existe un unique champ vectoriel qui prend en 1 la valeur D_1 .

On note $Lie(G)$ l'espace vectoriel des champs vectoriels invariants à gauche.

Grâce à la notion de différentiabilité sur les variétés nous pouvons aisément trouver la mesure de Haar sur G un groupe de Lie de dimension d . Si nous prenons une d -forme linéaire alternée non nulle ω_1 sur T_1 alors en prenant l'image par translation à gauche par g^{-1} de ω_1 nous obtenons une d -forme linéaire alternée sur T_g , pour tout g de G . En réalité, ce processus est le même que lorsque nous avons associé un champ vectoriel à tout vecteur tangent de T_1 .

Dès lors nous avons construit une d -forme linéaire sur G qui est invariante par translation à gauche et non nulle. Elle définit donc une mesure de volume invariante par translation à gauche. D'après le théorème 7, puisque G est localement compact, c'est la mesure de Haar de G .

4.2.3 Les sous-groupes de Lie de $GL_n(\mathbb{C})$

Dans la physique fondamental, nous utilisons les groupes de Lie de $GL_n(\mathbb{C})$ et surtout les groupes spéciaux unitaires $SU_n(\mathbb{C})$. Nous allons donc les étudier un peu plus en détail.

Tout d'abord notons que la multiplication matricielle est une combinaison polynomiale des coefficients des matrices et est donc de classe $C^\infty(M_n(\mathbb{C}))$. De même pour le passage à l'inverse dans $GL_n(\mathbb{C})$ car le passage aux cofacteurs est une combinaison polynomiale des coefficients des matrices, comme le déterminant et que :

$$\forall M \in GL_n(\mathbb{C}), M^{-1} = \frac{1}{\det(M)} {}^t \text{Comatrice}(M)$$

De plus, $GL_n(\mathbb{C})$ est dense dans $M_n(\mathbb{C})$ ce qui nous donne de suite, comme $M_n(\mathbb{C})$ est une variété de dimension $2n^2$ (car dimension n^2 sur $\mathbb{C}$), il vient que $GL_n(\mathbb{C})$ est une variété de dimension $2n^2$ et elle est clairement localement compact comme $M_n(\mathbb{C})$.

$GL_n(\mathbb{C})$ est donc bien un groupe de Lie.

Exemple 6

$$SU_n(\mathbb{C}) = \{A \in GL_n(\mathbb{C}) \mid AA^* = I \text{ et } \det(A) = 1\}$$

est clairement fermé car c'est l'intersection de l'image réciproque du fermé 1 par l'application continue déterminant, qui est donc fermé, avec l'ensemble des matrices unitaires ($AA^* = I$). De plus, c'est une variété de dimension $2(n^2 - 1)$. Nous avons donc ici une sous-groupe de Lie de $GL_n(\mathbb{C})$.

Nous allons étudier les propriétés intéressantes des sous-groupes de Lie de $GL_n(\mathbb{C})$ qui se démarquent surtout par leur espace tangent facilement identifiable. Notons que l'application exponentielle définit de la manière ci-dessous est une application de $M_n(\mathbb{C}) \rightarrow GL_n(\mathbb{C})$.

$$\forall M \in M_n(\mathbb{C}), \exp(X) = \sum_{k \in \mathbb{N}} \frac{M^k}{k!}$$

De plus, elle vérifie la propriété $\exp(M + N) = \exp(M)\exp(N)$ si M et N commutent (ce résultat s'obtenant par simple produit de convolution de Cauchy de séries absolument convergentes) ; nous donnant que $\exp$ est à valeurs dans $GL_n(\mathbb{C})$. La propriété suivante est très importante car elle va rapprocher $Lie(G)$, espace vectoriel des champs vectoriels sur G invariants à gauche, et l'exponentielle. Mais pour l'obtenir nous devons passer par ce lemme :

Lemme 1 Soient G un sous-groupe de Lie de $GL_n(\mathbb{C})$ et M dans $M_n(\mathbb{C})$

$$t \rightarrow \exp(tM) \text{ tangent à } G \text{ à } t = 0 \iff \forall t \in \mathbb{R}, \exp(tM) \subset G$$

Preuve Un résultat d'analyse nous donne que pour tout ensemble ouvert O de $\mathbb{R}^n$ et pour tout o de O , il existe une fonction de classe C^∞ sur $\mathbb{R}^n$ dans $\mathbb{R}$, à support compact contenu dans O qui ne s'annule pas en o et positive. Il suffit de voir que :

$$f(x) = \begin{cases} e^{-(1-\frac{|x|}{r})^{-1}} & \text{si } \|o - x\| \leq r \\ 0 & \text{sinon} \end{cases}$$

vérifie la propriété pour r suffisamment petit pour que $\overline{B(o, r)} \subset O$.

Le sens de droite à gauche est évident alors démontrons le sens gauche-droite du lemme.

Supposons alors qu'il existe $t_0 \neq 0$ tel que $\exp(t_0 M) \notin G$. On construit alors la fonction ϕ sur $GL_n(\mathbb{C})$ qui est C^∞ à support compact telle que ϕ est nulle sur G et $\phi(\exp(t_0 M)) \neq 0$.

On pose :

$$\forall h \in GL_n(\mathbb{C}), \psi(h) = \int_G \phi(hg) dg$$

Nous savons à présent qu'une mesure de Haar existe sur les groupes de Lie donc on choisit dg comme la mesure de Haar de G . Dès lors, il vient que ψ est constante sur hG , nulle sur G mais non nulle en $\exp(t_0M)$.

Considérons alors la fonction :

$$\forall t \in \mathbb{R}, f(t) = \psi(\exp(tM))$$

On a, pour tout t de $\mathbb{R}$:

$$f'(t) = \frac{d}{du}(\psi(\exp(tM)\exp(uM)))|_{u=0}$$

Or, $u \rightarrow \exp(tM)\exp(uM)$ est tangent à $\exp(tM)G$ pour u tendant vers 0, puisque $t \rightarrow \exp(tM)$ est tangent à G en 0. Donc il vient, comme ψ est constante sur hG que :

$$f'(t) = \frac{d}{du}(\psi(\exp(tM)\exp(uM)))|_{u=0} = 0$$

Donc au final f est constante et comme $f(0) = 0$ il vient que $f(t_0) = 0$ ce qui n'est pas, par construction. On obtient une absurdité et donc l'implication réciproque.

■

Nous sommes désormais en mesure d'énoncer la propriété forte suivante :

Proposition 13 *Soit G un sous-groupe de Lie de $GL_n(\mathbb{C})$.*

L'ensemble des M de $M_n(\mathbb{C})$ telles que : $\exp(tM) \in G, \forall t \in \mathbb{R}$ est isomorphe à l'ensemble des champs vectoriels sur G invariants à gauche.

Preuve Soit D un champ vectoriel sur G invariant à gauche. D'après la partie précédente, il est parfaitement déterminé par son élément en l'identité de G (ici I_n). Or, nous sommes dans $M_n(\mathbb{C})$ isomorphe à $\mathbb{R}^{2n^2}$ donc $D(I_n)$ est une dérivée locale par rapport à un vecteur M c'est-à-dire :

$$\begin{aligned} X(I_n) : Eqv_{I_n} &\rightarrow \mathbb{R} \\ f &\rightarrow \frac{d}{dt}(f(I_n + tM))|_{t=0} \end{aligned}$$

Or, en $t = 0$, $t \rightarrow I_n + tM$ est tangent à $t \rightarrow \exp(tM)$ et à G en $t = 0$. Cela nous donne donc que $t \rightarrow (\exp(tM))$ est tangent à G en $t = 0$. D'après le lemme il vient donc que :

$$\forall t \in \mathbb{R}, (\exp(tM)) \in G$$

Réciproquement, prenons M dans G tel que $\exp(tM) \in G$ pour tout t de $\mathbb{R}$.
On définit au voisinage de I_n dans G :

$$\begin{aligned} D_{I_n} : \text{Eqv}_{I_n} &\rightarrow \mathbb{R} \\ f &\rightarrow \frac{d}{dt}(f(\exp(tM)))|_{t=0} \end{aligned}$$

Or, par le même argument que précédemment, puisque f est continue dans un voisinage de I_n , on crée :

$$\begin{aligned} D'_{I_n} : \text{Eqv}_{I_n} &\rightarrow \mathbb{R} \\ f &\rightarrow \frac{d}{dt}(f(I_n + tM))|_{t=0} \end{aligned}$$

Qui est une dérivée locale à G en I_n et donc définit par isomorphisme un champ vectoriel invariant à gauche. ■

Une telle correspondance entre la carte exponentielle et l'espace tangent à G en I_n est très forte car elle va permettre, dans les sous-groupes de Lie de $GL_n(\mathbb{C})$, de décrire l'algèbre de Lie associée avec cette application. Terminons donc notre vision des sous-groupes de Lie de $GL_n(\mathbb{C})$ par ce corollaire :

Corollaire 1 *Si G est un sous-groupe de Lie de $GL_n(\mathbb{C})$ on a la correspondance suivante :*

$$\text{Lie}(G) = \{M \in M_n(\mathbb{C}), \forall t \in \mathbb{R}, \exp(tM) \in G\}$$

5 Vers les algèbres de Lie

Nous allons momentanément laisser de côté les groupes de Lie pour nous attacher aux algèbres de Lie qui sont des structures purement algébriques. Cependant, après avoir étudié leurs premières caractéristiques nous établirons un lien entre groupe de Lie et leur algèbre.

5.1 Les algèbres de Lie

La définition de telles algèbres est relativement abordable mais elle permet, comme on le verra au chapitre suivant, d'être parfaitement adaptée pour parler de systèmes de racines, qui permet la classification des algèbres de Lie. Une telle classification sera la base de la physique de certains particules fondamentales, ce qui est l'objet des dernières parties de ce travail.

Définition 15 Une algèbre de Lie $\mathfrak{g}$ est un espace vectoriel sur lequel il y a une application de $\mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$ bilinéaire $[\cdot, \cdot]$ qui vérifie :

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

$$\forall (X, Y, Z) \in \mathfrak{g}^3, [X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$$

Regardons un exemple important car il va définir une telle application $[\cdot, \cdot]$ que nous prendrons comme référence plus loin.

Exemple 7 Si A est une algèbre associative on définit aisément une application bilinéaire vérifiant les propriétés des algèbres de Lie en posant :

$$[X, Y] = XY - YX$$

Munie de cette opération A devient une algèbre de Lie.

Il existe une notation importante. Lorsque l'algèbre A de l'exemple précédent est $M_n(\mathbb{C})$ (resp. $M_n(\mathbb{R})$) on note l'algèbre de Lie construite ci-dessus $\mathfrak{gl}_n(\mathbb{C})$ (resp. $\mathfrak{gl}_n(\mathbb{R})$).

Prenons une algèbre de Lie $\mathfrak{g}$. Alors pour tout élément x de $\mathfrak{g}$ nous pouvons définir la représentation adjointe de x qui est un endomorphisme de $\mathfrak{g}$ en tant qu'espace vectoriel. On le note $ad(x)$.

$$\begin{aligned} ad(x) : \mathfrak{g} &\rightarrow \mathfrak{g} \\ y &\rightarrow [x, y] \end{aligned}$$

5.2 Du groupe de Lie à l'algèbre de Lie

Si l'espace $\mathfrak{g}$ est de dimension fini alors nous pouvons parler de :

$$\forall (x, y) \in \mathfrak{g}^2, K(x, y) = \text{Tr}(ad(x) \circ ad(y))$$

Cette forme linéaire sur $\mathfrak{g}$ s'appelle la forme de Killing et elle est clairement bilinéaire, symétrique et invariante par l'action des automorphismes. De plus on a l'identité suivante (qui se trouve par simple écriture) :

$$K(x, [y, z]) = B([x, y], z)$$

Cette forme de Killing est importante dans la théorie de représentation des algèbres de Lie et elle nous servira par la suite. Un exemple de son utilité est dans les algèbres semi-simples, que nous utiliserons.

Définition 16 *On dit que $\mathfrak{g}$ est une algèbre de Lie semi-simple si elle ne contient aucun idéal résoluble non trivial.*

Remarque 6 *La semi-simplicité d'une algèbre de Lie équivaut à la non dégénérescence de la forme de Killing. Ce résultat étant ici uniquement à but d'exemple nous ne le démontrerons pas.*

Il existe d'autres classifications des algèbres de Lie : algèbres de Lie nilpotentes, résolubles... Mais notre étude ne requiert que la semi-simplicité donc nous ne nous étendrons pas à décrire ces notions.

5.2 Du groupe de Lie à l'algèbre de Lie

Nous avons vu lors du chapitre précédent que l'espace tangent à un groupe de Lie G un son élément unité est très important. Il est isomorphe à l'espace des champs vectoriels invariants à gauche, que nous appelons $\text{Lie}(G)$ et de dimension la dimension de la variété différentielle G . Notre but ici est de montrer que cet espace tangent est une algèbre de Lie.

Théorème 8 *Pour tout groupe de Lie G , $\text{Lie}(G)$ muni de $[\cdot, \cdot]$ est une algèbre de Lie.*

Ce théorème n'est pas aisé de prime abord mais nous allons voir que les notions étudiées dans notre description des groupes de Lie et leur espace tangent nous a permis de démontrer la plus grande partie de ce théorème.

Nous avons vu que $\text{Lie}(G)$ s'identifie à l'espace tangent $T_1(G)$ et que pour les sous-groupe de Lie de $GL_n(\mathbb{C})$ il s'identifie également à

$$\{M \in M_n(\mathbb{C}), \forall t \in \mathbb{R}, \exp(tM) \in G\}$$

5.2 Du groupe de Lie à l'algèbre de Lie

Cette propriété supplémentaire nous pousse à étudier d'abord les sous-groupes de Lie de $GL_n(\mathbb{C})$ puis le cas général.

Rappelons que dans chacune de nos démonstrations nous notons $[\cdot, \cdot]$ le produit bilinéaire d'algèbre :

$$[X, Y] = XY - YX$$

Comme $Lie(G)$ est un espace vectoriel pour tout groupe de Lie G notre but est de montrer que $Lie(G)$ est stable par $[\cdot, \cdot]$.

5.2.1 A partir des sous-groupes de Lie de $GL_n(\mathbb{C})$

Démontrons la propriété suivante :

Proposition 14 *Pour G un sous-groupe de Lie de $GL_n(\mathbb{C})$ et X et Y de $Lie(G)$ on a $[X, Y] \in Lie(G)$.*

Preuve Par définition de $Lie(G)$ et par construction de l'exponentielle on a de suite que pour X dans $Lie(G)$:

$$\exp(tX) \in Lie(G) \implies \exp(tX)Y\exp(-tX) \in Lie(G)$$

Puis :

$$\forall t \in \mathbb{R}, \frac{1}{t}(\exp(tX)Y\exp(-tX) - Y) \in Lie(G)$$

Or nous avons :

$$\begin{aligned} \frac{1}{t}(\exp(tX)Y\exp(-tX) - Y) &= XY - YX + \frac{t}{2!}(X^2Y - 2XYX + YX^2) + \frac{t^2}{3!}(\dots) + \dots \\ &= XY - YX + tP(t) \end{aligned}$$

Où P est un polynôme en t . Donc par passage à la limite lorsque t tend vers 0 on obtient que $[X, Y]$ est dans $Lie(G)$. ■

Dès lors, nous avons que $Lie(G)$ est une sous-algèbre de Lie de l'algèbre de Lie associée à $M_n(\mathbb{C})$, $\mathfrak{gl}_n(\mathbb{C})$. Par densité de $GL_n(\mathbb{C})$ dans $M_n(\mathbb{C})$ on obtient de suite le théorème suivant.

Théorème 9

$$Lie(GL_n(\mathbb{C})) = \mathfrak{gl}_n(\mathbb{C})$$

Remarque 7 *Ce théorème est toujours vrai dans $\mathbb{R}$ car aucun de nos arguments ne prenaient en compte des propriétés de $\mathbb{C}$ non vérifiées dans $\mathbb{R}$.*

5.2 Du groupe de Lie à l'algèbre de Lie

Nous avons donc réussi, pour les sous-groupes de Lie de $GL_n(\mathbb{C})$, à associer une algèbre de Lie et nous avons prouvé que l'algèbre de Lie de $M_n(\mathbb{C})$ est l'algèbre de Lie de $GL_n(\mathbb{C})$. Nous voulons également trouver l'algèbre de Lie associée à $SU_n(\mathbb{C})$ puisque ces groupes sont à la base de la physique des particules élémentaires.

Définition 17 On note $\mathfrak{sl}_n(\mathbb{C})$ le sous-espace vectoriel de $\mathfrak{gl}_n(\mathbb{C})$ tel que la trace de ses éléments soit nulle.

$$\mathfrak{sl}_n(\mathbb{C}) = \{X \in \mathfrak{gl}_n(\mathbb{C}), \text{Tr}(X) = 0\}$$

La trace étant linéaire et comme $\text{Tr}(AB) = \text{Tr}(BA)$ on a de suite que pour X et Y dans $\mathfrak{sl}_n(\mathbb{C})$, $\text{Tr}([X, Y]) = 0$. Donc il vient que $\mathfrak{sl}_n(\mathbb{C})$ est une sous-algèbre de $\mathfrak{gl}_n(\mathbb{C})$.

Théorème 10 L'algèbre de Lie de $SU_n(\mathbb{C})$ est $\mathfrak{sl}_n(\mathbb{C})$.

Preuve

Toute matrice de $M_n(\mathbb{C})$ est trigonalisable supérieurement. Le produit de deux matrices triangulaires supérieures donne une matrice triangulaire supérieure. Dès lors il vient que :

$$\det(\exp(M)) = \exp(\text{Tr}(M))$$

Prenons donc un X dans $\text{Lie}(SU_n(\mathbb{C}))$ on a donc :

$$\forall t \in \mathbb{R}, \exp(tX) \in SU_n(\mathbb{C})$$

Dès lors on a :

$$\begin{aligned} \forall t \in \mathbb{R}, 1 &= \det(\exp(tX)) \\ &= \exp(t\text{Tr}(X)) \end{aligned}$$

Il vient donc $\text{Tr}(X) = 0$ et $X \in \mathfrak{sl}_n(\mathbb{C})$.

Réciproquement, en prenant une matrice Y dans $\mathfrak{sl}_n(\mathbb{C})$. Alors pour tout t de $\mathbb{R}$ nous avons :

$$\det(\exp(tY)) = 1$$

De plus,

$$[\exp(tY)]^{-1} = [\exp(-tY)] = {}^t[\exp(tY)]$$

Donc on a bien $\exp(tY) \in SU_n(\mathbb{C})$. ■

Ceci termine notre étude précise des algèbres et sous-algèbres dans $M_n(\mathbb{C})$. Nous allons démontrer le théorème d'association d'une algèbre de Lie à un groupe de Lie dans le cas général.

5.2.2 Le cas général

Dans le cas général, $Lie(G)$ est l'ensemble des champs vectoriels invariants à gauche. Cependant, il est identifiable à l'espace tangent $T_1(G)$. Cet espace est clairement stable par $[\cdot, \cdot]$ puisque la composition de dérivées locales sur Eqv_1 reste une dérivée locale, ainsi que l'addition et la multiplication par un scalaire. Nos propositions générales de la partie 3 donnent donc immédiatement le résultat.

Notre étude pour $GL_n(\mathbb{C})$ semble donc fastidieuse. Seulement, elle est très importante car elle met en jeu la carte de l'exponentielle qui nous a déjà servi à établir la correspondance entre $Lie(G)$ et l'ensemble des matrices telles que $\exp(tX) \in G$. Une réalité, la carte exponentielle joue un grand rôle dans la théorie des groupes de Lie sur lesquels il est possible de construire une application exponentielle. Cette théorie continue directement ce chapitre car elle permet de nombreuses correspondances. Nous allons cependant ne pas la développer car elle nécessite plus de connaissance en terme de variétés différentielles.

6 Etude des algèbres de Lie semi-simples

Après avoir présenté les algèbres de Lie, nous allons en étudier un cas particulier : les algèbres de Lie semi-simples. Dans cette section, L désignera une algèbre de Lie semi-simple. On va étudier **la structure** de L via sa représentation adjointe.

On utilisera l'exemple de $sl_2(\mathbb{R})$ pour avoir une idée de ce qui se passe en admettant qu'il s'agit d'une algèbre de Lie semi-simple.

Notre objectif est de montrer qu'il est possible d'associer un système de racines et un espace euclidien à toute algèbre de Lie semi-simple. Cela nécessite un développement assez conséquent et l'introduction de quelques notions algébriques. Nous n'en ferons qu'esquisser la démarche¹.

Rappels :

- On dit qu'une algèbre de Lie est semi-simple si elle ne contient pas d'idéal résoluble non trivial.
- On appelle représentation adjointe que l'on note ad l'application $L \rightarrow \mathfrak{gl}(L)$ qui à x fait correspondre $ad(x)$. $ad(x)$ désigne l'endomorphisme de L qui à y fait correspondre $ad(x)(y) = [xy]$.
- On appelle forme de Killing sur l'algèbre de Lie L la forme bilinéaire symétrique κ qui à $(x, y) \in L^2$ associe $\kappa(x, y) = Tr(ad(x) ad(y))$. Cette application de Killing est associative au sens où $\kappa(x, [yz]) = \kappa([xy], z)$.

6.1 Tore maximal et décomposition en système de racines

Définition 18 Soit V est un $\mathbb{R}$ -espace vectoriel. On dit qu'un élément $x \in End V$ est **semi-simple** si et seulement si x est diagonalisable.

Définition 19 On appelle **tore** de L une sous-algèbre de L non nulle constituée d'éléments semi-simples.

On peut montrer qu'un tore de L est abélien et qu'il existe bien un tore de L .

Considérons le tore maximal H de L , i.e. le tore qui est strictement inclus dans aucun autre tore.

Par exemple, dans le cas où $L = \mathfrak{sl}_n(\mathbb{R})$ (qui rappelons-le désigne l'ensemble des

1. On pourra lire les deux premiers chapitres de [Humphreys] pour plus de précision

6.1 Tore maximal et décomposition en système de racines

matrices $n \times n$ de trace nulle), on peut vérifier que le tore maximal H est l'ensemble des matrices diagonales de trace nulle.

Puisque H est abélien, $ad(H) = \{ad(x) \in \mathfrak{gl}(L) / x \in H\}$ est une famille d'endomorphismes de L (semi-simples par définition du tore) qui commutent entre eux.

En effet, soient x et y deux éléments de H . Soit $u \in L$. Alors on a $(ad(x)) \circ (ad(y))(u) = ad(x)(ad(y)(u)) = ad(x)([yu]) = [x[yu]]$. Mais par définition du commutateur d'une algèbre de Lie, on a $[ab] = -[ba]$ pour tout $a, b \in L$. Comme $xy = yx$, on a $[xy] = [yx] = 0$.

Dans les axiomes d'une algèbre de Lie, le commutateur doit également vérifier la relation (qu'on appelle identité de Jacobi) $[a[bc]] + [b[ca]] + [c[ab]] = 0$.

Dans notre cas, cela donne $[x[yu]] + [y[ux]] + [z[xy]] = 0$. Comme $[z[xy]] = 0$, on a

$$\begin{aligned} [x[yu]] + [y[ux]] &= 0 \\ \Leftrightarrow [x[yu]] &= -[y[ux]] \\ \Leftrightarrow [x[yu]] &= [(-y)(-xu)] \text{ car } [.,.] \text{ est bilinéaire} \\ \Leftrightarrow [x[yu]] &= [y[xu]] \end{aligned}$$

Ainsi, $ad(x)$ et $ad(y)$ commutent.

On peut alors montrer que $ad(H)$ est simultanément diagonalisable. En d'autres termes, il existe une base de L dans laquelle $ad(x)$ est diagonalisable pour tout $x \in H$. L est alors la somme directe des espaces

$$L_\alpha = \{x \in L / [hx] = \alpha(h)x, \text{ pour tout } h \in H\}$$

α est une forme linéaire sur H et donc un élément du dual de H noté H^* .

$L_0 = \{x \in L / [hx] = 0, \text{ pour tout } h \in H\}$ est ce qu'on appelle le **centralisateur** de H noté $C_L(H)$. Si $x \in H$ alors $[hx] = 0$ pour tout $h \in H$. On en déduit que $H \subset C_L(H)$.

L'ensemble des éléments non nuls $\alpha \in H^*$ pour lesquels $L_\alpha \neq 0$ est noté Φ . Les éléments de Φ sont appelés **les racines** de L relativement à H (elles sont en nombre fini).

6.1 Tore maximal et décomposition en système de racines

Avec cette notation, on a une **décomposition en système de racines** (appelée encore **décomposition de Cartan**) :

$$(*) \quad L = C_L(H) \oplus \coprod_{\alpha \in \Phi} L_\alpha$$

Si $L = \mathfrak{sl}_2(\mathbb{R})$ par exemple, $(*)$ correspond à la décomposition de L dans la base standard

$$x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3)$$

Le but est ensuite de montrer que Φ caractérise L complètement.

Proposition 15 *Pour tout $\alpha, \beta \in H^*$, $[L_\alpha L_\beta] \subset L_{\alpha+\beta}$. Si $\alpha, \beta \in H^*$ et $\alpha + \beta \neq 0$, alors L_α est orthogonal à L_β relativement à la forme de Killing κ de L .*

Preuve Soient $x \in L_\alpha$ et $y \in L_\beta$. Si $h \in H$ alors on a, d'après l'identité de Jacobi, $ad(H)([xy]) = [h[xy]] = [[yh]x] + [[hx]y] = [x[hy]] + [[hx]y] = \alpha(h)[xy] + \beta(h)[xy] = (\alpha + \beta)(h)[xy]$. Donc $[xy] \in L_{\alpha+\beta}$.

Montrons maintenant la deuxième partie de la proposition. Considérons $h \in H$ tel que $(\alpha + \beta)(h) \neq 0$. Alors, si $x \in L_\alpha$, $y \in L_\beta$, la bilinéarité et l'associativité de la forme de Killing permettent d'écrire $\kappa([hx], y) = -\kappa([xh], y) = -\kappa(x, [hy])$. Ceci implique alors $\alpha(h)\kappa(x, y) = -\beta(h)\kappa(x, y)$ soit $(\alpha + \beta)(h)\kappa(x, y) = 0$ et donc $\kappa(x, y) = 0$. Tout élément de L_α est orthogonal à tout élément de L_β donc L_α est orthogonal à L_β . ■

Nous avons vu que l'on a la relation $H \subset C_L(H)$. Nous admettrons la proposition suivante.

Proposition 16 *Soit H un tore maximal de L . Alors $H = C_L(H)$.*

Il est possible d'identifier H^* à H de la façon suivante : à chaque $\phi \in H^*$, on fait correspondre l'unique élément $t_\phi \in H$ satisfaisant $\phi(h) = \kappa(t_\phi, h)$ pour tout $h \in H$ (de façon analogue au théorème de représentation de riesz).

En particulier, à tout élément $\alpha \in \Phi$ (rappelons qu'alors α est dans H^*), on fait correspondre l'unique élément t_α tel que $\alpha(h) = \kappa(t_\alpha, h)$ pour tout $h \in H$. Φ correspond donc à l'ensemble $\{t_\alpha; \alpha \in \Phi\}$ de H .

6.2 Propriétés de la décomposition en système de racines

Nous présenterons quelques résultats de la décomposition en système de racines.

Proposition 17 (a) Φ engendre H^* .

(b) Si $\alpha \in \Phi$ alors $-\alpha \in \Phi$.

(c) Soit $\alpha \in \Phi$, $x \in L_\alpha$, $y \in L_{-\alpha}$. Alors $[xy] = \kappa(x, y)t_\alpha$.

(d) Si $\alpha \in \Phi$ alors $[L_\alpha L_{-\alpha}]$ est de dimension 1 et de base t_α .

Preuve (a) Si Φ n'engendre pas H^* , alors il existe un élément non nul du tore maximal H tel que $\alpha(h) = 0$, pour tout $\alpha \in H$ (raisonner par dualité pour obtenir ce résultat). Dans ce cas, si $x \in L_\alpha$ alors $[h x] = 0$. Ceci entraîne $[h L_\alpha] = 0$ pour tout $\alpha \in L_\alpha$ et donc en particulier, $[h H] = 0$. On peut montrer que ceci implique que h appartienne au centre de L qui est trivial dans le cas où L est semi-simple.

(b) Soit $\alpha \in \Phi$. Supposons $-\alpha \notin \Phi$ (i.e. $L_{-\alpha} = \{0\}$) alors pour tout $\beta \in H^*$, $\alpha + \beta$ est non nul dans H^* . On peut donc utiliser le résultat montré un peu plus haut pour dire qu'on a $\kappa(L_\alpha, L_\beta) = 0$ pour tout $\beta \in H^*$. D'après la décomposition de Cartan de L , ceci entraîne $\kappa(L_\alpha, L) = 0$. Or on peut montrer que la restriction de κ à H est **non dégénérée**, i.e. que son **radical** S est nul où $S = \{x \in L / \kappa(x, y) = 0 \text{ pour tout } y \in L\}$. Les deux propriétés sont contradictoires et notre hypothèse de départ était donc fausse. Ainsi, $-\alpha \in \Phi$.

(c) Soit $\alpha \in \Phi$. D'après ce qui précède, $-\alpha \in \Phi$. L_α et $L_{-\alpha}$ ne sont pas réduits à $\{0\}$. Soient $x \in L_\alpha$, $y \in L_{-\alpha}$. Prenons $h \in H$ quelconque. L'associativité, la bilinéarité et la symétrie de κ impliquent : $\kappa(h, [xy]) = \kappa([hx], y) = \alpha(h)\kappa(x, y) = \kappa(t_\alpha, h)\kappa(x, y) = \kappa(\kappa(x, y)t_\alpha, h) = \kappa(h, \kappa(x, y)t_\alpha)$. On a donc $\kappa(h, \kappa(x, y)t_\alpha - [xy]) = 0$. H est donc orthogonal à $\kappa(x, y)t_\alpha - [xy]$. Ceci impose $\kappa(x, y)t_\alpha - [xy] = 0$ et donc $\kappa(x, y)t_\alpha = [xy]$ car κ est non dégénérée.

(d) D'après (c), t_α engendre $[L_\alpha L_{-\alpha}]$. Montrons que $[L_\alpha L_{-\alpha}]$ est non réduit à 0. Soit $x \in L_\alpha$, $x \neq 0$. Si $\kappa(x, L_{-\alpha}) = 0$, on peut montrer qu'alors on a $\kappa(x, L) = 0$ (?). Encore une fois, ceci est absurde car κ est non dégénérée. Aussi, on peut trouver $y \in L_{-\alpha}$, $y \neq 0$ tel que $\kappa(x, y) \neq 0$. D'après (c), on a alors $[xy] \neq 0$. ■ Nous ne démontrerons pas d'autres résultats concernant cette décomposition en système de racines. L'objectif était de donner un aperçu de quelques propriétés de Φ .

Pour terminer cette partie de présentation de la structure d'une algèbre de Lie

6.2 Propriétés de la décomposition en système de racines

semi-simple, énonçons un théorème regroupant les propriétés caractéristiques d'un système de racines d'une algèbre semi-simple.

Théorème 11 *Soit donc L une algèbre de Lie semi-simple (sur un corps algébriquement clos F de caractéristique 0) de tore maximal H et de système de racine Φ . Il est possible de construire un espace euclidien E tel que :*

- (a) Φ engendre E et 0 n'appartient pas à Φ .
- (b) Si $\alpha \in \Phi$ alors $-\alpha \in \Phi$. Par contre, les autres multiples de α par des scalaires ne sont pas dans Φ .
- (c) Si $\alpha, \beta \in \Phi$ alors $\beta - \frac{2(\beta, \alpha)}{(\alpha, \alpha)}\alpha \in \Phi$.
- (d) Si $\alpha, \beta \in \Phi$ alors $\frac{2(\beta, \alpha)}{(\alpha, \alpha)} \in \mathbb{Z}$.

Dans cette partie, nous avons observé qu'il était possible d'établir une correspondance $(L, H) \mapsto (\Phi, E)$ entre une algèbre de Lie semi-simple muni de son tore maximale et son ensemble de racines. On construit ensuite un espace euclidien à partir de Φ . Nous allons maintenant nous attacher à décrire un système de racines quelconque dans l'optique de le représenter. Nous laisserons donc momentanément de côté la notion d'algèbre de Lie semi-simple en redéfinissant un système de racines à partir d'axiomes algébriques. Bien entendu, le système de racines d'une algèbre de Lie semi-simple vérifiera ces axiomes.)

7 Etude des systèmes de racines

Dans ce paragraphe, nous travaillerons dans un espace euclidien E , c'est à dire dans un $\mathbb{R}$ -espace vectoriel de dimension finie muni d'un produit scalaire que nous noterons $(\cdot, \cdot)$. Géométriquement, une réflexion dans E est une transformation linéaire inversible dont la restriction à un hyperplan est égale à l'identité et qui envoie tout vecteur normal à cet hyperplan en son opposé. Tout vecteur non nul α de E détermine une réflexion σ_α par rapport à l'hyperplan $P_\alpha = \{\beta \in E / (\beta, \alpha) = 0\}$. On peut montrer que, pour tout vecteur β de E , on a $\sigma_\alpha(\beta) = \beta - \frac{2(\beta, \alpha)}{(\alpha, \alpha)}\alpha$ (on a bien $\sigma_\alpha(\alpha) = -\alpha$ et $\sigma_\alpha(\beta) = \beta$ pour tout $\beta \in P_\alpha$). Nous noterons $\langle \beta, \alpha \rangle$ le rapport $\frac{2(\beta, \alpha)}{(\alpha, \alpha)}$ que nous retrouverons par la suite.

7.1 Définition d'un système de racines

Lemme 2 *Soit Φ une partie finie de E engendrant E . Pour tout $\alpha \in \Phi$ tel que $\alpha \neq 0$, il existe au plus une réflexion s de E telle que $s(\alpha) = -\alpha$ et $s(R) = R$. Nous la noterons σ_α .*

Preuve . Soit G le groupe des automorphismes sur E qui laissent Φ stable. Comme Φ engendre E , les éléments de G sont caractérisés par leur action sur les vecteurs de Φ . G est isomorphe à un sous-groupe du groupe symétrique de Φ . G est donc de cardinal fini.

Considérons deux réflexions s, s' de E telles que $s(\alpha) = s'(\alpha) = -\alpha$, $s(\Phi) = \Phi$, $s'(\Phi) = \Phi$. Alors $t = s \circ s'$ appartient à G , donc est d'ordre fini m . On a $t(\alpha) = \alpha$ et comme s et s' sont égales à l'identité sur l'hyperplan P_α , pour tout élément x de E , il existe $k \in \mathbb{R}$ tel que $t(x) = x + k\alpha$.

On peut alors construire une forme linéaire f telle que

$$t(x) = x + f(x)\alpha, \forall x \in E$$

On a alors

$$t^2(x) = t(x + f(x)\alpha) = t(x) + f(x)\alpha = x + 2f(x)\alpha, \forall x \in E$$

Par récurrence sur n on trouve

$$t^n(x) = x + nf(x)\alpha, \forall x \in E$$

En particulier, $x = t^m(x) = x + mf(x)\alpha, \forall x \in E$. Ceci implique $mf(x) = 0$ pour tout $x \in E$, d'où $f = 0$, $t = id$ et $s = s'$. ■

Définition 20 Soit Φ une partie de l'espace euclidien E . On dit que Φ est un système de racines dans E si les conditions suivantes sont vérifiées :

- (R1) Φ est fini, ne contient pas 0 et engendre E .
- (R2) Si $\alpha \in \Phi$, les seuls multiples de α qui sont dans Φ sont α et $-\alpha$.
- (R3) Si $\alpha \in \Phi$, la réflexion σ_α laisse stable Φ .
- (R4) Si $\alpha, \beta \in \Phi$ alors $\langle \beta, \alpha \rangle \in \mathbb{Z}$.

Les éléments de Φ sont appelés les **racines** (du système considéré). La dimension de l'espace euclidien E s'appelle le **rang** du système.

Notons W le sous-groupe des automorphismes de E engendré par les réflexions σ_α ($\alpha \in \Phi$). D'après (R3), les éléments de W permutent les vecteurs de Φ . W s'identifie donc à un sous-groupe du groupe symétrique de Φ . W est donc de cardinal fini. W est appelé le **groupe de Weyl** de Φ .

Proposition 18 Soit Φ un système de racines dans E de groupe de Weyl W . Si $\sigma \in GL(E)$ laisse Φ invariant alors $\sigma\sigma_\alpha\sigma^{-1} = \sigma_{\sigma(\alpha)}$ pour tout $\alpha \in \Phi$ et $\langle \beta, \alpha \rangle = \langle \sigma(\beta), \sigma(\alpha) \rangle$ pour tout $\alpha, \beta \in \Phi$.

Preuve $\sigma\sigma_\alpha\sigma^{-1}(\sigma(\beta)) = \sigma\sigma_\alpha(\beta) \in \Phi$ puisque $\sigma_\alpha(\beta) \in \Phi$. Comme $\sigma(\beta)$ parcourt Φ quand β parcourt Φ , on en déduit que $\sigma\sigma_\alpha\sigma^{-1}$ laisse Φ invariant.

Si $x \in \sigma(P_\alpha)$ alors $x = \sigma(y)$ avec $y \in P_\alpha$. On a alors

$$\sigma\sigma_\alpha\sigma^{-1}(x) = \sigma\sigma_\alpha\sigma^{-1}(\sigma(y)) = \sigma\sigma_\alpha(y) = \sigma(y) = x$$

La restriction de $\sigma\sigma_\alpha\sigma^{-1}$ à $\sigma(P_\alpha)$ est donc égale à l'identité.

Par ailleurs, $\sigma\sigma_\alpha\sigma^{-1}(\sigma(\alpha)) = \sigma\sigma_\alpha(\alpha) = \sigma(-\alpha) = -\sigma(\alpha)$. On peut donc appliquer le lemme précédent qui nous indique que $\sigma\sigma_\alpha\sigma^{-1} = \sigma_{\sigma(\alpha)}$.

Montrons maintenant la deuxième partie de la proposition.

On peut écrire $\sigma\sigma_\alpha\sigma^{-1}(\sigma(\beta)) = \sigma\sigma_\alpha(\beta) = \sigma(\beta - \langle \beta, \alpha \rangle \alpha) = \sigma(\beta) - \langle \beta, \alpha \rangle \sigma(\alpha)$.

Mais on a aussi $\sigma\sigma_\alpha\sigma^{-1}(\sigma(\beta)) = \sigma_{\sigma(\alpha)}(\sigma(\beta)) = \sigma(\beta) - \langle \sigma(\beta), \sigma(\alpha) \rangle \sigma(\alpha)$.

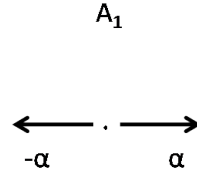
On obtient ainsi le résultat souhaité. ■

7.2 Exemples

- Système de racines de rang 1

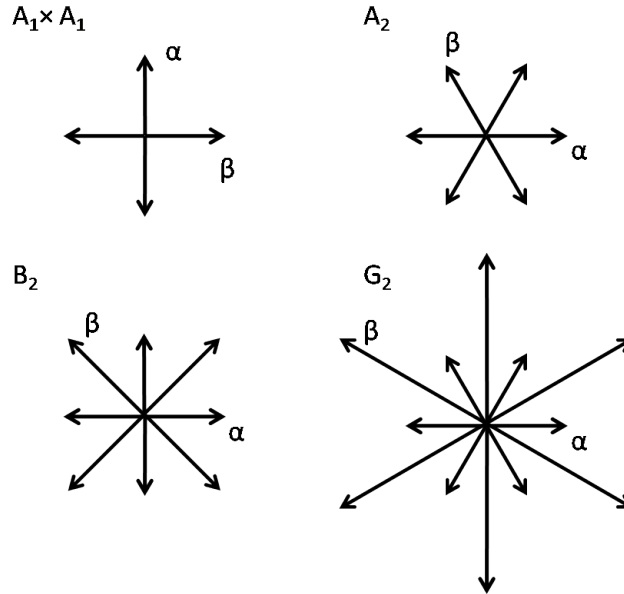
Cette situation se présente lorsque E est de dimension 1 (E est une droite

vectorielle). Si α est un vecteur non nul de E alors $\{\alpha, -\alpha\}$ est un système de racines de E . Le groupe de Weyl de ce système de racines est d'ordre 2.



– Système de racines de rang 2

E est de dimension 2 (E est un plan). Il y a dans ce cas plusieurs possibilités pour le système de racines de Φ . Nous verrons dans la suite qu'il n'y a en fait que ces quatre types de systèmes de racines de rang 2.



7.3 Paires de racines

L'axiome (R4) limite considérablement les valeurs pouvant être prises par les angles entre les paires de racines. Le cosinus de l'angle θ entre les vecteurs α et β est donné par la relation $\|\alpha\| \|\beta\| \cos \theta = (\alpha, \beta)$. Ainsi, on a $\langle \beta, \alpha \rangle = \frac{2(\beta, \alpha)}{(\alpha, \alpha)} = 2 \frac{\|\beta\|}{\|\alpha\|} \cos \theta$. Ceci implique $\langle \beta, \alpha \rangle \langle \alpha, \beta \rangle = 4 \cos^2 \theta$. Comme $\langle \beta, \alpha \rangle \langle \alpha, \beta \rangle$ doit être un entier,

7.4 Racines simples et groupe de Weyl

$\langle \beta, \alpha \rangle \langle \alpha, \beta \rangle$ ne peut prendre que quatre valeurs : 0, 1, 2, 3 ou 4. $\langle \beta, \alpha \rangle$ et $\langle \alpha, \beta \rangle$ ont même signe. Si on prend $\alpha \neq \pm \beta$ et $\|\beta\| \geq \|\alpha\|$, $\langle \beta, \alpha \rangle \langle \alpha, \beta \rangle \in \{0, 1, 2, 3\}$. Supposons par exemple, $\langle \beta, \alpha \rangle \langle \alpha, \beta \rangle = 2$. Alors, ou bien $\langle \beta, \alpha \rangle = 2$ et $\langle \alpha, \beta \rangle = 1$ ou bien $\langle \beta, \alpha \rangle = -2$ et $\langle \alpha, \beta \rangle = -1$. Dans le premier cas, on a $2 = \langle \beta, \alpha \rangle \langle \alpha, \beta \rangle = 4\cos^2\theta$, soit $\cos\theta = \frac{\sqrt{2}}{2}$. En se limitant à un angle $\theta \in]0; \pi[$, on a $\theta = \frac{\pi}{4}$.

En utilisant ce procédé, on peut construire la table suivante :

$\langle \alpha, \beta \rangle$	$\langle \beta, \alpha \rangle$	θ	$\frac{\ \beta\ ^2}{\ \alpha\ ^2}$
0	0	$\frac{\pi}{2}$	—
1	1	$\frac{\pi}{3}$	1
-1	-1	$\frac{2\pi}{3}$	1
1	2	$\frac{\pi}{4}$	2
-1	-2	$\frac{3\pi}{4}$	2
1	3	$\frac{\pi}{6}$	3
-1	-3	$\frac{5\pi}{6}$	3

Les quatre situations représentées sur la figure précédente correspondent donc à tous les systèmes de racines de rang 2.

Lemme 3 Soient α et β deux racines non proportionnelles. Si $(\alpha, \beta) > 0$ alors $\alpha - \beta$ est une racine. Si $(\alpha, \beta) < 0$ alors $\alpha + \beta$ est une racine.

Preuve (α, β) est positif si et seulement si $\langle \alpha, \beta \rangle$ l'est. Par conséquent, d'après la table que nous avons construite, on a nécessairement $\langle \alpha, \beta \rangle$ ou $\langle \beta, \alpha \rangle$ qui vaut 1. Si $\langle \alpha, \beta \rangle = 1$ alors $\sigma_\beta(\alpha) = \alpha - \langle \alpha, \beta \rangle \beta = \alpha - \beta \in \Phi$ (d'après (R3)). De la même façon, si $\langle \beta, \alpha \rangle = 1$ alors $\sigma_\alpha(\beta) = \beta - \alpha \in \Phi$.

La deuxième partie du lemme se montre en remarquant que $\alpha - \beta = \alpha + (-\beta)$. Comme $-\beta \in \Phi$ d'après (R2) et $(\alpha, -\beta) = -(\alpha, \beta) > 0$ si $(\alpha, \beta) < 0$, on a, d'après ce qui vient d'être montré, $\alpha - (-\beta) = \alpha + \beta \in \Phi$. ■

7.4 Racines simples et groupe de Weyl

Dans cette partie, nous considérerons Φ un système de racines de rang l dans un espace euclidien E , de groupe de Weyl W

7.5 Bases et chambres de Weyl

Définition 21 Un sous-ensemble Δ de Φ est une **base** si :

- (B1) Δ est une base de E
- (B2) toute racine β peut être écrite sous la forme $\beta = \sum_{\alpha \in \Delta} k_{\alpha} \alpha$ où les coefficients k_{α} sont dans $\mathbb{Z}$, tous positifs ou tous négatifs.

Les éléments de Δ sont appelés **racines simples**.

D'après (B1), on a $\text{card } \Delta = l$ et β s'exprime de façon unique comme combinaison des racines simples. On peut définir la **hauteur** d'une racine (relativement à Δ) $ht \beta = \sum_{\alpha \in \Delta} k_{\alpha}$. Si pour tout α dans Δ on a $k_{\alpha} \geq 0$ (respectivement $k_{\alpha} \leq 0$), nous dirons que β est une **racine positive** (respectivement **racine négative**) et nous écrirons $\beta \succ 0$ (respectivement $\beta \prec 0$).

Nous noterons Φ^+ l'ensemble des racines positives et Φ^- l'ensemble des racines négatives. On a la relation $\Phi^- = -\Phi^+$.

Lemme 4 Si Δ est une base de Φ alors $(\alpha, \beta) \leq 0$ pour $\alpha \neq \beta$ dans Δ . En outre, $\alpha - \beta$ n'est pas une racine.

Preuve Si $(\alpha, \beta) > 0$ alors on ne peut avoir $\alpha \neq -\beta$. D'autre part, par hypothèse, $\alpha \neq \beta$. Le lemme précédent nous assure alors que $\alpha - \beta$ est une racine. Ceci est contraire à l'hypothèse (B2). ■ Nous allons maintenant montrer que tout

système de racines possède une base.

Théorème 12 Φ possède une base.

Pour tout vecteur $\gamma \in E$, définissons $\Phi^+(\gamma) = \{\alpha \in \Phi \mid (\gamma, \alpha) > 0\}$ l'ensemble des racines situées du côté positif de l'hyperplan orthogonal à γ . L'union des hyperplans P_{α} ($\alpha \in \Phi$) ne peut pas recouvrir E . Nous dirons que $\gamma \in E$ est **régulier** si $\gamma \in E - \bigcup_{\alpha \in \Phi} P_{\alpha}$, **singulier** sinon.

Si γ est régulier, alors pour tout $\alpha \in \Phi$, $(\gamma, \alpha) \neq 0$ et donc α est dans $\Phi^+(\gamma)$ ou dans $-\Phi^+(\gamma)$. On en déduit $\Phi = \Phi^+(\gamma) \cup -\Phi^+(\gamma)$. Nous dirons que $\alpha \in \Phi^+(\gamma)$ est **décomposable** s'il existe β_1, β_2 éléments de $\Phi^+(\gamma)$ tels que $\alpha = \beta_1 + \beta_2$. Dans le cas contraire, nous dirons que α est **indécomposable**.

Théorème 13 Si γ est un élément régulier de E alors l'ensemble $\Delta(\gamma)$ de toutes les racines indécomposables de $\Phi^+(\gamma)$ est une base de Φ . En outre, toute base de Φ peut être obtenue de cette manière.

Preuve

1. Soit α une racine dans $\Phi^+(\gamma)$. Si α est indécomposable alors $\alpha \in \Delta(\Phi)$. Sinon $\alpha = \beta_1 + \beta_2$ avec $\beta_1, \beta_2 \in \Phi^+(\gamma)$. En redécomposant β_1 ou β_2 si nécessaire, on voit que toute racine de $\Phi^+(\gamma)$ est une $\mathbb{N}$ -combinaison linéaire d'éléments d'éléments de $\Delta(\Phi)$.
2. Si $\alpha, \beta \in \Delta(\gamma)$ alors $(\alpha, \beta) \leq 0$ sauf si $\alpha = \beta$. Dans le cas contraire (i.e. $(\alpha, \beta) > 0$), $\alpha - \beta$ serait une racine d'après le lemme démontré dans la partie sur les paires de racines. En effet, on ne peut avoir $\beta = -\alpha$ car alors il ne serait pas possible que les deux racines soient dans $\Phi^+(\gamma)$. $\alpha - \beta \in \Phi = \Phi^+(\gamma) \cup -\Phi^+(\gamma)$ donc $\alpha - \beta \in \Phi^+(\gamma)$ ou $\beta - \alpha \in \Phi^+(\gamma)$. Dans le premier cas, $\alpha = \beta + (\alpha - \beta)$ indique que α est décomposable, dans le second, $\beta = \alpha + (\beta - \alpha)$ indique que β est décomposable. Ceci est contraire à l'hypothèse de départ $\alpha, \beta \in \Delta(\gamma)$.
3. $\Delta(\gamma)$ est un ensemble linéairement indépendant. Supposons $\sum r_\alpha \alpha = 0$ ($\alpha \in \Delta(\gamma), r_\alpha \in \mathbb{R}$). Séparons les indices α pour lesquels $r_\alpha > 0$ de ceux pour lesquels $r_\alpha < 0$. Nous pouvons écrire $\sum s_\alpha \alpha = \sum r_\beta \beta$ avec $s_\alpha, t_\beta > 0$ et où les ensembles de racines α et β sont disjoints. Appelons $\epsilon = \sum s_\alpha \alpha$. On a $(\epsilon, \epsilon) = (\sum s_\alpha \alpha, \sum r_\beta \beta) = \sum_{\alpha, \beta} s_\alpha t_\beta (\alpha, \beta) \leq 0$ d'après l'étape précédente. $\|\epsilon\|^2 \leq 0 \Rightarrow \epsilon = 0$. On a alors $0 = (\gamma, \epsilon) = \sum s_\alpha (\gamma, \alpha) = \sum t_\beta (\gamma, \beta)$. Comme les α et les β sont dans $\Phi^+(\gamma)$ (donc $(\gamma, \alpha) > 0$ et $(\gamma, \beta) > 0$) et que les s_α et les t_β sont positifs, les s_α et les t_β sont nécessairement tous nuls. On vient là de montrer que tout ensemble de vecteurs situés strictement d'un même côté d'un hyperplan (le "strictement" veut simplement dire que les vecteurs de l'ensemble n'appartiennent pas à l'hyperplan) tel que toute paire de vecteurs de cet ensemble fasse un angle obtus (pour avoir $(\alpha, \beta) \leq 0$) est linéairement indépendant.
4. $\Delta(\gamma)$ est une base de Φ . En effet, puisque $\Phi = \Phi^+(\gamma) \cup -\Phi^+(\gamma)$, l'hypothèse (B2) pour que $\Delta(\gamma)$ soit une base de Φ est vérifiée d'après l'étape 1. Cette première étape nous indique également que $\Delta(\gamma)$ engendre E . L'étape 3 nous assure que $\Delta(\gamma)$ est bien une base E et donc que (B1) est aussi vérifiée.
5. Toute base Δ de Φ est de la forme $\Delta(\gamma)$ pour un vecteur $\gamma \in E$ régulier. Pour montrer cela, donnons-nous une base Δ et choisissons un vecteur $\gamma \in E$ tel que $(\gamma, \alpha) > 0$ pour tout $\alpha \in \Delta$. C'est possible car on peut montrer que l'intersection des demi-espaces "positifs" associés à une base de E est non vide. D'après (B2), γ est régulier et $\Phi^+ \subset \Phi^+(\gamma)$, $\Phi^- \subset -\Phi^+(\gamma)$. Comme $\Phi = \Phi^+(\gamma) \cup -\Phi^+(\gamma)$ et $\Phi = \Phi^+ \cup \Phi^-$, on a en fait $\Phi^+ = \Phi^+(\gamma)$. Si $x \in \Delta$ et $x = \alpha + \beta$ avec $\alpha, \beta \in \Phi^+(\gamma)$ alors $x = \alpha' + \beta'$ avec $\alpha', \beta' \in \Phi^+$. Ceci est contradictoire avec le fait que Δ doit être libre pour être une base de

E . x n'est pas donc pas décomposable et $\Delta \subset \Delta(\gamma)$. Mais comme $\text{card}\Delta = \text{card}\Delta(\gamma) = l$, $\Delta = \Delta(\gamma)$.

■

Nous allons introduire maintenant un peu de vocabulaire.

Soit H un hyperplan de E . $E - H$ a deux composantes connexes que l'on appelle les demi-espaces ouverts limités par H . Leurs adhérences s'appellent les demi-espaces fermés limités par H . Soient x, y deux éléments de E . On dit que x et y sont **situés strictement du même côté** de H s'ils sont contenus dans le même sous-espace ouvert limité par H . Ceci revient à dire que le segment $[x; y]$ ne rencontre pas H . On dit que x et y sont **situés de part et d'autre** de H si x appartient à l'un des demi-espaces ouverts limités par H , et y à l'autre.

Soit A une partie connexe non vide de E . Pour tout hyperplan H de E ne rencontrant pas A , on note $D_H(A)$ l'unique demi-espace ouvert limité par H qui contient A . Si M est un ensemble d'hyperplans de E dont aucun ne rencontre A , on pose

$$D_M(A) = \bigcap_{H \in M} D_H(A)$$

Soient x et y deux éléments de E et M une famille d'hyperplans de E . On définit la relation $xRy \Leftrightarrow$ Pour tout hyperplan $H \in M$, ou bien $x \in H$ et $y \in H$, ou bien x et y sont strictement du même côté de H .

On peut montrer que cette relation est une relation d'équivalence.

Définition 22 On appelle **facette** de E relativement à M toute classe selon la relation d'équivalence R .

Définition 23 On appelle **chambre** de E relativement à M toute facette de E relativement à M qui n'est contenue dans aucun hyperplan appartenant à M .

Soit U l'ensemble ouverts dans E formé des points qui n'appartiennent à aucun hyperplan de M . Comme un hyperplan de M ne peut rencontrer une facette sans la contenir, les chambres sont donc les facettes contenues dans U .

Définition 24 Soit C une chambre de E . On appelle **face** de C toute facette contenue dans l'adhérence de C et dont le support est un hyperplan. On appelle **mur** de C tout hyperplan qui est le support d'une face de C .

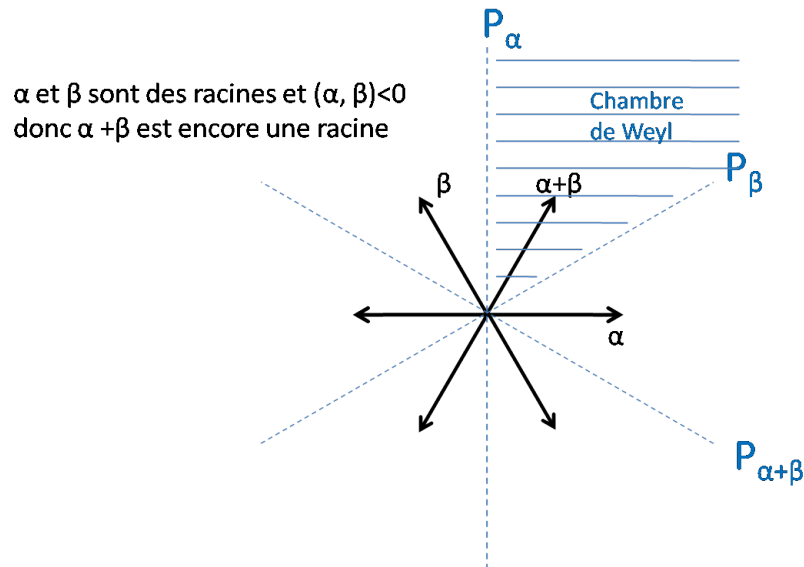
Tout mur de C appartient à M

Considérons la famille d'hyperplans $M = \{P_\alpha, (\alpha \in \Phi)\}$ (rappelons que P_α correspond à l'ensemble des éléments de E qui sont invariants par la réflexion σ_α). M divise E en un nombre fini de composantes connexes que l'on appelle **chambres de Weyl**.

Tout élément régulier $\gamma \in E$ appartient exactement à une chambre de Weyl, notée $C(\gamma)$. Dire $C(\gamma) = C(\gamma')$ signifie simplement que γ et γ' sont situés du même côté de chacun des hyperplans P_α ($\alpha \in \Phi$). Il est équivalent de dire $\Phi^+(\gamma) = \Phi^+(\gamma')$ ou $\Delta(\gamma) = \Delta(\gamma')$.

Le groupe de Weyl envoie les chambres de Weyl les unes dans les autres. Plus précisément, si $\sigma \in W$ et si γ est régulier alors $\sigma(C(\gamma)) = C(\sigma\gamma)$. D'autre part, W permute les bases : σ envoie Δ vers $\sigma(\Delta)$ qui est encore une base.

Représentons un exemple dans le plan vectoriel (système de racine d'ordre 2). Sur cet exemple, il y a 6 chambres de Weyl. Bien entendu, le nombre de chambre de weyl dépend du système de racines considéré. On peut le voir en représentant les chambres de Weyl pour les quatre types de systèmes de racines du plan.



7.6 Quelques résultats sur les racines simples

Soit Δ une base fixée de Φ .

Proposition 19 *Si α est une racine positive mais non simple alors $\alpha - \beta$ est une racine (nécessairement positive) pour un certain $\beta \in \Delta$.*

Preuve Si $(\alpha, \beta) \leq 0$ pour tout $\beta \in \Delta$ alors d'après ce qui a été montré dans l'étape 3 de la démonstration du théorème sur la construction des bases de Φ , l'ensemble $\{\alpha\} \cup \Delta$ est linéairement indépendant. Ceci n'est pas possible puisque Δ est une base de E . Ainsi donc, il existe $\beta \in \Delta$ tel que $(\alpha, \beta) > 0$. D'après le lemme démontré dans la partie sur les paires de racines, puisque α et β ne peuvent pas être proportionnels (Δ est une base de E), $\alpha - \beta$ est encore une racine et la proposition est vérifiée.

Ecrivons $\alpha = \sum_{\gamma \in \Delta} k_\gamma \gamma$ (avec tous les $k_\gamma \geq 0$, certains $k_\gamma > 0$ pour $\gamma \neq \beta$ car α et β ne peuvent pas être proportionnels). $\alpha - \beta$ est une racine avec au moins un coefficient strictement positif. Par conséquent, $\alpha - \beta$ est positive. ■

Corollaire 2 *Tout élément $\beta \in \Phi^+$ peut être écrit sous la forme $\alpha_1 + \alpha_2 + \dots + \alpha_k$ ($\alpha_i \in \Delta$, non nécessairement distincts) d'une façon telle que chaque somme partielle $\alpha_1 + \dots + \alpha_i$ soit une racine.*

Preuve $\beta \in \Phi^+$ donc $\beta = \sum_{\alpha \in \Delta} k_\alpha \alpha$ où les coefficients k_α sont dans $\mathbb{Z}$, tous positifs.

D'après la proposition précédente, il existe α_k dans Δ telle que $\beta - \alpha_k$ soit une racine positive. Dans ce cas, le coefficient associés à α_k dans la somme $\beta = \sum_{\alpha \in \Delta} k_\alpha \alpha$ est

forcément non nul car sinon $\beta - \alpha_k$ ne serait pas positive. En réitérant l'opération jusqu'à arriver à une racine simple, on obtient le résultat du corollaire. ■

Proposition 20 *Soit α une racine simple. Alors σ_α permute les racines positives autres que α .*

Preuve Soit $\beta \in \Phi^+ - \{\alpha\}$, $\beta = \sum_{\alpha \in \Delta} k_\alpha \alpha$ ($k_\gamma \in \mathbb{Z}^+$). D'après les axiomes d'un système de racines et puisque β est positive, β n'est pas proportionnelle à α . Ainsi, $k_\gamma \neq 0$ pour un certain $\gamma \neq \alpha$. Mais le coefficient de γ dans $\sigma_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha$ est encore k_γ . En d'autres termes, $\sigma_\alpha(\beta)$ a au moins un coefficient strictement positif

et donc, $\sigma_\alpha(\beta)$ est une racine positive. De plus, $\sigma_\alpha(\beta) \neq \alpha$ puisque α est l'image de $-\alpha$ et la réflexion σ_α est un automorphisme de E . ■

Proposition 21 Soient $\alpha_1, \alpha_2, \dots, \alpha_t \in \Delta$ (non nécessairement distinctes). Notons $\sigma_i = \sigma_{\alpha_i}$. Si $\sigma_1 \sigma_2 \dots \sigma_{t-1}(\alpha_t)$ est négative alors il existe un indice s , $1 \leq s < t$, tel que $\sigma_1 \sigma_2 \dots \sigma_t = \sigma_1 \sigma_2 \dots \sigma_{s-1} \sigma_{s+1} \dots \sigma_{t-1}$

Preuve Ecrivons $\beta_i = \sigma_{s+1} \dots \sigma_{t-1}(\alpha_t)$, $0 \leq i \leq t-2$ et $\beta_{t-1} = \alpha_t$. Par hypothèse, on a $\beta_0 \prec 0$ et $\beta_{t-1} \succ 0$. On peut alors trouver un plus petit indice s (≥ 1) pour lequel $\beta_s \succ 0$. On a alors $\sigma_s(\beta_s) = \beta_{s-1} \prec 0$. D'après le lemme précédent, on en déduit $\beta_s = \alpha_s$.

Dans un lemme précédent, nous avons vu que si $\sigma \in W$ alors $\sigma_{\sigma(\alpha)} = \sigma \sigma_\alpha \sigma^{-1}$. Ceci permet d'écrire $\sigma_s = \sigma_{\alpha_s} = \sigma_{\beta_s} = \sigma_{\sigma_{s+1} \dots \sigma_{t-1}(\alpha_t)} = (\sigma_{s+1} \dots \sigma_{t-1}) \sigma_{\alpha_t} (\sigma_{t-1} \dots \sigma_{s+1}) = (\sigma_{s+1} \dots \sigma_{t-1}) \sigma_t (\sigma_{t-1} \dots \sigma_{s+1})$.

On a alors :

$$\sigma_1 \dots \sigma_t = \sigma_1 \dots \sigma_{s-1} \sigma_s \sigma_{s+1} \dots \sigma_{t-1} = \sigma_1 \dots \sigma_{s-1} (\sigma_{s+1} \dots \sigma_{t-1}) \sigma_t (\sigma_{t-1} \dots \sigma_{s+1}) \sigma_{s+1} \dots \sigma_t$$

soit encore,

$$\sigma_1 \dots \sigma_t = \sigma_1 \dots \sigma_{s-1} \sigma_{s+1} \dots \sigma_{t-1}$$

■

7.7 Le groupe de Weyl

Nous allons maintenant énoncer les propriétés les plus intéressantes des groupes de Weyl. En particulier, W permute les bases de Φ (ou de façon équivalente les chambres de Weyl) et W est généré par les réflexions relatives à n'importe quelle base Δ (i.e. par les σ_α pour $\alpha \in \Delta$). Nous ne démontrerons pas ces résultats².

Théorème 14 Soit Δ une base de Φ .

- (a) If $\gamma \in E$, γ régulière, il existe $\sigma \in W$ telle que $(\sigma(\gamma), \alpha) > 0$ pour tout $\alpha \in \Delta$.
- (b) Si Δ' est une autre base de Φ alors $\sigma(\Delta') = \Delta$ pour un certain $\sigma \in W$.
- (c) Si α est une racine alors il existe $\sigma \in W$ telle que $\sigma(\alpha) \in \Delta$.
- (d) W est généré par les σ_α ($\alpha \in \Delta$).
- (e) Si $\sigma(\Delta) = \Delta$, $\sigma \in W$ alors $\sigma = 1$.

2. Pour plus de détails, le lecteur pourra se référer à [Humphreys]

7.8 Systèmes de racines irréductibles

Ce théorème montre que le nombre de chambres de Weyl est égal à l'ordre du groupe de Weyl.

On appelle expression réduite de $\sigma \in W$ la forme $\sigma_{\alpha_1} \dots \sigma_{\alpha_t}$ ($\alpha_i \in \Delta$, t minimal). On note $l(\sigma) = t$ et on appelle cet entier la **longueur** de σ relativement à Δ . Par définition, $l(id) = 0$. Il est possible de caractériser la longueur d'une autre façon. Définissons $n(\sigma)$ = nombre de racines positives α telles que $\sigma(\alpha) \prec 0$. On peut montrer par récurrence que $l(\sigma) = n(\sigma)$.

7.8 Systèmes de racines irréductibles

On dit qu'un système de racines Φ est **irréductible** s'il n'est pas égal à l'union de deux ensembles propres tels que les racines de l'un soient orthogonales à toutes les racines de l'autre. A_1, A_2, B_2, G_2 (voir les représentations précédentes) sont irréductibles tandis que $A_1 \times A_1$ ne l'est pas.

Considérons une base Δ de Φ . On dit que Φ est irréductible si et seulement si Δ n'est pas égale à l'union de deux ensembles propres comme précédemment.

Proposition 22 *Soit Φ un système de racines irréductible. Relativement à l'ordre partiel $\prec$ ($a \prec b$ si et seulement si $b - a$ est une somme de racines positives ou $a = b$), il existe une unique racine maximale β (en particulier, $\alpha \neq \beta \Rightarrow ht \alpha < ht \beta$ pour tout $\alpha \in \Phi$, et $(\beta, \alpha) \geq 0$ pour tout $\alpha \in \Delta$). Si $\beta = \sum k_\alpha \alpha$ ($\alpha \in \Delta$) alors tous les k_α sont strictement positifs.*

Preuve Soit $\beta = \sum k_\alpha \alpha$ ($\alpha \in \Delta$) un élément maximal par rapport à l'ordre partiel $\prec$. $\beta \succ 0$ car Φ possède des racines positives et car si u est positive, $u \succ 0$. Posons $\Delta_1 = \{\alpha \in \Delta | k_\alpha > 0\}$ et $\Delta_2 = \{\alpha \in \Delta | k_\alpha = 0\}$. $\Delta = \Delta_1 \cup \Delta_2$ est une partition de Δ . Supposons Δ_2 non vide. Soit $\alpha \in \Delta_2$. On a $(\alpha, \beta) = \sum_{u \in \Delta_1} k_u (\alpha, u) \leq 0$ car les k_u sont strictement positifs et les (α, u) sont négatifs d'après le lemme suivant la définition d'une base d'un système de racines.

Comme Φ est irréductible, on ne doit pas avoir $(\Delta_1, \Delta_2) = 0$ donc il existe au moins un $\alpha \in \Delta_2$ qui n'est pas orthogonal à Δ_1 . On a donc nécessairement $(\alpha, \alpha') < 0$ pour au moins un $\alpha' \in \Delta_1$. On a alors $(\alpha, \beta) < 0$.

Ceci implique que $\beta + \alpha$ soit une racine d'après le lemme démontré dans la partie

7.8 Systèmes de racines irréductibles

concernant l'étude des paires de racines. Ceci est contraire à l'hypothèse de maximalité de β . Ainsi, Δ_2 est vide et tous les k_α sont strictement positifs.

Ceci montre également que $(\alpha, \beta) \geq 0$ pour tout $\alpha \in \Delta$ (avec $(\alpha, \beta) > 0$ pour au moins une racine simple α puisque Δ engendre E).

Maintenant, considérons β' une autre racine maximale. L'argument précédent s'applique également à β' : il existe $\alpha \in \Delta$ tel que $(\alpha, \beta') > 0$. On a alors $(\beta, \beta') = (\sum k_\alpha \alpha, \sum k'_\alpha \alpha) = \sum k_\alpha (\alpha, \sum k'_\alpha \alpha) > 0$ car les k_α sont strictement positifs. $\beta - \beta'$ est alors une racine (toujours d'après le lemme de la partie sur l'étude des paires de racines) sauf si $\beta = \beta'$. Mais si $\beta - \beta'$ est une racine alors $\beta \prec \beta'$ ou $\beta' \prec \beta$ ce qui n'est pas possible puisque β et β' sont deux racines maximales. On en déduit $\beta = \beta'$ et l'unicité de la racine maximale. ■

Proposition 23 *Soit Φ un système de racines irréductible. Alors il y a au plus deux longueurs de racines (la longueur d'une racine est égale à sa norme) dans Φ . D'autre part, toutes les racines de même longueur sont conjuguées sous l'action de W .*

Preuve Si α et β sont des racines alors les $\sigma(\alpha)$ ($\sigma \in W$) ne peuvent pas être toutes orthogonales à β puisque qu'on peut montrer que les $\sigma(\alpha)$ ($\sigma \in W$) engendrent E . Si $(\alpha, \beta) \neq 0$, nous savons (cf la table des caractéristiques des paires de racines) que $\frac{\|\alpha\|^2}{\|\beta\|^2} = 1, 2, 3, \frac{1}{2}$ ou $\frac{1}{3}$. S'il y avait trois longueurs de racine, alors nécessairement il y aurait un rapport de $\frac{3}{2}$. On a donc bien la première partie de la proposition.

Maintenant considérons α, β deux racines de même longueur. D'après la table des caractéristiques des paires de racines, on a

$$\langle \alpha, \beta \rangle = \langle \beta, \alpha \rangle = \pm 1$$

Comme $-\beta = \sigma_\beta(\beta)$, on peut se ramener au cas $\langle \alpha, \beta \rangle = \langle \beta, \alpha \rangle = 1$. On a alors

$$(\sigma_\alpha \sigma_\beta \sigma_\alpha)(\beta) = \sigma_\alpha \sigma_\beta(\beta - \alpha) = \sigma_\alpha(\sigma_\beta(\beta) - \sigma_\beta(\alpha)) = \sigma_\alpha(-\beta - \alpha + \beta) = \alpha$$

α et β sont bien conjuguées sous l'action de W . ■

Dans le cas où Φ est irréductible avec deux longueurs de racines, on parle de **racines longues** et de **racines courtes**.

7.8 Systèmes de racines irréductibles

On peut montrer que la racine maximale d'un système de racines irréductible est longue (par convention, si toutes les racines ont même longueur, on dira que ce sont des racines longues).

Nous allons maintenant utiliser ces propriétés pour représenter et classer les systèmes de racines.

8 Classification des systèmes de racines

Dans cette partie, Φ désigne un système de racines de rang l , de groupe de Weyl W . Δ est une base de Φ .

8.1 Matrice de Cartan de Φ

Donnons-nous un ordre pour les $(\alpha_1, \alpha_2, \dots, \alpha_l)$ pour les racines simples. La matrice $(\langle \alpha_i, \alpha_j \rangle)$ est appelée la **Matrice de Cartan** de Φ . Ces éléments sont les **entiers de Cartan**.

Exemples Pour les systèmes de racines de rang 2, en reprenant la table construite précédemment et en n'oubliant pas que $\langle \alpha_i, \alpha_i \rangle = \frac{2(\alpha, \alpha)}{(\alpha, \alpha)} = 2$, on trouve que les matrices de Cartan sont :

$$A_1 \times A_1 \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}; A_2 \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}; B_2 \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}; G_2 \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}; \quad (4)$$

Bien entendu, les matrices de Cartan dépendent de l'ordre que l'on s'est donné pour la base Δ . Par contre, et c'est important, la matrice de Cartan est indépendante du choix de la base Δ du système de racines. Ceci vient du (b) du théorème précédent (si Δ' est une autre base de Φ alors il existe $\sigma \in W$ tel que $\sigma(\Delta') = \Delta$) et du fait que les éléments de W conservent le produit scalaire. Les matrices de Cartan vont donc caractériser les systèmes de racines.

Proposition 24 *Soit $\Phi' \subset E'$ un autre système de racines de base $\Delta' = \{\alpha'_1, \dots, \alpha'_l\}$. Si $\langle \alpha_i, \alpha_j \rangle = \langle \alpha'_i, \alpha'_j \rangle$ pour $1 \leq i, j \leq l$ alors la bijection $\alpha_i \mapsto \alpha'_i$ se prolonge (de façon unique) en un isomorphisme $\phi : E \longrightarrow E'$ satisfaisant $\langle \phi(\alpha), \phi(\beta) \rangle = \langle \alpha, \beta \rangle$ pour tous $\alpha, \beta \in \Phi$ et $\phi(\Phi) = \Phi'$. Ainsi, la matrice de Cartan de Φ détermine Φ à un isomorphisme près.*

Preuve Puisque Δ (resp. Δ') est une base de E (resp. E'), il existe un unique isomorphisme d'espaces vectoriels $\phi : E \longrightarrow E'$ envoyant α_i en α'_i ($1 \leq i \leq l$).

Si $\alpha, \beta \in \Delta$, on a alors $\sigma_{\phi(\alpha)}(\phi(\beta)) = \sigma_{\alpha'}(\beta') = \beta' - \langle \beta', \alpha' \rangle \alpha' = \phi(\beta) - \langle \beta, \alpha \rangle \phi(\alpha) = \phi(\beta - \langle \beta, \alpha \rangle \alpha) = \phi(\sigma_\alpha(\beta))$ (car par hypothèse $\langle \alpha, \beta \rangle = \langle \alpha', \beta' \rangle$). Comme les groupes de Weyl W et W' sont engendrés par les réflexions orthogonales aux racines simples, l'application $\sigma \mapsto \phi \circ \sigma \circ \phi^{-1}$ est un isomorphisme de W dans W' qui à σ_α fait correspondre $\sigma_{\phi(\alpha)}$, ($\alpha \in \Delta$).

8.2 Graphes de Coxeter et diagrammes de Dynkin

Mais comme pour tout racine $\beta \in \Phi$, il existe une racine simple α est un élément $\sigma \in E$ tel que $\beta = \sigma(\alpha)$ (d'après le théorème sur les groupes de Weyl), on a

$$\phi(\beta) = \phi(\sigma(\alpha)) = (\phi \circ \sigma \circ \phi^{-1})(\phi(\alpha)) \in \Phi'$$

Puisque $\phi \circ \sigma \circ \phi^{-1} \in W'$. Donc $\phi(\Phi) \subset \Phi'$. Comme ϕ est un isomorphisme de E dans E' , on a bien $\phi(\Phi) = \Phi'$. La relation $\langle \phi(\alpha), \phi(\beta) \rangle = \langle \alpha, \beta \rangle$ se montre en travaillant avec la formule des réflexions (on l'a déjà vérifiée pour les racines simples). ■

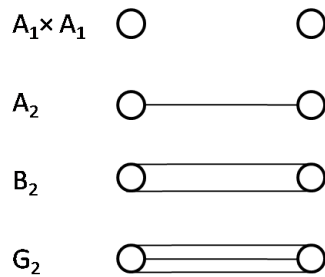
Cette proposition montre qu'il est possible de construire le système de racines à partir de sa matrice de Cartan.

8.2 Graphes de Coxeter et diagrammes de Dynkin

Si α et β sont des racines positives distinctes alors nous savons que $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 0, 1, 2$ ou 3. Définissons le **graphe de Coxeter** de Φ comme le graphe possédant l sommets (l désignant le rang du système de racines égal à la dimension de l'espace euclidien E), le i^{ieme} lié au j^{ieme} par $\langle \alpha_i, \alpha_j \rangle \langle \alpha_j, \alpha_i \rangle$ branches.

Exemples : En reprenant la table caractérisant les paires de racines, on peut construire facilement les graphes de Coxeter des systèmes de racines de rang 2.

Dans le cas où toutes les racines ont même longueur, on a $\langle \alpha_i, \alpha_j \rangle = \langle \alpha_j, \alpha_i \rangle$



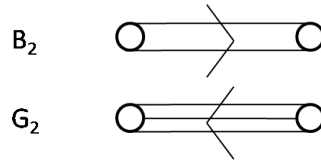
et on peut déterminer les entiers de Cartan. Dans le cas où il y a plusieurs longueurs de racines dans le système de racines (par exemple pour B_2 (pour lequel $\frac{\|\beta\|^2}{\|\alpha\|^2} = 2$ ou $\frac{1}{2}$) ou G_2 (pour lequel $\frac{\|\beta\|^2}{\|\alpha\|^2} = 3$ ou $\frac{1}{3}$)), on ne peut pas savoir lequel des sommets correspond à une racine simple courte ou à une racine simple longue.

Quand dans un graphe de Coxeter, deux sommet sont liés par deux ou trois

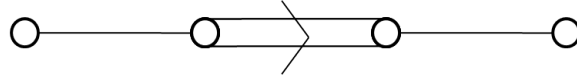
8.2 Graphes de Coxeter et diagrammes de Dynkin

branches (cas où les racines simples n'ont pas la même longueur), on peut ajouter une flèche pointant vers la plus courte des deux racines. Cette information supplémentaire permet de retrouver les entiers de Cartan. On appelle **diagramme de Dynkin** le graphe résultant (qui dépend lui aussi de l'ordre que l'on s'est donné pour le système de racines).

On a par exemple :



On peut reconstruire la matrice de Cartan à partir du diagramme de Dynkin.



Ce graphe de Dynkin comporte quatre noeuds. Le système de racines associés est donc de rang 4 et la matrice de Cartan est une matrice 4×4 . Les termes diagonaux de la matrice de Cartan valent toujours deux. Les angles entre les vecteurs d'une base de Φ sont obtus (d'après le lemme démontré après la définition d'une base d'un système de racines). Par conséquent $\langle \alpha_1, \alpha_2 \rangle = \langle \alpha_2, \alpha_1 \rangle = 1$ (une seule branche) $\Rightarrow \langle \alpha_1, \alpha_2 \rangle = \langle \alpha_2, \alpha_1 \rangle = -1$. De même, $\langle \alpha_3, \alpha_4 \rangle = \langle \alpha_4, \alpha_3 \rangle = -1$.

$\langle \alpha_2, \alpha_3 \rangle \langle \alpha_3, \alpha_2 \rangle = 2$ donc $\langle \alpha_2, \alpha_3 \rangle = -1$ ou -2 . Comme α_2 est plus longue que α_3 , on en déduit $\langle \alpha_2, \alpha_3 \rangle = \frac{2\langle \alpha_2, \alpha_3 \rangle}{\langle \alpha_3, \alpha_3 \rangle} = -2$. On a alors enfin $\langle \alpha_3, \alpha_2 \rangle = -1$.

La matrice de Cartan de ce diagramme de Dynkin est donc :

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -2 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$$

8.3 Composantes irréductibles

Rappelons qu'on dit qu'un système de racines Φ est irréductible si et seulement si Φ (ou de façon équivalente, Δ) ne s'écrit pas comme l'union de deux sous-ensembles propres (distincts de Φ) orthogonaux entre eux. On peut sentir que Φ est irréductible si et seulement si son graphe de Coxeter est connexe (au sens "classique" du terme, c'est-à-dire d'un seul tenant). De façon générale, un graphe de Coxeter se divisera en un nombre fini de composantes connexes. Soit $\Delta = \Delta_1 \cup \dots \cup \Delta_t$ la partition de Δ correspondante à cette division du graphe de Coxeter en composantes connexes. Si E_i est le sous-espace vectoriel de E engendré par Δ_i , on va avoir $E = E_1 \oplus \dots \oplus E_t$ (espaces en somme directe).

Par ailleurs, toute combinaison $\mathbb{Z}$ -linéaire des éléments de Δ_i (appelons cet ensemble Φ_i) forme un système de racines de E_i dont le groupe de Weyl est le sous-groupe de W engendré par les σ_α ($\alpha \in \Delta_i$). On peut alors montrer que $\Phi = \Phi_1 \cup \dots \cup \Phi_t$.

Proposition 25 *Φ se décompose de manière unique comme l'union de systèmes de racines irréductibles Φ_i (engendrant les sous-espaces E_i de E) tels que $E = E_1 \oplus \dots \oplus E_t$.*

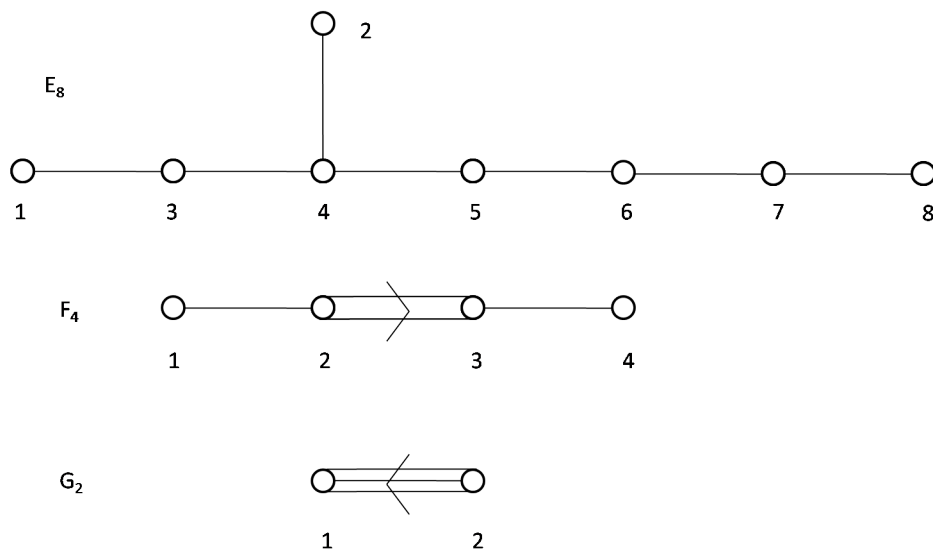
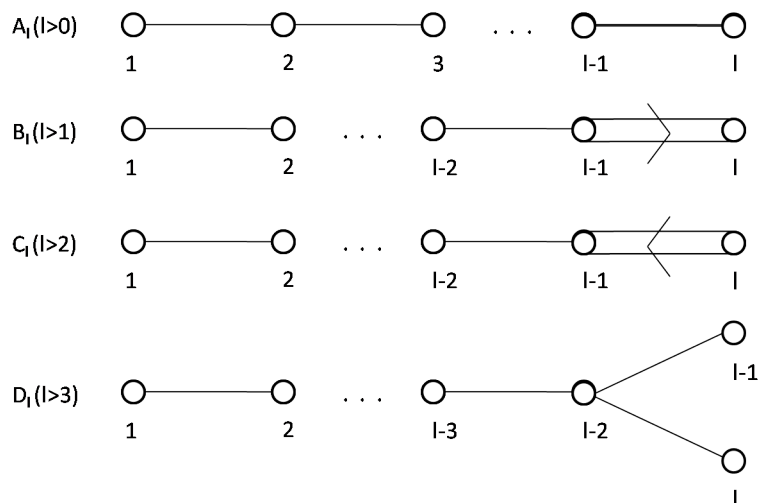
8.4 Théorème de classification

Pour classer les algèbres de Lie semi-simples, il suffit de classer les systèmes de racines et plus particulièrement, d'après ce qui précède, les systèmes de racines irréductibles ou, ce qui revient au même, les diagrammes de Dynkin connexes. Nous ne démontrerons pas le théorème suivant³.

Théorème 15 *Si Φ est un système de racines irréductible de rang l , son diagramme de Dynkin est l'un des diagrammes suivants :*

3. Les détails de la preuve de ce théorème sont dans [Bourbaki]

8.4 Théorème de classification



Nous sommes donc maintenant capables de représenter et classer les systèmes de racines. Comme les systèmes de racines des algèbres de Lie semi-simples vérifient les axiomes de la définition d'un système de racines, on peut représenter les systèmes de racines des algèbres semi-simple. Ceci devient intéressant car il est possible de montrer⁴ que la correspondance $(L, H) \mapsto (\Phi, E)$ entre une algèbre de Lie semi-simple munie de son tore maximale et son ensemble de racines est bijective. En particulier, cela signifie qu'un système de racines caractérise son algèbre de Lie semi-simple. En sachant classer les systèmes de racines, on sait donc classer les algèbres de Lie semi-simples.

4. Voir chapitres IV-V de [Humphreys] par exemple.

9 Des symétries et des groupes : pourquoi et comment utiliser des groupes de Lie en physique ?

9.1 Comment décrire une transformation continue en physique ?

Commençons par des considérations physiques très générales, qui vont nous montrer que les transformations infinitésimales de la physique peuvent être caractérisées par des groupes de Lie.

Pour ce faire, imaginons un espace, ou un système physique quelconque, que nous noterons Ω , pouvant être décrit par un nombre fini de n paramètres distincts, que nous appellerons "coordonnées" (en référence à l'espace). Nous noterons par $M(x_1, x_2, \dots, x_n)$ un point (si l'ensemble que nous considérons est un espace), ou un état (si l'ensemble que nous considérons est un système physique). Nous allons nous intéresser aux transformations G qui permettent de passer de $M(x_1, x_2, \dots, x_n)$ donné $M'(x'_1, x'_2, \dots, x'_n)$ autre élément de l'ensemble à n paramètres. Pour des raisons physiques, il apparaît aberrant de penser que la transformation G permettant ce passage dépende d'une infinité de paramètres (il nous serait dès lors impossible de la décrire théoriquement). Par conséquent, nous allons supposer que G dépend de r paramètres finis, i.e.

$$G = G(a_1, a_2, \dots, a_r)$$

Les coordonnées x'_i sont donc données par un jeu de n fonctions g_i , dépendantes de $n + r$ paramètres :

$$x'_i = g_i(x_1, x_2, \dots, x_n; a_1, a_2, \dots, a_r)$$

La question consiste à savoir à quel espace appartiennent les transformations G ? Nous allons nous intéresser aux transformations préservant une symétrie donnée, i.e. intuitivement laissant quelque chose d'invariant. Nous noterons leur ensemble Γ .

Logiquement si nous introduisons un nouveau M'' (toujours un autre état ou un autre point, cela dépend de l'ensemble de départ), nous voulons que le passage de M à M' par G puis de M' à M'' par G' puisse se faire directement par une transformation G'' , qui ne soit rien d'autre que la composition des deux précédentes transformations. Cette transformation préservera la symétrie, par composition,

9.1 Comment décrire une transformation continue en physique ?

donc appartiendra toujours à Γ . Ceci nous montre donc que la composition

$$G'(a'_1, a'_2, \dots, a'_r)G(a_1, a_2, \dots, a_r) = G''(a''_1, a''_2, \dots, a''_r)$$

où les a''_i dépendent des $2r$ paramètres continus :

$$a''_i = \phi_i(a_1, a_2, \dots, a_r; a'_1, a'_2, \dots, a'_r)$$

Intuitivement, ce "produit" de transformation sera associatif, car le passage de M à un hypothétique $M^{(3)}$ en suivant le "chemin" $MM'M''$ peut aussi bien s'effectuer en "s'arrêtant" sur M' que sur M'' . Qui plus est, comme les transformations que nous considérons sont symétriques, nous avons envie que quelque chose ne soit pas brisé, i.e. qu'il nous soit possible de revenir de M' à M en laissant la symétrie conservée. Intuitivement, si l'on fait l'analogie avec la thermodynamique, le fait que nos transformations soient continues nous indiquent qu'elles s'opèrent de proche en proche de manière infinitésimales, donc qu'elles sont d'une certaine manière réversibles. Par suite, nous allons supposer que si G appartient à Γ , alors son inverse y appartient également. Enfin avant d'arriver à notre première conclusion quant à la nature de Γ , remarquons que la continuité des transformations implique que nous puissions passer de manière infinitésimale d'une transformation à une autre, i.e. que le groupe des transformations est continu. Toute transformation identique préservant l'ensemble Ω et donc les symétries présupposées, nous sommes donc arrivés à la première constatation de notre éclairage :

Postulat 1 *L'ensemble des transformations continues Γ préservant une symétrie donnée sur Ω forme un groupe continu*

C'est déjà bien, mais ce n'est pas suffisant pour pouvoir espérer mener à bien l'ensemble de nos calculs. En effet, en bons physiciens, intéressons nous à une transformation infinitésimale

$$M(x_1, x_2, \dots, x_n) \rightarrow M(x_1 + dx_1, x_2 + dx_2, \dots, x_n + dx_n)$$

D'après les relations précédentes,

$$x_i + dx_i = g_i(x_1, x_2, \dots, x_n; da_1, da_2, \dots, da_r)$$

Comme nous avons vraiment très envie de pouvoir dériver les g_i en fonction des paramètres a_i afin de faire intervenir un développement limité, nous allons supposer les fonctions g_i différentiables par rapport aux a_i . Cette construction nous montre que notre groupe serait très bien décrit si l'on pouvait en plus lui attribuer une

9.1 Comment décrire une transformation continue en physique ?

structure de variété différentiable de dimension r , puisque nous pourrions alors dériver les g_i . Nous voyons donc déjà que mettre sur Γ une structure de groupe de Lie serait très intéressante. Pour en être sûr, et voir l'intérêt d'une telle description de Γ , prenons une fonction $F(x_1, x_2, \dots, x_n)$ quelconque 2-fois dérivable par rapport aux x_i et notons par $\Theta(G(a_1, a_2, \dots, a_r))$ l'opérateur induit par la transformation $G(a_1, a_2, \dots, a_r)$:

$$F(x'_1, x'_2, \dots, x'_n) = \Theta(G(a_1, a_2, \dots, a_r))F(x_1, x_2, \dots, x_n)$$

Si nous reconsidérons la transformation infinitésimale précédente, en nommant $G_{da} = G(da_1, da_2, \dots, da_r)$ et nous limitons à l'ordre deux, en introduisant par dF_0^2 la hessienne de F et en supposant que l'application dF_{a0} correspondante lorsqu'on remplace les dx par des da reste bilinéaire, nous avons :

$$\begin{aligned} \Theta(G_{da})F(x_1, x_2, \dots, x_n) &= F(x_1 + dx_1, x_2 + dx_2, \dots, x_n + dx_n) \\ &= F(x_1, x_2, \dots, x_n) + \sum_{i=1}^n \frac{\partial F}{\partial x_i} dx_i + \frac{1}{2} dF_0^2(dx, dx) + \dots \\ &= \left(1 + \sum_{i=1}^n \sum_{j=1}^r \left(\frac{\partial g_j}{\partial a_j} \right)_{a_j=0} \frac{\partial}{\partial x_i} da_j \right) F(x_1, x_2, \dots, x_n) \\ &\quad + \frac{1}{2} (dF_{a0}(a, a)) + \dots \\ &= \left[1 + \sum_{j=1}^r X_j da_j + \frac{1}{2} \sum_{i,j=1}^r dF_{ij} da_i da_j \right] F(x_1, x_2, \dots, x_n) + \dots \end{aligned}$$

Comme l'ensemble Γ est un groupe, si nous faisons une autre transformation élémentaire $G(d\bar{a})$, alors le produit de ces deux transformation s'écrit

$$G(d\bar{a})G(da) = G(f(d\bar{a}, da))$$

La fonction $f = (\phi_1, \dots, \phi_r)$ où les ϕ_i ont été définis précédemment est supposé être différentiable, i.e. comme $f(da, 0) = f(0, da) = da$,

$$f(d\bar{a}, da) = d\bar{a} + da + \sum_{i,j=1}^r \Phi_{ij} da_i d\bar{a}_j + \dots$$

est le seul développement qui satisfasse les conditions précédentes. Il en ressort qu'en identifiant le produit :

$$G(d\bar{a})G(da) = G(f(d\bar{a}, da))$$

il vient :

$$\begin{aligned} & \left[1 + \sum_{j=1}^r X_j da_j + \frac{1}{2} \sum_{i,j=1}^r dF_{ij} da_i da_j \right] \left[1 + \sum_{j=1}^r X_j d\bar{a}_j + \frac{1}{2} \sum_{i,j=1}^r dF_{ij} d\bar{a}_i d\bar{a}_j \right] \\ &= \left[1 + \sum_{j=1}^r (d\bar{a}_j + da_j) X_j + \sum_{i,j,k=1}^r \Phi_{ij}^k da_i d\bar{a}_j X_k + \frac{1}{2} \sum_{i,j=1}^r dF_{ij} (da_i + d\bar{a}_i) (da_j + d\bar{a}_j) \right] \end{aligned}$$

Comme on veut que ceci soit vrai pour tous "déplacements élémentaires" da et $d\bar{a}$, l'identification des coefficients en $da_i d\bar{a}_j$ nous fournit la relation

$$X_j X_i = dF_{ij} + \sum_{k=1}^r \Phi_{ij}^k X_k \quad (5)$$

Or, dF_{ij} étant symétrique, ceci équivaut en inversant i et j à la relation :

$$X_j X_i - X_i X_j = \sum_{k=1}^r \xi_{ij}^k X_k \quad (6)$$

avec

$$\xi_{ij}^k = \Phi_{ij}^k - \Phi_{ji}^k$$

Cette analogie de la description d'un voisinage de l'application identité par des générateurs infinitésimaux qui obéissent à la même règle qu'une base d'une algèbre de Lie nous pousse à la constatation suivante :

Postulat 2 Γ est un groupe de Lie de dimension r

Les X_j sont appelés générateurs infinitésimaux du groupe de Lie.

Forts de notre constatation précédente sur la description d'un groupe de transformations symétriques à l'aide des groupes de Lie, nous allons nous intéresser aux deux groupes de Lie SU(2) et SU(3). Avant d'expliquer pourquoi et comment on peut expliquer l'intérêt de ces groupes en physique des particules, nous allons nous attacher à décrire leur représentation et leur système de racine.

9.2 Etude de SU(2)

Rappelons déjà la définition de SU(2) :

$$SU(2) = \{A \in GL(2, \mathbb{C}) \mid AA^* = I \text{ et } \det(A) = 1\} \quad (7)$$

Nous allons étudier sa représentation à l'aide de son algèbre de Lie. Nous allons donc nous ramener à l'étude des représentations de $\mathfrak{sl}(2, \mathbb{C}) = \{X \in \mathfrak{gl}(2, \mathbb{C}) \mid \text{Tr}(X) = 0\}$. Pour ce faire, on a besoin de la définition générale d'une représentation finie d'une algèbre de Lie (en regard de celle que nous avons donné dans les paragraphes précédent pour un groupe) :

Définition 25 Soit L une algèbre de Lie sur un corps $\mathbb{K}$, E un espace vectoriel complexe de dimension finie. Une représentation (finie) de L est une application $\mathbb{K}$ -linéaire R de L dans $\mathfrak{gl}(E)$ satisfaisant $\forall X, Y \in L$:

$$[R(X), R(Y)] = R([X, Y])$$

C'est donc un morphisme d'algèbre de Lie complexe de L dans $\mathfrak{gl}(E)$.

9.2.1 Représentations irréductibles de $\mathfrak{sl}(2, \mathbb{C})$

Considérons les matrices

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, X_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, X_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad (8)$$

Il est évident de vérifier que ces trois matrices forment une base de $\mathfrak{sl}(2, \mathbb{C})$ et qu'elles satisfont les relations de commutation

$$[H, X_{\pm}] = \pm 2X_{\pm}, [X_+, X_-] = H \quad (9)$$

Considérons une représentation (E, R) irréductible de dimension finie de $\mathfrak{sl}(2, \mathbb{C})$. Comme $R(H)$ est non nul, il admet au moins une valeur propre λ et un vecteur propre v correspondant. On établit alors la proposition suivante

Proposition 26 Les représentations irréductibles de dimension finie de $\mathfrak{sl}(2, \mathbb{C})$ sont les $(E^{(n)}, R^{(n)})$, $n \in \mathbb{N}$, où $E^{(n)}$ est un espace de dimension $n+1$ dont la base $v_0, \dots, v_n$ satisfait les relations :

$$\begin{cases} R^{(n)}(H)v_k &= (n - 2k)v_k \\ R^{(n)}(X_-)v_k &= v_{k+1} \\ R^{(n)}(X_+)v_k &= k(n - k + 1)v_{k-1} \end{cases} \quad (10)$$

Preuve

Servons nous tout d'abord des relations de commutations. $R([H, X_{\pm}]) = [R(H), R(X_{\pm})]$ implique

$$R(H)R(X_{\pm}) - R(X_{\pm})R(H) = \pm 2R(X_{\pm})$$

i.e. en l'appliquant à v :

$$R(H)R(X_{\pm})v = (\lambda \pm 2)R(X_{\pm})v \quad (11)$$

Cette relation implique que si v est un vecteur propre de $R(H)$ de valeur propre λ , alors $R(X_{\pm})v$ est aussi un vecteur propre de $R(H)$ de valeur propre $\lambda + 2$. Comme l'espace E est de dimension fini, on voit qu'il y a nécessairement un moment où ce processus doit s'arrêter, i.e. qu'il existe un v_0 vecteur propre de $R(H)$ de valeur propre λ_0 avec $R(X_{\pm})v_0 = 0$. Posons dès lors $v_k = R(X_-)^k v_0$. En reportant dans (11) dans le cas $-$ nous obtenons :

$$R(H)R(X_-)v_0 = (\lambda_0 - 2)R(X_-)v_0$$

i.e.

$$R(H)v_1 = (\lambda_0 - 2)v_1$$

De même,

$$R(H)R(X_-)v_1 = (\lambda_0 - 2 - 2)R(X_-)v_1$$

i.e.

$$R(H)v_2 = (\lambda_0 - 4)v_2$$

... de proche en proche on en déduit la relation triviale :

$$R(H)v_k = (\lambda_0 - 2k)v_k \quad (12)$$

Montrons désormais, par récurrence sur k , que, pour $k \in \mathbb{N}$,

$$R(X_+)v_k = k(\lambda_0 - k + 1)v_{k-1}$$

- Cette relation est vérifiée pour $k = 1$. En effet comme d'après les relations de commutation,

$$R(X_+)R(X_-) = R(X_-)R(X_+) + R(H)$$

i.e. en l'appliquant à v_0 ,

$$R(X_+)v_1 = R(X_-)R(X_+)v_0 + R(H)v_0 = R(H)v_0 = \lambda_0 v_0$$

- Supposons la relation de récurrence vraie au rang k . Alors en utilisant la même relation que pour le cas $k = 1$, il vient immédiatement,

$$\begin{aligned} R(X_+)v_k &= R(X_-)R(X_+)v_k + R(H)v_k = R(X_-)k(\lambda_0 - k + 1)v_{k-1} + R(H)v_k \\ &= k(\lambda_0 - k + 1)v_k + (\lambda_0 - 2k)v_k \\ &= (k + 1)(\lambda_0 - k)v_k \end{aligned}$$

ce qui achève la preuve par récurrence.

Or, rappelons que $R(X_-) \in \mathfrak{gl}(E)$. Par conséquent, d'après le théorème de Kayleigh-Hamilton, il existe un entier n tel que pour $i = 0..n$, les v_i sont linéairement indépendants et tel que $v_{n+1} = 0$ (car E est de dimension finie, donc le nombre de vecteurs propres non nul est fini). Ceci nous montre directement que $\lambda_0 = n$. Les vecteurs $v_0, \dots, v_n$ engendrent donc un sous-espace invariant par R . Comme nous voulons que la représentation soit irréductible, il faut nécessairement (mais ce n'est pas suffisant) que $v_0, \dots, v_n$ engendrent E , et donc que E soit un espace de dimension $n + 1$. Vérifions que la représentation ainsi définie est bien irréductible. Démontrons que tout vecteur d'un sous espace vectoriel F invariant par R est dans E . Soit u non nul de F , F invariant par R . En décomposant u sur la base $(v_0, \dots, v_n)$, quitte à réorganiser les indices, supposons que

$$u = \sum_{k=0}^n u_k v_k$$

avec $u_{k_0} \neq 0$ et $u_i = 0$ pour $i > k_0$. Alors $R^{k_0}(X_+)u = \mu v_0$ où $\mu \in \mathbb{C}$. Par conséquent, v_0 appartient à F , donc l'ensemble des v_i également. D'où $F = E$. ■

Définition 26 *En physique, on considère généralement la représentation $D^j = (E^{(2j)}, R^{(2j)})$, avec pour base standard de $E^{(2j)}$ les $|j, m\rangle$ ($-j \leq m \leq j$) définies par*

$$|j, m\rangle = (-1)^{j+m} \sqrt{\frac{(j-m)!}{(j+m)!}} v_{j+m}$$

et pour base de $\mathfrak{sl}(2)$ les $J_3 = -\frac{H}{2}$, $J_{\pm} = -X_{\pm}$.

9.2.2 Représentations irréductibles de $SU(2)$

Commençons par la définition de la différentielle d'une représentation d'un groupe de Lie (nous nous restreignons ici aux sous groupes fermés de $GL(n, \mathbb{C})$).

Définition 27 *Soit (E, ρ) une représentation d'un groupe sous groupe fermé de $GL(n, \mathbb{C})$. La différentielle de la représentation ρ , notée $D\rho$, est une représentation de l'algèbre de Lie $\mathfrak{g}$ de G définie par :*

$$\begin{aligned} \mathfrak{g} &\longrightarrow \mathfrak{gl}(E) \\ (D\rho)(X) &= \frac{d}{dt} \rho(\exp(tX))_{t=0} \end{aligned}$$

Preuve

Nous allons prouver que cette définition est cohérente, i.e. de la dérivation est bien définie et que $D\rho$ définit une représentation de l'algèbre de Lie. Tout d'abord faisons remarquer que l'application $t \mapsto \rho(\exp(tX))$ est continue et de classe C^∞ pour tout X dans $\mathfrak{g}$, par définition de l'exponentielle de matrice. Qui plus est,

$$\begin{aligned} \frac{d}{dt}\rho(\exp(tX))_t &= \lim_{s \rightarrow 0} \frac{\rho(\exp((t+s)X)) - \rho(\exp(tX))}{s} \\ &= \rho(\exp(tX)) \cdot \frac{d}{dt}\rho(\exp(tX))_{t=0} \\ &= \rho(\exp(tX))(D\rho)(X) \end{aligned}$$

Par suite, comme ρ est un morphisme de groupe de Lie, $\rho(I_G) = I_{GL(E)}$, donc l'équation différentielle précédente se résout simplement par

$$\rho(\exp(tX)) = \exp(t(D\rho)(X)) \quad (13)$$

(on dit que $(D\rho)$ est le générateur infinitésimal de $t \mapsto \rho(\exp(tX))$) Montrons à l'aide de cette relation que $(D\rho)$ est un morphisme $\mathbb{K}$ -linéaire de $\mathfrak{g}$ dans $\mathfrak{gl}(E)$ préservant le crochet de Lie afin de justifier notre définition.

Déjà, comme par définition, $X \in \mathfrak{g}$ ssi $\forall t \in \mathbb{R}$, $\exp(tX) \in G$, nous voyons que $\forall t \in \mathbb{R}$, $\exp(t(D\rho)(X)) \in GL(E)$, i.e. $(D\rho)$ est bien une application de $\mathfrak{g}$ dans $\mathfrak{gl}(E)$.

Soit $\lambda \in \mathbb{K}$. Alors par un simple changement de variables dans la dérivation ($t \rightarrow \lambda t$), on voit directement par la définition de $(D\rho)$ que $(D\rho)(\lambda X) = \lambda(D\rho)(X)$. Démontrons désormais qu'elle est additive. D'après la relation précédente,

$$\begin{aligned} \exp(t(D\rho)(X+Y)) &= \rho(\exp(tX)\exp(tY)) = \lim_{n \rightarrow \infty} \left[\rho\left(\exp\left(\frac{t}{n}X\right)\exp\left(\frac{t}{n}Y\right)\right) \right]^n \\ &= \lim_{n \rightarrow \infty} \left[\exp\left(\frac{t}{n}(D\rho)(X)\right)\exp\left(\frac{t}{n}(D\rho)(Y)\right) \right]^n \end{aligned}$$

[car $\frac{1}{n}(D\rho)(X)$ est le générateur infinitésimal de $t \mapsto \rho(\exp(\frac{t}{n}X))$ d'après ce qui vient de précéder] D'où

$$\exp(t(D\rho)(X+Y)) = \exp(t((D\rho)(X) + (D\rho)(Y)))$$

ce qui nous montre l'additivité de $(D\rho)$.

Enfin toujours en exploitant (13), ainsi qu'en utilisant la stabilité de $\mathfrak{g}$ par conjugaison par des éléments de G , il vient $\forall g \in G, \forall X \in \mathfrak{g}$:

$$\begin{aligned} \exp(t(D\rho)(gXg^{-1})) &= \rho(\exp(g(tX)g^{-1})) = \rho(g \exp(tX) g^{-1}) \\ &= \rho(g)\rho(\exp(tX))\rho(g)^{-1} = \rho(g)\exp(t(D\rho)(X))(\rho(g))^{-1} \\ &= \exp(t\rho(g)(D\rho)(X)(\rho(g))^{-1}) \end{aligned}$$

où nous nous sommes servis à la première et la troisième ligne de la définition de l'exponentielle $\exp(A) = \sum_n \frac{A^n}{n!}$ pour sortir/renter les termes g et g^{-1} . Par conséquent, $\forall Y \in \mathfrak{g}$,

$$(D\rho)(\exp(tY)X\exp(-tY)) = \rho(\exp(tY)(D\rho)(X)\exp(-tY))$$

donc en dérivant par rapport à t en 0, on obtient directement :

$$(D\rho)(YX - XY) = (D\rho)(Y)(D\rho)(X) - (D\rho)(X)(D\rho)(Y)$$

Par conséquent, l'application $(D\rho)$ définie bien une représentation de $\mathfrak{g}$ ■

Nous allons maintenant établir une importante proposition qui va nous montrer que nous pouvons étudier certaines représentations de SU(2) par l'intermédiaire des D^j de $\mathfrak{su}(2)$

Proposition 27 *Il existe des représentations irréductibles de SU(2), notés S^j , dont la différentielle précédemment définie s'identifie avec les D^j de l'algèbre de Lie $\mathfrak{su}(2)$.*

Preuve Soit $j \in \frac{1}{2}\mathbb{N}$. On note V^j l'espace vectoriel des polynômes homogènes de deux variables (z_1, z_2) de degré $2j$ à coefficients complexes. Il est facile de voir que cet espace vectoriel est de dimension $2j + 1$, avec pour base les $z_1^k z_2^{2j-k}$ pour k variant de 0 à $2j$. Soit

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

un élément de SU(2). Comme son déterminant $AA^* = I$, il est clair que son inverse est

$$A^{-1} = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$$

Il est facile de montrer que l'action des éléments A de SU(2) sur V^j , définie par $(\rho(A)P) = P \circ (A^{-1})$, i.e.

$$(\rho(A)P)(z_1, z_2) = (P \circ (A^{-1}(z_1, z_2))) = P(\bar{a}z_1 - bz_2, \bar{b}z_1 + az_2)$$

est une représentation de SU(2). Nous allons démontrer que cette représentation ρ de SU(2) dans V^j est irréductible et de différentielle équivalente à D^j .

Faisons tout d'abord remarquer que les matrices

$$\alpha_1 = \frac{1}{2} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \alpha_2 = \frac{1}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \alpha_3 = \frac{1}{2} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

forment une base de l'algèbre de Lie $\mathfrak{su}(2) = \{X \in \mathfrak{gl}(2, \mathbb{C}) \mid X^* + X = 0, \text{Tr}(X) = 0\}$ de SU(2). On remarque que les éléments J_3 et $J_{\pm}$ que nous avons introduit dans la définition 26 peuvent s'écrire en fonction des α_k :

$$J_{\pm} = i\alpha_1 \pm \alpha_2, J_3 = i\alpha_3$$

Nous allons donc chercher à montrer que les opérateurs $DS^j(J_{\pm})$ et $DS^j(J_3)$ coïncident respectivement avec $D^j(J_{\pm})$ et $D^j(J_3)$. Pour ce faire, par définition de la différentielle, il est facile de voir que pour tout élément $X = \begin{pmatrix} e & f \\ g & h \end{pmatrix} \in \mathfrak{sl}(2, \mathbb{C})$

en notant $L(t) = e^{tX} = \begin{pmatrix} a(t) & b(t) \\ c(t) & d(t) \end{pmatrix}$ alors on a :

$$\begin{aligned} ((D\rho)(-X)P)(z_1, z_2) &= \left[\frac{d}{dt} (\rho(e^{-tX})f)(z_1, z_2) \right]_{t=0} \\ &= \left[\frac{d}{dt} (f \circ (L(t)))(z_1, z_2) \right]_{t=0} \\ &= \left[\frac{d}{dt} (a(t)z_1 + b(t)z_2) \right]_{t=0} \partial_{z_1} f(z_1, z_2) \\ &\quad + \left[\frac{d}{dt} (c(t)z_1 + d(t)z_2) \right]_{t=0} \partial_{z_2} f(z_1, z_2) \\ &= [ez_1 + fz_2] \partial_{z_1} f(z_1, z_2) + [gz_1 + hz_2] \partial_{z_2} f(z_1, z_2) \end{aligned}$$

Un rapide calcul à l'aide des expressions des α_k et de $J_{\pm}$ et J_3 en fonction de ces derniers nous donne donc les résultats suivants :

$$\begin{aligned} (D\rho)(\alpha_1) &= -\frac{i}{2}(z_1 \partial_{z_1} - z_2 \partial_{z_2}) \\ (D\rho)(\alpha_2) &= -\frac{i}{2}(z_2 \partial_{z_1} + z_1 \partial_{z_2}) \\ (D\rho)(\alpha_3) &= \frac{1}{2}(z_2 \partial_{z_1} - z_1 \partial_{z_2}) \end{aligned}$$

9.3 SU(2) et le groupe d'isospin : une immersion dans les propriétés de l'interaction nucléaire forte

et par la suite

$$\begin{aligned}(D\rho)(J_+) &= z_1 \partial_{z_2} \\ (D\rho)(J_-) &= z_2 \partial_{z_1} \\ (D\rho)(J_3) &= \frac{1}{2}(z_1 \partial_{z_1} - z_2 \partial_{z_2})\end{aligned}$$

En faisant agir ces opérateurs sur les éléments $f_m^j(z_1, z_2) = z_1^{j+m} z_2^{j-m}$, qui forment une base de V^j on trouve de manière triviale

$$\begin{aligned}(D\rho)(J_+)f_m^j &= (j-m)f_{m+1}^j \\ (D\rho)(J_-)f_m^j &= (j+m)f_{m-1}^j \\ (D\rho)(J_3)f_m^j &= mf_m^j\end{aligned}$$

Par conséquent en notant $|j, m\rangle = \frac{1}{\sqrt{(j-m)!(j+m)!}} f_m^j$ on voit que les actions précédentes se transforment en

$$(D\rho)(J_+) |j, m\rangle = \sqrt{(j-m)(j+m+1)} |j, m+1\rangle \quad (14)$$

$$(D\rho)(J_-) |j, m\rangle = \sqrt{(j+m)(j-m+1)} |j, m-1\rangle \quad (15)$$

$$(D\rho)(J_3) |j, m\rangle = m |j, m\rangle \quad (16)$$

Pour ce rendre compte de l'intérêt de ces relations, il suffit de remarquer que l'on a exactement les mêmes relations pour D^j avec la base que nous avons défini pour $\mathfrak{sl}(2, \mathbb{C})$. Par conséquent la différentielle de la représentation $(V^j, S^j := \rho)$ et la représentation D^j du paragraphe précédent sont les restrictions d'une même représentation de $\mathfrak{sl}(2, \mathbb{C})$, et sont donc équivalentes. Démontrons désormais l'irréductibilité de (V^j, S^j) . Soit F un sous espace de V^j invariant par S^j . Alors $\forall X \in \mathfrak{sl}(2, \mathbb{C})$, F est invariant par $S^j(e^{tX})$, donc invariant par $(DS^j)(X)$. Par conséquent DS^j étant irréductible (car D^j l'est), (V^j, S^j) est irréductible. ■

9.3 SU(2) et le groupe d'isospin : une immersion dans les propriétés de l'interaction nucléaire forte

9.3.1 SU(2) et l'interaction forte

Les physiciens, après la découverte du neutron au tout début des années 1930 (3 séries d'expériences de l'époque ont mis en évidence son existence, dont une menée par les français Irène et François Joliot-Curie en 1931), ont rapidement été intrigués par la similitude très forte du neutron (n) avec le proton (p), tous deux

9.3 SU(2) et le groupe d'isospin : une immersion dans les propriétés de l'interaction nucléaire forte

sensibles de manière exactement semblable à l'interaction nucléaire forte. Ils eurent donc très tôt l'idée de décrire ces particules comme deux états particuliers d'une même particule, le nucléon N . Bien sûr, comme on le sait, leurs masses ne sont pas tout à fait égales : $1,672623.10^{-27}kg$ pour le proton, $1,674928.10^{-27}kg$ pour le neutron. Les physiciens de l'époque l'interprétèrent alors comme la manifestation d'une brisure de symétrie : en l'absence d'interaction électromagnétique, on ne pourrait pas faire de distinction entre un proton et un neutron, car tous les deux auraient la même masse et seraient indiscernables. En revanche en présence d'un champ électromagnétique, de la même manière que la symétrie entre les états de spin d'un électron est brisée lorsqu'on le place dans un champ magnétique, les deux états se différencient.

Mettons en place le formalisme quantique adéquat. En physique quantique, rappelons que les particules sont décrites par des fonctions d'onde Ψ qui vivent dans des espaces hilbertiens. La densité de probabilité de trouver une particule à un temps t avec une position $\mathbf{r}$ et un spin s (sorte de moment angulaire intrinsèque qui n'apparaît qu'en physique quantique) est alors donnée par $|\Psi(\mathbf{r}, t, s)|^2$. Notons par $|p\rangle$ et $|n\rangle$ les fonctions d'ondes respectives du proton et du neutron. L'état général d'un nucléon peut donc se voir comme une combinaison linéaire de $|p\rangle$ et $|n\rangle$. Nous aimerions avoir un moyen de passer de l'état $|p\rangle$ à l'état $|n\rangle$, et vice-versa. Nous remarquons que ce type de transformation pourrait s'opérer naturellement si nous nous plaçons dans la représentation $D^{1/2}$ de $SU(2)$ (nous avons ici fait l'identification). En effet, le fait que :

$$\begin{aligned} D^j(J_+) \left| \frac{1}{2}, m \right\rangle &= \sqrt{\left(\frac{1}{2} - m\right)\left(\frac{1}{2} + m + 1\right)} \left| j, m + 1 \right\rangle \\ D^j(J_-) \left| \frac{1}{2}, m \right\rangle &= \sqrt{\left(\frac{1}{2} + m\right)\left(\frac{1}{2} - m + 1\right)} \left| j, m - 1 \right\rangle \\ D^j(J_3) \left| \frac{1}{2}, m \right\rangle &= m \left| \frac{1}{2}, m \right\rangle \end{aligned}$$

nous montre, comme dans ce cas $m = \pm\frac{1}{2}$, que

9.3 SU(2) et le groupe d'isospin : une immersion dans les propriétés de l'interaction nucléaire forte

$$\begin{aligned}
D^j(J_+) \left| \frac{1}{2}, \frac{1}{2} \right\rangle &= 0 \\
D^j(J_-) \left| \frac{1}{2}, \frac{1}{2} \right\rangle &= \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \\
D^j(J_3) \left| \frac{1}{2}, \frac{1}{2} \right\rangle &= \frac{1}{2} \left(\left| \frac{1}{2}, \frac{1}{2} \right\rangle \right) \\
D^j(J_+) \left| \frac{1}{2}, -\frac{1}{2} \right\rangle &= \left| \frac{1}{2}, \frac{1}{2} \right\rangle \\
D^j(J_-) \left| \frac{1}{2}, -\frac{1}{2} \right\rangle &= 0 \\
D^j(J_3) \left| \frac{1}{2}, -\frac{1}{2} \right\rangle &= -\frac{1}{2} \left(\left| \frac{1}{2}, -\frac{1}{2} \right\rangle \right)
\end{aligned}$$

En d'autres termes, représenter les états $|p\rangle$ et $|n\rangle$ par les deux vecteurs de base $\left| \frac{1}{2}, \frac{1}{2} \right\rangle$ et $\left| \frac{1}{2}, -\frac{1}{2} \right\rangle$ respectivement nous permet :

- de passer de $|p\rangle$ à $|n\rangle$ et de $|n\rangle$ à $|p\rangle$
- d'identifier à l'aide de $D^j(J_3)$ les neutrons et les protons avec les valeurs observables $\pm \frac{1}{2}$
- de déterminer la charge de la particule considérée à l'aide des valeurs propres de l'opérateur qualifié "d'hypercharge" Q défini par

$$Q = e \left(D^j(J_3) + \frac{1}{2} I_{V^j} \right) \quad (17)$$

L'espace $V^{\frac{1}{2}}$ dans lequel "vivent" les nucléons est appelé isoespace ou espace "isotopique". En fait les physiciens sont allés encore plus loin dans ces analogies et ne se sont pas cantonnés à la description des nucléons à l'aide de cet espace d'isospin. Par exemple, d'autres particules sensibles à l'interaction nucléaire forte, les pions, furent découvertes à partir 1947. Comme le proton et le neutron, les pions, qui sont au nombre de trois, π^+ , π^- , π^0 , présentent une masse quasi semblable ($m_{\pi^+} = 139,59 MeV$, $m_{\pi^-} = 139,59 MeV$, $m_{\pi^0} = 135,00 MeV$) et une charge respective de e , $-e$ et 0 . Et de la même manière que nous avons interprété un nucléon N comme étant constitué de deux états p et n , nous pouvons supposer que les trois particules précédentes ne sont que les représentants d'une même et unique particule constituée de trois états distincts, le pion Π . Cette fois nous n'allons

9.3 SU(2) et le groupe d'isospin : une immersion dans les propriétés de l'interaction nucléaire forte

donc pas considérer la représentation irréductible $D^{\frac{1}{2}}$, mais la représentation D^1 de SU(2). En effet il vient :

$$\begin{aligned} D^j(J_+) |1, m\rangle &= \sqrt{(1-m)(m+2)} |1, m+1\rangle \\ D^j(J_-) |1, m\rangle &= \sqrt{(m+1)(2-m)} |1, m-1\rangle \\ D^j(J_3) |1, m\rangle &= m |1, m\rangle \end{aligned}$$

i.e. en groupant trois par trois les transformations en fonction de la base de $\mathfrak{sl}(2)$

$$\begin{aligned} D^j(J_+) |1, 1\rangle &= 0 \\ D^j(J_+) |1, 0\rangle &= \sqrt{2} |1, 1\rangle \\ D^j(J_+) |1, -1\rangle &= \sqrt{2} |1, 0\rangle \end{aligned}$$

$$\begin{aligned} D^j(J_-) |1, 1\rangle &= \sqrt{2} |1, 0\rangle \\ D^j(J_-) |1, 0\rangle &= \sqrt{2} |1, -1\rangle \\ D^j(J_-) |1, -1\rangle &= 0 \end{aligned}$$

$$\begin{aligned} D^j(J_3) |1, 1\rangle &= |1, 1\rangle \\ D^j(J_3) |1, 0\rangle &= 0 \\ D^j(J_3) |1, -1\rangle &= - |1, -1\rangle \end{aligned}$$

Comme pour les nucléons, nous allons donc poser

$$|\pi^+\rangle = |1, 1\rangle \quad |\pi^0\rangle = |1, 0\rangle \quad |\pi^-\rangle = |1, -1\rangle \quad (18)$$

Pour aller plus loin dans la description de certains effets de l'interaction nucléaire forte, nous allons démontrer la formule de Clebsch-Gordan pour le produit tensoriel des représentations, afin de pouvoir décrire des systèmes formés de pions et de nucléons et caractériser certaines transformations/collisions de particules

9.3 SU(2) et le groupe d'isospin : une immersion dans les propriétés de l'interaction nucléaire forte

constituées de nucléons et de pions. Remarquons ici que nous nous bornons à l'étude des pions et des nucléons, mais que les hadrons (particules sensibles à l'interaction forte) connus actuellement sont bien plus nombreux, comme le résumant les deux tableaux en fin de ce rapport (baryons et mésons, nous verrons ce que signifient ces termes lorsque nous introduirons les quarks).

9.3.2 Décomposition de Clebsch-Gordan

On commence par deux définitions. Soit $g \in G$, G un groupe de Lie.

Définition 28 On définit la somme directe $(E_1 \oplus E_2, \rho_1 \oplus \rho_2)$ de deux représentations (E_1, ρ_1) et (E_2, ρ_2) en posant pour tout $x_1 \in E_1$ et $x_2 \in E_2$,

$$((\rho_1 \oplus \rho_2)(g))(x_1 \oplus x_2) = (\rho_1(g)x_1, \rho_2(g)x_2) \quad (19)$$

Définition 29 On définit le produit tensoriel $(E_1 \otimes E_2, \rho_1 \otimes \rho_2)$ de deux représentations (E_1, ρ_1) et (E_2, ρ_2) en posant pour tout $x_1 \in E_1$ et $x_2 \in E_2$,

$$((\rho_1 \otimes \rho_2)(g))(x_1 \otimes x_2) = \rho_1(g)x_1 \otimes \rho_2(g)x_2 \quad (20)$$

Comme cette définition est valable pour un groupe, il s'en suit que pour une algèbre de Lie, à l'aide de l'application différentielle précédemment introduite qui permet de passer d'une représentation ρ de G à une représentation $D\rho$ de $\mathfrak{g}$, nous avons $\forall X \in \mathfrak{g}$

$$D((\rho_1 \otimes \rho_2))(X)(x_1 \otimes x_2) = D(\rho_1)(X)x_1 \otimes x_2 + x_1 \otimes D(\rho_2)(X)x_2 \quad (21)$$

Nous considérons à nouveau les représentations D^j de SU(2) des parties précédentes. Alors nous avons la formule dite de Clebsch-Gordan :

Proposition 28 Soit $j_1, j_2 \in \frac{1}{2}\mathbb{N}$. Alors

$$D^{j_1} \otimes D^{j_2} \approx D^{|j_2-j_1|} \oplus D^{|j_2-j_1|+1} \oplus \dots \oplus D^{j_1+j_2} \quad (22)$$

Preuve Nous avons démontré précédemment que toute représentation de degré fini dans un espace hilbertien est semi-simple, i.e. décomposable en sommes directes hilbertienne irréductible. Par conséquent ceci est valable aussi pour $\mathfrak{sl}(2)$. Considérons l'action de $(R_{j_1} \otimes R_{j_2})(H)$ sur la représentation $D^{j_1} \otimes D^{j_2}$. Alors les valeurs propres de $(R_{j_1} \otimes R_{j_2})(H)$ sont les $j_1 - 2p + j_2 - 2q$ avec $0 \leq p \leq j_1$ et $0 \leq q \leq j_2$. Supposons $j_1 \leq j_2$. Ecrivons les valeurs propres de $(R_{j_1} \otimes R_{j_2})(H)$ sous la forme $j_1 + j_2 - 2k$. Il est assez rapide de voir que l'on est face à trois cas distincts :

9.3 SU(2) et le groupe d'isospin : une immersion dans les propriétés de l'interaction nucléaire forte

- $0 \leq k \leq j_1$. Alors la multiplicité de $j_1 + j_2 - 2k$ est égale à $k + 1$
- $j_1 \leq k \leq j_2$. Alors la multiplicité de $j_1 + j_2 - 2k$ est égale à $j_1 + 1$
- $j_2 \leq k \leq j_1 + j_2$. Alors la multiplicité de $j_1 + j_2 - 2k$ est égale à $j_1 + j_2 - k + 1$

Or en considérant $(R_{|j_2-j_1|} \oplus \dots \oplus R_{j_2+j_1})(H)$ agissant sur la représentation $D^{|j_2-j_1|} \oplus D^{|j_2-j_1+1|} \oplus \dots \oplus D^{j_1+j_2}$ on obtient exactement les mêmes valeurs propres, avec le même ordre de multiplicité. Par conséquent les deux représentations sont équivalentes. ■

9.3.3 Hypothèse d'invariance de l'interaction forte par le groupe d'isospin

Il existe un grand nombre de conséquences qui découlent directement de la description des nucléons et des pions par les représentations $D^{\frac{1}{2}}$ et D^1 . On peut par exemple démontrer que l'hypothèse d'invariance de l'interaction forte par isospin mène directement à l'indépendance de charge des forces nucléaires. Avant de démontrer ce résultat, qu'entend on en mécanique quantique par l'invariance par rapport à un groupe de symétrie. Pour le comprendre, nous noterons par $\hat{U}(\alpha)$ les opérateurs d'un groupe de symétrie (i.e. les éléments de $GL(E)$ issus de la représentation).

Que nous dit la mécanique quantique ?

L'équation d'évolution des fonctions d'onde représentant les particules de la mécanique quantique est représentée par l'équation de Schrödinger :

$$i\hbar \frac{\partial}{\partial t} |\Psi\rangle = \hat{H} |\Psi\rangle \quad (23)$$

où $\hat{H}$, l'opérateur hamiltonien, qualifie l'énergie du système physique. Si le système représenté est invariant par $\hat{U}(\alpha)$, alors si

$$|\Psi'\rangle = \hat{U}(\alpha) |\Psi\rangle$$

l'équation de Schrödinger doit toujours être vérifiée par $|\Psi'\rangle$. En d'autres termes, en multipliant (23) par $\hat{U}(\alpha)$, nous avons :

$$i\hbar \frac{\partial \hat{U}(\alpha) |\Psi\rangle}{\partial t} = \hat{U}(\alpha) \hat{H} \hat{U}^{-1}(\alpha) \hat{U}(\alpha) |\Psi\rangle$$

donc l'invariance par $\hat{U}(\alpha)$ implique :

$$\hat{H} = \hat{U}(\alpha) \hat{H} \hat{U}^{-1}(\alpha)$$

9.3 SU(2) et le groupe d'isospin : une immersion dans les propriétés de l'interaction nucléaire forte

ce qui se traduit directement par :

$$[\hat{U}(\alpha), \hat{H}] = 0 \quad (24)$$

Pour que ce résultat soit valable pour les éléments du groupe de symétrie, il doit être valable pour les générateurs de ce groupe. Dans notre cas pour le groupe d'isospin, les $D^j(J_\pm)$ et $D^j(J_3)$ doivent également vérifier cette relation de commutativité. Considérons alors l'opérateur J défini par :

$$\hat{J} = D^j J_+ + D^j J_-$$

Nous en déduisons par linéarité du crochet de Lie que :

$$[\hat{J}, \hat{H}_{fort}] = 0$$

donc que

$$\hat{H} = \hat{J} \hat{H} \hat{J}^{-1}$$

Pourquoi s'intéresser à $\hat{J}$?

Tout simplement parce que d'après les relations du chapitre précédent, il est facile de voir que $\hat{J} | p \rangle = | n \rangle$ donc que

$$\hat{J} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

i.e.

$$\hat{J}^{-1} = \hat{J} = \hat{J}^t$$

Or, en mécanique quantique, si nous voulons décrire l'interaction entre deux particules (1) et (2), nous considérons qu'elles vivent sur le produit tensoriel $E_1 \otimes E_2$. Par conséquent, la valeur de l'interaction forte entre deux protons sera donnée par le produit scalaire, en notant $\hat{J}_\otimes = \hat{J} \otimes \hat{J}$:

$$\begin{aligned} \langle p(1)p(2) | \hat{H}_{fort} p(1)p(2) \rangle &= \langle \hat{J}_\otimes n(1)n(2) | \hat{H}_{fort} | \hat{J}_\otimes n(1)n(2) \rangle \\ &= \langle \hat{J}n(1) \otimes \hat{J}n(2) | \hat{H}_{fort} | \hat{J}n(1) \otimes \hat{J}n(2) \rangle \\ &= \langle n(1)n(2) | \hat{J}_\otimes^{-1} \hat{H}_{fort} \hat{J}_\otimes | n(1)n(2) \rangle \\ &= \langle n(1)n(2) | \hat{H}_{fort} | n(1)n(2) \rangle \end{aligned}$$

En d'autres termes, l'interaction forte entre deux neutrons et deux protons est identique, donc indépendante de la charge. Ce qui semble observé expérimentalement.

9.3 SU(2) et le groupe d'isospin : une immersion dans les propriétés de l'interaction nucléaire forte

9.3.4 Description des états à plusieurs particule : prévoir la probabilité de certaines réactions nucléaires

Nous allons dans cette partie nous pencher sur la description des états possibles à plusieurs particules. Nous allons tout d'abord nous attacher à décrire l'isospin d'un système de deux nucléons ainsi que l'isospin d'un système pion-nucléon, puis allons discuter de la faisabilité de certaines réactions nucléaires.

Isospin d'un système de deux nucléons

Rappelons encore une fois qu'en physique quantique, lorsqu'on cherche à représenter un système de deux particules, on utilise le produit tensoriel des fonctions d'ondes associées. Par exemple, si l'on considère l'évolution quantique d'un système proton-proton, nous nous intéresserons à la fonction d'onde $|\Psi\rangle = |p\rangle \otimes |p\rangle$. Il s'en suit que l'on décrira un système de deux nucléons par la représentation

$$D^{\frac{1}{2}} \otimes D^{\frac{1}{2}} \approx D^0 \oplus D^1$$

d'après la formule de Clebsch-Gordan. Essayons de déterminer une base des deux représentations irréductibles de droite en fonction de la base ($|p\rangle \otimes |p\rangle$, $|p\rangle \otimes |n\rangle$, $|n\rangle \otimes |p\rangle$, $|n\rangle \otimes |n\rangle$) de $D^{\frac{1}{2}} \otimes D^{\frac{1}{2}}$. Pour ce faire, notons par $|\Psi_{ij}\rangle$, $i, j = 0, 1$ les éléments de base de $D^0 \oplus D^1$. Alors en faisant agir l'opérateur $L_- = (D^0 \oplus D^1)(J_-)$ sur $|\Psi_{ij}\rangle$, d'après les formules que nous avons obtenu au chapitre précédent :

$$\begin{aligned} L_- |\Psi_{11}\rangle &= \sqrt{2} |\Psi_{10}\rangle \\ L_- |\Psi_{10}\rangle &= \sqrt{2} |\Psi_{1-1}\rangle \\ L_- |\Psi_{1-1}\rangle &= 0 \\ L_- |\Psi_{00}\rangle &= 0 \end{aligned}$$

Or, en raison de l'isomorphisme, en identifiant $|\Psi_{11}\rangle$ avec $|p\rangle \otimes |p\rangle$ et en supposant que L_- agit sur le même espace

$$\begin{aligned} \sqrt{2} |\Psi_{10}\rangle &= L_- |\Psi_{11}\rangle = ((D^{\frac{1}{2}})(J_-) |p\rangle) \otimes |p\rangle + |p\rangle \otimes ((D^{\frac{1}{2}})(J_-) |p\rangle) \\ &= |n\rangle \otimes |p\rangle + |p\rangle \otimes |n\rangle \end{aligned}$$

D'où

$$\begin{aligned} L_- |\Psi_{10}\rangle &= \frac{1}{\sqrt{2}} L_- (|n\rangle \otimes |p\rangle + |p\rangle \otimes |n\rangle) = \\ &= |n\rangle \otimes |p\rangle + |p\rangle \otimes |n\rangle = \frac{2}{\sqrt{2}} |n\rangle \otimes |n\rangle \\ &= \sqrt{2} |\Psi_{1-1}\rangle \end{aligned}$$

9.3 SU(2) et le groupe d'isospin : une immersion dans les propriétés de l'interaction nucléaire forte

Enfin en cherchant $|\Psi_{00}\rangle$ comme étant orthogonal à $|\Psi_{10}\rangle$, on obtient au final la base :

$$|\Psi_{00}\rangle = \frac{1}{\sqrt{2}} [|p\rangle \otimes |n\rangle - |n\rangle \otimes |p\rangle] \quad (25)$$

$$|\Psi_{11}\rangle = |p\rangle \otimes |p\rangle \quad (26)$$

$$|\Psi_{10}\rangle = \frac{1}{\sqrt{2}} [|n\rangle \otimes |p\rangle + |p\rangle \otimes |n\rangle] \quad (27)$$

$$|\Psi_{1-1}\rangle = |p\rangle \otimes |p\rangle \quad (28)$$

Les physiciens appellent l'état $j = 0$ l'état "singulet", et l'état $j = 1$ l'état "triplet". Les bases des représentations irréductibles D^0 et D^1 ainsi trouvées vont représenter les états possibles du système à deux nucléons, également appelé deutéron. On voit que les états $j = 1$ et $m = \pm 1$ sont les deux états formés respectivement de deux protons et de deux neutrons. On s'attend à ce que ces états soient instables. En fait, l'état triplet $j = 1$ est toujours instable, alors que le singulet $j = 0$ représente l'état fondamental du deutéron. On peut en donner l'interprétation suivante : on voit clairement que $|\Psi_{00}\rangle$ est antisymétrique par changement $|p\rangle \leftrightarrow |n\rangle$ alors que $|\Psi_{10}\rangle$ est symétrique par ce même changement. Or, le principe d'exclusion de Pauli en mécanique quantique nous dit que la fonction d'onde décrivant le deutéron doit être antisymétrique par $|p\rangle \leftrightarrow |n\rangle$. Par conséquent, comme l'expression de la fonction d'onde du deutéron contient à la fois une partie spatiale décrivant le mouvement relatif des nucléons, une partie dans l'espace des spins et une autre dans l'espace des isospins, la partie appartenant à l'espace des isospins doit aussi être antisymétrique, i.e. seul l'état $|\Psi_{00}\rangle$ pourra représenter un état "lié" du deutéron, et donc être stable et observer dans la nature. L'état triplet ne pourra que se manifester dans des réactions de haute énergie.

Isospin d'un système de pion-nucléon

Par un calcul tout à fait similaire à celui que nous venons d'effectuer, on peut chercher les différents états possibles d'un système pion-nucléon. D'après la formule de Clebsch-Gordan, nous avons :

$$D^1 \otimes D^{\frac{1}{2}} \approx D^{\frac{3}{2}} \oplus D^{\frac{1}{2}}$$

Un même calcul que précédemment nous amènerait à trouver les deux bases des représentations irréductibles associées en fonction des $|\Pi\rangle \otimes |N\rangle$ (que nous noterons $|\Pi N\rangle$ par souci de simplifications :

9.3 SU(2) et le groupe d'isospin : une immersion dans les propriétés de l'interaction nucléaire forte

– Base de $D^{\frac{3}{2}}$: alors on a :

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

– Base de $D^{\frac{1}{2}}$: alors on a :

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

Application de la théorie à la production de pion par diffusion proton-deutéron

Nous allons essayer de prévoir de manière qualitative les rapports de probabilité dans la production de pion par diffusion proton-deutéron, sans aborder les phénomènes d'interaction au cours de cette réaction nucléaire. La formule de Clebsch-Gordan et les calculs que nous avons réalisés précédemment, nous montrent que de manière générale, un état d'isospin d'un système à plusieurs particules peut se mettre sous la forme :

$$|\Psi_{kl}\rangle = \sum_{m+m'=l} \beta_{(jj'k)(mm'l)} |j, m\rangle \otimes |j', m'\rangle \quad (29)$$

Les coefficients $\beta_{(jj'k)(mm'l)}$, que nous avons calculé pour un deutéron et un système pion-nucléon, sont dénommés coefficients de Clebsch-Gordan. Imaginons maintenant que l'état d'isospin $|\Psi_{kl}\rangle$ représente une particule qui, au cours d'une réaction quelconque, se désintègre en deux particules distinctes (1) et (2). En mécanique quantique, on peut relier l'état initial, à savoir $|\Psi_{kl}\rangle$, à l'état représentant

9.3 SU(2) et le groupe d'isospin : une immersion dans les propriétés de l'interaction nucléaire forte

les particules finales $|j, m\rangle \otimes |j', m'\rangle$ grâce à l'opérateur $\hat{S}$. Ainsi la probabilité de transition est proportionnelle à :

$$|\langle jm; j'm' | \hat{S} | \Psi_{kl} \rangle|^2$$

où l'on a adopté le formalisme quantique habituel à l'aide des bras-kets : $\langle \phi | \phi' \rangle$ représente le produit scalaire dans E de ϕ par ϕ' . Nous allons imposer que ce produit scalaire sur E soit tel que la base formée par les $|jm\rangle$ soit orthonormale. Alors les coefficients de Clebsch-Gordan peuvent être interprétés comme les coefficients d'une matrice de passage entre deux bases orthonormales, donc comme les coefficients d'une matrice unitaire. La relation d'orthogonalité des coefficients de Clebsch nous permet alors d'expliciter la relation inverse de (29) :

$$|j, m\rangle \otimes |j', m'\rangle = \sum_{m+m'=l} \beta_{(jj'k')(mm'l')} | \Psi_{k'l'} \rangle \quad (30)$$

et par conséquent

$$|\langle jm; j'm' | \hat{S} | \Psi_{kl} \rangle|^2 = \sum_{m+m'=l} \beta_{(jj'k')(mm'l')} \langle \Psi_{kl} | \hat{S} | \Psi_{k'l'} \rangle$$

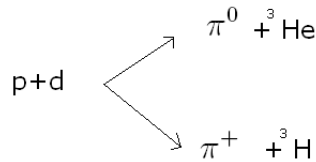
Or, en mécanique quantique, le théorème de Wigner-Eckart, que nous n'établirons pas ici (c.f. Greiner), nous montre que sous l'hypothèse de l'invariance de l'isospin par l'interaction forte, alors

$$\langle \Psi_{kl} | \hat{S} | \Psi_{k'l'} \rangle = \delta_{kk'} \delta_{ll'} \langle k | \hat{S} | k \rangle$$

où nous avons délibérément enlevé la dépendance en l de l'élément de droite, qualifié d'"élément de matrice réduit", qui ne dépend que de k , j et j' . En d'autres termes, si nous faisons le rapport des probabilités pour des combinaisons de charge distinctes apparaissant dans l'état final, alors l'élément de matrice réduit s'en va et il reste le rapport (en indiquant par 1 et 2 les deux types d'états finaux) :

$$r_{\text{reac}} = \frac{\beta_{(jj'k)(m_1 m'_1 l)}^2}{\beta_{(jj'k)(m_2 m'_2 l)}^2} \quad (31)$$

Appliquons dès lors ce résultat à la réaction de diffusion suivante :



9.3 SU(2) et le groupe d'isospin : une immersion dans les propriétés de l'interaction nucléaire forte

Nous savons que dans l'état fondamental le deuteron a l'isospin $j = 0$ alors que le proton est caractérisé par un isospin de 1. Nous savons également que l'isospin des pions vaut 1. Par conséquent, nous avons l'état initial et les états finaux respectifs (en admettant les valeurs de charge et d'isospin de 3H et 3He) :

- état initial : $|\frac{1}{2}\frac{1}{2}\rangle \otimes |00\rangle$
- état final 1 (π^+ et 3H) : $|11\rangle \otimes |\frac{1}{2} - \frac{1}{2}\rangle$
- état final 2 (π^0 et 3He) : $|10\rangle \otimes |\frac{1}{2}\frac{1}{2}\rangle$

Par suite,

$$r_{reac} = \frac{\beta_{(1\frac{1}{2}\frac{1}{2})(1-\frac{1}{2}\frac{1}{2})}^2}{\beta_{(11\frac{1}{2})(0\frac{1}{2}\frac{1}{2})}^2}$$

Or, nous avons explicité au chapitre précédent quand nous considérons les systèmes pion-nucléon :

$$|\Psi_{\frac{1}{2}\frac{1}{2}}\rangle = \frac{\sqrt{2}}{\sqrt{3}} |\pi^+ n\rangle - \frac{1}{\sqrt{3}} |\pi^0 p\rangle$$

En d'autres termes,

$$\beta_{(1\frac{1}{2}\frac{1}{2})(1-\frac{1}{2}\frac{1}{2})}^2 = \frac{2}{3}$$

et

$$\beta_{(11\frac{1}{2})(0\frac{1}{2}\frac{1}{2})}^2 = \frac{1}{3}$$

d'où

$$r_{reac} = 2$$

Ce rapport signifie que la première réaction s'observe plus fréquemment dans la nature d'un facteur 2. Or expérimentalement, les physiciens ont trouvé des valeurs de l'ordre de $r_{reac} \approx 1,91 \pm 0.25$ et $r_{reac} \approx 2,26 \pm 0.11$, montrant que l'invariance de la force nucléaire par l'isospin n'est valable qu'à 10% près.

9.4 Représentations de $SU(3)$

Nous allons maintenant nous intéresser au groupe $SU(3)$ et à son algèbre de Lie (complexifiée) $\mathfrak{sl}(3, \mathbb{C})$. Avant d'expliciter différentes représentations de $SU(3)$, rappelons et introduisons quelques notions mathématiques supplémentaires qui vont nous être utiles.

9.4.1 Quelques notions mathématiques supplémentaires

Lorsque nous avons étudié les systèmes de racines, nous avons fait remarquer le couple (L, H) où L est une algèbre de Lie semi-simple, et H son tore maximal, pouvait être entièrement caractérisée par le couple (Φ, E) , où Φ est le système de racine et E l'espace de la représentation. En fait on peut introduire une notion plus générale que la notion de racine : la notion de poids. Il est alors facile de voir que les racines sont les poids associés à la représentation adjointe.

Avant d'introduire cette notion nous avons besoin d'utiliser un nouveau langage pour les représentations, en introduisant la notion de L -modules (L une algèbre de Lie)

Définition 30 *Un espace vectoriel V muni d'une opération $L \times V \rightarrow V$ (notée $x.v$ pour $x \in L$) est appelé L -module si les conditions suivantes sont satisfaites pour tous $x, y \in L$, $v, w \in V$ et $a, b \in \mathbb{K}$:*

- (M1) $(ax + by).v = a(x.v) + b(y.v)$
- (M2) $x.(av + bw) = a(x.v) + b(x.w)$
- (M3) $[xy].v = x.y.v - y.x.v$

Par conséquent, vu la définition d'une représentation (finie) $\phi : L \rightarrow \mathfrak{gl}(V)$ d'une algèbre de Lie L , on voit que V peut être considéré comme un L -module via l'action $x.v = \phi(x)(v)$. Et réciproquement si V est de dimension finie.

Venons en maintenant à la définition des poids. Soit V un L -module de dimension finie, H un tore maximal de L , L une algèbre de Lie semi-simple. Nous avons fait remarquer précédemment que H commutait avec l'ensemble des éléments de L .

Pour établir notre définition, plaçons nous dans le cas particulier où les algèbres de Lie que nous considérons sont incluses dans $\mathfrak{gl}(n, \mathbb{K})$. Alors il est clair que la représentation $A \mapsto \phi(A) = A$ permet de considérer le L -module $\mathbb{R}^n$ par l'action, pour tout $A \in \mathfrak{gl}(n, \mathbb{K})$ et $x \in \mathbb{R}^n$: $A.x = A \cdot x$ où $\cdot$ désigne le produit usuel d'une matrice carré de dimension n par un élément de $\mathbb{R}^n$. Par conséquent nous voyons que l'ensemble des éléments de H sont simultanément diagonalisables et

que $V = \mathbb{R}^n$ se décompose comme la somme directe des V_λ :

$$V = \Pi_\lambda V_\lambda$$

où $\lambda \in H^*$ et $V_\lambda = \{v \in V | h.v = \lambda(h)v\}$.

Définition 31 *Les éléments V_λ pour λ non nul sont appelés les espace associés aux **poids** de la représentation ϕ . Les λ non nuls sont les **poids** de cette représentation.*

Nous avons volontairement laissé une notation générale pour V (et ne l'avons pas remplacée par $\mathbb{R}^n$). En fait, les poids peuvent se définir dans un cadre plus général, en introduisant la notion de sous-algèbre de Cartan (le tore maximal est un cas particulier de sous-algèbre de Cartan). A tout L -module de dimension fini V , on peut alors associer des poids et leurs espaces correspondant, qui satisferont exactement la même définition que ci-dessus en remplaçant H par une sous-algèbre de Cartan.

Dans tous les raisonnements suivants, nous allons nous baser sur la semi-simplicité de $\mathfrak{sl}(3, \mathbb{C})$. Les affirmations précédentes nous montrent que l'on peut caractériser entièrement l'espace associé à la représentation par les poids de cette dernière. Comme en physique, nous souhaitons décrire les particules comme des bases des espaces de représentations irréductibles, nous voyons tout l'intérêt d'associer les racines et les poids de $\mathfrak{sl}(3, \mathbb{C})$ à de nouvelles particules

9.4.2 Etude des racines de $\mathfrak{sl}(3, \mathbb{C})$

Notons $\mathfrak{h}$ la sous algèbre de Lie abélienne de $\mathfrak{sl}(3, \mathbb{C})$ formée des matrices diagonales de trace nulle. Il est facile de voir que cette sous-algèbre est maximale, car toute sous algèbre de $\mathfrak{sl}(3, \mathbb{C})$ comportant une matrice non diagonale de trace nulle A ne pourrait être commutative (A ne pourrait commuter avec toute les matrices diagonales de trace nulle). $\mathfrak{h}$ est une sous-algèbre de Cartan et permettra donc de caractériser à l'aide des poids et des racines l'espace de la représentation associée. Il est alors facile de voir que les matrices

$$H_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, H_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

forment une base de $\mathfrak{h}$. Posons alors

$$H_3 = H_1 + H_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

et notons par E_{ij} , $i, j = 1..3$ les matrices élémentaires. Les E_{ij} , $i, j = 1..3$ pour $j \neq i$ et les $H_i = E_{ii} - E_{i+1, i+1}$ forment une base de $\mathfrak{sl}(3, \mathbb{C})$. Notons par ailleurs, $\forall A \in \mathfrak{h}$, $diag_i : A \mapsto diag_i(A) = A_{ii}$ l'application linéaire de $\mathfrak{h}$ dans $\mathbb{C}$. Alors il est facile de voir que $\forall A \in \mathfrak{h}$,

$$[A, E_{ij}] = (diag_i(A) - diag_j(A)) E_{ij} \quad (32)$$

9.4.3 Représentation 8

Considérons les formes linéaires sur $\mathfrak{h}$, α_k et α'_k , $k = 1..3$, définies par :

$$\alpha_1 = diag_1 - diag_2, \alpha_2 = diag_2 - diag_3, \alpha_3 = diag_1 - diag_3 \quad (33)$$

$$\alpha'_k = -\alpha_k \quad (34)$$

Ce sont les racines de $\mathfrak{sl}(3, \mathbb{C})$. En effet, d'après (32), les $\alpha_k(H_i)$ pour $k, i = 1..3$ sont les valeurs propres des endomorphismes ad_{H_i} correspondants aux vecteurs propres respectifs

$$E_{k=1} = E_{12}, E_{k=2} = E_{23}, E_{k=3} = E_{13}$$

De même, les $\alpha'_k(H_i)$ pour $k, i = 1..3$ sont les valeurs propres des endomorphismes ad_{H_i} correspondant aux vecteurs propres respectifs

$$F_{k=1} = E_{21}, F_{k=2} = E_{32}, F_{k=3} = E_{31}$$

Les physiciens ne se servent en fait pas directement de la base (H_1, H_2) de $\mathfrak{h}$, mais plutôt des deux bases (I_3, Y) et (I_3, T_8) de $\mathfrak{h}$ définies par :

$$I_3 = \frac{1}{2}H_1, Y = \frac{1}{3}(H_1 + 2H_2) \quad (35)$$

$$T_8 = \frac{\sqrt{3}}{2}Y \quad (36)$$

Alors on peut représenter les racines α_k et α'_k dans la base (I_3, T_8) . On aperçoit que dans cette base les racines se répartissent sur les sommets d'un hexagone. Notons également que comme 0 est valeur propre double de l'endomorphisme ad_H pour tout $H \in \mathfrak{h}$, et bien que ce ne soit pas une racine proprement dite, on le représente également sur le graphe des racines. La représentation que l'on obtient est alors de dimension 8 et est qualifiée de "représentation de l'octet **8**" :

9.4 Représentations de $SU(3)$

racine	valeurs	base (I_3, Y)	base (I_3, T_8)
α_1	$\alpha_1(I_3) = 1, \alpha_1(Y) = 0$	$(1, 0)$	$(1, 0)$
α_2	$\alpha_2(I_3) = -\frac{1}{2}, \alpha_2(Y) = 1$	$(-\frac{1}{2}, 1)$	$(-\frac{1}{2}, \frac{\sqrt{3}}{2})$
α_3	$\alpha_3(I_3) = \frac{1}{2}, \alpha_3(Y) = 1$	$(\frac{1}{2}, 1)$	$(\frac{1}{2}, \frac{\sqrt{3}}{2})$
α'_1	$\alpha'_1(I_3) = -1, \alpha'_1(Y) = 0$	$(-1, 0)$	$(-1, 0)$
α'_2	$\alpha'_2(I_3) = \frac{1}{2}, \alpha'_2(Y) = -1$	$(\frac{1}{2}, -1)$	$(\frac{1}{2}, -\frac{\sqrt{3}}{2})$
α'_3	$\alpha'_3(I_3) = -\frac{1}{2}, \alpha'_3(Y) = -1$	$(-\frac{1}{2}, -1)$	$(-\frac{1}{2}, -\frac{\sqrt{3}}{2})$

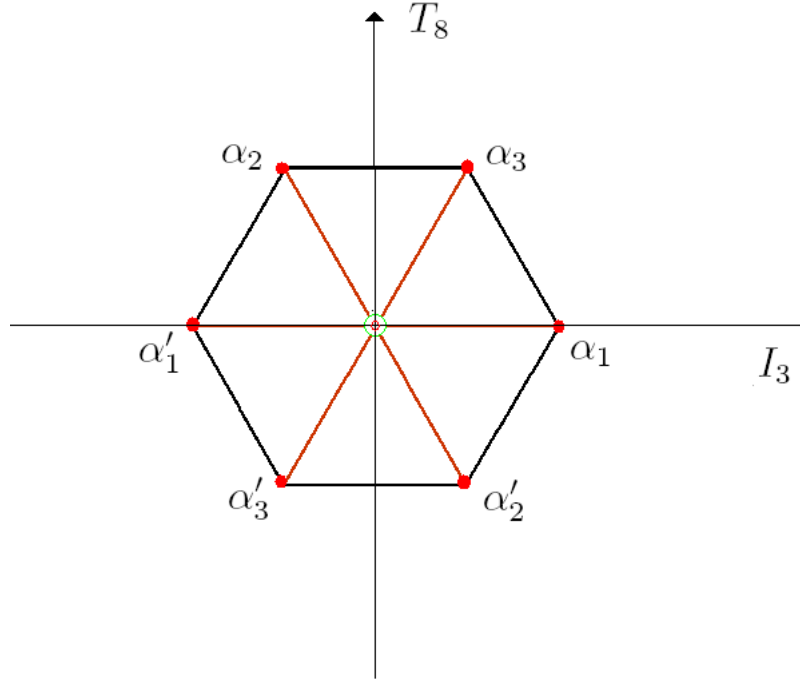


FIGURE 1 – Racines de $\mathfrak{sl}(3, \mathbb{C})$ et représentation **8**

9.4.4 Représentations **3** et $\bar{\mathbf{3}}$

Les représentations **3** et $\bar{\mathbf{3}}$ sont très simples à établir. Notons par e_1, e_2 et e_3 les vecteurs de base canonique de $\mathbb{R}^3$. Définissons les applications λ_i pour $i = 1..3$ par : $\forall H \in \mathfrak{h}$,

$$\begin{aligned} H e_i &= \lambda_i(H) e_i \\ -{}^t H e_i &= \lambda'_i(H) e_i \end{aligned}$$

9.4 Représentations de $SU(3)$

Il en résulte facilement que dans les bases (I_3, Y) et (I_3, T_8) , nous avons :

λ_i	valeurs	base (I_3, Y)	base (I_3, T_8)
λ_1	$\lambda_1(I_3) = \frac{1}{2}, \lambda_1(Y) = \frac{1}{3}$	$(\frac{1}{2}, \frac{1}{3})$	$(\frac{1}{2}, \frac{1}{2\sqrt{3}})$
λ_2	$\lambda_2(I_3) = -\frac{1}{2}, \lambda_2(Y) = \frac{1}{3}$	$(-\frac{1}{2}, \frac{1}{3})$	$(-\frac{1}{2}, \frac{1}{2\sqrt{3}})$
λ_3	$\lambda_3(I_3) = 0, \lambda_3(Y) = -\frac{2}{3}$	$(0, -\frac{2}{3})$	$(0, -\frac{1}{\sqrt{3}})$
λ'_1	$\lambda'_1(I_3) = -\frac{1}{2}, \lambda'_1(Y) = -\frac{1}{3}$	$(-\frac{1}{2}, -\frac{1}{3})$	$(-\frac{1}{2}, -\frac{1}{2\sqrt{3}})$
λ'_2	$\lambda'_2(I_3) = \frac{1}{2}, \lambda'_2(Y) = -\frac{1}{3}$	$(\frac{1}{2}, -\frac{1}{3})$	$(\frac{1}{2}, -\frac{1}{2\sqrt{3}})$
λ'_3	$\lambda'_3(I_3) = 0, \lambda'_3(Y) = \frac{2}{3}$	$(0, \frac{2}{3})$	$(0, \frac{1}{\sqrt{3}})$

Comme précédemment pour la représentation **8**, on peut représenter graphiquement les poids λ_i et λ'_i . Nous noterons **3** la représentation associée aux λ_i , $\bar{\mathbf{3}}$ celle associée aux λ'_i .

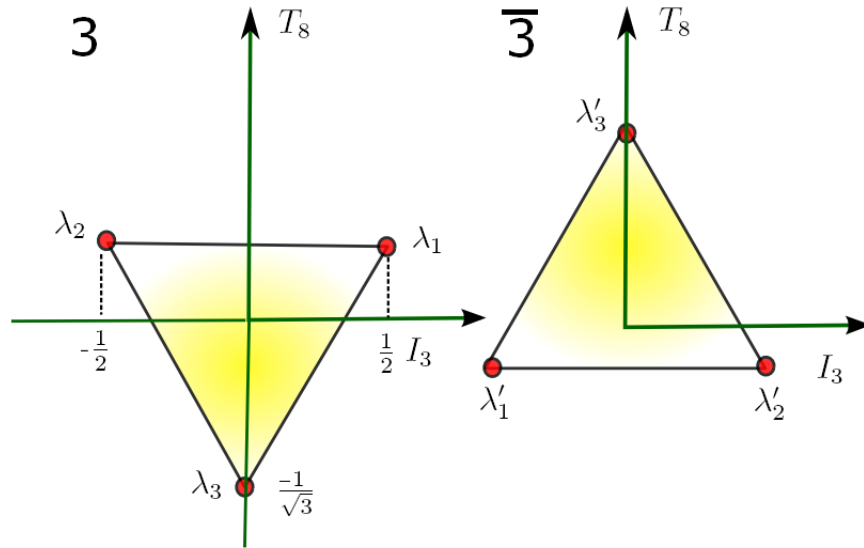


FIGURE 2 – Diagramme des poids des représentations **3** et $\bar{\mathbf{3}}$

9.4.5 Représentation 10

Avant de passer à l'analogie entre les poids d'une représentation de $SU(3)$ et les particules élémentaires, nous allons développer une dernière représentation, dénommée la représentation 10, car elle comporte 10 poids. Cette représentation nous sera utile par la suite pour classer les "baryons". Il est facile de voir que tout poids de $\mathfrak{sl}(3), \mathbb{C}$, d'après la structure de $\mathfrak{h}$, se décompose comme une combinaison linéaire des $diag_i$ précédemment définis, i.e. si l'on note w un poids d'une représentation quelconque,

$$w = m_1 diag_1 + m_2 diag_2 + m_3 diag_3$$

Le triplet (m_1, m_2, m_3) n'est pas unique, car on peut remarquer que comme les éléments de $\mathfrak{h}$ sont de trace nulle, on a $diag_1 + diag_2 + diag_3 = 0$, donc on peut toujours se ramener à annuler un des m_i avec les autres positifs ou nuls. Dans les exemples que nous avons précédemment développé, il est facile de voir par exemple que l'on peut représenter les λ_i par :

$$\begin{aligned}\lambda_1 &= (1, 0, 0) \\ \lambda_2 &= (0, 1, 0) \\ \lambda_3 &= (0, 0, 1)\end{aligned}$$

λ_1 est un poids maximal de la représentation **3** : un vecteur poids maximal (ici en l'occurrence e_1) lui est associé, avec e_1 vérifiant, en notant $\rho(H_i)(e_k) = H_i e_k$ et en reprenant la notation E_i précédente pour $i = 1..3$:

- $\rho(H_i)(e_1) = \lambda_1(H_i)e_1$
- $\rho(E_i)(e_1) = 0$
- les images de e_1 par les $\rho(F_i)$ (c.f. notations précédentes) engendrent l'espace de la représentation.

On pourrait de même trouver un vecteur poids maximal pour toutes les représentations de $\mathfrak{sl}(3, \mathbb{C})$ que nous venons d'évoquer. En fait ceci découle d'un résultat plus général, qui stipule que toute représentation irréductible ρ de dimension finie de $\mathfrak{sl}(3, \mathbb{C})$ admet un vecteur poids maximal (qui satisfasse les critères précédents). La représentation 10 est la représentation qui admet pour poids maximal $w = 3diag_1$. A partir du poids maximal, on peut alors calculer l'ensemble des autres vecteurs poids de la représentation correspondante. Par exemple, en notant par v le vecteur poids maximal correspondant, comme $Hv = w(H)v$, il vient :

$$H(F_i v) = F_i(Hv) + [H, F_i](v) = (w(H) - \alpha_i(H))(F_i v) = (3diag_1 - \alpha_i(H))(F_i v)$$

En d'autres termes, en ne notant que les termes non égaux, nous obtenons les poids :

$$v_1 = H(F_1 v) = (3diag_1(H) - \alpha_1(H))(F_1 v) = (2diag_1(H) + diag_2(H))(F_1 v)$$

$$v'_1 = H(F_1 v_1) = (diag_1(H) + 2diag_2(H))(F_1 v_1)$$

$$v''_1 = H(F_1 v'_1) = (3diag_2(H))(F_1 v''_1)$$

$$v_3 = H(F_3 v) = (3diag_1(H) - \alpha_3(H))(F_3 v) = (diag_1(H) - diag_2(H))(F_3 v)$$

$$v'_3 = H(F_3 v_3) = (diag_1(H) + 2diag_3(H))(F_3 v_3)$$

$$v''_3 = H(F_3 v'_3) = (3diag_3(H))(F_3 v''_3)$$

$$v_2 = H(F_2 v''_1) = (3diag_2(H) - \alpha_2(H))(F_2 v''_1) = (2diag_2(H) + diag_3(H))(F_2 v''_1)$$

$$v'_2 = H(F_2 v_2) = (diag_2(H) + 3diag_3(H))(F_2 v'_2)$$

$$v''_2 = H(F_1 v_3) = 0$$

Les autres combinaisons que nous obtiendrions seraient identiques ou multipliées par un scalaire non nul. Par suite, les poids que nous venons de trouver peuvent, comme précédemment, se classer sur un diagramme de représentation dans la base (I_3, T_8) . En faisons cela, nous obtiendrions le diagramme (nous avons omis ici les noms, nous verrons ensuite quand nous ferons l'analogie avec les particules élémentaires comment les noter) :

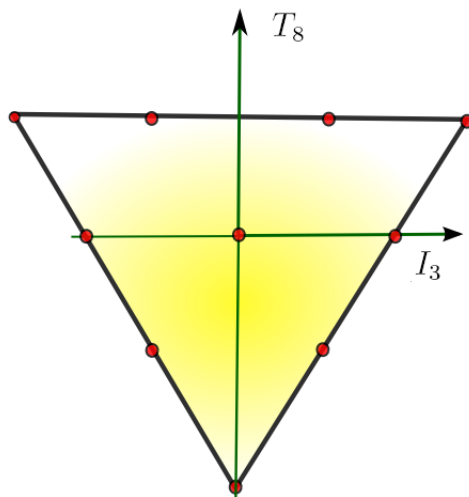


FIGURE 3 – Diagramme des poids de la représentation 10

9.5 SU(3) et les quarks

Lorsque nous nous étions penchés sur l'analogie entre protons et neutrons, nous avons remarqué que les transformations induites par la représentation $D^{1/2}$ de $SU(2)$ permettaient de passer de l'un à l'autre à l'aide des opérateurs associés. La valeur propre $\pm \frac{1}{2}$ de l'opérateur $D^{1/2}(J_3)$, l'isospin, nous permettait notamment de différencier protons et neutrons. Qui plus est, nous avons également vu qu'un grand nombre de particules dans la nature semblaient posséder des propriétés analogues, notamment du point de vue de leur masse. Et pouvaient par conséquent être décrites par une représentation D^j . C'est le cas des pions que nous avons introduits précédemment ($j = 1$), mais nous aurions pu aussi citer des particules plus exotiques, dénommées sous des noms étranges, comme les groupements $(\Sigma^+, \Sigma^0, \Sigma^-)$ ($j = 1$) ou encore $(K^+, K^0, K^-, \bar{K}^0)$ ($j = \frac{3}{2}$). En fait ces groupements semblent pouvoir s'interpréter comme des sous-ensembles d'un groupement plus large. En effet, en s'intéressant au spin de ces particules ("leur moment angulaire intrinsèque") ainsi qu'à leur parité (l'opérateur de parité P en mécanique quantique inverse les variables de positions et a donc seul de valeurs propres possibles : ± 1 (car $P^2 = Id$)), les physiciens des années 60 ont remarqué qu'ils pouvaient regrouper la plupart des particules obéissant à l'interaction forte (connues à l'époque) dans deux ensembles distincts : celui des particules de spin 1/2 et de parité positive,

les baryons, et celui des particules de spin nul et de parité négative, les mésons scalaires :

baryons		mésons scalaires	
Particule	Masse (MeV)	Particule	Masse
p	$938,28 \pm 0,005$	π^+	$139,58 \pm 0,01$
n	$939,57 \pm 0,005$	π^0	$134,97 \pm 0,01$
Λ^0	$115,6 \pm 0,006$	π^-	$139,58 \pm 0,01$
Σ^+	$1189,4 \pm 0,1$	K^+	$493,8 \pm 0,1$
Σ^0	$1192,5 \pm 0,1$	K^-	$493,8 \pm 0,1$
Σ^-	$1197,4 \pm 0,07$	K^0	$497,8 \pm 0,1$
Θ^0	$1314,7 \pm 0,7$	$\bar{K}^0$	$497,8 \pm 0,1$
Θ^-	$1321,3 \pm 0,2$	η^0	$548,8 \pm 0,6$

FIGURE 4 – Classification en octet des baryons et des mésons

De la même manière que l'on peut décrire les protons et les neutrons comme le doublet d'isospin représenté par $j = 1/2$, ils eurent l'idée de représenter l'ensemble des baryons comme un supermultiplet, et l'ensemble des mésons comme un autre supermultiplet. Pour différencier les différentes particules au sein de ces multiplets, les physiciens de l'époque introduisirent deux nouveaux nombres quantiques : l'hypercharge, dont nous avons déjà un peu parlé, et la charge baryonique. En remarquant que la charge des nucléons est donnée par $Q = e \left(D^{\frac{1}{2}}(J_3) + \frac{1}{2}Id \right)$ et celle des pions par $Q = e D^1(J_3)$, on peut définir le nouveau nombre quantique Y (l'hypercharge à fortiori) par :

$$Q = D^j(J_3) + \frac{Y}{2}$$

Et de la même manière que l'on peut décrire la charge de l'ensemble des particules élémentaires à l'aide de la charge élémentaire e , les physiciens donnèrent au proton la charge baryonique $B(p) = 1$. En supposant cette charge conservée lors des phénomènes (collisions/créations de particules) impliquant l'action de l'interaction forte, ils purent alors assimiler à l'ensemble des particules connues réagissant à l'interaction forte une charge baryonique.

Or, en représentant les baryons et les mésons précédemment cités dans le plan $(Y, D^j(J_3))$, ils eurent la bonne surprise de découvrir les diagrammes suivants :

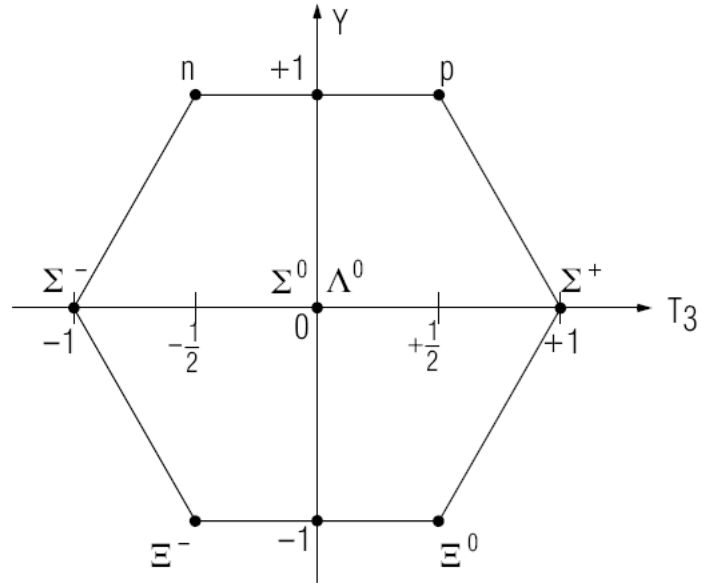
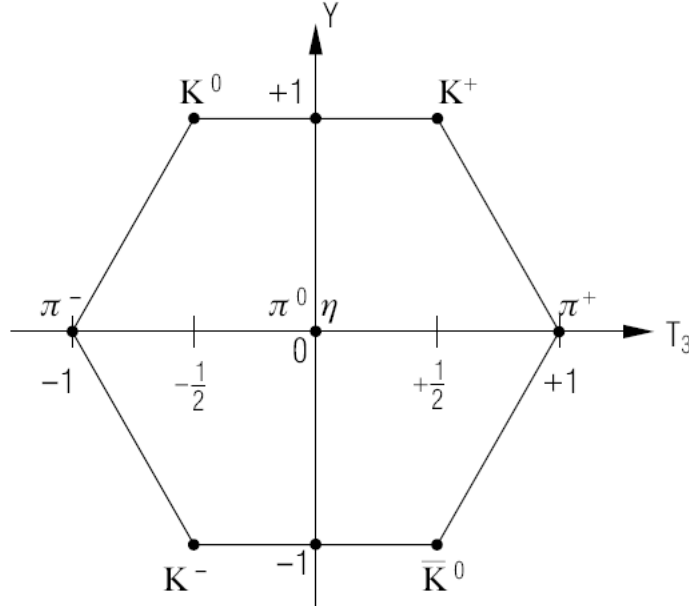


FIGURE 5 – Représentation des baryons dans le diagramme $(Y, D^j(J_3) = T_3)$


 FIGURE 6 – Représentation des mésons dans le diagramme $(Y, D^j(J_3) = T_3)$

En d'autres termes, nous voyons que la représentation dans $(Y, D^j(J_3))$ des baryons et des mésons coïncide exactement avec la représentation des racines α_i et α'_i dans la base (I_3, Y) . Ces observations conduisirent Gell-Man et Ne'eman à identifier, au début des années 60, les particules élémentaires avec les poids de différentes représentations de $\mathfrak{sl}(3, \mathbb{C})$. Si ce système marchait plutôt bien, les représentations $\mathbf{3}$ et $\bar{\mathbf{3}}$ ne correspondaient encore à aucune particule observée. Pourtant il faut remarquer une chose fondamentale : nous avons fait observer dans les paragraphes précédents qu'un système à plusieurs particules était représenté par un produit tensoriel. En clair, si jamais les représentations $\mathbf{3}$ et $\bar{\mathbf{3}}$ correspondaient à des particules, nous pourrions décrire un système constitué de ce type de particules par un produit tensoriel de représentations de $SU(3)$. Or, rappelons qu'au paragraphe précédent nous avons montré que la différentielle de la représentation (au sens des groupes, car rappelons que nous avons bijection entre les représentations irréductibles de $\mathfrak{sl}(3, \mathbb{C})$ et celles de $SU(3)$) $\rho_1 \otimes \rho_2$ était donnée par $(D\rho_1) \otimes \rho_2 + (D\rho_2) \otimes \rho_1$. Par conséquent, nous voyons que nous avons la construction suivante pour les diagrammes de poids d'une représentation (au sens des algèbres de Lie) :

Proposition 29 *Si v_1 et v_2 sont les vecteurs poids respectifs des représentations*

R_1 et R_2 de $\mathfrak{sl}(3, \mathbb{C})$ avec pour poids w_1 et w_2 , alors $v_1 \otimes v_2$ est un vecteur poids de la représentation $R_1 \otimes R_2$ de poids $w_1 + w_2$

Par conséquent, pour représenter graphiquement le diagramme de poids de la représentation $R_1 \otimes R_2$ de $\mathfrak{sl}(3, \mathbb{C})$, il suffit additionner les vecteurs poids. En faisant cela avec $\mathbf{3} \otimes \bar{\mathbf{3}}$ nous obtenons :

$$\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1}$$

où $\mathbf{1}$ est un singlet à l'origine.

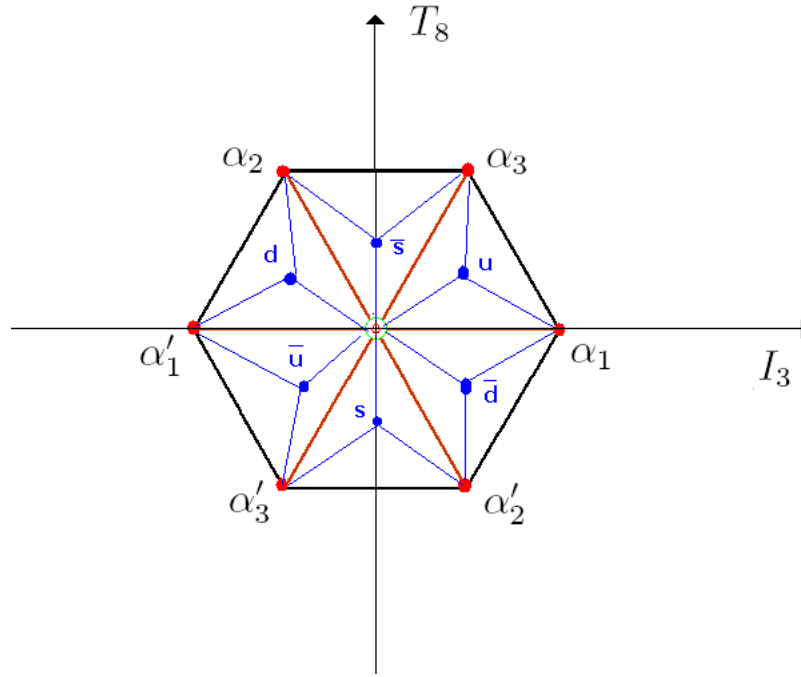


FIGURE 7 – En additionnant vectoriellement les poids de $\mathbf{3}$ et $\bar{\mathbf{3}}$ on obtient la représentation $\mathbf{3} \otimes \bar{\mathbf{3}}$

Sur le diagramme précédent nous avons volontairement changé les λ_i et λ'_i en u , d et s pour les λ_i et $\bar{u}$, $\bar{d}$ et $\bar{s}$ pour les λ'_i . Les hypothétiques particules u , d et s furent dénommées quarks, avec les surnoms respectifs de *up*, *down* et *strange*. Les particules barrées sont les antiquarks correspondants. Un rapide calcul nous montre en effet que, d'après la définition de l'hypercharge, les particules u , d et s ont pour charges respectives $Q_u = \frac{2e}{3}$, $Q_d = -\frac{e}{3}$ et $Q_s = -\frac{e}{3}$ et que les particules $\bar{u}$,

$\bar{d}$ et $\bar{s}$ ont une charge opposée. On est alors très surpris de voir que ces particules ont une charge en valeur absolue inférieure à $|e|$, alors que e est censée être indivisible. Existes-ils donc vraiment ?

A l'heure actuelle, les quarks n'ont jamais été observés séparément dans un accélérateur de particules. Pourtant il semble bien que l'ensemble des particules élémentaires sensibles à l'interaction forte soit constitué de quarks. En effet, si jamais les mésons scalaires précédemment introduits étaient constitués d'un quark et d'un antiquark, alors il existerait une autre particule, le méson η' qui serait situé en $(0,0)$ comme η et π^0 (correspondant au singulet **1**). Or ce méson fut découvert en 1962, juste après l'identification faite par Gell-Man et Ne'eman. Et la théorie des quarks va bien plus loin.

Par exemple, on peut s'amuser à représenter le diagramme $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$, i.e. l'ensemble des particules élémentaires constituées de trois quarks. En faisant $\mathbf{3} \otimes \mathbf{3}$, il est facile de voir que nous obtenons graphiquement :

$$\mathbf{3} \otimes \mathbf{3} = \mathbf{6} \oplus \bar{\mathbf{3}}$$

donc

$$\begin{aligned} \mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} &= (\mathbf{3} \otimes \mathbf{6}) \oplus (\mathbf{3} \otimes \bar{\mathbf{3}}) \\ &= \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1} \end{aligned}$$

Le résultat graphique est représenté ci-dessous :

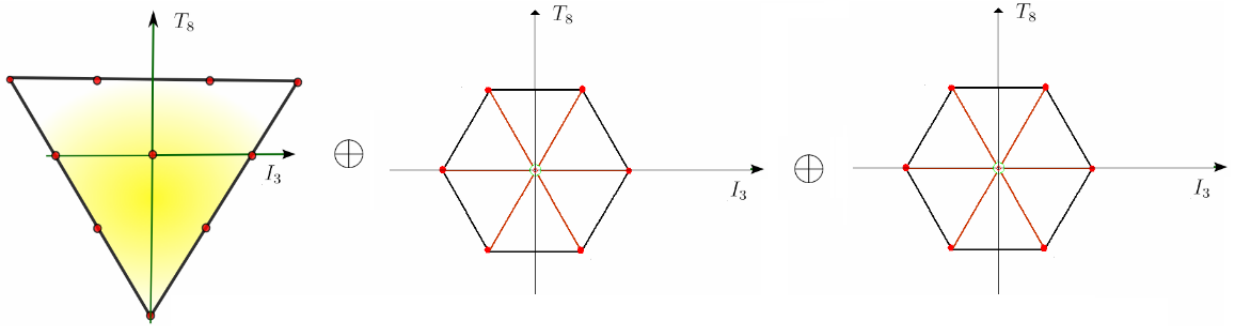


FIGURE 8 – Représentation $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$

Or les physiciens trouvèrent que l'ensemble des baryons actuellement connus de spin $\frac{3}{2}$ et de parité positive, et ceux de spin $\frac{1}{2}$ et de parité positive se représentaient sur leurs graphes (Y, I_3) par :

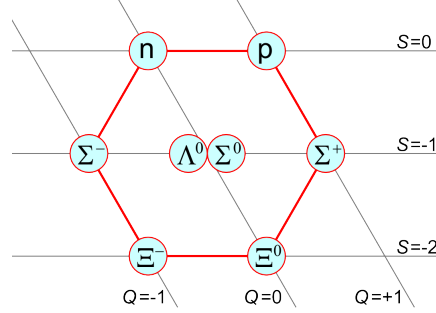


FIGURE 9 – Octet (spin $\frac{1}{2}$) des baryons de parité positive

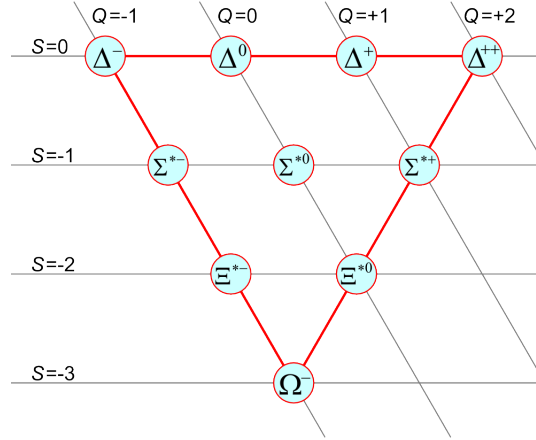


FIGURE 10 – décuplet (spin $\frac{3}{2}$) des baryons de parité positive

Par conséquent, le fait que ces baryons soient constitués de trois combinaisons de quarks u , d , et s , semble donc ne faire aucun doute. Ce résultat est ainsi largement admis dans la communauté scientifique actuelle, et l'existence des quarks semble aujourd'hui ne plus pouvoir être remise en cause, même si les physiciens n'ont jamais encore réussi à les observer séparément. Ce défi sera peut être rendu possible grâce au nouveau accélérateur de particules, le LHC, qui va entrer en service à la fin de cette année.

9.6 Pour aller plus loin ...

En fait les trois quarks up, down et strange ne sont pas les seuls quarks apparemment existants. En effet, des considérations théoriques ont amenées les physiciens, dont Glashow à partir de 1964, à supposer l'existence d'autres quarks, au nombre de trois : les quarks charme, top et bottom. En plus de la charge hadronique, de la charge, de l'isospin, ils introduisirent un nouveau nombre quantique, la charge de couleur. Dans cette description, les particules sensibles à l'interaction forte ne sont plus alors décrites par $SU(3)$ mais par $SU(4)$, ce qui complique encore les représentations (il faut se placer en trois dimensions). Ici encore, la découverte de nombreuses particules se plaçant sur les poids des représentations de $SU(4)$ dans les accélérateurs semble confirmer l'existence de ces six quarks.

Récemment, un physicien amateur du nom de Garrett Lisi, a proposé de représenter l'ensemble des champs de particules élémentaires connus (particules + particules virtuelles véhiculant les interactions) à l'aide du système de racine du groupe E_8 , ce en se basant sur le modèle standard (dont nous avons donné dans ce rapport un très court aperçu). Sa théorie permettrait ainsi de décrire l'ensemble des champs constituant l'univers dans un même principe unificateur. Sa théorie, bien que très esthétique sur le papier, est cependant encore hautement spéculative et très discutée dans le milieu des physiciens théoriciens.

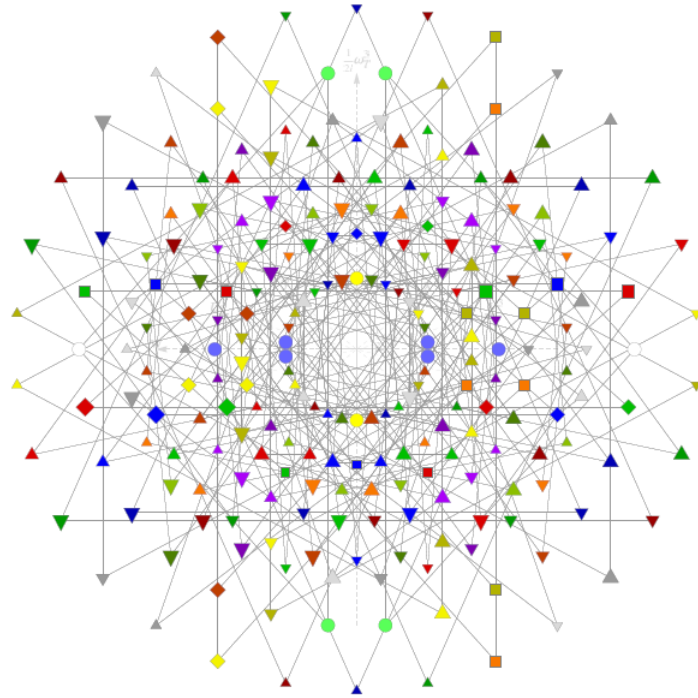


FIGURE 11 – Représentation du système de racine du groupe E_8

Conclusion

Après s'être penchés sur les groupes topologiques et la mesure de Haar, notre incursion sur le terrain des groupes de Lie nous a montré que nous pouvions associer à tout sous-groupe fermé de $GL(n, \mathbb{C})$ une algèbre de Lie à l'aide de l'application exponentielle. Comme nous l'avions alors fait remarquer, ce résultat se généralise à tout groupe de Lie. Cette constatation générale nous a alors permis d'introduire les systèmes de racines définis pour des algèbres de Lie semi-simples, que nous avons pu représenter à l'aide des diagrammes de Dynkin.

Comme nous l'avions alors énoncé sans pour autant pouvoir le démontrer par manque de temps, la correspondance biunivoque entre les couples (L, H) et (Φ, E) où L est une algèbre de Lie semi-simple, H son tore maximal et Φ son système de racine et E l'espace euclidien associé nous montre que nous pouvons classifier les algèbres de Lie semi-simples, et donc les groupes de Lie correspondants, à l'aide de leurs systèmes de racine.

Forts de ces résultats, nous avons tenté, dans la partie faisant le lien entre les groupes de Lie et la physique, de montrer que les symétries de la nature pouvaient être décrites à l'aide des groupes de Lie.

Le fait notamment que les particules sensibles à l'interaction fortes puissent être vues comme des poids de représentations de $SU(3)$ nous montre que le groupe de Lie $SU(3)$ joue un rôle fondamental en physique nucléaire. Sa présence comme groupe de symétrie implique la conservation de nouveaux nombres quantiques lors des réactions nucléaires, à savoir l'isospin, l'hypercharge et le nombre baryonique. On ne peut alors s'empêcher d'être émerveillé face à la vision si esthétique qui consisterait à identifier l'ensemble des particules de la nature aux poids d'une représentation d'un groupe de Lie particulier.

La recherche de cette beauté intrinsèque pourrait peut être la voie vers une nouvelle physique....

Références

- [Bump] D.Bump, *Lie groups*, Springer, 2004
- [Bekka] B.Bekka, *Théorème du point fixe de Markov-Kakutani : applications à l'existence de mesures de Haar et aux chaines de Markov*, article, 2005
- [Bourbaki] N.Bourbaki, *Eléments de mathématiques, fascicule XXXIV, Groupes et algèbres de Lie Chapitres 4, 5 et 6*, Hermann, 1968
- [Greiner] W.Greiner, B.Müller, *Mécanique quantique. Symétries*, Springer, 1999
- [Hladik] J.Hladik, *La théorie des groupes en physique et chimie quantique*, Masson, 1995
- [Humphreys] James E.Humphreys, *Introduction to Lie Algebras and Representation Theory*, Springer, 1972
- [Kosmann] Y.Kosmann-Schwarzbach, *Groupes et symétries*, Editions de l'école polytechnique, 2003
- [Lisi] G.Lisi, *An Exceptionally Simple Theory of Everything*, article, 2007
- [Mneimné] R.Mneimné, F.Testard, *Introduction à la théorie des groupes de Lie classiques*, Hermann, 2005
- [Pichon] G.Pichon, *Groupes de Lie - Représentations linéaires et applications*, Hermann, 1973
- [Sternberg] S.Sternberg, *Lie algebras*, article, 2004