MP059 : Invariances in Physics and Group Theory

Jean-Bernard Zuber
Foreword

The following notes cover the content of the course “Invariances en Physique et Théorie des Groupes” given in the fall 2011. Additional material was presented during the week of “prérentrée” in lectures given by Adel Bilal, Paul Windey and myself. Their content is contained in the chapters 0 to 000 (in French), available on the web site.

Chapters 1 to 5 also contain, in sections in smaller characters and Appendices, additional details that were not treated in the oral course.

General bibliography


• [BDm] J.D. Bjorken and S. Drell: Relativistic Quantum Mechanics, McGraw Hill.


• [DFMS] P. Di Francesco, P. Mathieu et D. Sénéchal, Conformal Field Theory, Springer,


• [FH] W. Fulton and J. Harris, Representation Theory, Springer.


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


• [M] A. Messiah, Mécanique Quantique, 2 tomes, Dunod.


# Contents

1 Groups. Lie groups and Lie algebras

1.1 Generalities on groups ............................... 1
1.1.1 Definitions and first examples ..................... 1
1.1.2 Conjugacy classes of a group ..................... 3
1.1.3 Subgroups .................................. 3
1.1.4 Homomorphism of a group $G$ into a group $G'$ .... 4
1.1.5 Cosets with respect to a subgroup ................. 4
1.1.6 Invariant subgroups ............................ 5
1.1.7 Simple, semi-simple groups ....................... 5
1.2 Continuous groups. Topological properties. Lie groups. 6
1.2.1 Connectivity ................................ 6
1.2.2 Simple connectivity. Homotopy group. Universal covering . 6
1.2.3 Compact and non compact groups .................. 9
1.2.4 Invariant measure ............................. 9
1.2.5 Lie groups .................................. 10
1.3 Local study of a Lie group. Lie algebra .............. 11
1.3.1 Algebras and Lie algebras de Lie. Definitions .... 11
1.3.2 Tangent space in a Lie group .................... 12
1.3.3 Relations between the tangent space $\mathfrak{g}$ and the group $G$ 13
1.3.4 The tangent space as a Lie algebra ............... 13
1.3.5 An explicit example: the Lie algebra of $\text{SO}(n)$ .......... 15
1.3.6 An example of infinite dimension: the Virasoro algebra 16
1.4 Relations between properties of $\mathfrak{g}$ and $G$ ....... 16
1.4.1 Simplicity, semi-simplicity ..................... 16
1.4.2 Compact. Complexification ....................... 17
1.4.3 Connectivity, simple-connectivity ................. 18
1.4.4 Structure constants. Killing form. Cartan criteria .... 18
1.4.5 Casimir operator(s) .......................... 20

2 Linear representations of groups .......................... 33
2.1 Basic definitions and properties ..................... 33
2.1.1 Basic definitions ................................ 33
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1.2</td>
<td>Equivalent representations. Characters</td>
<td>34</td>
</tr>
<tr>
<td>2.1.3</td>
<td>Reducible and irreducible representations</td>
<td>35</td>
</tr>
<tr>
<td>2.1.4</td>
<td>Schur lemma</td>
<td>38</td>
</tr>
<tr>
<td>2.1.5</td>
<td>Tensor product of representations. Clebsch-Gordan decomposition</td>
<td>39</td>
</tr>
<tr>
<td>2.1.6</td>
<td>Decomposition into irreducible representations of a subgroup of a group representation</td>
<td>42</td>
</tr>
<tr>
<td>2.2</td>
<td>Representations of Lie algebras</td>
<td>42</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Definition. Universality</td>
<td>42</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Representations of a Lie group and of its Lie algebra</td>
<td>43</td>
</tr>
<tr>
<td>2.3</td>
<td>Representations of compact Lie groups</td>
<td>44</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Orthogonality and completeness</td>
<td>44</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Consequences</td>
<td>48</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Case of finite groups</td>
<td>48</td>
</tr>
<tr>
<td>2.3.4</td>
<td>Recap</td>
<td>50</td>
</tr>
<tr>
<td>2.4</td>
<td>Projective representations. Wigner theorem.</td>
<td>50</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Definition</td>
<td>50</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Wigner theorem</td>
<td>51</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Invariances of a quantum system</td>
<td>53</td>
</tr>
<tr>
<td>2.4.4</td>
<td>Transformations of observables. Wigner–Eckart theorem</td>
<td>54</td>
</tr>
<tr>
<td>2.4.5</td>
<td>Infinitesimal form of a projective representation. Central extension</td>
<td>56</td>
</tr>
<tr>
<td>2.4.6</td>
<td>Orthogonality, completeness, characters</td>
<td>58</td>
</tr>
<tr>
<td>2.4.7</td>
<td>Special functions. Spherical harmonics</td>
<td>59</td>
</tr>
<tr>
<td>2.4.8</td>
<td>Physical applications</td>
<td>63</td>
</tr>
</tbody>
</table>

3 Simple Lie algebras. Classification and representations. Roots and weights 71

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Cartan subalgebra. Roots. Canonical form of the algebra</td>
<td>71</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Cartan subalgebra</td>
<td>71</td>
</tr>
<tr>
<td>3.1.2</td>
<td>Canonical basis of the Lie algebra</td>
<td>72</td>
</tr>
<tr>
<td>3.2</td>
<td>Geometry of root systems</td>
<td>75</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Scalar products of roots. The Cartan matrix</td>
<td>75</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Root systems of simple algebras. Cartan classification</td>
<td>79</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Chevalley basis</td>
<td>79</td>
</tr>
<tr>
<td>3.2.4</td>
<td>Coroots. Highest root. Coxeter number and exponents</td>
<td>80</td>
</tr>
<tr>
<td>3.3</td>
<td>Representations of semi-simple algebras</td>
<td>80</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Weights. Weight lattice</td>
<td>80</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Roots and weights of su(n)</td>
<td>84</td>
</tr>
<tr>
<td>3.4</td>
<td>Tensor products of representations of su(n)</td>
<td>88</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Littlewood-Richardson rules</td>
<td>88</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Explicit tensor construction of representations of SU(2) and SU(3)</td>
<td>89</td>
</tr>
<tr>
<td>3.5</td>
<td>Young tableaux and representations of GL(n) and SU(n)</td>
<td>91</td>
</tr>
</tbody>
</table>
4 Global symmetries in particle physics

4.1 Global exact or broken symmetries. Spontaneous breaking

4.1.1 Overview. Exact or broken symmetries

4.1.2 Chiral symmetry breaking

4.1.3 Quantum symmetry breaking. Anomalies

4.2 The SU(3) flavor symmetry and the quark model

4.2.1 Why SU(3)?

4.2.2 Consequences of the SU(3) symmetry

4.2.3 Electromagnetic breaking of the SU(3) symmetry

4.2.4 “Strong” mass splittings. Gell-Mann–Okubo mass formula

4.2.5 Quarks

4.2.6 Hadronic currents and weak interactions

4.3 From SU(3) to SU(4) to six flavors

4.3.1 New flavors

4.3.2 Introduction of color

5 Gauge theories. Standard model

5.1 Gauge invariance. Minimal coupling. Yang–Mills Lagrangian

5.1.1 Gauge invariance

5.1.2 Non abelian Yang–Mills extension

5.1.3 Geometry of gauge fields

5.1.4 Yang–Mills Lagrangian

5.1.5 Quantization. Renormalizability

5.2 Massive gauge fields

5.2.1 Weak interactions and intermediate bosons

5.2.2 Spontaneous breaking of gauge symmetry. Brout–Englert–Higgs mechanism

5.3 The standard model

5.3.1 The strong sector

5.3.2 The electro-weak sector, a sketch

5.4 Complements

5.4.1 Standard Model and beyond

5.4.2 Grand-unified theories or GUTs

5.4.3 Anomalies
Chapter 1

Groups. Lie groups and Lie algebras

1.1 Generalities on groups

1.1.1 Definitions and first examples

Let us consider a group \( G \), with an operation denoted \( \cdot \), \( \times \) or \( + \) depending on the case, a neutral element (or “identity”) \( e \) (or 1 or \( I \) or 0), and an inverse \( g^{-1} \) (or \( -a \)). If the operation is commutative, the group is called abelian. If the group is finite, i.e. has a finite number of elements, we call that number the order of the group and denote it by \( |G| \). In these lectures we will be mainly interested in infinite groups, discrete or continuous.

Examples (that the physicist may encounter . . . )

- Finite groups
  - the cyclic group \( \mathbb{Z}_p \) of order \( p \), considered geometrically as the invariance rotation group of a circle with \( p \) marked equidistant points, or as the multiplicative group of \( p \)-th roots of the unity, \( \{ e^{2i\pi q/p} \}, q = 0, 1, \ldots, p - 1 \), or as the additive group of integers modulo \( p \);
  - the groups of rotation invariance and the groups of rotation and reflection invariance of regular solids or of regular lattices, of great importance in solid state physics and crystallography;
  - the permutation group \( S_n \) of \( n \) objects, called also the symmetric group, of order \( n! \);
  - the homotopy groups, to be encountered soon, are other examples, and there are many others.

- Discrete infinite groups.
  The simplest example is the additive group \( \mathbb{Z} \). Let us also mention the translation groups of regular lattices.

Also the groups generated by reflections in a finite number of hyperplanes of the Euclidean space \( \mathbb{R}^n \), that are finite or infinite depending on the arrangement of these hyperplanes, see Weyl groups in Chap. 4.
Another important example is the modular group $PSL(2, \mathbb{Z})$ of matrices $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with integer coefficients, of determinant 1, $ad - bc = 1$, with matrices $A$ and $-A$ identified. Given a 2-dimensional lattice in the complex plane generated by two complex numbers $\omega_1$ and $\omega_2$ of non real ratio (why?), this group describes the changes of basis (i.e. Continuous groups. We shall be dealing only with matrix groups of finite dimension, in physics, in string theory and conformal field theory . . . group plays an important role in mathematics in the study of elliptic functions, modular forms, etc, and in physics, in string theory and conformal field theory . . .

- Continuous groups. We shall be dealing only with matrix groups of finite dimension, i.e. subgroups of the linear groups $GL(n, \mathbb{R})$ ou $GL(n, \mathbb{C})$, for some $n$. In particular

- $U(n)$, the group of complex unitary matrices, $UU^\dagger = I$, which is the invariance group of the sesquilinear form $(x, y) = \sum x^i y^i$ ;
- $SU(n)$ its unimodular subgroup, of unitary matrices of determinant $\det U = 1$;
- $O(n)$ and $SO(n)$ are orthogonal groups of invariance of the symmetric bilinear form $\sum_{i=1}^n x_i y_i$. Matrices of $SO(n)$ have determinant 1 ;
- $U(p, q)$, $SU(p, q)$, resp. $O(p, q)$, $SO(p, q)$, invariance groups of a sesquilinear, resp. bilinear form, of signature $((+)^p, (-)^q)$ (e.g. the Lorentz group $O(3,1)$).

Most often one considers groups $O(n, \mathbb{R})$, $SO(n, \mathbb{R})$ of matrices with real coefficients but groups $O(n, \mathbb{C})$, $SO(n, \mathbb{C})$ of invariance of the same bilinear form over the complex numbers may also play a role.

- $\text{Sp}(2n, \mathbb{R})$: Let $Z$ be the matrix $2n \times 2n$ made of a diagonal of $n$ blocks $i\sigma_2$:

\[
Z = \text{diag} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},
\]

and consider the bilinear skew-symmetric form

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

The symplectic group $\text{Sp}(2n, \mathbb{R})$ is the group of real $2n \times 2n$ matrices $B$ that preserve that form: $B^T Z B = Z$. That form appears naturally in Hamiltonian mechanics with the symplectic 2-form $\omega = \sum_{i=1}^n dp_i \wedge dq_i = \frac{1}{2} \sum_{i,j} Z_{ij} \xi_i \wedge d\xi_j$ in the coordinates $\xi = (p_1, q_1, p_2, \cdots, q_n)$ ; $\omega$ is invariant by action of $\text{Sp}(2n, \mathbb{R})$ on $\xi$. For $n = 1$, verify that $\text{Sp}(2, \mathbb{R}) = \text{SL}(2, \mathbb{R})$.

One may also consider the complex symplectic group $\text{Sp}(2n, \mathbb{C})$. A related group, often denoted $\text{Sp}(n)$ but that I shall denote $\text{USp}(n)$ to avoid confusion with the previous ones, the unitary symplectic group, is the invariance group of a Hermitian quaternionic form, $\text{USp}(n) = U(2n) \cap \text{Sp}(2n, \mathbb{C})$ See Appendix A.

- the group of displacements in $\mathbb{R}^3$, and groups obtained by adjoining dilatations, and then inversions with respect to a point;
- the group of conformal transformations, i.e. angle preserving, in $\mathbb{R}^n$ (see Problem at the end of this chapter).
- the Galilean group of transformations $x' = Rx + vt + x_0$, $t' = t + t_0$,
- $O(3,1)$, the Lorentz group, and the Poincaré group,
- etc etc.
1.1. GENERALITIES ON GROUPS

1.1.2 Conjugacy classes of a group

In a group $G$ we define the following equivalence relation:

$$a \sim b \iff \exists g \in G : a = g.b.g^{-1}$$

and the elements $a$ et $b$ are said to be *conjugate*.

The equivalence classes (*conjugacy classes*) that follow provide a partition of $G$, since any element belongs to a unique class. For finite group, the different classes generally have different orders (or cardinalities). For instance, the class of the neutral element $e$ has a unique element, $e$ itself.

We have already noted that in the rotation group $SO(3)$, a conjugacy class is characterized by the rotation angle $\psi$ (around some unitary vector $\mathbf{n}$). But this notion is also familiar in the group $U(n)$, where a class is characterized by an unordered $n$-tuple of eigenvalues $(e^{i\alpha_1}, \ldots, e^{i\alpha_n})$. This notion of class plays an important role in the discussion of representations of groups and will be abundantly illustrated in the following.

What are the conjugacy classes in the symmetric group $S_n$? One proves easily that any permutation $\sigma$ of $S_n$ decomposes into a product of cycles (cyclic permutations) on distinct elements. (To show that, construct the cycle $(1, \sigma(1), \sigma^2(1), \cdots)$; then, once back to 1, construct another cycle starting from a number not yet reached, etc.). Finally if $\sigma$ is made of $p_1$ cycles of length 1, $p_2$ of length 2, etc, with $\sum ip_i = n$, one writes $\sigma \in [1^{p_1}2^{p_2} \cdots]$, and one may prove that this decomposition into cycles characterizes the conjugacy classes: two permutations are conjugate iff they have the same decomposition into cycles.

1.1.3 Subgroups

The notion of subgroup $H$, subset of a group $G$ itself endowed with a group structure, is familiar. The subgroup is *proper* if it is not identical to $G$. If $H$ is a subgroup, for any $a \in G$, the set $a^{-1}.H.a$ of elements of the form $a^{-1}.h.a$, $h \in H$ forms also a subgroup, called *conjugate subgroup* to $H$.

Examples of particular subgroups are provided by:

- the center $Z$:
  
  In a group $G$, the *center* is the subset $Z$ of elements that commute with all other elements of $G$:
  
  $$Z = \{a \mid \forall g \in G, a.g = g.a\}$$

  $Z$ is a subgroup $G$, and is proper if $G$ is nonabelian. Examples: the center of the group $GL(2, \mathbb{R})$ of regular $2 \times 2$ matrices is the set of matrices multiple of $I$; the center of $SU(2)$ is the group $\mathbb{Z}_2$ of matrices $\pm I$ (check by direct calculation).

- the centralizer of an element $a$:
  
  The *centralizer* (or *commutant*) of a given element $a$ of $G$ is the set of elements of $G$ that commute with $a$.
  
  $$Z_a = \{g \in G \mid a.g = g.a\}$$
The commutant $Z_a$ is never empty: it contains at least the subgroup generated by $a$. The center $Z$ is the intersection of all commutants. Example: in the group $GL(2, \mathbb{R})$, the commutant of the Pauli matrix $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ is the abelian group of matrices of the form $aI + b\sigma_1$, $a^2 - b^2 \neq 0$.

- More generally, given a subset $S$ of a group $G$, its centralizer $Z(S)$ and its normalizer $N(S)$ are the subgroups commuting respectively individually with any element of $S$ or globally with $S$ as a whole:
  \[ Z(S) = \{ y : \forall s \in S \quad y.s = s.y \} \]
  \[ N(S) = \{ x : x^{-1}.S.x = S \}. \]

### 1.1.4 Homomorphism of a group $G$ into a group $G'$

An *homomorphism* of a group $G$ into a group $G'$ is a map $\rho$ of $G$ into $G'$ which respects the composition law:

\[ \forall g, h \in G, \quad \rho(g.h) = \rho(g).\rho(h) \]  

(1.7)

In particular, $\rho$ maps the neutral element of $G$ onto that of $G'$, and the inverse of $g$ onto that of $g' = \rho(g)$: $\rho(g^{-1}) = (\rho(g))^{-1}$.

An example of homomorphism that we shall study in great detail is that of a linear representation of a group, whose definition has been given in Chap. 00 and that we return to in Chap. 2.

The *kernel* of the homomorphism, denoted $\ker \rho$, is the set of preimages of the neutral element in $G'$: $\ker \rho = \{ x \in G : \rho(x) = e' \}$. It is a subgroup of $G$.

For example, the parity (or signature) of a permutation of $S_n$ defines an homomorphism from $S_n$ into $\mathbb{Z}_2$. Its kernel is made of even permutations: this is alternating group $A_n$ of order $n!/2$.

### 1.1.5 Cosets with respect to a subgroup

Consider a subgroup $H$ of a group $G$. We define the following relation between elements of $G$:

\[ g \sim g' \iff g.g'^{-1} \in H, \]  

(1.8)

which may also be rewritten as

\[ g \sim g' \iff \exists h \in H : g = h.g' \quad \text{ou encore} \quad g \in H.g'. \]

(1.9)

This is an equivalence relation (check !), called the *right equivalence* One defines in a similar way the *left equivalence* by

\[ g \sim_R g' \iff g'^{-1}.g' \in H \iff g \in g'.H. \]

(1.10)

This relation (say, right) defines equivalence classes that give a partition of $G$; if $g_j$ is a representative of class $j$, that class, called *right-coset*, may be denoted $H.g_j$. The elements of $H$ form by themselves a coset. The set of (say right) cosets is denoted $G/H$ and called the (right) coset “space”. If $H$ is of finite order $|H|$, all cosets have $|H|$ elements, and if $G$ is itself of finite
order $|G|$, it is partitioned into $|G|/|H|$ classes, and one obtains the Lagrange theorem as a corollary: the order $|H|$ of any subgroup $H$ divides that of $G$, and the ratio $|G|/|H|$ is the order of the coset space $G/H$.

The left equivalence gives rise in general to a different partition. For example, the group $S_3$ has a $\mathbb{Z}_2$ subgroup generated by the permutation of the two elements 1 et 2. Exercise: check that the left and right cosets do not coincide.

### 1.1.6 Invariant subgroups

Consider a group $G$ with a subgroup $H$. $H$ is an invariant subgroup (one also says normal) if one of the following equivalent properties holds true:

- $\forall g \in G, \forall h \in H, ghg^{-1} \in H$.
- Left and right cosets coincide;
- $H$ is equal to all its conjugates $\forall g \in G, gHg^{-1} = H$.

Exercise: check the equivalence of these three assertions.

The important property to remember is the following:

**If $H$ is an invariant subgroup of $G$, the coset space $G/H$ may be given a group structure, and is called the quotient group.**

Note that in general one cannot consider the quotient group $G/H$ as a subgroup of $G$.

Let us sketch the proof. If $g_1 \sim g'_1$ and $g_2 \sim g'_2$, $\exists h_1, h_2 \in H : g_1 = h_1 g'_1$, $g_2 = g'_2 h_2$, hence $g_1 g_2 = h_1 (g'_1 g'_2) h_2$ i.e. $g_1 g_2 \sim g'_1 g'_2$ et $g_1^{-1} = g'_1^{-1} h_1^{-1} \sim g'_2^{-1}$. The equivalence relation is thus compatible with the composition and inverse operations, and if $[g_1]$ and $[g_2]$ denote two cosets, one defines $[g_1], [g_2] = [g_1 g_2]$ where on the right hand side (rhs), one takes any representative $g_1$ of the coset $[g_1]$ and $g_2$ of $[g_2]$ ; and likewise for the inverse. Thus the group structure passes to the coset space. The coset made by $H$ is the neutral element in the quotient group.

Example of an invariant subgroup: The kernel of an homomorphism $\rho$ of $G$ into $G'$ is an invariant subgroup: show that the quotient group is isomorphic to the image $\rho(G) \subset G'$ of $G$ by $\rho$.

### 1.1.7 Simple, semi-simple groups

A group is simple if it has no non-trivial invariant subgroup (non trivial, i.e. different from $\{e\}$ and from $G$ itself). A group is semi-simple if it has no non-trivial abelian invariant subgroup. Any simple group is obviously semi-simple.

This notion is important in representation theory and in the classification of groups.

Examples: The rotation group in two dimensions is not simple, and not even semi-simple (why?). The group $SO(3)$ is simple (non trivial proof, see below, section 1.2.2). The group $SU(2)$ is neither simple nor semi-simple, as it contains the invariant subgroup $\{I, -I\}$. The group $S_n$ is not simple, for $n > 2$ (why?).
1.2 Continuous groups. Topological properties. Lie groups.

A continuous group (one also says a topological group) is a topological space (hence endowed with a basis of neighbourhoods that allows us to define notions of continuity etc\(^1\)) with a group structure, such that the composition and inverse operations \((g, h) \mapsto g \cdot h\) et \(g \mapsto g^{-1}\) are continuous functions. In other words, if \(g'\) is nearby \(g\) (in the sense of the topology of \(G\)), and \(h'\) nearby \(h\), then \(g' \cdot h'\) is nearby \(g \cdot h\) and \(g'^{-1}\) is nearby \(g^{-1}\).

The matrix groups presented at the beginning of this chapter all belong to this class of topological groups, but there are also groups of “infinite dimension” like the group of diffeomorphisms invoked in General Relativity.

Let us first study some topological properties of these continuous groups.

1.2.1 Connectivity

A group may be connected or not. If \(G\) is not connected, the connected component of the identity (i.e. of the neutral element) is an invariant subgroup.

One may be interested in the connectivity in the general topological sense (a space is connected if it is both open and closed), but we shall be mainly concerned by the arc connectivity: for any pair of points, there exists a continuous path in the space (here the group) that joins them. Show that the connected component of the identity is an invariant subgroup for both definitions. Ref. [K-S, Po].

Examples. \(O(3)\) is disconnected and the connected component of the identity is \(SO(3)\); for the Lorentz group \(L=O(3,1)\), the connected component of \(I\) is its proper orthochronous subgroup \(L_+\), (see Chap. 0), the other “sheets” then result from the application on it of parity \(P\), of time reversal \(T\) and their product \(PT\)…

1.2.2 Simple connectivity. Homotopy group. Universal covering

This notion should not be mistaken for the previous one.

As it does not apply only to groups, consider first an arbitrary topological space \(E\). Let us consider closed paths drawn in the space (or “loops”), with a fixed end-point \(x_0\), i.e. continuous maps \(x(t)\) from \([0, 1]\) into \(E\) such that \(x(0) = x(1) = x_0\). Given two such closed paths \(x_1(\cdot)\) et \(x_2(\cdot)\) from \(x_0\) to \(x_0\), can one deform them continuously into one another? In other words, is there a continuous function \(f(t, \xi)\) of two variables \(t, \xi \in [0, 1]\), taking its values in the space \(E\), such that

\[
\forall \xi \in [0, 1] \quad f(0, \xi) = f(1, \xi) = x_0 : \text{closed paths} \tag{1.11}
\]

\[
\forall t \in [0, 1] \quad f(t, 0) = x_1(t) \quad f(t, 1) = x_2(t) : \text{interpolation}.
\tag{1.12}
\]

If this is the case, one says that the paths \(x_1\) and \(x_2\) are homotopic (this is an equivalence relation between paths), or equivalently that they belong to the same homotopy class, see fig. 1.1.

\(^1\)See Appendix B for a reminder of some points of vocabulary…
1.2. CONTINUOUS GROUPS. TOPOLOGICAL PROPERTIES. LIE GROUPS.

Figure 1.1: The paths $x_1$ et $x_2$ are homotopic. But none of them is homotopic to the “trivial” path that stays at $x_0$. The space is not simply connected.

One may also compose paths: If $x_1(\cdot)$ and $x_2(\cdot)$ are two paths from $x_0$ to $x_0$, their product $x_2 \circ x_1$ also goes from $x_0$ to $x_0$ by following first $x_1$ and then $x_2$. The inverse path of $x_1(\cdot)$ for that composition is the same path but followed in the reverse direction: $x_1^{-1}(t) := x_1(1 - t)$. Both the composition and the inverse are compatible with homotopy: if $x_1 \sim x_1'$ and $x_2 \sim x_2'$, then $x_2 \circ x_1 \sim x_2' \circ x_1'$ and $x_1^{-1} \sim x_1'^{-1}$. These operations thus pass to classes, giving the set of homotopy classes a group structure: this is homotopy group $\pi_1(E, x_0)$. Hence, a representative of the identity class is given by the “trivial” path, $x(t) = x_0$, $\forall t$. One finally shows that homotopy groups relative to different end-points $x_0$ are isomorphic (in a connected space); if $E$ is a connected group, see below, one may take for example the base point to be the identity $x_0 = e$. One may thus talk of the homotopy group (or fundamental group) $\pi_1(E)$. For more details, see for example [Po], [DNF].

If all paths from $x_0$ to $x_0$ may be continuously contracted into the trivial path $\{x_0\}$, $\pi_1(E)$ is trivial, and $E$ is said to be simply connected. In the opposite case, one may prove and we shall admit that one may construct a space $\tilde{E}$, called the universal covering space of $E$, such that $\tilde{E}$ is simply connected and that locally, $E$ and $\tilde{E}$ are homeomorphic. This means that there exists a continuous and surjective mapping $p$ from $\tilde{E}$ to $E$ such that any point $x$ in $\tilde{E}$ has a neighborhood $V_x$ and that $V_x \to p(V_x)$ is a homeomorphism, i.e. a bicontinuous bijection. The universal covering space $\tilde{E}$ of $E$ is unique (up to a homeomorphism).

Let us now restrict ourselves to the case where $E = G$, a topological group. Then one shows that its covering $\tilde{G}$ is also a group, the universal covering group, and moreover, that the map $p$ is a homomorphism of $\tilde{G}$ into $G$. Its kernel which is an invariant subgroup of $\tilde{G}$, is proved to be isomorphic to the homotopy group $\pi_1(G)$ ([Po], sect. 51). The quotient group is isomorphic to $G$

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

(1.13)

(according to a general property of the quotient group by the kernel of an homomorphism, cf. sect.1.6).

One may construct the universal covering group $\tilde{G}$ by considering paths that join the identity $e$ to a point $g$, and their equivalence classes under continuous deformation with fixed ends. $\tilde{G}$ is the set of these equivalence classes. It is a group for the multiplication of paths defined as follows: if two paths $g_1(t)$ and $g_2(t)$ join $e$ to $g_1$ and to $g_2$ respectively, the path $g_1(t).g_2(t)$ joins $e$ to $g_1.g_2$. This composition law is compatible with equivalence

---

2“bicontinuous” means that the map and its inverse are both continuous.
and gives $\tilde{G}$ a group structure and one shows that $\tilde{G}$ is simply connected (cf. [Po] sect. 51). The projection $p$ of $\tilde{G}$ into $G$ associates with any class of paths their common end-point. One may verify that this is indeed a local homeomorphism and a group homomorphism, and that its kernel is the homotopy group $\pi_1(G)$.

Example: The group $G = U(1)$ of complex numbers of modulus 1, seen as the unit circle $S^1$, is non simply connected: a path from the identity 1 to 1 may wind an arbitrary number of times around the circle and this (positive or negative) winding number characterizes the different homotopy classes: the homotopy group is $\pi_1(U(1)) = \mathbb{Z}$. The group $\tilde{G}$ is nothing else than the additive group $\mathbb{R}$ and may be visualised as a helix above $U(1)$. The quotient is $\mathbb{R}/\mathbb{Z} \simeq U(1)$, which must be interpreted as the fact that a point of $U(1)$, i.e. an angle, is a real number modulo an integer multiple of $2\pi$. One may also say that $\pi_1(S^1) = \mathbb{Z}$. More generally one may convince oneself that for spheres, $\pi_1(S^n)$ is trivial (all loops are contractible) as soon as $n > 1$.

Another fundamental example: The rotation group $SO(3)$ is not simply connected, as foreseen in Chap. 00. To see this fact, represent the rotation $R_n(\psi)$ by the point $x = \tan \frac{\psi}{4} n$ of an auxiliary space $\mathbb{R}^3$; all these points are in the ball $B^3$ of radius 1, with the identity rotation at the center and rotations of angle $\pi$ on the sphere $S^2 = \partial B^3$, but because of $R_n(\pi) = R_{-n}(\pi)$, (see Chap 00, sect. 1.1), diametrically opposed points must be identified. It follows that there exists in $SO(3)$ closed loops that are non contractible: a path from $x$ to $x$ passing through two diametrically opposed points on the sphere $S^2$ must be considered as closed but is not contractible (Fig. 1.2.b). There exist two classes of non homotopic closed loops from $x$ to $x$ and the group $SO(3)$ is doubly connected. Its homotopy group is $\pi_1(SO(3)) = \mathbb{Z}_2$. In fact, we already know the universal covering group of $SO(3)$: it is the group $SU(2)$, which has been shown to be homeomorphic to the sphere $S^3$, hence is simply connected, and for which there exists a homomorphism mapping it to $SO(3)$, according to $\pm U_n(\psi) = \pm (\cos \frac{\psi}{2} - i \sin \frac{\psi}{2} \sigma \cdot n) \mapsto R_n(\psi)$.
This property of SO(3) to be non simply connected may be illustrated by various home experiments, the precise interpretation of which may not be obvious, such as “Dirac’s belt” and “Feynman’s plate”, see http://gregegan.customer.netspace.net.au/APPLETS/21/21.html and http://www.math.utah.edu/ palais/links.html for nice animations, and V. Stojanoska et O. Stoytchev, Mathematical Magazine, 81, 2008, 345-357, for a detailed discussion involving the braid group.

The same visualisation of rotations by the interior of the unit ball also permits to understand the above assertion that the group SO(3) is simple. Suppose it is not, and let \( R = R_n(\psi) \) be an element of an invariant subgroup of SO(3), which also contains all the conjugates of \( R \) (by definition of an invariant subgroup). These conjugates are represented by points of the sphere of radius \( \tan \psi/4 \). The invariant subgroup containing \( R_n(\psi) \) and points that are arbitrarily close to its inverse \( R_{-n}(\psi) \) contains also points that are arbitrarily close to the identity, which by conjugation, fill a small ball in the vicinity of the identity. It remains to show that the products of such elements allow to fill all the bowl, i.e. that the invariant subgroup may only be SO(3) itself; this is in fact true for any connected Lie group, as we shall see below.

Other examples: classical groups. One may prove that

- the groups SU\( (n) \) are all simply connected, for any \( n \), whereas \( \pi_1(U(n)) = \mathbb{Z} \);
- for the group SO\( (2) \cong U(1) \), we have seen that \( \pi_1(SO(2)) = \mathbb{Z} \);
- for any \( n > 2 \), SO\( (n) \) is doubly connected, \( \pi_1(SO(n)) = \mathbb{Z}_2 \), and its covering group is called Spin\( (n) \). Hence Spin\( (3) = SU(2) \).

The notion of homotopy, i.e. of continuous deformation, that we have just applied to paths, i.e. to maps of \( S^1 \) into a manifold \( V \) (a group \( G \) here), may be extended to maps of a sphere \( S^n \) into \( V \). Even though the composition of such maps is less easy to visualise, it may be defined and is again compatible with homotopy, leading to the definition of the homotopy group \( \pi_n(V) \). For example \( \pi_n(S^n) = \mathbb{Z} \). This notion is important in physics to describe topological defects, solitons, instantons, monopoles, etc. See [DNF] for more details and the determination of these groups \( \pi_n \).

### 1.2.3 Compact and non compact groups

If the domain \( D \) in which the parameters of the group \( G \) take their values is compact, \( G \) is said to be a compact group.

Recall some of the many equivalent characterizations of a compact set \( E \). From any infinite sequence one may extract a converging subsequence. Given a covering of \( E \) by a set of open sets \( U_i \), \( E \) may be covered by a finite number of them. Any continuous function on \( E \) is bounded, etc For a subset \( D \) of \( \mathbb{R}^d \), being compact is equivalent to being closed and bounded.

Examples. The unitary groups \( U(n) \) and their subgroups SU\( (n) \), O\( (n) \), USp\( (n/2) \) (\( n \) even), are compact. The groups SL\( (n, \mathbb{R}) \) or SL\( (n, \mathbb{C}) \), Sp\( (n, \mathbb{R}) \) or Sp\( (n, \mathbb{C}) \), the translation group in \( \mathbb{R}^n \), the Galilean group, the Lorentz and Poincaré groups are not, why?

### 1.2.4 Invariant measure

When dealing with a finite group, one often considers sums over all elements of the group and makes use of the “rearrangement lemma”, in which one writes

\[
\sum_{g \in G} f(g'g) = \sum_{h=g'g \in G} f(g') = \sum_{g \in G} f(g),
\]
(left invariance), the same thing with \( g'g \) changed into \( gg' \) (right invariance), and also

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

Can one do similar operations in continuous groups, the finite sum being replaced by an integral, which converges and enjoys the same invariances? This requires the existence of an integration measure, with left and right invariance,

\[
d\mu(g) = d\mu(g',g) = d\mu(g.g') = d\mu(g^{-1})
\]

such that \( \int d\mu(g)f(g) \) be finite for any continuous function \( f \) on the group.

One may prove (and we admit) that

• if the group is compact, such a measure exists and is unique up to a normalization.

This is the Haar measure.

For example, in the unitary group \( U(n) \), one may construct explicitly the Haar measure, using the method proposed in chap. 0, §2.3: one first defines a metric on \( U(n) \) by writing \( ds^2 = \text{tr } dU.dU^\dagger \) in any parametrization; this metric is invariant under \( U \rightarrow UU' \) or \( U \rightarrow U'U \) and by \( U \rightarrow U^{-1} = U^\dagger \); the measure \( d\mu(U) \) that follows has the same properties. See Appendix B for the explicit calculation for \( SU(2) \) and \( U(n) \), and more details in the TD.

Conversely if the group is non compact, left and right measures may still exist, they may even coincide, (non compact abelian or semi-simple groups) but their integral over the group diverges.

Thus, if \( G \) is locally compact, (i.e. any point has a basis of compact neighbourhoods), one proves that there exists a left invariant measure, unique up to a multiplicative constant. There exists also a right invariant measure, but they may not coincide. For example, take

\[
G = \left\{ \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \big| x, y \in \mathbb{R}, y > 0 \right\}
\]

one easily checks that \( d\mu_L(g) = y^{-2}dxdy \), \( d\mu_R(g) = y^{-1}dxdy \) are left and right invariant measures, respectively, and that their integrals diverge. See [Bu].

1.2.5 Lie groups

Imposing more structure on a continuous group leads us in a natural way to the notion of Lie groups.

According to the usual definition, a Lie group is a topological group which is also a differentiable manifold and such that the composition and inverse operations \( G \times G \rightarrow G \) and \( G \rightarrow G \) are infinitely differentiable functions. One sometimes also requests them to be analytic real functions, i.e. functions for which the Taylor series converges to the function. That the two definitions coexist in the literature is a hint that the weakest (infinite differentiability) implies the strongest. In fact, according to a remarkable theorem (Montgomery et Zippen, 1955), much weaker hypotheses suffice to ensure the property of Lie group.

A topological connected group which is locally homeomorphic to \( \mathbb{R}^d \), for some finite \( d \), is a Lie group. In other words, the existence of a finite number of local coordinates, together with the properties of being a topological group (continuity of the group operations), are sufficient to imply the analyticity properties! ³ This shows that the structure of Lie group is quite powerful and rigid.

³For an elementary example of such a phenomenon, consider a function \( f \) of one real variable, satisfying \( f(x)f(y) = f(x+y) \). Under the only assumption of continuity, show that \( f(x) = \exp kx \), hence that it is analytic!
To avoid a mathematical discussion unnecessary for our purpose, we shall restrict ourselves to continuous groups of finite size matrices. In such a group, the matrix elements of \( g \in G \) depend continuously on real parameters \((\xi^1, \xi^2, \cdots, \xi^d) \in D \subset \mathbb{R}^d\), and in the group operations \( g(\xi'') = g(\xi)'g(\xi), \) and \( g(\xi)^{-1} = g(\xi'')\), the \( \xi''_i \) are continuous (in fact analytic) functions of the \( \xi_j \) (and \( \xi'_j \)). Such a group is called a Lie group, and \( d \) is its dimension.

More precisely, in the spirit of differential geometry, one has in general to introduce several domains \( D_j \), with continuous (in fact analytic) transition functions between coordinate charts, etc.

Examples: all the matrix groups presented in §1.1 are Lie groups. Check that the dimension of \( U(n) \) is \( n^2 \), that of \( SU(n) \) is \( n^2 - 1 \), that \( O(n) \) or \( SO(n) \) is \( n(n - 1)/2 \). What is the dimension of \( \text{Sp}(2n, \mathbb{R}) \) ? of the Galilean group in \( \mathbb{R}^3 \)? of the Lorentz and Poincaré groups?

Show that \( \dim(\text{Sp}(2n, \mathbb{R})) = \dim(\text{USp}(n)) = \dim(\text{SO}(2n + 1)) \), and we shall see below in chap. 3 that this is not an accident.

The study of a Lie group and of its representations involves two steps: first a local study of its tangent space in the vicinity of the identity (its Lie algebra), and then a global study of its topology, i.e. an information not provided by the local study.

### 1.3 Local study of a Lie group. Lie algebra

#### 1.3.1 Algebras and Lie algebras de Lie. Definitions

Let us first recall the definition of an algebra. An algebra is a vector space over a field (for physicists, \( \mathbb{R} \) or \( \mathbb{C} \)), endowed with a product denoted \( X * Y \), (not necessarily associative), bilinear in \( X \) and \( Y \):

\[
(\lambda_1 X_1 + \lambda_2 X_2) * Y = \lambda_1 X_1 * Y + \lambda_2 X_2 * Y \tag{1.14}
\]
\[
X * (\mu_1 Y_1 + \mu_2 Y_2) = \mu_1 X * Y_1 + \mu_2 X * Y_2 . \tag{1.15}
\]

Examples: the set \( M(n, \mathbb{R}) \) or \( M(n, \mathbb{C}) \) of \( n \times n \) matrices with real, resp. complex coefficients, is an associative algebra for the usual matrix product. The set of vectors of \( \mathbb{R}^3 \) is a (non associative !) algebra for the vector product (denoted \( \wedge \) in the French literature, and \( \times \) in the anglo-saxon one).

A Lie algebra is an algebra in which the product denoted \([X, Y]\) has the additional properties of being antisymmetric and of satisfying the Jacobi identity

\[
[X, Y] = -[Y, X] \tag{1.16}
\]
\[
[X_1, [X_2, X_3]] + [X_2, [X_3, X_1]] + [X_3, [X_1, X_2]] = 0 . \tag{1.17}
\]

Examples: Any associative algebra for a product denoted *, in particular any matrix algebra, is a Lie algebra for a product, the Lie bracket, defined by the commutator

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

The bilinearity and antisymmetry properties are obvious, and verifying the Jacobi identity takes one line. Another example: the space \( \mathbb{R}^3 \) with the above-mentionned vector product is in
fact a Lie algebra, with the Jacobi identity following from the “double vector product” formula,
\[ u \times (v \times w) = (u.w)v - (u.v)w. \]

### 1.3.2 Tangent space in a Lie group

Consider a Lie group \( G \) and a one-parameter subgroup \( g(t) \), where \( t \) is a real parameter taking values in a neighborhood of 0, with \( g(0) = e \); in other words, \( g(t) \) is a curve in \( G \), assumed to be differentiable, and passing through the identity, and one assumes that (for \( t \) near 0)

\[
g(t_1)g(t_2) = g(t_1 + t_2) \quad g^{-1}(t) = g(-t). \tag{1.18}
\]

The composition law in this subgroup locally amounts to the addition of parameters \( t \); thus, locally, this one-parameter subgroup is isomorphic to the abelian group \( \mathbb{R} \). It is then natural to differentiate

\[
g(t + \delta t) = g(t)g(\delta t) \iff g^{-1}(t)g(t + \delta t) = g(\delta t). \tag{1.19}
\]

As we have chosen to restrict ourselves to matrix groups, (with \( e \equiv I \), the identity matrix), we may write the linear tangent map in the form

\[
g(\delta t) = I + \delta tX + \cdots
\]

which defines a vector \( X \) in the tangent space. One may also write

\[
X = \left. \frac{d}{dt} g(t) \right|_{t=0}, \tag{1.20}
\]

this is the velocity at \( t = 0 \) (or at \( g = e \)) along the curve. Equation (1.19) thus reads

\[
g'(t) = g(t)X. \tag{1.21}
\]

As usual in differentiable geometry, (see Appendix B.3), the tangent space \( T_eG \) at \( e \) to the group \( G \), which we denote from now on \( \mathfrak{g} \), is the vector space generated by the tangent vectors to all one-parameter subgroups (i.e. all velocity vectors at \( t = 0 \)). If coordinates \( \xi^\alpha \) of \( G \) have been chosen in the vicinity of \( e \) (\( \equiv I \)), a tangent vector is a differential operator \( X = X^\alpha \frac{\partial}{\partial \xi^\alpha} \). The dimension (as a vector space) of this tangent space is equal to the dimension of the (group) manifold defined above as the number of (real) parameters \( \dim \mathfrak{g} = \dim G \).

In the case of a group \( G \subset \text{GL}(n, \mathbb{R}) \) to which we are restricting ourselves, \( X \in \mathfrak{g} \subset M(n, \mathbb{R}) \), the set of real \( n \times n \) matrices, and one may carry out all calculations in that algebra. In particular, one may integrate (1.21) as

\[
g(t) = \exp tX = \sum_{n=0}^{\infty} \frac{t^n}{n!}X^n, \tag{1.22}
\]
a converging sum. (In fact, the assumption that the group is a matrix group may be relaxed, provided one makes sense of the map \( \exp X \), a map that enjoys some of the usual properties of the exponential, see Appendix B.4.)
1.3.3 Relations between the tangent space $\mathfrak{g}$ and the group $G$

1. If $G$ is the linear group $\text{GL}(n, \mathbb{R})$, $\mathfrak{g}$ is the algebra of real $n \times n$ matrices, denoted $M(n, \mathbb{R})$. If $G$ is the group of unitary matrices $\text{U}(n)$, $\mathfrak{g}$ is the space of anti-Hermitian $n \times n$ matrices. Moreover they are traceless if $G = \text{SU}(n)$. Likewise, for the orthogonal $\text{O}(n)$ group, $\mathfrak{g}$ is made of skew-symmetric, hence traceless, matrices. For the symplectic group $G = \text{USp}(n)$, $\mathfrak{g}$ is generated by “anti-selfdual” quaternionic matrices, see Appendix A.

2. The exponential map plays an important role in the reconstruction of the Lie group $G$ from its tangent space $\mathfrak{g}$. One may prove, and we admit, that

- the map $X \in \mathfrak{g} \mapsto e^X \in G$ is bijective in the neighborhood of the identity;
- it is surjective (= every element in $G$ is reached) if $G$ is connected and compact;
- it is injective (any $g \in G$ has only one antecedent) only if $G$ is simply connected. An example of non-injectivity is provided by $G = \text{U}(1)$, for which $\mathfrak{g} = i\mathbb{R}$ and all the $i(x + 2\pi k)$, $k \in \mathbb{Z}$ have the same image by exp. The converse is in general wrong: for example in $\text{SU}(2)$ which is simply connected, if $n$ is a unit vector, $e^{i\pi n \cdot \sigma} = -I$, hence all elements $\pi in \cdot \sigma \mathfrak{g} = \text{su}(2)$ have the same image!

- Example of a non-compact group for which the exp map is non surjective: $G = \text{SL}(2, \mathbb{R})$, for which $\mathfrak{g} = \text{sl}(2, \mathbb{R})$, the set of real traceless matrices. For any $A \in \mathfrak{g}$, hence traceless, use its characteristic equation to show that $\text{tr} A^{2n+1} = 0$, $\text{tr} A^{2n} = 2(-\det A)^n$, hence $\text{tr} e^A = 2 \cosh \sqrt{-\det A} \geq -2$. There exist in $G$, however, matrices of trace $< -2$, for instance diag $(-2, -\frac{1}{2})$.

- For a non compact group, the exp map may still be useful. One may prove that any element of a matrix group may be written as the product of a finite number of exponentials of elements in its Lie algebra. [Cornwell p 151].

Observe that one still has $\det e^X = e^{\text{tr} X}$, a property easily established if $X$ belongs to the set of diagonalizable matrices. As the latter are dense in $M(d, \mathbb{R})$, the property holds true in general.

1.3.4 The tangent space as a Lie algebra

Let us now show that the tangent space $\mathfrak{g}$ of $G$ at $e \equiv I$ has a Lie algebra structure. Given two one-parameter groups generated by two independent vectors $X$ et $Y$ of $\mathfrak{g}$, we measure their lack of commutativity by constructing their commutator (in a sense different from the usual one!) $g = e^{tX}e^{uY}e^{-tX}e^{-uY}$; for small $t \sim u$, that $g$ is close to the identity, and may be written $g = \exp Z$, $Z \in \mathfrak{g}$. Compute $Z$ to the first non trivial order

$$e^{tX}e^{uY}e^{-tX}e^{-uY} = (I + tX + \frac{1}{2}t^2X^2)(I + uY + \frac{1}{2}u^2Y^2)(I - tX + \frac{1}{2}t^2X^2)(I - uY + \frac{1}{2}u^2Y^2)$$

$$= I + (XY - YX)tu + O(t^3).$$

(1.23)

The computation has been carried out in the associative algebra of matrices, the neutral element being denoted $I$. All the neglected terms are of third order since $t \sim u$. To order 2, one thus
sees the appearance of the commutator in the usual sense, \(XY - YX\), \textit{i.e.} the Lie bracket of matrices \(X\) and \(Y\). In general, for an arbitrary Lie group, the bracket is defined by
\[
e^{tX}e^{uY}e^{-tX}e^{-uY} = e^Z, \quad Z = tu[X,Y] + O(t^3)
\] (1.24)
and one proves that this bracket has the properties (1.16) of a Lie bracket.

This fundamental result follows from a detailed discussion of the local form of the group operations in a Lie group ("Lie equations", see for example [OR]).

- **Adjoint map in the Lie algebra \(\mathfrak{g}\). Baker-Campbell-Hausdorff formula**

Let us introduce a handy notation. For any \(X \in \mathfrak{g}\), let \(\text{ad} X\) be the linear operator in the Lie algebra defined by
\[
Y \mapsto (\text{ad} X)Y := [X,Y],
\] (1.25)

hence
\[
(\text{ad}^p X)Y = [X,[X,\ldots[X,Y]\ldots]],
\] with \(p\) brackets (commutators).

Given two elements \(X\) and \(Y\) in \(\mathfrak{g}\), and \(e^X\) et \(e^Y\) the elements they generate in \(G\), does there exist a \(Z \in \mathfrak{g}\) such that \(e^Xe^Y = e^Z\)? The answer is yes, at least for \(X\) et \(Y\) small enough.

Note first that if \([X,Y] = 0\), the ordinary rules of computation apply and \(Z = X + Y\). In general, the \textit{Baker-Campbell-Hausdorff formula}, that we admit, gives an explicit expression of \(Z\).

\[
e^Xe^Y = e^Z \quad Z = X + \int_0^1 dt\psi(\exp \text{ad} X \exp t \text{ad} Y)Y
\] (1.26)

where \(\psi(.)\) is the function
\[
\psi(u) = \frac{u \ln u}{u - 1} = 1 + \frac{1}{2}(u-1) - \frac{1}{6}(u-1)^2 + \cdots, \quad (1.27)
\]
which is regular at \(u = 1\). The first terms in the expansion in powers of \(X\) and \(Y\) read explicitly
\[
Z = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}([X,[X,Y]] + [Y,[Y,X]]) + \cdots
\] (1.28)

This complicated formula has some useful particular cases. Hence if \(X\) et \(Y\) commute with \([X,Y]\), (1.26) simplifies to
\[
e^Xe^Y = e^{X+Y+\frac{1}{2}[X,Y]} = e^{X+Y}e^{\frac{1}{2}[X,Y]},
\] (1.29)
a formula that one may prove directly using the general identity
\[
e^{XY}e^{-X} = \sum_{n=0}^{\infty} \frac{1}{n!}\text{ad}^nXY
\] (1.30)
(which is nothing else than the Taylor expansion at \(t = 0\) of \(e^{XY}e^{-tX}\) evaluated at \(t = 1\), and writing and solving the differential equation satisfied by \(f(t) = e^{tX}e^{tY}\), \(f(0) = 1\)
\[
f'(t) = (X + e^{tX}e^{-tX})f(t)
\] (1.31)
\[
= (X + Y + t[X,Y])f(t).
\] (1.32)
On the other hand, to first order in $Y$, one may replace the argument of $\psi$ in (1.26) by $\exp(\text{ad}X)$ and then

$$Z = X + \sum_{n=0}^{\infty} \frac{B_n}{n!} (-1)^n \text{ad}^n X Y + O(Y^2) \quad (1.33)$$

where the $B_n$ are the Bernoulli numbers: $\sum_{n=0}^\infty B_n \frac{t^n}{n!} = \frac{t}{e^t - 1}$, $B_0 = 1, B_2 = \frac{1}{6}, B_4 = -\frac{1}{30}$ and, beside $B_1 = -\frac{1}{2}$, all $B$ of odd index vanish. Still to first order in $Y$, one has also

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

which is obtained by writing and solving the differential equation satisfied by $F(t) = \exp(t(X+Y)).\exp(-tX)$.

The convergence of these expressions may be proven for $X$ and $Y$ small enough. Note that this BCH formula makes only use of the ad map in the Lie algebra, and not of the ordinary matrix multiplication in $GL(d, \mathbb{R})$. This is what makes it a canonical and universal formula.

### 1.3.5 An explicit example: the Lie algebra of $SO(n)$

From the definition of elements of $\mathfrak{g}$ as tangent vectors in $G$ at $e \equiv I$, or else from the construction of one-parameter subgroups associated with each $X \in \mathfrak{g}$, follows the interpretation of $X$ as “infinitesimal generator” of the Lie group $G$. The actual determination of the Lie algebra of a given Lie group $G$ may be done in several ways, depending on the way the group is defined or represented.

If one has an explicit parametrization of the elements of $G$ in terms of $d$ real parameters, infinitesimal generators are obtained by differentiation wrt these parameters. See in Chap. 00, the explicit cases of $SO(3)$ and $SU(2)$ treated in that way.

If the group has been defined as the invariance group of some quadratic form in variables $x$, one may derive an expression of the infinitesimal generators as differential operators in $x$. Let us illustrate it on the group $O(n)$, the invariance group of the form $\sum_{i=1}^n x_i^2$ in $\mathbb{R}^n$. The most general linear transformation leaving that form invariant is $x \rightarrow x' = Ox$, with $O$ orthogonal. In an infinitesimal form, $O = I + \omega$, and $\omega = -\omega^T$ is an arbitrary skew-symmetric real matrix. An infinitesimal transformation of the form $\delta x^i = \omega^j x^j$ may also be written

$$\delta x^i = \omega^j x^j = -\frac{1}{2} \omega^{kl} J_{kl} x^i$$

$$(1.34)$$

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

$$(1.35)$$

(note that we allow to raise and lower freely the indices, thanks to the signature $(+)^n$ of the metric). This yields an explicit representation of infinitesimal generators of the $\text{so}(n)$ algebra. It is then a simple matter to compute the commutation relations$^4$

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

$$(1.36)$$

(In other words, the only non-vanishing commutators are of the form $[J_{ij}, J_{ik}] = -J_{jk}$ for any triplet $i \neq j \neq k \neq i$, and those that follow by antisymmetry in the indices.)

$^4$Note that wrt to the calculation carried out in the $O(3,1)$ group in Chap. 00, § 6.1, we have changed our conventions and use here anti-Hermitian generators.
One may proceed in a different way, by using a basis of matrices in the Lie algebra, regarded as the space of skew-symmetric $n \times n$ matrices. Such a basis is provided by matrices $A_{ij}$ labelled by pairs of indices $1 \leq i < j \leq n$, with matrix elements

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

Hence the matrix $A_{ij}$ has only two non vanishing (and opposite) elements, at the intersection of the $i$-th row and $j$-th column and vice versa. Check that these matrices $A_{ij}$ have commutation relations given (1.36).

**Exercise**: repeat this discussion and the computation of commutation relations for the group $SO(p,q)$ of invariance of the form $\sum_{i=1}^p x_i^2 - \sum_{i=p+1}^{p+q} x_i^2$. It is useful to introduce the metric tensor $g = \text{diag } ((+1)^p, (-1)^q)$.

### 1.3.6 An example of infinite dimension: the Virasoro algebra

In these notes, we are restricting our attention to Lie groups and algebras of finite dimension. Let us give here an example of infinite dimension. One considers diffeomorphisms $z \mapsto z' = f(z)$ where $f$ is an analytic (holomorphic) function of its argument except maybe at 0 and at infinity. (One also speaks of the “diffeomorphisms of the circle”.) This is obviously a group and an infinite dimensional manifold, and it manifests itself in the algebra of infinitesimal diffeomorphisms $z \mapsto z + \epsilon(z)$, generated by differential operators $\ell_n$

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

which satisfy

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

This Lie algebra is the Witt algebra. A modified form of this algebra, with a central extension (see. Chap. 2), i.e. with an additional “central” generator $c$ commuting with all generators, is called Virasoro algebra and appears naturally in physics. Calling $L_n$ et $c$ the generators of that algebra

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

(The $L_n$ may be thought of as quantum realizations of the operators $\ell_n$, with the $c$ term resulting from quantum effects...)

Check that the Jacobi identity is indeed satisfied by this algebra. One proves that this is the most general central extension of (1.38) that respects the Jacobi identity. Show that the subalgebra generated by $L_{\pm 1}$, $L_0$ is not affected by the central term. What is the geometric interpretation of the corresponding transformations?

The Virasoro algebra plays a central role in the construction of conformal field theories in 2d and in their application to two-dimensional critical phenomena and to string theory. More details in [DFMS].

### 1.4 Relations between properties of $g$ and $G$

Let us examine how properties of $G$ translate in $g$.

#### 1.4.1 Simplicity, semi-simplicity

Let us define the infinitesimal version of the notion of invariant subgroup. An ideal (also sometimes called an invariant subalgebra) in a Lie algebra $g$ is a subspace $\mathcal{I}$ of $g$ which is stable under multiplication (defined by the Lie bracket) by any element of $g$, i.e. such that $[\mathcal{I}, g] \subset \mathcal{I}$. The ideal is called abelian si $[\mathcal{I}, \mathcal{I}] = \{0\}$. 


A Lie algebra \( g \) is \textit{simple} if \( g \) has no other ideal than \( \{0\} \). It is \textit{semi-simple} if \( g \) has no other abelian ideal than \( \{0\} \).

Example. Consider the Lie algebra of SO(4), denoted so(4), see the formulae given in (1.36) for so(\( n \)). It is easy to check that the combinations

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

commute with

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

and that

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

where one sees two commuting copies of so(3). One writes so(4)=so(3)⊕ so(3). Obviously the algebra so(4) is not simple, but it is semi-simple.

Notice the difference between this case of so(4) and the case of the algebra so(3,1) studied in Chap. 00, §6.1. There, the indefinite signature forced us to complexify the algebra to “decouple” the two copies of the algebra so(3).

One has the following relations

\[
G \text{ simple } \implies g \text{ simple} \\
G \text{ semi-simple } \implies g \text{ semi-simple}
\]

but the converse is not true! Several different Lie groups may have the same Lie algebra, \textit{e.g.} SO(3) which is simple, and SU(2) which is not semi-simple, as seen above in §1.1.7.  

### 1.4.2 Compactness. Complexification

The Lie algebra of a compact Lie group is said to be compact.

At first sight, this definition looks non intrinsic to the algebra and seems to depend on the Lie group from which it derives. We shall see below that a condition (Cartan criterion) allows to remove this dependance.

One should also examine the issue of complexification. Several distinct groups may have different Lie algebras, that become isomorphic when the parameters are complexified. For instance, the groups O(3) et O(2,1), the first compact, the second non compact, have Lie algebras

\[
\begin{align*}
\mathfrak{o}(3) & \quad \begin{cases} 
X_1 = z\partial_y - y\partial_z \\
X_2 = x\partial_z - z\partial_x \\
X_3 = y\partial_x - x\partial_y \\
X_1, X_2 = y\partial_x - x\partial_y = X_3 \text{ etc}
\end{cases} \\
\mathfrak{o}(2,1) & \quad \begin{cases} 
\tilde{X}_1 = z\partial_y + y\partial_z \\
\tilde{X}_2 = x\partial_z + z\partial_x \\
\tilde{X}_3 = y\partial_x - x\partial_y \\
\tilde{X}_1, \tilde{X}_2 = y\partial_x - x\partial_y = \tilde{X}_3
\end{cases}
\end{align*}
\]

\[(1.40)\]

\footnote{Beware! Some authors call “simple” any Lie group, whose Lie algebra is simple. This amounts to making a distinction between the concept of simple group and simple Lie group. The latter is such that it has no non trivial invariant Lie group. Thus the Lie group SU(2) is a simple Lie group but not a simple group, as it has an invariant subgroup which is not Lie...}
that non isomorphic on the real numbers, but \(i\tilde{X}_1, i\tilde{X}_2\) et \(-\tilde{X}_3\) verify the \(o(3)\) algebra. The algebras \(o(3)\) and \(o(2,1)\) are said to have the same complexified form \(g_c\), or else, to be two real forms of \(g_c\), but only one of them, namely \(o(3)\), (or \(so(3)\)=\(su(2)\)), of this complexified form is compact. This complexified form is the \(sl(2,\mathbb{C})\) algebra, of which \(sl(2,\mathbb{R})\) is another non compact real form. (See Exercise B).

The algebras \(so(4)\) and \(so(3,1)\) studied above and in Chap. 00 provide another example of two algebras, which are two non-isomorphic real forms of the same complexified form. Autre exemple, \(sp(2n,\mathbb{R})\) et \(usp(n)\). (Voir Appendice A).

More generally, one may prove ([FH] p. 130) that

- any semi-simple complex Lie algebra has a unique real compact form.

To summarize, local topological properties of the Lie group are transcribed in the Lie algebra. The Lie algebra, however, is unable to capture global topological properties of the group.

### 1.4.3 Connectivity, simple-connectivity

- If \(G\) is non connected and \(G'\) is the subgroup of the connected component of the identity, the Lie algebras of \(G\) and \(G'\) coincide: \(g = g'\).
- If \(G\) is non simply connected, let \(\tilde{G}\) be its universal covering group. \(G\) et \(\tilde{G}\) being locally isomorphic, they have the same Lie algebra. Examples: \(U(1)\) and \(\mathbb{R}\); \(SO(3)\) and \(SU(2)\); \(SO(3,1)\) and \(SL(2,\mathbb{C})\).

To summarize:

Given a Lie group \(G\), we have constructed its Lie algebra. Conversely, a theorem by Cartan asserts that any Lie algebra is the Lie algebra of some Lie group [Ki, p.99]. More precisely, with every Lie algebra \(\mathfrak{g}\) corresponds a unique connected and simply connected Lie group \(G\), whose Lie algebra is \(\mathfrak{g}\). Any other connected Lie group \(G'\) with the same Lie algebra \(\mathfrak{g}\) has the form \(G' = G/H\) with \(H\) a finite or discrete invariant subgroup of \(G\). This agrees with what we saw above: if \(G\) is the covering group of \(G'\), \(G' = G/\pi_1(G')\). For example \(U(1) = \mathbb{R}/\mathbb{Z}\), \(SO(3) = SU(2)/\mathbb{Z}_2\). If \(G'\) is non connected, the previous property applies to the connected component of the identity.

### 1.4.4 Structure constants. Killing form. Cartan criteria

Given a basis \(\{t_\alpha\}\) in a \(d\)-dimensional Lie algebra \(\mathfrak{g}\), any element \(X\) of \(\mathfrak{g}\) reads \(X = \sum_{\alpha=1}^d x^\alpha t_\alpha\).

The structure constants of \(\mathfrak{g}\) (in that basis), defined by

\[
[t_\alpha, t_\beta] = C_{\alpha\beta}^\gamma t_\gamma,
\]

are clearly antisymmetric in their two lower indices \(C_{\alpha\beta}^\gamma = -C_{\beta\alpha}^\gamma\). Consider then

\[
\text{ad } X Y = [X, Y] = \sum x^\alpha y^\beta C_{\alpha\beta}^\gamma t_\gamma,
\]
1.4. RELATIONS BETWEEN PROPERTIES OF $\mathfrak{g}$ AND $G$

with $\text{ad}$ the operator defined above in (1.25), and consider the linear operator $\text{ad} X \text{ad} Y$ which acts in the Lie algebra according to

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

Exercises (easy !): show that the Jacobi identity is equivalent to the identity

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

(1.42)

(note the structure : a cyclic permutation on the three indices $\alpha, \beta, \gamma$ with $\epsilon$ fixed and summation over the repeated $\delta$) ; and show that this identity may also be expressed as

$$[\text{ad} X, \text{ad} Y] Z = \text{ad} [X, Y] Z .$$

(1.43)

Taking the trace of this linear operator $\text{ad} X \text{ad} Y$ defines the Killing form

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

(1.44)

a symmetric bilinear form (a scalar product) on vectors of the Lie algebra. The symmetric tensor $g_{\alpha\beta}$ is thus given by

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

(Symmetry in $\alpha, \beta$ is manifest on the 1st expression, it follows from the cyclicity of the trace in the 2nd.)

Note that this Killing form is invariant under the action of any $\text{ad} Z$ :

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

(1.45)

(think of $\text{ad} Z$ as an infinitesimal generator acting like a derivative, either on the first term, or on the second). Indeed the first term equals $\text{tr} (\text{ad} Z \text{ad} X \text{ad} Y - \text{ad} X \text{ad} Y \text{ad} Z)$ while the second is $\text{tr} (\text{ad} X \text{ad} Z \text{ad} Y - \text{ad} X \text{ad} Y \text{ad} Z)$, and they cancel thanks to the cyclicity of the trace. One may prove that in a simple Lie algebra, an invariant symmetric form is necessarily a multiple of the Killing form.

One may then use the tensor $g_{\alpha\beta}$ to lower the 3d label of $C_{\alpha\beta\gamma}^\gamma$, thus defining

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

Let us then show that this $C_{\alpha\beta\gamma}$ is completely antisymmetric in $\alpha, \beta, \gamma$. Given the already known antisymmetry in $\alpha, \beta$, it suffices to show that $C_{\alpha\beta\gamma}$ is invariant by cyclic permutations. This follows from (1.45) which may be written in a more symmetric form as

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

(1.46)

thus proving the announced property.
A quite remarkable theorem of E. Cartan states that:

- (i) A Lie algebra is semi-simple iff the Killing form is non-degenerate, i.e. \( \det g \neq 0 \).
- (ii) A real semi-simple algebra is compact iff the Killing form is negative definite.

Those are the Cartan criteria.

In one way, property (i) is easy to prove. Suppose that \( g \) is not semi-simple and let us show that \( \det g = 0 \). Let \( J \) be an ideal of \( g \), choose a basis of \( g \) made of a basis of \( J, \{t_1\}, \ i = 1, \cdots r \), complemented by \( t_a \), \( a = r + 1, \cdots d \). For \( 1 \leq i, j \leq r \), compute \( g_{ij} = \sum_{\alpha, \beta} c_{i\alpha}^{\beta} C_{\alpha\beta} \). By definition of an ideal, \( \alpha \) and \( \beta \) are themselves between 1 and \( r \), \( g_{ij} = \sum_{1 \leq k, l \leq r} C_{ik}^{\ell} C_{\ell j}^{\kappa} \). Hence the restriction of the Killing form of \( g \) to \( J \) is the Killing form of \( J \). If moreover the ideal is assumed to be abelian, \( g_{ij} = 0 \) and \( g_{a0} = 0 \) (Exercise: check that point!). The form is obviously degenerate (\( \det g = 0 \)). The reciprocal, \( \det g = 0 \Rightarrow g \) non semi-simple, is more delicate to prove.

Likewise, property (ii) is relatively easy to prove in the sense compacity \( \Rightarrow \) definite negative form. Start from an arbitrary positive definite symmetric bilinear form; for example in a given basis \( \{t_\alpha\} \), consider \( (X,Y) = \sum x^\alpha y^\beta \). For a compact group \( G \), one can make this form invariant by averaging over \( G : \varphi(X,Y) := \int d\mu(g)(gXg^{-1},gYg^{-1}) \). It is invariant \( \varphi(gXg^{-1},gYg^{-1}) = \varphi(X,Y) \), or in infinitesimal form, \( \varphi([X,Y]_\alpha, [X,Y]) = 0, (\text{cf.}(1.45)) \). It is also positive definite. Let \( e_\alpha \) be a basis which diagonalises it, \( \varphi(e_\alpha, e_\beta) = \delta_{\alpha\beta} \).

Let us calculate in that basis the matrix of the \( \text{ad}X \) operator and show that it is antisymmetric, \( (\text{ad}X)_{\alpha\beta} = -(\text{ad}X)_{\beta\alpha} \):

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

Hence the Killing form

\[
(X,X) = \text{tr}(\text{ad}X) = \sum (\text{ad}X)_{\alpha\beta}(\text{ad}X)_{\beta\alpha} = -\sum (\text{ad}X)_{\alpha\beta}^2 \leq 0
\]

is negative semi-definite, and if the algebra is semi-simple, it is negative definite, q.e.d.

Example. The case of \( \text{SO}(3) \) or \( \text{SU}(2) \) is familiar. The structure constants are given by the completely antisymmetric tensor \( C_{\alpha\beta\gamma} = \epsilon_{\alpha\beta\gamma} \). The Killing form is \( g_{\alpha\beta} = -2\delta_{\alpha\beta} \). Exercise : compute the Killing form for the algebra \( so(2,1) \), (see Exercise B).

A last important theorem (again by Cartan !) states that

- Any semi-simple Lie algebra \( g \) is a direct sum of simple Lie algebras \( g_i \):

\[
g = \bigoplus g_i .
\]

This is a simple consequence of (1.46). Consider a semi-simple algebra \( g \) with an ideal \( J \) and call \( C \) the complement of \( J \) wrt the Killing form, i.e. \( (J,C) = 0 \). By (1.46), \( ([C,J],J) = (C,[J,J]) = (C,J) = 0 \) (since \( J \) is a subalgebra), and \( ([C,J],C) = (J,C) = 0 \) (since \( J \) is an ideal), hence \( [C,J], \) orthogonal to any element of \( g \) for the non-degenerate Killing form, vanishes, \( [C,J] = 0 \), which means that \( g = J \oplus C \). Iterating the argument on \( C \), one gets the announced property.

Cartan made use of these properties to classify the simple complex and real Lie algebras. We return to this classification in Chap. 3.

### 1.4.5 Casimir operator(s)

With previous notations, given a semi-simple Lie algebra \( g \), hence with an invertible Killing form, and a basis \( \{t_\alpha\} \) of \( g \), we define

\[
C_2 = \sum_{\alpha,\beta} g^{\alpha\beta} t_\alpha t_\beta
\]  
(1.47)
where $g^{\alpha\beta}$ is the inverse of $g_{\alpha\beta}$, i.e. $g_{\alpha\gamma}g^{\gamma\beta} = \delta^\beta_\alpha$.

Formally, this combination of the $t$'s which does not make use of the bracket, does not live in Lie algebra but in its universal enveloping algebra $U\mathfrak{g}$, defined as the associative algebra of polynomials in elements of $\mathfrak{g}$. Here, since we restricted ourselves to $\mathfrak{g} \subset M(n, \mathbb{R})$, $U\mathfrak{g}$ may also be considered as a subalgebra of $M(n, \mathbb{R})$.

Let us now show that $C_2$ has a vanishing bracket (commutator) with any $t_\gamma$ hence with any element of $\mathfrak{g}$. This is the quadratic Casimir operator.

$$[C_2, t_\gamma] = \sum_{\alpha, \beta} g^{\alpha\beta} [t_\alpha t_\beta, t_\gamma]$$

$$= \sum_{\alpha, \beta} g^{\alpha\beta} (t_\alpha [t_\beta, t_\gamma] + [t_\alpha, t_\gamma] t_\beta)$$

$$= \sum_{\alpha, \beta, \delta} g^{\alpha\beta} C_{\beta\gamma}^\delta (t_\alpha t_\delta + t_\delta t_\alpha) \quad (1.48)$$

$$= \sum_{\alpha, \beta, \delta, \kappa} g^{\alpha\beta} g^{\delta\kappa} C_{\beta\gamma}^\delta (t_\alpha t_\delta + t_\delta t_\alpha) .$$

The term $\sum_{\beta, \kappa} g^{\alpha\beta} g^{\delta\kappa} C_{\beta\gamma}^\delta$ is antisymmetric in $\alpha \leftrightarrow \delta$, while the term in parentheses is symmetric. The sum thus vanishes, qed.

One shows that in a simple Lie algebra, (more precisely in its universal enveloping algebra), a quadratic expression in $t$ that commutes with all the $t$’s is proportional to the Casimir operator $C_2$. In other words, the quadratic Casimir operator is unique up to a factor.

Example. In the Lie algebra $so(3) \cong su(2)$, the Casimir operator $C_2$ is (up to a sign) $J^2$, which, as everybody knows, commutes with the infinitesimal generators $J^i$ of the algebra. In a non simple algebra, there are as many quadratic operators as there are simple components, see for example the two Casimir operators $J^2$ and $K^2$ in the (complexified) $so(3,1) \cong su(2) \oplus su(2)$ algebra of the Lorentz group (see Chap 00 §6.1); or $P^2$ and $W^2$ in the Poincaré algebra, see Chap. 00, §6.4.

There may exist other, higher degree Casimir operators. Check that

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

has a vanishing bracket with any $t_\gamma$. What is that $C_3$ in $su(2)$? See Bourbaki ([Bo], chap.) for a discussion of these general Casimir operators. See also exercise C below.

If one remembers that infinitesimal generators (vectors of the Lie algebra) may be regarded as differential operators in the group coordinates, one realizes that the Casimir operators yield invariant (since commuting with the generators) differential operators. In particular, the quadratic Casimir operator corresponds to an invariant Laplacian on the group (see Chap. 00, §2.3 for the case of $SO(3)$).

These Casimir operators will play an important role in the study of group representations.
A short bibliography

Mathematics books

[D] J. Dieudonné, *Éléments d’analyse*, Gauthier-Villars, in particular volumes 5-7 (comprehensive but difficult!).

A recent book is close to the spirit of the present course:


Group theory for physicists

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

See also lecture notes of group theory by and for physicists (in French), available on the CCSD server http://cel.ccsd.cnrs.fr/, in particular

J.-B. Z., *Introduction à la théorie des groupes et de leurs représentations*, (Notes de cours au Magistère MIP 1994), which focuses mainly on finite groups.
Appendix A. Quaternion field and symplectic groups

A.1 Quaternions

The set of quaternions is the algebra over \( \mathbb{C} \) generated by 4 elements, \( e_i, i = 1, 2, 3, \)

\[
q = q^{(0)}1 + q^{(1)}e_1 + q^{(2)}e_2 + q^{(3)}e_3 \quad q^{(i)} \in \mathbb{C}
\]  

(A-1)

with multiplication \( e_i^2 = e_1e_2e_3 = -1 \), from which it follows that

\[
e_1e_2 = -e_2e_1 = e_3
\]

and cyclic permutations. One may represent the \( e_i = -i\sigma_i \) in terms of Pauli matrices :

\[
e_i \mapsto -i\sigma_i.
\]

The conjugate of \( q \) is the quaternion \( \bar{q} = q^{(0)}1 - q^{(1)}e_1 - q^{(2)}e_2 - q^{(3)}e_3 \).

(A-2)

Note que la conjugation and Hermitian conjugation reverse the order of factors

\[
\overline{q_1q_2} = \bar{q}_2\bar{q}_1 \quad (q_1q_2)^\dagger = q_2^\dagger q_1^\dagger
\]

(A-5)

A real quaternion is a quaternion of the form (A-1) with \( q^{(i)} \in \mathbb{R} \), hence identical with its complex conjugate.

The set of real quaternions forms a field, which is also a space of dimension 4 over \( \mathbb{R} \). It is denoted \( H \) (from Hamilton).

A.2 Quaternionic matrices

Let us consider matrices \( Q \) with quaternionic elements \( (Q)_{ij} = q_{ij} \), ou \( Q = (q_{ij}) \). One may apply to \( Q \) the conjugations defined above. One may also transpose \( Q \). The Hermitian conjugate of \( Q \) is

\[
(Q^\dagger)_{ij} = q_{ji}^\dagger.
\]

(A-6)

The dual \( Q^R \) of a quaternionic matrix \( Q \) is the matrix

\[
(Q^R)_{ij} = \bar{q}_{ji}.
\]

(A-7)

(It plays for quaternionic matrices the same role as Hermitian conjugates for complex matrices.) A quaternionic matrix is self-dual if

\[
Q^R = Q = (q_{ij}) = (\bar{q}_{ji}),
\]

(A-8)

it is real quaternionic if

\[
Q^R = Q^\dagger \quad \text{hence} \quad q_{ij} = q_{ij}^*.
\]

(A-9)

i.e. if its elements are real quaternions.
A.3 Symplectic groups $\text{Sp}(2n, \mathbb{R})$ and $\text{USp}(n)$, and the Lie algebras $\mathfrak{sp}(2n)$ et $\mathfrak{usp}(n)$

Consider the $2n \times 2n$ matrix
\begin{equation}
S = \begin{pmatrix}
0 & 1_N \\
-1_N & 0
\end{pmatrix}
\end{equation}
and the associated “skew-symmetric” bilinear form
\begin{equation}
(X,Y) = X^T SY = \sum_{i=1}^{n} (x_i y_{i+n} - y_i x_{i+n}) .
\end{equation}

The symplectic group $\text{Sp}(2n, \mathbb{R})$ is the group of real $2n \times 2n$ matrices that preserve that form
\begin{equation}
B^T SB = S .
\end{equation}

In the basis where $X^T = (x_1, x_{n+1}, x_2, x_{n+2}, \cdots)$, the matrix $S = \text{diag} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \text{diag} (-e_2)$ in terms of quaternions, and the symplectic group is then generated by quaternionic $n \times n$ matrices $Q$ satisfying $Q^R Q = I$, (check !); the matrix $B$ being real, however, the elements of $Q$ are such that $q^{(\alpha)}_i$ are real for $\alpha = 0, 2$ and purely imaginary for $\alpha = 1, 3$. This group is non compact. Its Lie algebra $\mathfrak{sp}(2n, \mathbb{R})$ is generated by real matrices $A$ such that $A^T S + SA = 0$. The dimension of that group or of its Lie algebra is $n(2n+1)$. For $n = 1$, $\text{Sp}(2, \mathbb{R}) = \text{SL}(2, \mathbb{R})$.

A related group is $\text{USp}(n)$, generated by unitary real quaternionic $n \times n$ matrices $Q$ satisfying $Q^\dagger = Q^{-1}$. This is the invariance group of the quaternionic Hermitian form $\sum \bar{x}_i y_i$, $x,y \in \mathbb{H}^n$. It is compact since it is a subgroup of $\text{U}(2n)$. Its Lie algebra $\mathfrak{usp}(n)$ is generated by antiselfdual real quaternionic matrices $A = -A^R = -A^\dagger$ (check !). Its dimension is again $n(2n+1)$. For $n = 1$, $\text{USp}(1) = \text{SU}(2)$.

Expressing the condition on matrices $A$ of $\mathfrak{sp}(n, \mathbb{R})$ in terms of quaternions, one sees that the two algebras $\mathfrak{sp}(2n, \mathbb{R})$ and $\mathfrak{usp}(n)$ have the same complexified algebra, namely $\mathfrak{sp}(2n, \mathbb{C})$. Only $\mathfrak{usp}(n)$ is compact.

Appendix B. A short reminder of topology and differential geometry.

B.1 A lexicon of some concepts of topology used in these notes

*Topological space*: set $E$ with a collection of *open subsets*, with the property that the union of open sets and the intersection of a finite number of them is an open subset, and that $E$ and $\emptyset$ are open.

*Closed subset* of $E$: complement of an open subset of $E$.

*Neighborhood of a point* $x$: subset $E$ that contains an open set containing $x$. Let $V(x)$ be the set of neighborhoods of $x$.

A topological space is *separated* (or Hausdorff) if two distinct points have distinct neighborhoods. This will always be assumed in these notes.

*Basis of neighborhoods* $B(x)$ of a point $x$: subset of $V(x)$ such that any $V \in V(x)$ contains a $W \in B(x)$. (Intuitively, a basis is made of “enough” neighborhoods.)

*Continuous function*: a function $f$ from topological space $E$ to topological space $F$ is called continuous if the inverse image of every open set in $F$ is open in $E$.

*Compact space* $E$: topological (separated) space such that from any covering of $E$ by open sets,
one may extract a finite covering. on peut extraire un recouvrement fini;

Consequences :
– any infinite sequence of points in E has an accumulation point;
– if E is compact and \( f : E \mapsto F \) is continuous, \( f(E) \) is compact;
– any continuous function on a compact space E is bounded.

If E is a subspace of \( \mathbb{R}^n \), E compact \( \iff \) E closed and bounded (Borel-Lebesgue).

**Locally compact space** : (separated) space in which any point has at least one compact neighborhood. Examples : \( \mathbb{R} \) is not compact but is locally compact ; \( \mathbb{Q} \) is neither compact nor locally compact.

### B.2 Notion of manifold

A manifold \( M \) of dimension \( n \) is a space which locally, in the vicinity of each point, “resembles” \( \mathbb{R}^n \) or \( \mathbb{C}^n \). Counter-examples are given by two secant lines, or by \( \circ \). More precisely, there exists a basis of neighborhoods \( U_i \) covering \( M \), with *charts* \( f_i \), i.e. invertible and bicontinuous (homeomorphisms) functions between \( U_i \) and an open set of \( \mathbb{R}^n \) : \( f_i(U_i) \subset \mathbb{R}^n \). Let \( m \) be a point of \( M \), \( m \in U_i \), and \( f_i(m) = (x_1, x_2, \ldots, x_n) \) its image in \( \mathbb{R}^n : (x_1, x_2, \ldots, x_n) \) are the *local coordinates* of \( m \), which depend on the map. It is fundamental to know how to change the coordinate chart. The manifold is said to be differentiable of class \( C^k \) if for any pair of open sets \( U_i \) and \( U_j \) with a non-empty intersection, \( f_j \circ f_i^{-1} \) which maps \( f_i(U_i \cap U_j) \subset \mathbb{R}^n \) onto \( f_j(U_i \cap U_j) \) is of class \( C^k \).

Example : the sphere \( S^2 \) is an analytic manifold of dimension 2. One may choose as two open sets the sphere with its North, resp. South, pole removed, with a map to \( \mathbb{R}^2 \) given by the stereographic projection (see Problem below) from that pole.

### B.3 Tangent space

In differential geometry, a tangent vector \( X \) to a manifold \( M \) at a point \( x_0 \) is a linear differential operator, of first order in the derivatives in \( x_0 \), acting on functions \( f \) on \( M \). In local coordinates \( x^i \),

\[
  f(x) \mapsto \sum_i X^i \frac{\partial}{\partial x^i} \bigg|_{x_0} f(x)
\]

and under a change of coordinates \( \{x^i\} \mapsto \{y^j\} \), these operators transform by the Jacobian matrix \( \frac{\partial}{\partial y^j} = \sum_i \frac{\partial x^i}{\partial y^j} \bigg|_{x_0} \frac{\partial}{\partial x^i} \) with the transformation of \( X^i \mapsto Y^j \) that follows from it.

**Tangent vector to a curve** : if a curve \( C(t) \) passes through the point \( x_0 \) at \( t = 0 \), one may differentiate a function \( f \) along that curve

\[
  f \mapsto \frac{d f(C(t))}{dt} \bigg|_{t=0}
\]

which defines the tangent vector to the curve \( C \) at point \( x_0 \), also called velocity vector and denoted \( C'(0) \).

The tangent space to \( M \) at \( x_0 \), denoted \( T_{x_0} M \), is the vector space generated by the velocity vectors of all curves passing through \( x_0 \). The space \( T_{x_0} M \) has a basis made of \( \frac{\partial}{\partial x^i} \bigg|_{x_0} \) : it has the same dimension as \( M \).

If a vector \( X_x \) tangent to \( M \) at \( x \) is defined for any \( x \), this defines a *vector field* on the manifold \( M \).
CHAPTER 1. GROUPS. LIE GROUPS AND LIE ALGEBRAS

B.4 Lie group. Exponential map

Take a group $G$, $e$ its identity. Let $C(t)$ be a curve passing through $C(0) = e$, and let $X_e = (C'(t))_{t=0}$ be its velocity vector at $e$. For $g \in G$, one defines the left translate $g.C(t)$ of $C$ by $g$. Its velocity at $g$, $X_g = (g.C(t))'_{t=0}$, is called a left translated vector of $X_e$. The vector field $g \mapsto X_g$ is said to be left-invariant, it is the set of left translated vectors of $X_e$. The tangent space at $e$ and the space of invariants vector fields are thus isomorphic, and are both denoted $g$.

Conversely, given a tangent vector $X_e$ at $e$, let $C(t) = \exp tX_e$ be the unique solution to the differential equation

$$C'(t) = X_{C(t)}$$

which expresses that the curve $C(t)$ is tangent at any of its points to the left-invariant vector field, that equation being supplemented by the initial condition that $C(0) = e$. (This first-order differential equation has a solution, determined up to a constant (in the group), and that constant is fixed uniquely by the initial condition.)

Let us now prove that the function $\exp$ defined by (B-1) satisfies property (1.18). Note that $C(t)$ satisfies (B-2), and so does $C(t + t')$. Thus $C(t + t') = k.C(t)$, (with $k$ constant in the group), and that constant is fixed by taking $t = 0$, $C(t') = k$, hence $C(t + t') = C(t')C(t)$ and $C(-t) = C(t)^{-1}$, qed.

In the case of matrix groups considered in this course, the function $\exp$ is of course identical to the exponential function defined by its Taylor series (1.22).

Appendix C. Invariant measure on SU(2) and on U($n$)

The group SU(2) being isomorphic to a sphere $S^3$ is compact and one may thus integrate a function on the group with a wide variety of measures $d\mu(g)$. The invariant measure, such that $d\mu(g.g_1) = d\mu(g_1.g) = d\mu(g^{-1}) = d\mu(g)$, is, on the other hand, unique up to a factor.

A possible way to determine that measure is to consider the transformation $U \to U' = U.V$ where $U, V$ and hence $U'$ are unitary of the form (00-1.8) (i.e. $U = u_0 I - u_\sigma$, $u \in S^3$ etc) ; if the condition $u_0^2 + u^2 = 1$ is momentarily relaxed (but $v_0^2 + v^2 = 1$ maintained), this defines a linear transformation $u \to u'$ which conserves the norm $\det U = u_0^2 + u^2 = u_0'^2 + u'^2 = \det U'$. This is thus a rotation of the space $\mathbb{R}^4$ which preserves the natural measure $d^4 u \delta(u^2 - 1)$ on the unit sphere $S^3$ of equation $\det U = 1$. In other terms, that measure on the sphere $S^3$ gives
a right invariant measure: $d\mu(U) = d\mu(U.V)$. One may prove in a similar way that it is left invariant: $d\mu(U) = d\mu(VU)$. It is also invariant under $U \rightarrow U^{-1}$, since inversion in SU(2) amounts to the restriction to $S^3$ of the orthogonal transformation $u_0 \rightarrow u_0$, $u \rightarrow -u$ in $\mathbb{R}^4$, which preserves of course the natural measure on $S^3$:

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

The explicit form of the measure depends on the chosen parametrization. If one uses the direction $\hat{n}$ (or its two polar angles $\theta$ et $\phi$) and the rotation angle $\psi$, one finds

$$d\mu(U) = \frac{1}{2} \sin^2 \frac{\psi}{2} \sin \theta \, d\psi \, d\theta \, d\phi$$

normalized for SU(2) to

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

which is the “area” of the unit sphere $S^3$ and the volume of SU(2). For SO(3) where the angle $\psi$ has a range restricted to $(0, \pi)$, one finds instead $v(SO(3)) = \int_{SO(3)} d\mu(g) = \pi^2$.

The expression in any other coordinate system, like the Euler angles, is then obtained by computing the adequate Jacobian

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

(Note that $0 \leq \gamma \leq 4\pi$ for SU(2), whereas $0 \leq \alpha \leq 2\pi$ and $0 \leq \beta \leq \pi$).

Another method to derive these results appeals to the introduction of an invariant metric on the group; a square distance between two elements $U$ et $U + dU$ is defined by $ds^2 = \frac{1}{2} tr dU dU^\dagger$, it is invariant by $U \rightarrow UV$, $U \rightarrow VU$ or $U \rightarrow U^{-1}$, and an invariant integration measure then follows (cf Chap. 0, §2.3). With the parametrization $(\hat{n} = (\theta, \phi), \psi)$, one finds

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

which leads indeed to (C-1). In the parametrization by Euler angles,

$$U = e^{-i\alpha_2} e^{-i\beta_2} e^{-i\gamma_2}$$

whence

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

and with $\sqrt{g} = \sin \beta$ one does recover (C-3) (check it!).

**Case of U(n).**

Let us discuss rapidly the case of U(n). Any unitary matrix $U \in U(n)$ may be diagonalized in the form

$$U = VAV^\dagger,$$  \hfill (C-7)

with $\Lambda = \text{diag} (\lambda_1, \cdots, \lambda_n)$ and the $\lambda_i$ are in fact of modulus one $\lambda_j = e^{i\alpha_j}$. These $\lambda_i$ may be regarded as “radial” variables, while $V$ represents the “angular” variables. Note that $V$ has to be restricted not to commute with the diagonal matrix $\Lambda$. If the latter is generic, with distinct eigenvalues $\lambda_i$, $V$ lives in $U(n)/U(1)^n$. The natural metric, invariant under $U \leftrightarrow U'U$ or $\leftrightarrow U'V'$, reads $\text{tr} (dUdU^\dagger)$. But $dU = V (d\Lambda + [dX, \Lambda]) V^\dagger$, where $dX := V^\dagger dV$ is anti-Hermitian (and with no diagonal elements, why?). Thus $\text{tr} (dUdU^\dagger) = \sum_i |d\alpha_i|^2 + \sum_j |d\lambda_j|^2$.
2 \sum_{i<j} |dX_{ij}|^2 |\lambda_i - \lambda_j|^2 \] which defines the metric tensor $g_{\alpha \beta}$ in coordinates $\xi^\alpha = (\alpha_i, \Re X_{ij}, \Im X_{ij})$ and determines the integration measure

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

(C-8)

Here $\Delta(\lambda)$ is the Vandermonde determinant

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

(C-9)

The “radial” part of the integration measure is thus given by $|\Delta(e^{i\alpha})|^2 \prod d\alpha_i$ up to a factor, or equivalently

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

(C-10)
1.4. RELATIONS BETWEEN PROPERTIES OF $G$ AND $G^*$

Exercises for chapter 1.

A. Action of a group on a set
A group $G$ is said to act on a set $E$ if there exists a homomorphism $\beta$ of $G$ into the group of bijections of $E$ into itself.

1. Write explicitly the required conditions.
One then defines the orbit $O(x)$ of a point $x \in E$ as the set of images $\beta(g)x$ for all $g \in G$.

2. Show that belonging to the same orbit is an equivalence relation.
3. Example: action of $O(n)$ on $\mathbb{R}^n$. What are the orbits?
4. A space is homogeneous if it has only one orbit. Show that a trivial example is given by the action of translations on $\mathbb{R}^n$. More generally, what can be said of the left action of $G$ on itself, with $E = G$? Give other examples of homogeneous spaces for $G = O(3)$ or $\mathcal{L} = O(3,1)$.

5. One also defines the isotropy group $S(x)$ of the element $x \in E$, (also called stabilizer, or, by physicists, little group): this is the subgroup of $G$ leaving $x$ invariant:

$$S(x) = \{ g \in G | \beta(g)x = x \} . \quad (1-49)$$

Show that if $x$ and $y$ belong to the same orbit, their isotropy groups are conjugate. What is the isotropy group of a point $x \in \mathbb{R}^n$ under the action of $SO(n)$? of a time-like vector $p$ in Minkowski space under the action of the Lorentz group? Is $S(x)$ an invariant subgroup?

6. Show that there exists a bijection between points of the orbit $O(x)$ and the coset space $G/S(x)$. For a finite group $G$, deduce from it a relation between the orders (cardinalities) of $G$, $O(x)$ and $S(x)$. Is this set $G/S(x)$ homogeneous for the action of $G$?

Chap. 2 will be devoted to the particular case where $E$ is a vector space, with the linear transformations of $GL(E)$ acting as bijections: one then speaks of representations of $G$ in $E$.

B. Lie groups and algebras of dimension 3.

1. Recall the definition of the group $SU(1,1)$. What is its dimension?
2. Which equation defines its Lie algebra? What does that imply on the matrix elements of $X \in su(1,1)$? Prove that one may write a basis of $su(1,1)$ in terms of 3 Pauli matrices and compute their commutation relations. Is this algebra isomorphic to the $so(3)$ algebra?
3. One now considers the linear group $SL(2,\mathbb{R})$. What is its definition? How is its Lie algebra defined? Give a basis in terms of Pauli matrices. 4. Prove the isomorphism of the two algebras $su(1,1)$ et $sl(2,\mathbb{R})$.
5. Same questions with the algebra $so(2,1)$: definition, dimension, commutation relations, isomorphism with one of the previous algebras?
6. Using the Cartan criteria, discuss the semi-simplicity and the compacity of these various algebras. What is their relationship with $su(2)$?

(For the geometric relationship between the groups $SU(1,1)$, $SL(2,\mathbb{R})$ et $SO(1,2)$), see §13 and §24, vol. 1 in [DNF].)
C. Casimir operators in $u(n)$. 1. Prove that the $n^2$ matrices $t_{(ij)}$ of size $n \times n$, $1 \leq i, j \leq n$, with elements $(t_{(ij)})_{ab} = \delta_{ia}\delta_{jb}$ form a basis of the algebra $u(n)$. Compute their commutation relations and the structure constants of the algebra.

2. Compute the Killing form in that basis and check that the properties related to Cartan criteria are satisfied.

3. Show that the elements in the envelopping algebra $\mathbb{C}[t_{(ij)}] = \bigoplus_{1 \leq i_1, i_2, \cdots, i_r \leq n} t_{(i_1 i_2)} t_{(i_2 i_3)} \cdots t_{(i_r i_1)}$ commute with all $t_{(ij)}$ and are thus Casimir operators of degree $r$.

4. How to modify this discussion for the $su(n)$ algebra? ([Bu], chap 10).

Problem: Conformal transformations

I-1. We recall that in a (classical) local, translation invariant field theory, one may define a stress-energy tensor $\Theta_{\mu\nu}(x)$ such that

- under an infinitesimal change of coordinates $x^\mu \to x'^\mu = x^\mu + a^\mu(x)$, the action has a variation
  \[ \delta S = \int d^d x (\partial_\mu a_\nu) \Theta^{\mu\nu}(x) ; \]  
  \[ (1-50) \]
- $\Theta_{\mu\nu}$ is conserved: $\partial_\mu \Theta^{\mu\nu}(x) = 0$ ;
- we assume that $\Theta_{\mu\nu}$ is symmetric in $\mu$, $\nu$.

Prove that if $\Theta$ is traceless, $\Theta_{\mu}^{\mu} = 0$, the action is also invariant under dilatations, $x^\mu \to x'^\mu = (1 + \delta\lambda) x^\mu$.

2. In a Riemannian or pseudo-Riemannian of dimension $d$, with a metric tensor $g_{\mu\nu}(x)$ of signature $\{(+1)^p, (-1)^{d-p}\}$, a conformal transformation is a coordinate transformation $x^\mu \to x'^\mu$ which is a local dilatation of lengths

\[ ds'^2 = g_{\mu\nu}(x)dx^\mu dx^\nu \to ds'^2 = g_{\mu\nu}(x')dx'^\mu dx'^\nu = \alpha(x) ds^2 \]  
\[ (1-51) \]

a) Write the infinitesimal form of that condition, when $x^\mu \to x'^\mu = x^\mu + a^\mu(x)$. (Hint : One may relate the dilatation parameter $1 + \delta\alpha$ to $a^\mu$ by taking some adequate trace.)

b) Prove that for an Euclidean or pseudo-Euclidean space of metric $g_{\mu\nu} = \text{diag} \{(+1)^p, (-1)^{d-p}\}$, that condition may be recast as

\[ \partial_\mu a_\nu + \partial_\nu a_\mu = \frac{2}{d} g_{\mu\nu} \partial_\rho a^\rho . \]  
\[ (1-52) \]

3. Prove, using (1-50,1-52), that under the conditions of 1. and 2.b, any field theory invariant under translations, rotations and dilatations is also invariant under conformal transformations.

4. We now study consequences of (1-52).We set $D := \frac{1}{d} \partial_\rho a^\rho$.

- a) Differentiating (1-52) with respect to $x^\nu$, prove that
  \[ \partial^2 a_\mu = (2 - d) \partial_\mu D. \]  
  \[ (1-53) \]
1.4. RELATIONS BETWEEN PROPERTIES OF $G$ AND $G$

- b) Differentiating (1-53) w.r.t. $x^\mu$, prove that in dimension $d > 1$, $D$ is a harmonic function: $\partial^2 D = 0$.

- c) We assume in the following that $d \geq 2$. Differentiating (1-53) w.r.t. $x^\nu$, symmetrizing it in $\mu$ and $\nu$ and using (1-52), prove that if $d > 2$, then $\partial_\mu \partial_\nu D = 0$. Show that it implies the existence of a constant scalar $h$ and of a constant vector $k$ such that $D = k_\mu x^\mu + h$.

- d) Differentiating (1-52) w.r.t. $x^\sigma$ and antisymmetrizing it in $\nu$ and $\sigma$, prove that

$$\partial_\mu (\partial_\sigma a_\nu - \partial_\nu a_\sigma) = 2(g_{\mu\nu} k_\sigma - g_{\mu\sigma} k_\nu) = \partial_\mu (2k_\sigma x^\nu - 2k_\nu x^\sigma). \quad (1-54)$$

- e) Show that it implies the existence of a constant skew-symmetric tensor $l_{\sigma\nu}$ such that

$$\partial_\sigma a_\nu - \partial_\nu a_\sigma = (2k_\sigma x^\nu - 2k_\nu x^\sigma) + 2l_{\sigma\nu}, \quad (1-55)$$

which, together with (1-52), gives

$$\partial_\sigma a_\nu = x_\nu k_\sigma - x_\sigma k_\nu + l_{\sigma\nu} + g_{\nu\sigma} k_\rho x^\rho + hg_{\nu\sigma}. \quad (1-56)$$

- f) Conclude that the general expression of an infinitesimal conformal transformation in dimension $d > 2$ reads

$$a_\nu = k_\sigma x^\sigma x_\nu - \frac{1}{2} x_\sigma x^\sigma k_\nu + l_{\sigma\nu} x^\sigma + hx_\nu + c_\nu \quad (1-56)$$

with $c$ a constant vector\(^6\). On how many independent real parameters does such a transformation depend in dimension $d$?

II-1. One learns in geometry that in the (pseudo-)Euclidean space of dimension $d > 2$, conformal transformations are generated by translations, rotations, dilatations and “special conformal transformations”, obtained by composition of an inversion $x^\mu \rightarrow x^\mu / x^2$, a translation and again an inversion. Write the finite and the infinitesimal forms of special conformal transformations, and show that this result is in agreement with (1-56), which justifies the previous assertion.

2. Write the expression of infinitesimal generators $P_\mu$ of translations, $J_{\mu\nu}$ of rotations, $D$ of dilatations and $K_\mu$ of special transformations, as differential operators in $x$.

3. Write with the minimum of calculations the commutation relations of these generators (Hint: use already known results on the generators $P_\mu$ and $J_{\mu\nu}$, and make use of homogeneity and of the definition of special conformal transformations to reduce the only non-trivial computation to that of $[K_\mu, P_\nu]$). Check that these commutation “close” on this set of generators $P, J, D$ et $K$.

4. What is the dimension of the conformal group in the Euclidean space $\mathbb{R}^d$?

III-1. To understand better the nature of the conformal group, one now maps the space $\mathbb{R}^d$, completed by the point at infinity and endowed with its metric $x^2 = x_1^2 + \cdots + x_d^2$, on the

\(^6\)This pretty argument is due to Michel Bauer.
CHAPTER 1. GROUPS. LIE GROUPS AND LIE ALGEBRAS

Figure 1.4: Stereographic projection from the North pole

sphere $S^d$. This sphere is defined by the equation $r^2 + r_{d+1}^2 = 1$ in the space $\mathbb{R}^{d+1}$, and the mapping is the stereographic projection from the “North pole” $r = 0$, $r_{d+1} = 1$ (see fig. 1.4). Prove that

$$r = \frac{2x}{x^2 + 1} \quad r_{d+1} = \frac{x^2 - 1}{x^2 + 1}.$$  

What is the image of the point at infinity? What is the effect of the inversion in $\mathbb{R}^d$ on the point $r = (r, r_{d+1}) \in S^d$?

2. The previous sphere is in turn regarded as the section of the light-cone $C$ in Minkowski space $\mathcal{M}_{d+1,1}$ with metric $z_0^2 - z^2 - z_{d+1}^2 = 0$ by the hyperplane $z_0 = 1$. Prove that this establishes a one-to-one correspondence between points of $\mathbb{R}^d \cup \{\infty\}$ and rays of the light-cone (i.e. vectors up a dilatation) and that the expression of $x \in \mathbb{R}^d$ as a function of $z = (z_0, z, z_{d+1}) \in C$ is

$$x = \frac{z}{z_0 - z_{d+1}}.$$  

3. We now want to prove that the action of the conformal group in $\mathbb{R}^d$ follows from linear transformations in $\mathcal{M}_{d+1,1}$ that preserve the light-cone. Without any calculation, show that these transformations must then belong to the Lorentz group in $\mathcal{M}_{d+1,1}$, that is $O(d + 1, 1)$.

a) What are the linear transformations of $z$ corresponding to rotations of $x$ in $\mathbb{R}^d$. Show that dilatations of $x$ correspond to “boosts” of rapidity $\beta$ in the plane $(z_0, z, z_{d+1})$, by giving the relation between the dilatation parameter and the rapidity.

b) Let us now consider transformations of $O(d + 1, 1)$ that preserve $z_0 - z_{d+1}$. Write the matrix $T_a$ of such an infinitesimal transformation acting on coordinates $(z_0, z, z_{d+1})$, and such that $\delta z = a(z_0 - z_{d+1})$ (to first order in $a$). To which transformation of $x \in \mathbb{R}^d$ does it correspond? Compute by exponentiation of $T_a$ the matrix of a finite transformation (Hint: compute the first powers $T_a^2, T_a^3$ ...).

c) What is finally the interpretation of the inversion in $\mathbb{R}^d$ in the Lorentz group of $\mathcal{M}_{d+1,1}$? What can be said about special conformal transformations? What is the dimension of the group $O(d + 1, 1)$? What can be concluded about the relation between the Lorentz group in Minkowski space $\mathcal{M}_{d+1,1}$ and the conformal group $\mathbb{R}^d$?

IV. Last question: Do you know conformal transformations in the space $\mathbb{R}^2$ that are not of the type discussed in II.1?
Chapter 2

Linear representations of groups

The action of a group in a set has been mentioned in the previous chapter (see exercise A and TD). We now focus our attention on the linear action of a group in a vector space. This situation is frequent in geometry and in physics (quantum mechanics, statistical physics, field theory, ...). One should keep in mind, however, that other group actions may have some physical interest: for instance the rotation group SO(n) acts on the sphere $S^{n-1}$ in a non-linear way, and this is relevant for example in models of ferromagnetism and field theories called non linear $\sigma$ models, see the course of F. David.

2.1 Basic definitions and properties

2.1.1 Basic definitions

A group $G$ is said to be represented in a vector space $E$ (on a field which for us is always $\mathbb{R}$ or $\mathbb{C}$), or stated differently, $E$ carries a representation of $G$, if one has a homomorphism $D$ of the group $G$ into the group $\text{GL}(E)$ of linear transformations of $E$:

$$\forall g \in G \quad g \mapsto D(g) \in \text{GL}(E)$$  \hspace{1cm} (2-1)

$$\forall g, g' \in G \quad D(g.g') = D(g).D(g')$$  \hspace{1cm} (2-2)

$$D(e) = I$$  \hspace{1cm} (2-3)

$$\forall g \in G \quad D(g^{-1}) = (D(g))^{-1}$$  \hspace{1cm} (2-4)

where $I$ denotes the identity operator in $\text{GL}(E)$. If the representation space $E$ is of finite dimension $p$, the representation itself is said to be of dimension $p$. The representation which to any $g \in G$ associates 1 (considered as $\in \text{GL}(\mathbb{R})$) is called trivial or identity representation; it is of dimension 1.

The representation is said to be faithful if $\ker D = \{e\}$, or equivalently if $D(g) = D(g') \iff g = g'$. Else, the kernel of the homomorphism is an invariant subgroup $H$, and the representation of the quotient group $G/H$ in $E$ is faithful (check!). Consequently, any non trivial representation
of a simple group is faithful. Conversely, if $G$ has an invariant subgroup $H$, any representation of $G/H$ gives a degenerate (i.e. non faithful) representation of $G$.

If $E$ is of finite dimension $p$, one may choose a basis $e_i$, $i = 1, \ldots, p$, and associate with any $g \in G$ the representative matrix of $D(g)$, denoted with a curly letter:

$$D(g)e_j = e_i D_{ij}(g)$$

with, as (almost) always in these notes, the convention of summation over repeated indices. The setting of indices ($i$: row index, $j$: column index) is dictated by (2-1). Indeed we have

$$D(g,g')e_k = e_i D_{ik}(g,g') = D(g)D(g')e_k = D(g)e_j D_{jk}(g')$$

$$D_{ik}(g,g') = D_{ij}(g)D_{jk}(g').$$

**Examples**: The group $SO(2)$ of rotations in the plane admits a dimension 2 representation, with matrices

$$
\begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}
$$

which describe indeed rotations of angle $\theta$ around the origin.

The group $SU(3)$ is defined as the set of unitary, unimodular $3 \times 3$ matrices $U$. These matrices form by themselves a representation of $SU(3)$, it is the “defining representation”. Show that the complex conjugate matrices form another representation of $SU(3)$.

Of which group the matrices

$$
\begin{pmatrix}
1 & a \\
0 & 1
\end{pmatrix}
$$

form a representation?

### 2.1.2 Equivalent representations. Characters

Take two representations $D$ et $D'$ of $G$ in spaces $E$ and $E'$, and suppose that there exists a linear operator $V$ from $E$ into $E'$ such that

$$\forall g \in G \quad V D(g) = D'(g)V.$$ 

Such a $V$ is called an *intertwining operator*, or “intertwiner” in short. If $V$ is invertible (and hence if $E$ and $E'$ have equal dimension, if finite), we say that the representations $D$ et $D'$ are equivalent. (It is an equivalence relation between representations!).

In the case of finite dimension, where one identifies $E$ and $E'$, the representative matrices of $D$ et $D'$ are related by a similarity transformation and may be considered as differing by a change of basis. There is thus no fundamental distinction between two equivalent representations, and in representation theory, one strives to study inequivalent representations.

One calls *character* of a finite dimension representation the trace of the operator $D(g)$:

$$\chi(g) = \text{tr} D(g).$$

It is a function of $G$ in $\mathbb{R}$ or $\mathbb{C}$ which satisfies the following properties (check!):
• The character is independent of the choice of basis in $E$.

• Two equivalent representations have the same character.

• The character takes the same value for all elements of a same conjugacy class of $G$: one says that the character is a class function: $\chi(g) = \chi(hgh^{-1})$.

The converse property, namely whether any class function may be expressed in terms of characters, is true for any finite group, and for any compact Lie group and continuous (or $L^2$) function on $G$: this is the Peter-Weyl theorem, see below §2.3.1.

Note also that the character, evaluated for the identity element in the group, gives the dimension of the representation 
\[ \chi(e) = \dim D. \] (2-13)

### 2.1.3 Reducible and irreducible representations

Another redundancy is due to direct sums of representations. Assume that we have two representations $D_1$ and $D_2$ of $G$ in two spaces $E_1$ and $E_2$. One may then construct a representation in the space $E = E_1 \oplus E_2$ and the representation is called direct sum of representations $D_1$ and $D_2$ and denoted $D_1 \oplus D_2$. (Recall that any vector of $E_1 \oplus E_2$ may be written in a unique way as a linear combination of a vector of $E_1$ and of a vector of $E_2$). The two subspaces $E_1$ and $E_2$ of $E$ are clearly left separately invariant by the action of $D_1 \oplus D_2$.

Inversely, if a representation of $G$ in a space $E$ leaves invariant a subspace of $E$, it is said to be reducible. Else, it is irreducible. If $D$ is reducible and leaves both the subspace $E_1$ and its complementary subspace $E_2$ invariant, one says that the representation est completely reducible (or decomposable); one may then consider $E$ as the direct sum of $E_1$ and $E_2$ and the representation as a direct sum of representations in $E_1$ and $E_2$.

If $E$ is finite dimensional, this means that the matrices of the representation take the following form (in a basis adapted to the decomposition!) with blocks of dimensions $\dim E_1$ and $\dim E_2$
\[ \forall g \in G \quad D(g) = \begin{pmatrix} D_1(g) & 0 \\ 0 & D_2(g) \end{pmatrix}. \] (2-14)

If the representation is reducible but not completely reducible, its matrix takes the following form, in a basis made of a basis of $E_1$ and a basis of some complementary subspace
\[ D(g) = \begin{pmatrix} D_1(g) & D'(g) \\ 0 & D_2(g) \end{pmatrix}. \] (2-15)

This is the case of representations of the translation group in one dimension. The representation
\[ D(a) = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \] (2-16)
is reducible, since it leaves invariant the vectors \((X, 0)\) but it has no invariant supplementary subspace.

On the other hand, if the reducible representation of \(G\) in \(E\) leaves invariant the subspace \(E_1\), there exists a representation in the subspace \(E_2 = E/E_1\). In the notations of equ. (2-15), its matrix is \(\mathcal{D}_2(g)\).

One should stress the importance of the number field in that discussion of irreducibility. For instance the representation (2-10) which is irreducible on a space over \(\mathbb{R}\) is not over \(\mathbb{C}\): it may be rewritten by a change of basis in the form

\[
\begin{pmatrix}
e^{-i\theta} & 0 \\
0 & e^{i\theta}
\end{pmatrix}.
\]  

(2-17)

**Conjugate and contragredient representations**

Given a representation \(D\), \(\mathcal{D}\) its matrix in some basis, the complex conjugate matrices \(\mathcal{D}^*\) form another representation \(D^*\), called conjugate representation, since they also satisfy (2-6)

\[
\mathcal{D}_{ik}^*(g,g') = \mathcal{D}_{ij}^*(g)\mathcal{D}_{jk}^*(g').
\]

(2-18)

The representation \(D\) is said to be real if there exists a basis where \(\mathcal{D} = \mathcal{D}^*\). This implies that its character \(\chi\) is real. Conversely if \(\chi\) is real, the representation \(D\) is equivalent to its conjugate \(D^*\) \(^1\). If the representations \(D\) and \(D^*\) are equivalent but if there no basis where \(\mathcal{D} = \mathcal{D}^*\), the representations are called pseudoreal. (This is for example the case of the spin \(\frac{1}{2}\) representation of SU(2).) For alternative and more canonical definitions of these notions of real representations, see the Problem III.

This concept plays a key role in the study of the "chiral non-singlet anomaly" in gauge theories: if fermions belong to a real or pseudoreal representation of the gauge group, their potential anomaly cancels, which is determinant for the consistency of the theory. In the standard model, this comes from a balance between contributions of quarks and leptons, see chap 5.

The contragredient representation of \(D\) is defined by

\[
\mathcal{D}(g) = D^{-1\text{T}}(g)
\]

(2-19)

or alternatively, \(\mathcal{D}_{ij}(g) = \mathcal{D}_{ji}(g^{-1})\), which does satisfy (2-6). For a unitary representation, see next paragraph, \(\mathcal{D}_{ij}(g) = \mathcal{D}_{ij}^*(g)\), and the contragredient equals the conjugate. The representations \(D, D^*\) and \(D\) are simultaneously reducible or irreducible.

**Unitary representations**

Suppose that the vector space \(E\) is “prehilbertian”, i.e. is endowed with a scalar product, (i.e. a form \(J(x, y) = \langle x|y \rangle = \langle y|x \rangle^*\), bilinear symmetric if we work on \(\mathbb{R}\), or sesquilinear on \(\mathbb{C}\),

\(^1\)This is true at least for the irreducible representations of finite and compact groups, for which we see below (§2.3) that two non irreducible representations are equivalent iff they have the same character.
such that the norm be positive definite: \( x \neq 0 \Rightarrow \langle x|x \rangle > 0 \). If the dimension of \( E \) is finite, one may find an orthonormal basis where the matrix of \( J \) reduces to \( I \) and then define unitary operators \( U \) such that \( U^\dagger U = I \). If the space is finite dimensional, (and is assumed to be a separable prehilbertian space\(^2\), one proves that one may find a countable orthonormal basis, thus labelled by a discrete index. A representation of \( G \) in \( E \) is called unitary if for any \( g \in G \), the operator \( D(g) \) is unitary. Then for any \( g \in G \) and \( x, y \in E \)

\[
\langle x|y \rangle = \langle D(g)x|D(g)y \rangle
\]

hence

\[
D(g)^\dagger D(g) = I
\]

and

\[
D(g^{-1}) = D^{-1}(g) = D^\dagger(g) .
\]

The following important properties hold:

(i) Any unitary reducible representation is completely reducible.

Proof: let \( E_1 \) be an invariant subspace, its complementary subspace \( E_2 = (E_1)_\perp \) is invariant since for all \( g \in G \), \( x \in E_1 \) et \( y \in E_2 \)

\[
\langle x|D(g)y \rangle = \langle D(g^{-1})x|y \rangle = 0
\]

whic proves that \( D(g)y \in E_2 \).

(ii) Any representation of a finite or compact group on a prehilbertian space is “unitarisable”, i.e. equivalent to a unitary representation.

Proof: consider first a finite group and define

\[
Q = \sum_{g' \in G} D^\dagger(g') D(g')
\]

which satisfies

\[
D^\dagger(g)QD(g) = \sum_{g' \in G} D^\dagger(g'.g) D(g'.g) = Q
\]

where the “rearangement lemma” (see §1...) \( \sum_{g'} \text{par} \sum_{g'.g} \) has been used. The self-adjoint operator \( Q \) is positive definite (why?) and may thus be written

\[
Q = V^\dagger V
\]

with \( V \) invertible. (For example, by diagonalisation of the operator \( Q \) by a unitary operator, \( Q = U\Lambda U^\dagger \), with \( \Lambda \) diagonal real, one may construct the “square root” \( V = UGLU^\dagger \).) The intertwiner \( V \) defines a representation \( D' \) equivalent to \( D \) and unitary:

\[
D'(g) = VD(g)V^{-1}
\]

\[
D'^\dagger(g)D'(g) = V^{-1}D^\dagger(g)V D(g)V^{-1}
\]

\[
= V^{-1}D^\dagger(g)Q D(g)V^{-1} = V^{-1}QV^{-1} = I
\]

\(^2\)A space is separable if it contains a dense countable subset.
In the case of a continuous compact group, the existence of the invariant Haar measure (see §1.2.4) allows to repeat the same argument with 

$$Q = \int d\mu(g') D^\dagger(g') D(g').$$

As a corollary of the two previous properties, any reducible representation of a finite or compact group on a prehilbertian is equivalent to a unitary completely reducible representation. It thus suffices to construct and classify unitary irreducible representations. We show below that, for a finite or compact group, these irreducible representations are finite dimensional.

Counter-example for a non compact group: the matrices

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$$

form an indecomposable (=non completely reducible) representation of the group $\mathbb{R}$.

### 2.1.4 Schur lemma

Consider two irreducible representations $D$ in $E$ and $D'$ in $E'$ and an intertwiner between them, as defined in (2-11). We then have the important

**Schur lemma:** *either $V = 0$, or $V$ is a bijection and the representations are equivalent.*

Proof: Suppose $V \neq 0$. Then $VD(g) = D'(g)V$ implies that ker $V$ is a subspace of $E$ invariant under $D$; by the assumption of irreducibility, it reduces to 0 (it cannot be equal to the whole $E$ otherwise $V$ would vanish). Likewise, the image of $V$ is a subspace of $E'$ invariant under $D'$, it cannot be $\{0\}$ and thus equals $E'$. A classical theorem on linear operators between vector spaces then asserts that $V$ is a bijection from $E$ to $E'$ and the representations are thus equivalent. q.e.d.

Note that if the two representations are not irreducible, this is generally false. A counter-example is given by the representation (2-16) which commutes with matrices $V = \begin{pmatrix} 0 & b \\ b & 0 \end{pmatrix}$.

**Corollary 1.** Any intertwining operator of an irreducible representation on $\mathbb{C}$ with itself, i.e. any operator that commutes with all the representatives of the group, is a multiple of the identity.

Proof: on $\mathbb{C}$, $V$ has at least one eigenvalue $\lambda$; $\lambda \neq 0$ since $V$ is invertible by Schur lemma). The operator $V - \lambda I$ is itself an intertwining operator, but it is singular and thus vanishes.

**Corollaire 2.** An irreducible representation on $\mathbb{C}$ of an abelian group is necessarily of dimension 1.

Proof: take $g' \in G$, $D(g')$ commutes with all $D(g)$ since $G$ is abelian. Thus (corollary 1) $D(g') = \lambda(g')I$. The representation decomposes into dim $D$ copies of the representation of dimension 1: $g \mapsto \lambda(g)$, and irreducibility imposes that dim $D = 1$.

Let us insist on the importance of the property of the complex field $\mathbb{C}$ to be algebraically closed, in contrast with $\mathbb{R}$, in these two corollaries. The representation on $\mathbb{R}$ of the group $SO(2)$ by matrices $D(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ provides counterexamples to both propositions: any matrix $D(\alpha)$ commutes with $D(\theta)$ but has no real eigenvalue (for $\theta \neq 0, \pi$) and the representation is irreducible on $\mathbb{R}$, although of dimension 2.

Application of Corollary 1: in the Lie algebra of a Lie group, the Casimir operators defined in Chap. 1 commute with all infinitesimal generators and thus with all the group elements. Anticipating a little bit on a
forthcoming discussion of representations of a Lie algebra, in a unitary representation these Casimir operators may be chosen hermitian hence diagonalisable, which allows to apply the argument of Corollary 1: in an irreducible representation, they are multiples of the identity. Thus for SU(2), $J^2 = j(j + 1)I$ in the spin $j$ representation.

2.1.5 Tensor product of representations. Clebsch-Gordan decomposition

Tensor product of representations

A very commonly used method to construct irreducible representations of a given group consists in building the tensor product of known representations and decomposing it into irreducible representations. This is the situation encountered in Quantum Mechanics, when the transformation properties of the components of a system are known and one wants to know how the system transforms as a whole (a system of two particles of spins $j_1$ and $j_2$ for example).

Let $E_1$ et $E_2$ be two vector spaces vectoriels carrying representations $D_1$ and $D_2$ of a group $G$. The tensor product $E = E_1 \otimes E_2$ is the space generated by linear combinations of (tensor) "products" of a vector of $E_1$ and a vector of $E_2$: $z = \sum_i x^{(i)} \otimes y^{(i)}$. The space $E$ carries also a representation, denoted $D = D_1 \otimes D_2$, the tensor product (one says also direct product) of representations $D_1$ and $D_2$. On the vector $z$ above

$$D(g)z = \sum_i D_1(g)x^{(i)} \otimes D_2(g)y^{(i)}. \quad (2-30)$$

One readily checks that the character of representation $D$ is the product of characters $\chi_1$ et $\chi_2$ de $D_1$ et $D_2$

$$\chi(g) = \chi_1(g)\chi_2(g) \quad (2-31)$$

In particular, evaluating this relation for $g = e$, one has for finite dimensional representations

$$\dim D = \dim(E_1 \otimes E_2) = \dim E_1 \cdot \dim E_2 = \dim D_1 \cdot \dim D_2 \quad (2-32)$$

as is well known for a tensor product.

Clebsch-Gordan decomposition

The tensor product representation of two irreducible representations $D$ et $D'$ is in general not irreducible. If it is fully reducible (as is the case for the unitary representations that are our chief concern), one performs the Clebsch-Gordan decomposition into irreducible representations

$$D \otimes D' = \oplus_j D_j \quad (2-33)$$

where in the right hand side, certain irreducible representations $D_1, \cdots$ appear. The notation $\oplus_j$ encompasses very different situations: summation over a finite set (for finite groups), on a

---

3The reader will find in Appendix A a short summary on tensor products and tensors.
finite subset of an a priori infinite but discrete set (compact groups) or on possibly continuous variables (non compact groups).

If $G$ is finite or compact and if its inequivalent irreducible representations are classified and labelled: $D^{(\rho)}$, one may rather rewrite (2-33) in a way showing which of these inequivalent representations appear, and with which multiplicity

$$D \otimes D' = \bigoplus_{\rho} m_{\rho} D^{(\rho)} .$$

(2-34)

A more correct expression would be $E \otimes E' = \bigoplus_{\rho} F_{\rho} \otimes E^{(\rho)}$ where $F_{\rho}$ is a vector space of dimension $m_{\rho}$, the “multiplicity space”.

The integers $m_{\rho}$ are all non negative. The equations (2-33) and (2-34) imply simple rules on characters and dimensions

$$\chi_{D} \cdot \chi_{D'} = \sum_{j} \chi_{j} = \sum_{\rho} m_{\rho} \chi^{(\rho)}$$

(2-35)

$$\text{dim } D \cdot \text{dim } D' = \sum_{j} \text{dim } D_{j} = \sum_{\rho} m_{\rho} \text{dim } D^{(\rho)} .$$

(2-36)

Example: the tensor product of two copies of the euclidean space $\mathbb{R}^3$ does not form an irreducible representation of the rotation group $\text{SO}(3)$. This space is generated by tensor products of vectors $\vec{x}$ and $\vec{y}$ and one may construct the scalar product $\vec{x} . \vec{y}$ which is invariant under the group (trivial representation), a skew-symmetric rank 2 tensor

$$A_{ij} = x_{i} y_{j} - x_{j} y_{i}$$

which transforms as a dimension 3 irreducible representation (spin 1),\footnote{Such a tensor is “dual” of a vector: $A_{ij} = \epsilon_{ijk} z_{k}$, $z = x \times y$} and a symmetric traceless tensor

$$S_{ij} = x_{i} y_{j} + x_{j} y_{i} - \frac{2}{3} \delta_{ij} \vec{x} . \vec{y}$$

which transforms as an irreducible representation of dimension 5 (spin 2); thus we may always decompose

$$x_{i} y_{j} = \frac{1}{3} \delta_{ij} \vec{x} . \vec{y} + \frac{1}{2} A_{ij} + \frac{1}{2} S_{ij} ;$$

(2-37)

the total dimension is of course $9 = 3 \times 3 = 1 + 3 + 5$, and labelling in that simple case the representations by their dimension, we write

$$D^{(3)} \otimes D^{(3)} = D^{(1)} \oplus D^{(3)} \oplus D^{(5)} .$$

(2-38)

Or equivalently, in a “spin” notation

$$(1) \otimes (1) = (0) \oplus (1) \oplus (2)$$

in which one recognizes the familiar rules of “addition of angular momentum”

$$(j) \otimes (j') = \bigoplus_{j''=|j-j'|}^{j+j'} (j'') .$$

(2-39)
By iteration, one finds

\[ D^{(3)} \otimes D^{(3)} \otimes D^{(3)} = D^{(1)} \oplus 3D^{(3)} \oplus 2D^{(5)} \oplus D^{(7)}, \]

with now multiplicities.

**Invariants.** A frequently encountered problem consists in counting the number of linearly independent invariants (under the action of a group \( G \)) in the tensor product of certain prescribed representations. This is an information contained in the decompositions into irr representations like \((2-34, 2-38, 2-40)\), where the multiplicity of the identity representation provides this number of invariants in the product of the considered representations. Exercise : interpret in terms of classical geometric invariants the multiplicities \( m_0 = 1, m_1, m_3 \) of the identity representation that appear in tensor products \((1) \otimes (1), (1) \otimes (1) \otimes (1), (1) \otimes (1) \otimes (1) \otimes (1)\) of \( \text{SO}(3) \). See also Problem II at the end of this chapter.

**Clebsch-Gordan coefficients**

Formula \((2-33)\) describes how the representation matrices decompose into irreducible representations under a group transformation. It is also often important to know how vectors of the representations at hand decompose. Let \( e^{(\rho)}_\alpha, \alpha = 1, \cdots, \dim D^{(\rho)} \), be a basis of vectors of representation \( \rho \). One wants to expand the product of two such basis vectors, that is \( e^{(\rho)}_\alpha \otimes e^{(\sigma)}_\beta \), on some \( e^{(\tau)}_\gamma \). As representation \( \tau \) may appear \( m_\tau \) times, one must introduce an extra index, \( i = 1, \cdots, m_\tau \). One writes

\[ e^{(\rho)}_\alpha \otimes e^{(\sigma)}_\beta = \sum_{\tau,\gamma,i} C^{\rho,\alpha;\sigma,\beta}_{\tau,\gamma,i} e^{(\tau)}_i. \]

or with notations borrowed from Quantum Mechanics

\[ |\rho, \alpha; \sigma, \beta \rangle \equiv |\rho \alpha \rangle |\sigma \beta \rangle = \sum_{\tau,\gamma,i} \langle \tau \gamma | \rho, \alpha; \sigma, \beta \rangle |\tau \gamma \rangle. \]

The coefficients \( C_{\rho,\alpha;\sigma,\beta|\tau,\gamma,i} = \langle \tau \gamma | \rho, \alpha; \sigma, \beta \rangle \) are the Clebsch-Gordan coefficients. In contrast with the multiplicities \( m_\rho \) in \((2-34)\), they have no reason of being integers, as we saw in Chap. 00 on the case of the rotation group, nor even real in general. Suppose that we consider unitary representations and that the bases have been chosen orthonormal. Then C.-G. coefficients which represent a change of orthonormal basis in the space \( E_1 \otimes E_2 \) satisfy orthonormality and completeness conditions

\[ \sum_{\tau,\gamma,i} \langle \tau \gamma | \rho, \alpha; \sigma, \beta \rangle \langle \tau \gamma | \rho, \alpha'; \sigma, \beta' \rangle^* = \delta_{\alpha,\alpha'} \delta_{\beta,\beta'} \]

\[ \sum_{\alpha,\beta} \langle \tau \gamma | \rho, \alpha; \sigma, \beta \rangle \langle \tau' \gamma' | \rho, \alpha; \sigma, \beta \rangle^* = \delta_{\tau,\tau'} \delta_{\gamma,\gamma'} \delta_{i,j}. \]

This enables us to invert relation \((2-42)\) into

\[ |\tau_i \gamma \rangle = \sum_{\alpha,\beta} \langle \tau_i \gamma | \rho, \alpha; \sigma, \beta \rangle^* |\rho, \alpha; \sigma, \beta \rangle. \]
and justifies the notation
\[ \langle \rho, \alpha; \sigma, \beta | \tau \gamma \rangle = \langle \tau \gamma | \rho, \alpha; \sigma, \beta \rangle^* \] (2-46)

\[ |\tau \gamma \rangle = \sum_{\alpha, \beta} \langle \rho, \alpha; \sigma, \beta | \tau \gamma \rangle |\rho, \alpha; \sigma, \beta \rangle. \] (2-47)

Finally, applying a group transformation on both sides of (2-42) and using these relations, one decomposes the product of matrices \( D^{(\rho)} \) and \( D^{(\sigma)} \) in a quite explicit way
\[ D^{(\rho)}_{\alpha \alpha'} D^{(\sigma)}_{\beta \beta'} = \sum_{\tau, \gamma, \gamma', i} \langle \tau \gamma | \rho, \alpha; \sigma, \beta \rangle^* \langle \tau \gamma' | \rho, \alpha'; \sigma, \beta' \rangle D^{(\tau_i)}_{\gamma \gamma'}. \] (2-48)

We shall see below (§2.4.3) an application of these formulae to Wigner-Eckart theorem.

### 2.1.6 Decomposition into irreducible representations of a subgroup of a group representation

Let \( H \) be a subgroup of a group \( G \), then any representation \( D \) of \( G \) may be restricted to \( H \) and yields a representation \( D' \) of the latter
\[ \forall h \in H \quad D'(h) = D(h). \] (2-49)

This is a very common method to build representations of \( H \), once those of \( G \) are known. In general, if \( D \) is irreducible, \( D' \) is not, and once again the question arises of its decomposition into irreducible representations. For example, given a finite subgroup of SU(2), one wants to set up the (finite, as we see below) list of its irreducible representations, starting from those of SU(2).

Another instance frequently encountered in physics: a symmetry group \( G \) is “broken” into a subgroup \( H \); how do the representations of \( G \) decompose into representations of \( H \)? Examples: in solid state physics, the “point group” \( G \subset SO(3) \) of symmetry (of rotations and reflexions) of a crystal is broken in \( H \) by an external field; in particle physics, we shall encounter in Chap. 4 and 5 the instances of \( SU(2) \subset SU(3); U(1) \times SU(2) \times SU(3) \subset SU(5) \), etc.

### 2.2 Representations of Lie algebras

#### 2.2.1 Definition. Universality

The notion de representation also applies to Lie algebras.

A representation of a Lie algebra \( \mathfrak{g} \) in a vector space \( E \) is by definition an homomorphism of \( \mathfrak{g} \) into the Lie algebra of linear operators on the space \( E \), i.e. a map \( X \in \mathfrak{g} \rightarrow d(X) \in \text{End} E \) which respects linearity and Lie bracket \( X, Y \in \mathfrak{g}, [X, Y] \rightarrow d([X, Y]) = [d(X), d(Y)] \in \text{End} V \). A corollary of this definition is that in any representation of the algebra, the (representatives of) generators satisfy the same commutation relations. In other words, in well chosen bases,
the structure constants are the same in all representations. More precisely, if \( t_i \) is a basis of \( g \), with \([t_i, t_j] = C_{ij}^k t_k\), and if \( T_i = d(t_i) \) is its image by the representation \( d \)

\[
[T_i, T_j] = [d(t_i), d(t_j)] = d([t_i, t_j]) = C_{ij}^k d(t_k) = C_{ij}^k T_k.
\]

Thus calculations carried in some particular representation and involving only commutation rules of the Lie algebra remain valid in any representation. We have seen in Chap. 00, §2.2, an illustration of this universality property. In contrast, Casimir operators take different values in different irreducible representations.

In parallel with the definitions of sect 2.1, one defines the notions of faithful representation of a Lie algebra (its kernel \( \text{ker} d = \{X | d(X) = 0\} \) reduces to the element 0 of \( g \)), of reducible or irreducible representation (existence or not of an invariant subspace), etc.

### 2.2.2 Representations of a Lie group and of its Lie algebra

Any differentiable representation \( D \) of \( G \) in a space \( E \) gives a map \( d \) of the Lie algebra \( g \) into the algebra of operators on \( E \). It is obtained by taking the infinitesimal form of \( D(g) \), with \( g(t) = I + tX \) (or \( g = e^{tX} \))

\[
d(X) := \left. \frac{d}{dt} D(g(t)) \right|_{t=0}, \tag{2-50}
\]

or, for \( t \) infinitesimal,

\[
D(e^{tX}) = e^{td(X)}. \tag{2-51}
\]

Let us show that this map is indeed compatible with the Lie bracket, thus giving a representation of the Lie algebra. For this purpose, we repeat the discussion of chap. 1, §1.3.4, to build the commutator in a natural way. Let \( g(t) = e^{tX} \) and \( h(u) = e^{uY} \) be two one-parameter subgroups, for \( t \) and \( u \) infinitesimally small and of same order. We have

\[
e^{tX}e^{uY}e^{-tX}e^{-uY} = e^Z \quad \text{avec} \quad Z = ut[X, Y] + \cdots,
\]

whence

\[
e^{d(Z)} = D(e^Z) = D(e^{tX}e^{uY}e^{-tX}e^{-uY}) = D(e^{tX})D(e^{uY})D(e^{-tX})D(e^{-uY})
\]

\[
= e^{td(X)}e^{ud(Y)}e^{-td(X)}e^{-ud(Y)} = e^{ud([X, Y]) + \cdots}, \tag{2-53}
\]

and by identification of the leading terms, \( d([X, Y]) = [d(X), d(Y)] \), qed.

This connection between a representation of \( G \) and a representation of \( g \) applies in particular to a representation of \( G \) which plays a special role, the adjoint representation of \( G \) into its Lie algebra \( g \). It is defined by the following action

\[
X \in g \quad D^{\text{adj}}(g)(X) = gXg^{-1}, \tag{2-54}
\]

which we denote \( \text{Ad} g X \). (The right hand side of (2-54) must be understood either as resulting from the derivative at \( t = 0 \) of \( g e^{tX}g^{-1} \), or, following the standpoint of these notes, as a matrix multiplication, since then the matrices \( g \) and \( X \) act in the same space.)

The adjoint representation of \( G \) gives rise to a representation of \( g \) in the space \( g \), also called adjoint representation. It is obtained by taking the infinitesimal form of (2-54), formally
g = I + tY, or by considering the one-parameter subgroup generated by Y ∈ g, g(t) = exp tY and by calculating Ad g(t)X = g(t)Xg(t)^{-1} = X + t[Y, X] + O(t^2) (cf. chap.1 equ. (1.3.15)), whence
\[
\left. \frac{d}{dt} \text{Ad} g(t)X \right|_{t=0} = [Y, X] = \text{ad} Y X.
\] (2-55)
where we recover (and justify) our notation ad of chap. 1.

Exercise: show that matrices T_i defined by (T_i)_{jk} = -C_{ij} \delta^k_j satisfy commutation relations of the Lie algebra as a consequence of the Jacobi identity, and thus form a basis of generators in the adjoint representation.

Remark. To a unitary representation of G corresponds a representation of g by antihermitian operators (or matrices). Physicists, who like Hermitian operators, usually include an “i” in front of the infinitesimal generators: for example e^{-i\psi J}, \[J_a, J_b]\] = i\epsilon^{abc} J_c, etc.

Conversely, a representation of a Lie algebra g generates a representation of the unique connected and simply connected group whose Lie algebra is g. In other words if X ↦ d(X) is a representation of the algebra, e^X ↦ e^{d(X)} is a representation of the group. Indeed, the BCH formula being “universal”, i.e. involving only linear combinations of brackets in the Lie algebra, and being thus insensitive to the representation of g, we have:
\[e^X e^Y = e^Z \mapsto e^{d(X)} e^{d(Y)} = e^{d(Z)},\]
showing that the homomorphism of Lie algebras integrates into a group homomorphism in the neighbourhood of the identity. One finally proves that such a local homomorphism of a connected and simply connected group G into a group G' (here, the linear group GL(E)) extends in a unique way into an infinitely differentiable homomorphism of the whole G into G'. To summarize, in order to find the (possibly unitary) representations of the group G it is sufficient to find the representations by (possibly antihermitian) operators of its Lie algebra.

This fundamental principle has already been illustrated in Chap. 00 on the concrete cases of SU(2) and SL(2, C).

### 2.3 Representations of compact Lie groups

In this section, we study the representations of compact groups on the field \( \mathbb{C} \) of complex numbers. Most of the results rely on the fact that one may integrate over the group with the Haar measure \( d\mu(g) \). Occasionally, we will compare with the non compact case. It is thus useful to have in mind two archetypical cases: the compact group U(1) = \{ e^{ix} \} with \( x \in \mathbb{R}/2\pi\mathbb{Z} \) (an angle modulo 2\pi), and the non compact group \( \mathbb{R} \), the additive group of real numbers. The case of finite groups, very close to that of compact groups, will be briefly mentionned at the end.

#### 2.3.1 Orthogonality and completeness

Let G be a compact group. We shall admit\(^5\) that its inequivalent irreducible representations are labelled by a discrete index, written in upper position: \( D^{(\rho)} \). These representations are a priori

\(^5\)See Kirillov p 135-137
of finite or infinite dimension, but we shall see below that the dimension \( n_\rho \) of \( D^{(\rho)} \) is in fact finite. In a finite or countable basis, the matrices \( D^{(\rho)}_{\alpha\beta} \) may be assumed to be unitary, according to the result of §2.1.2. (In contrast, a generic representation of a non compact compact depends on a continuous parameter. And we shall see that its unitary representations are necessarily of infinite dimension.)

In our two cases of reference, the irreducible representations of U(1) (hence of dimension 1 for this abelian group) are such that \( D^{(k)}(x)D^{(k)}(x') = D^{(k)}(x + x') \), they are of the form \( D^{(k)}(x) = e^{ikx} \) with \( k \in \mathbb{Z} \), the latter condition to make the representation single valued when one changes the determination \( x \to x + 2\pi n \). For \( G = \mathbb{R} \), one may also take \( x \to e^{ikx} \), but nothing restricts \( k \in \mathbb{C} \), except unitarity which imposes \( k \in \mathbb{R} \).

**Theorem:** For a compact group, the matrices \( D^{(\rho)}_{\alpha\beta} \) satisfy the following orthogonality properties

\[
\int \frac{d\mu(g)}{v(G)} D^{(\rho)}_{\alpha\beta}(g) D^{(\rho')\dagger}_{\alpha'\beta'}(g) = \frac{1}{n_\rho} \delta_{\rho\rho'} \delta_{\alpha\alpha'} \delta_{\beta\beta'}
\]

(2-56)

and their characters satisfy thus

\[
\int \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g) \chi^{(\rho')\dagger}(g) = \delta_{\rho\rho'} .
\]

(2-57)

In these formulae, \( d\mu(g) \) denotes the Haar measure and \( v(G) = \int d\mu(g) \) is the “volume of the group”.

**Proof:** Take \( M \) an arbitrary matrix of dimension \( n_\rho \times n_{\rho'} \) and consider the matrix

\[
V = \int d\mu(g') D^{(\rho)}(g') M D^{(\rho')\dagger}(g') .
\]

(2-58)

The left hand side of (2-56) is (up to a facteur \( v(G) \)) the derivative with respect to \( M_{\beta\beta'} \) of \( V_{\alpha\alpha'} \). The representations being unitary, \( D^{\dagger}(g) = D(g^{-1}) \), it is easy, using the left invariance of the measure \( d\mu(g') = d\mu(g g') \), to check that \( V \) satisfies

\[
V D^{(\rho')\dagger}(g) = D^{(\rho)}(g)V
\]

(2-59)

for all \( g \in G \). By Schur lemma, the matrix \( V \) is thus vanishing if representations \( \rho \) and \( \rho' \) are different, and a multiple of the identity if \( \rho = \rho' \).

a) In the former case, choosing a matrix \( M \) whose only non vanishing element is \( M_{\beta\beta'} = 1 \) and identifying the matrix element \( V_{\alpha\alpha'} \), one finds the orthogonality condition (2-56).

b) If \( \rho = \rho' \), choose first \( M_{11} = 1 \), the other \( M_{\beta\beta'} \) vanishing. One has \( V = c_1 I \), where the coefficient \( c_1 \) is determined by taking the trace: \( c_1 n_\rho = v(G) D_{11}(I) = v(G) \), which proves that the dimension \( n_\rho \) is finite.

c) Repeating the argument with an arbitrary matrix \( M \), one finds again \( V = c_M I \) and one computes \( c_M \) by taking the trace: \( c_M n_\rho = v(G) \text{tr} M \), which, upon differentiation wrt \( M_{\beta\beta'} \), leads to the orthonormality (2-56), qed.

The proposition (2-57) then follows simply from the previous one by taking the trace on \( \alpha = \beta \) and \( \alpha' = \beta' \).

Let us stress two important consequences of that discussion:
• we just saw that any irreducible (and unitary) representation of a compact group is of finite dimension;

• the relation (2-57) implies that two irreducible representations $D^{(\rho)}$ and $D^{(\sigma)}$ are equivalent (in fact identical, according to our labelling convention) iff their characters are equal:
  \[ \chi^{(\rho)} = \chi^{(\sigma)} \iff \rho = \sigma. \]

Case of a non compact group

A large part of the previous calculation still applies to a non compact group, provided it has a left invariant (which holds true for a wide class of groups, cf chap. 1, end of §1.2.4) and if the representation is in a prehilbertian separable space, hence with a discrete basis, and is square integrable: $D_{\alpha\beta} \in L^2(G)$. Choosing $M$ as in b), one finds again $\int d\mu(g) = c_1 \text{tr } I$. In the lhs, the integral over the group (“volume of the group” $G$) diverges. In the rhs, $\text{tr } I$, the dimension of the representation, is thus infinite.

More generally, one may assert

Any unitary square integrable representation of a non compact group is of infinite dimension. Of course, the trivial representation $g \mapsto 1$ (which is not in $L^2(G)$) evades the argument.

Let us test these results on the two cases $U(1)$ and $\mathbb{R}$. For the unitary representation $e^{ikx}$ of $U(1)$, the relation (2-56) (or (2-57), which makes no difference for these representations of dimension 1) expresses that

\[ \int_0^{2\pi} \frac{dx}{2\pi} e^{ikx} e^{-ik'x} = \delta_{kk'}, \]

as is well known. On the other hand on $\mathbb{R}$ it would lead to

\[ \int_{-\infty}^{\infty} dx e^{ikx} e^{-ik'x} = 2\pi \delta(k - k') \]

with a Dirac function. Of course this expression is meaningless for $k = k'$, the representation is not square integrable.

Completeness.

We return to compact groups. One may prove that the matrices $D_{\alpha\beta}^{(\rho)}(g)$ also satisfy a completeness property

\[ \sum_{\rho,\alpha,\beta} n_{\rho} D_{\alpha\beta}^{(\rho)}(g) D_{\alpha\beta}^{(\rho)*}(g') = v(G) \delta(g, g'), \quad (2-60) \]

or stated differently

\[ \sum_{\rho,\alpha,\beta} n_{\rho} D_{\alpha\beta}^{(\rho)}(g) D_{\beta\alpha}^{(\rho\dagger)}(g') = \sum_{\rho} n_{\rho} \chi^{(\rho)}(g) g^{-1} = v(G) \delta(g, g'), \quad (2-60)' \]

where $\delta(g, g')$ is the Dirac distribution adapted to the Haar measure, i.e. such that $\int d\mu(g') f(g') \delta(g, g') = f(g)$ for any sufficiently regular function $f$ on $G$. 
This completeness property is important: it tells us that any \( \mathbb{C} \)-valued function on the group, continuous or square integrable, may be expanded on the functions \( D_{\alpha \beta}^{(\rho)}(g) \)

\[
f(g) = \int d\mu(g') \delta(g, g') f(g') = \sum_{\rho, \alpha, \beta} n_{\rho} D_{\alpha \beta}^{(\rho)}(g) \int \frac{d\mu(g')}{v(G)} D_{\beta \alpha}^{(\rho)\dagger}(g') f(g') =: \sum_{\rho, \alpha, \beta} n_{\rho} D_{\alpha \beta}^{(\rho)}(g) f_{\alpha \beta}^{(\rho)} .
\]

(2.61)

This is the Peter–Weyl theorem, a non trivial statement that we admit\(^6\). A corollary then asserts that characters \( \chi^{(\rho)} \) of a compact group form a complete system of class functions, i.e. invariant under \( g \to hgh^{-1} \). In other words, any continuous (or \( L^2 \)) class function can be expanded of irreducible characters.

Let us prove the latter assertion. If \( f \) is a continuous class function, \( f(g) = f(hgh^{-1}) \), we apply the Peter-Weyl theorem and examine the integral appearing in (2.61):

\[
f_{\alpha \beta}^{(\rho)} = \int \frac{d\mu(g')}{v(G)} f(g') D_{\beta \alpha}^{(\rho)\dagger}(g') = \int \frac{d\mu(g')}{v(G)} f(hgh^{-1}) D_{\beta \alpha}^{(\rho)\dagger}(hgh^{-1}) \quad \forall h
\]

(2.62)

\[
= \int \frac{d\mu(h)}{v(G)} \frac{d\mu(g')}{v(G)} f(g') D_{\beta \gamma}^{(\rho)\dagger}(h) D_{\gamma \alpha}^{(\rho)\dagger}(g') D_{\delta \alpha}^{(\rho)\dagger}(h^{-1})
\]

(2.63)

\[
= \int \frac{d\mu(g')}{v(G)} f(g') D_{\beta \gamma}^{(\rho)\dagger}(g') \frac{1}{n_\rho} \delta_{\alpha \beta} \delta_{\gamma \delta}
\]

(2.64)

\[
= \frac{1}{n_\rho} \int \frac{d\mu(g')}{v(G)} f(g') \chi^{(\rho)*}(g') \delta_{\alpha \beta}
\]

(2.65)

from which follows that (2.61) reduces to an expansion on characters, qed.

Let us test these completeness relations again in the case of \( U(1) \). They express that

\[
\sum_{k=\pm \infty} e^{ikx} e^{-ikx'} = 2\pi \delta_P(x - x')
\]

(2.66)

where \( \delta_P(x - x') = \sum_{\ell=\pm \infty} \delta(x - x' - 2\pi \ell) \) is the periodic Dirac distribution (alias “Dirac’s comb”). Then (2.61) means that any \( 2\pi \)-periodic function (with adequate regularity conditions) may be represented by its Fourier series

\[
f(x) = \sum_{k=-\infty}^{\infty} e^{ikx} f_k \\
f_k = \int_{-\pi}^{\pi} \frac{dx}{2\pi} f(x) e^{-ikx} .
\]

(2.67)

For the non compact group \( \mathbb{R} \), the completeness relation (which is still true in that case) amounts to a Fourier transformation

\[
f(x) = \int_{-\infty}^{\infty} dk \tilde{f}(k) e^{ikx} \\
\tilde{f}(k) = \int_{-\infty}^{\infty} \frac{dx}{2\pi} f(x) e^{-ikx} .
\]

(2.68)

The Peter–Weyl theorem for an arbitrary group is thus a generalization of Fourier decompositions.

The \( SO(2) \) rotation group in the plane is isomorphic to the \( U(1) \) group. If we look at real representations, their dimension is no longer equal to 1 but to 2 (except the trivial representation)

\[
D^{(k)}(\alpha) = \begin{pmatrix} \cos k\alpha & -\sin k\alpha \\ \sin k\alpha & \cos k\alpha \end{pmatrix} , \quad k \in \mathbb{N}^* , \quad \chi^{(k)}(\alpha) = 2 \cos k\alpha
\]

(2.69)

What are now the orthogonality and completeness relations?

---

\(^6\)For a proof, see for example, T. Bröcker and T. tom Dieck, see bibliography at the end of this chapter.
2.3.2 Consequences

For a compact group,
(i) any representation being completely reducible, its character reads

\[ \chi = \sum \rho m_\rho \chi^{(\rho)} \]  

(2-70)

and multiplicities may be computed by

\[ m_\rho = \int \frac{d\mu(g)}{v(G)} \chi(g) \chi^{(\rho)^*}(g). \]  

(2-71)

One also has \( \|\chi\|^2 = \frac{1}{v(G)} \int |\chi(g)|^2 \, d\mu(g) = \sum \rho m_\rho^2 \), an integer greater or equal to 1. Thus a representation is irreducible iff its character satisfies the condition \( \frac{1}{v(G)} \int |\chi(g)|^2 \, d\mu(g) = 1 \). And the computation of \( \|\chi\|^2 \) gives indications on the number of irreducible representations appearing in the decomposition of the representation of character \( \chi \), a very useful information in the contexts mentionned in §2.1.5 et 2.1.6.

More generally, any class function may be expanded on irreducible characters (Peter-Weyl).

(ii) In a similar way one may determine multiplicities in the Clebsch-Gordan decomposition of a direct product of two representations by projecting the product of their characters on irreducible characters. Let us illustrate this on the product of two irreducible representations \( \rho \) et \( \sigma \)

\[ D^{(\rho)} \otimes D^{(\sigma)} = \bigoplus_\tau m_\tau D^{(\tau)} \]  

(2-72)

\[ \chi^{(\rho)} \chi^{(\sigma)} = \sum_\tau m_\tau \chi^{(\tau)} \]  

(2-73)

\[ m_\tau = \int \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g) \chi^{(\sigma)}(g) \chi^{(\tau)^*}(g), \]  

(2-74)

and hence the representation \( \tau \) appears in the product \( \rho \otimes \sigma \) with the same multiplicity as \( \sigma^* \) in \( \rho \otimes \tau^* \).

Case of SU(2)

It is a good exercise to understand how the different properties discussed in this section are realized by representation matrices of SU(2). This will be discussed in detail in TD and in App. E.

2.3.3 Case of finite groups

We discuss only briefly the case of finite groups. Theorems (2-56, 2-57, 2-60) and their consequences (2-70, 2-71, 2-72), which are based on the existence of an invariant measure, remain of course valid. It suffices to replace in these theorems the group volume \( v(G) \) by the order \( |G| \).
(=number of elements) of \( G \), and \( \int d\mu(g) \) by \( \sum_{g \in G} \) :

\[
\frac{1}{|G|} \sum_{g \in G} \mathcal{D}^{(\rho)}_{\alpha\beta}(g) \mathcal{D}^{(\rho')*}_{\alpha\beta}(g) = \frac{1}{n_{\rho}} \delta_{\rho\rho'} \delta_{\alpha\alpha'} \delta_{\beta\beta'}
\]

(2-75)

\[
\sum_{\rho,\alpha,\beta} \frac{n_{\rho}}{|G|} \mathcal{D}^{(\rho)}_{\alpha\beta}(g) \mathcal{D}^{(\rho')*}_{\alpha\beta}(g') = \delta_{g,g'}.
\]

(2-76)

But representations of finite groups enjoy additional properties. Let us show that the dimensions of inequivalent irreducible representations verify

\[
\sum_{\rho} n_{\rho}^2 = |G|.
\]

(2-77)

This follows from the fact that the system of equations (2-75) expresses that the matrix \( \mathcal{U}_{\rho,\alpha;\beta;g} := \left( \frac{n_{\rho}}{|G|} \right)^{\frac{1}{2}} \mathcal{D}^{(\rho)}_{\alpha\beta}(g) \) of dimensions \( \sum_{\rho} n_{\rho}^2 \times |G| \) satisfies \( \mathcal{U} \mathcal{U}^\dagger = I \), \( \mathcal{U}^\dagger \mathcal{U} = I \), which is possible only if it is a square matrix, qed.

Moreover

**Proposition.** The number \( r \) of inequivalent irreducible representations is finite and is equal to the number \( m \) of classes \( C_i \) in the group.

Proof: Denoting \( \chi^{(\rho)}_j \) the value of character \( \chi^{(\rho)} \) in class \( C_i \), one may rewrite the orthogonality and completeness properties of characters as

\[
\frac{1}{|G|} \sum_{i=1}^{m} |C_i| \chi^{(\rho)}_i \chi^{(\rho')*}_i = \delta_{\rho\rho'}
\]

(2-78)

\[
\frac{|C_i|}{|G|} \sum_{\rho=1}^{r} \chi^{(\rho)}_i \chi^{(\rho')*}_i = \delta_{ij}.
\]

(Exercise : derive the second relation from (2-61) and (2-62), applied to a finite group.) But once again, these relations mean that the matrix \( \mathcal{K}_{\rho,i} := \left( \frac{|C_i|}{|G|} \right)^{\frac{1}{2}} \chi^{(\rho)}_i \) of dimensions \( r \times m \) satisfies \( \mathcal{K} \mathcal{K}^\dagger = I \), \( \mathcal{K}^\dagger \mathcal{K} = I \), thus is square (and unitary), \( m = r \), qed.

The character table of a finite group is the square table made of the (real or complex) numbers \( \chi^{(\rho)}_i \), \( \rho = 1, \cdots r \), \( i = 1, \cdots, m = r \). Its rows and columns satisfy the orthogonality properties (2-78).

We illustrate it on the example of the group \( T \), subgroup of the rotation group \( \text{SO}(3) \) leaving invariant a regular tetrahedron. This group of order 12 has 4 conjugacy classes \( C_i \), that of the identity, that of the 3 rotations of \( \pi \) around an axis joining the middles of opposite edges, that of the 4 rotations of \( 2\pi/3 \) around an axis passing through a vertex, and that of the 4 rotations of \( -2\pi/3 \), see Fig. 2.1.

This group has 4 irreducible representations, and one easily checks using (2-77) that their dimensions can only be \( n_{\rho} = 1, 1, 1 \) and 3. The character table is thus a \( 4 \times 4 \) table, of which one row is already known, that of the identity representation \( D_1 \), and one column, that of dimensions \( n_{\rho} \). The spin 1 representation of \( \text{SO}(3) \) yields a dimension 3 representation of \( T \) whose character \( \chi \) takes the values \( \chi_i = 1 + 2 \cos \theta_i = (3, -1, 0, 0) \) in the four classes ; according to the criterion of §2.3.2, \( \| \chi \|^2 = \sum_i \frac{|C_i|}{|G|} |\chi_i|^2 = 1 \) and this character is irreducible. This gives a second row (called \( D_2 \)). The spin 2 representation of \( \text{SO}(3) \) gives a representation of dimension 5 which is reducible (same criterion) into a sum of 3 irreps, and is orthogonal to \( D_1 \). This is the sum of rows \( D_2 \), \( D_3 \) et \( D_4 \) in which \( j = e^{2\pi i/3} \), with \( j + j^2 = -1 \).
Check that relations (2-78) are satisfied. Explain why the group $T$ is nothing else than the alternate group $A_4$ of even permutations of 4 objects.

2.3.4 Recap

For a compact group, any irreducible representation is of finite dimension and equivalent to unitary representation. Its matrix elements and characters satisfy orthogonality and completeness relations. The set irreducible representations is discrete.

For a finite group, (a case very superficially treated in this course), the same orthogonality and completeness properties are satisfied. And one has additional properties, for example the number of inequivalent irreducible representations is finite, and equal to the number of conjugacy classes of the group.

For a non compact group, the unitary representations are generally of infinite dimension. (On the other hand there may exist non unitary finite dimensional representations, see for instance SL(2,$\mathbb{C}$)). The set of irreducible representations is indexed by discrete and continuous parameters.

2.4 Projective representations. Wigner theorem.

2.4.1 Definition

A projective representation of a group $G$ is a linear representation up to a phase of that group (here we restrict ourselves to unitary representations). For $g_1, g_2 \in G$, one has

$$U(g_1)U(g_2) = e^{i\zeta(g_1, g_2)}U(g_1g_2).$$  

(2-79)
2.4. PROJECTIVE REPRESENTATIONS. WIGNER THEOREM.

One may always choose $U(e) = I$, and thus $\forall g \quad \zeta(e, g) = \zeta(g, e) = 0$. One may also redefine $U(g) \to U'(g) = e^{i\alpha(g)}U(g)$, which changes

$$\zeta(g_1, g_2) \to \zeta'(g_1, g_2) = \zeta(g_1, g_2) + \alpha(g_1) + \alpha(g_2) - \alpha(g_1g_2). \quad (2-80)$$

The function $\zeta(g_1, g_2)$ of $G \times G$ in $\mathbb{R}$ is what is called a 2-cochain. It is closed (and it is thus called 2-cocycle) because of the associativity property:

$$\forall g_1, g_2, g_3 \quad (\partial \zeta)(g_1, g_2, g_3) := \zeta(g_1, g_2) + \zeta(g_1g_2, g_3) - \zeta(g_2, g_3) - \zeta(g_1, g_2g_3) = 0 \quad (2-81)$$

(check it). In general, for a $n$-cochain $\varphi(g_1, \ldots, g_n)$, one defines the operator $\partial$ which takes $n$-cochains to $n+1$-cochains:

$$(\partial \varphi)(g_1, \ldots, g_{n+1}) = \sum_{i=1}^{n} (-1)^{i+1} \varphi(g_1, g_2, \ldots, g_i g_{i+1}, \ldots, g_{n+1}) - \varphi(g_2, \ldots, g_{n+1}) + (-1)^n \varphi(g_1, \ldots, g_n).$$

For a 1-cochain $\alpha(g)$, $\partial \alpha(g_1, g_2) = \alpha(g_1, g_2) - \alpha(g_1) - \alpha(g_2)$, and hence (2-80) reads $\zeta' = \zeta - \partial \alpha$.

Check that $\partial^2 = 0$.

The questions whether representation $U(g)$ is intrinsically projective, or may be brought back to an ordinary representation by a change of phase amounts to knowing if the cocycle $\zeta$ is trivial, i.e. if there exists $\alpha(g)$ such that in (2-80), $\zeta' = 0$.

In other words, is the 2-cocycle $\zeta$, which is closed ($\partial \zeta = 0$) by (2-81), also exact, i.e. of the form $\zeta = \partial \alpha$? This is a typical problem of cohomology. Cohomology of Lie groups is a broad and much studied subject, ... on which we won't dwell in these lectures.

One may summarize a fairly long and complex discussion (sketched below in §2.4.4) by saying that for a semi-simple group $G$, such as $\text{SO}(n)$, the origin of the projective representations is to be found in the non simple-connectivity of $G$. Indeed, in the case of a non simply connected group $G$, the unitary representations of $\tilde{G}$, its universal covering, give representations up to a phase of $G$. For example, one recovers that the projective representations of $\text{SO}(3)$ (up to a sign) are representations of $\text{SU}(2)$. This is also the case of the Lorentz group $\text{O}(1,3)$, the universal covering of which is $\text{SL}(2, \mathbb{C})$.

Before we proceed, it is legitimate to ask the question: why are projective representations of interest for the physicist? The reason is that transformations of a quantum system make use of them, as we shall now see.

2.4.2  Wigner theorem

Consider a quantum system, the (pure) states of which are represented by rays\textsuperscript{7} of a Hilbert space $\mathcal{H}$, and in which the observables are auto-adjoint operators on $\mathcal{H}$. Suppose there exists a transformation $g$ of the system (states and observables) which leave unchanged the observable quantities $|\langle \phi | A | \psi \rangle|^2$, i.e.

$$|\psi \rangle \to |^g \psi \rangle, \quad A \to ^g A \quad \text{such that} \quad |\langle \phi | A | \psi \rangle|^2 = |\langle ^g \phi | ^g A | ^g \psi \rangle|^2. \quad (2-82)$$

One then proves the following theorem

\textsuperscript{7}ray = vector up to scalar, up to a phase if normalized
**Wigner theorem** If a bijection between rays and between auto-adjoint operators of a Hilbert space $\mathcal{H}$ preserves the modules of scalar products

$$|\langle \phi | A | \psi \rangle| = |\langle ^g\phi | ^gA | ^g\psi \rangle| ,$$  

(2-83)

then this bijection is realized by an operator $U(g)$, linear or antilinear, unitary on $\mathcal{H}$, and unique up to a phase, i.e.

$$|^g\psi \rangle = U(g)|\phi \rangle, \quad ^gA = U(g)AU^\dagger(g); \quad U(g)U^\dagger(g) = U(g)^\dagger U(g) = I.$$  

(2-84)

Recall first what is meant by antilinear operator. Such an operator satisfies

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

(2-85)

and its adjoint is defined by

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

(2-86)

so as to be consistent with linearity

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

(2-87)

If it is also unitary,

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

(2-88)

hence $\langle U\phi | U\psi \rangle = \langle \psi | \phi \rangle$.

The proof of the theorem is a bit cumbersome. It consists in showing that given an orthonormal basis $|\psi_k \rangle$ in $\mathcal{H}$, one may find representatives $|^g\psi_k \rangle$ of the transformed rays such that a representative of the transformed ray of $\sum c_k|\psi_k \rangle$ is $\sum c'_k|^g\psi_k \rangle$ with either all the $c'_k = c_k$, or all the $c'_k = c_k^*$. Stated differently, the action $|\psi \rangle \rightarrow |^g\psi \rangle$ is on the whole $\mathcal{H}$ either linear, or antilinear.

Once the transformation of states by the operator $U(g)$ is known, one determines that of observables: $^gA = U(g)AU^\dagger(g)$ so as to have

$$\langle ^g\phi | ^gA | ^g\psi \rangle = \langle U\phi | UAU^\dagger U | \psi \rangle$$  

(2-89)

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

(2-90)

$$= \langle \phi | A | \psi \rangle^*$$  

(2-91)

with $^\# = $ nothing or $^* $ depending on whether $U$ is linear or antilinear.

The antilinear case is not of academic interest. One encounters it in the study of *time reversal.*

The $T$ operation leaves unchanged the position operator $x$ inchangé, but changes the sign of velocities, hence of the momentum vector $p$

$$x' = U(T)xU^\dagger(T) = x$$  

(2-92)

$$p' = U(T)pU^\dagger(T) = -p .$$  

(2-93)
2.4. PROJECTIVE REPRESENTATIONS. WIGNER THEOREM.

The canonical commutation relations are consistent with $T$ only if $U(T)$ is antilinear

$$[x_j', p_k'] = -[x_j, p_k] = -i\hbar \delta_{jk} \tag{2-94}$$

$$= U(T)[x_j, p_k]U^\dagger(T) = U(T) i \hbar \delta_{jk} U^\dagger(T) \tag{2-95}$$

Another argument: $U(T)$ commutes with time translations, the generator of which is the Hamiltonian: $U(T)iHU^\dagger(T) = -iH$ (since $t \rightarrow -t$). If $U$ were linear, one would conclude that $UHU^\dagger = -H$, something embarrassing if the spectrum of $H$ is bounded from below, $\text{Spec}(H) \geq E_{\text{min}}$!

The transformations of a quantum system, i.e. the bijections of Wigner theorem, form a group $G$: if $g_1$ and $g_2$ are two such bijections, their composition $g_1g_2$ is another one, and so is $g_1^{-1}$ etc. By virtue of the unicity up to a phase of $U(g)$ in the theorem, the operators $U(g)$ (that will be assumed linear in the following) thus form a representation up to a phase, i.e. a projective representation of $G$.

**An important point of terminology**

Up to this point, we have been discussing transformations of a quantum system without an assumption on its possible invariance under these transformations, i.e. on the way they affect (or not) its dynamics. These transformations may be considered from an active standpoint: the original system is compared with the transformed system, or from a passive standpoint: the same system is examined in two different coordinate systems obtained from one another by the transformation.

2.4.3 Invariances of a quantum system

Suppose now that under the action of some group of transformations $G$, the system is invariant, in the sense that its dynamics, controlled by its Hamiltonian $H$, is unchanged. Let us write

$$H = U(g)HU^\dagger(g)$$

or alternatively

$$[H, U(g)] = 0 . \tag{2-96}$$

An invariance (or symmetry) of a quantum system under the action of a group $G$ is thus defined as the existence of a unitary projective (linear or antilinear) representation of that group in the space of states, that commutes with the Hamiltonian.

- This situation implies the existence of conservation laws. To see that, note that any observable function of the $U(g)$ commutes with $H$, and is thus a conserved quantity

$$i\hbar \frac{\partial \mathcal{F}(U(g))}{\partial t} = [\mathcal{F}(U(g)), H] = 0 \tag{2-97}$$

and each of its eigenvalues is a “good quantum number”: if the system is in an eigenspace $\mathcal{V}$ of $\mathcal{F}$ at time $t$, it stays in $\mathcal{V}$ in its time evolution. If $G$ is a Lie group, take $g$ an infinitesimal transformation and denote by $T$ the infinitesimal generators in the representation under study,

$$U(g) = I - i \delta \alpha^j T_j$$
(where one chose self-adjoint \( T \) to have \( U \) unitary), the \( T_j \) are observables that commute with \( H \), hence conserved quantities.

Examples.

Translation group \( \rightarrow P_\mu \) energy–momentum; rotation group \( \rightarrow M_{\mu \nu} \) angular momentum. Note also that these operators \( T_i \) which realise in the quantum theory the infinitesimal operations of the group \( G \) form a representation of the Lie algebra \( \mathfrak{g} \). One may thus state that they satisfy the commutation relations

\[
[T_i, T_j] = iC_{ij}^k T_k
\]  

(with an “i” because one chose to consider Hermitian operators). The maximal number of these operators that may be simultaneously diagonalised, hence of these conserved quantities that may be fixed, depends on the structure of \( \mathfrak{g} \) and of these commutation relations.

On the other hand, the assumption of invariance made above has another consequence, of frequent and important application. If the space of states \( \mathcal{H} \) which “carries” a representation of a group \( G \) is decomposed into irreducible representations, in each space \( E^{(\rho)} \), assumed first to be of multiplicity 1, the Hamiltonian is a multiple of the identity operator, by Schur’s lemma.

One has thus a complete information on the nature of the spectrum: eigenspace \( E^{(\rho)} \) and multiplicity of the eigenvalue \( \mathcal{E}_\rho \) of \( H \) equal to \( \dim E^{(\rho)} \). If some representation spaces \( E^{(\rho)} \) appear with a multiplicity \( m_\rho \) larger than 1, one has still to diagonalise \( H \) in the sum of these spaces \( \oplus_i E^{(\rho_i)} \), which is certainly easier that the original diagonalisation problem in the initial space \( \mathcal{H} \). We shall see below that the Wigner-Eckart theorem allows to simplify further the complexity of this last step. Group theory has thus considerably simplified our task, although it does not give the values of the eigenvalues \( \mathcal{E}_\rho \).

### 2.4.4 Transformations of observables. Wigner–Eckart theorem

According to (2-84), the transformation of an operator on \( \mathcal{H} \) obeys: \( A \rightarrow U(g)A U(g)^\dagger \). Suppose we are given a set of such operators, \( A_\alpha, \alpha = 1, 2, \cdots \), transforming linearly among themselves, \( i.e. \) forming a representation:

\[
A_\alpha \rightarrow U(g)A_\alpha U(g)^\dagger = \sum_{\alpha'} A_{\alpha'} D_{\alpha'\alpha}(g).
\]  

(2-99)

If the representation \( D \) is irreducible, the operators \( A_\alpha \) form what is called an irreducible operator (or “tensor”).

For example, in atomic physics, the angular momentum \( \vec{J} \) and the electric dipole moment \( \sum_i q_i \vec{r}_i \) are operators transforming like vectors under rotations.

Using the notations of section 2.2, suppose that the \( A_\alpha \) transform by the irreducible representation \( D^{(\rho)} \) and apply them on states \( |\sigma\beta\rangle \) transforming according to the irreducible representation \( D^{(\sigma)} \). The resulting state transforms as

\[
U(g)A_\alpha |\sigma\beta\rangle = U(g) A_\alpha U(g)^\dagger U(g)|\sigma\beta\rangle = D_{\alpha'\alpha}(g) D^{(\sigma)}_{\beta'\beta}(g) A_{\alpha'} |\sigma'\beta'\rangle
\]  

(2-100)

---

\( \text{It may happen that the multiplicity of some eigenvalue } \mathcal{E}_\rho \text{ of } H \text{ is higher than } m_\rho, \text{ either because of the existence of a symmetry group larger than } G, \text{ or because some representations come in complex conjugate pairs, or for some “accidental” reason.} \)
that is, according to the tensor product of representations $D^{(\rho)}$ and $D^{(\sigma)}$. Following (2-101), one may decompose on irreducible representations

$$D_{\alpha'\alpha}(g)D_{\beta'\beta}(g) = \sum_{\tau,\gamma,i} \langle \tau\gamma|\rho,\alpha;\sigma,\beta\rangle \langle \tau\gamma'|\rho,\alpha';\sigma,\beta'\rangle^{*} D_{\gamma'\gamma}(g).$$

(2-101)

Suppose now that the group $G$ is compact (or finite). The representation matrices satisfy the orthogonality property (2-56). One may thus write

$$\langle \tau\gamma|A_{\alpha}|\sigma\beta\rangle = \langle \tau\gamma|U(g)^{\dagger}U(g)A_{\alpha}|\sigma\beta\rangle \quad \forall g \in G$$

(2-102)

$$= \int \frac{d\mu(g)}{v(G)} \langle \tau\gamma|U(g)^{\dagger}U(g)A_{\alpha}|\sigma\beta\rangle$$

(2-103)

$$= \int \frac{d\mu(g)}{v(G)} \sum_{\alpha',\beta',\gamma'} D^{(\tau)}_{\gamma'\gamma}(g) \langle \tau\gamma'|A_{\alpha'}|\sigma\beta'\rangle D^{(\rho)}_{\alpha'\alpha}(g)D^{(\sigma)}_{\beta'\beta}(g)$$

(2-104)

$$= \frac{1}{n_{\tau}} \sum_{\alpha',\beta',\gamma,i} \langle \tau\gamma|\rho,\alpha;\sigma,\beta\rangle \langle \tau\gamma'|\rho,\alpha';\sigma,\beta'\rangle^{*} \langle \tau\gamma'|A_{\alpha}|\sigma\beta'\rangle.$$

(2-105)

Introduce the notation

$$\langle \tau \parallel A \parallel \sigma \rangle_{i} = \frac{1}{n_{\tau}} \sum_{\alpha'<\beta',\gamma'} \langle \tau\gamma'|\rho,\alpha';\sigma,\beta'\rangle^{*} \langle \tau\gamma'|A_{\alpha}|\sigma\beta'\rangle.$$

(2-106)

It follows that (Wigner–Eckart theorem):

$$\langle \tau\gamma|A_{\alpha}|\sigma\beta\rangle = \sum_{i=1}^{m_{\tau}} \langle \tau \parallel A \parallel \sigma \rangle_{i} \langle \tau\gamma|\rho,\alpha;\sigma,\beta\rangle$$

(2-107)

in which the “reduced matrix elements” $\langle . \parallel A \parallel . \rangle_{i}$ are independent of $\alpha,\beta,\gamma$. The matrix element of the lhs in (2-107) vanishes if the Clebsch-Gordan coefficient is zero, (in particular if the representation $\tau$ does not appear in the product of $\rho$ and $\sigma$). This theorem has many consequences in atomic and nuclear physics, where it gives rise to “selection rules”. See for example in Appendix E.3 the case of the electric multipole moment operators.

This theorem enables us also to simplify the diagonalisation problem of the Hamiltonian $H$ mentioned at the end of §2.4.3, when a representation space appears with a multiplicity $m_{\rho}$. Labelling by an index $i = 1, \cdots m_{\rho}$ the various copies of representation $\rho$, one has thanks to (2-107)

$$\langle \rho\alpha i|H|\rho\alpha' i'\rangle = \delta_{\alpha\alpha'} \langle \rho i \parallel H \parallel \rho i' \rangle$$

(2-108)

and the problem boils down to the diagonalisation of a $m_{\rho} \times m_{\rho}$ matrix.

**Exercise.** For the group SO(3), let $K_{1}^{m}$ be the components of an irreducible vector operator (for example, the electric dipole moment of Appendix E.3). Using Wigner-Eckart theorem show that

$$\langle j,m_{1}|K_{1}^{m}|j,m_{2}\rangle = \langle j,m_{1}|J^{m}|j,m_{2}\rangle \frac{\langle \tilde{J}\tilde{K} \rangle}{\sqrt{(j+1)}}$$

(2-109)

where $\langle \tilde{J}\tilde{K} \rangle$ denotes the expectation value of $\tilde{J}\tilde{K}$ in state $j$. In other terms, one may replace $\tilde{K}$ by its projection $\frac{\tilde{J}\tilde{K}}{\sqrt{j(j+1)}}$. 

2.4. PROJECTIVE REPRESENTATIONS. WIGNER THEOREM. 55
2.4.5 Infinitesimal form of a projective representation. Central extension

If $G$ is a Lie group of Lie algebra $\mathfrak{g}$, let $t_a$ be a basis of $\mathfrak{g}$

$$[t_a, t_b] = C_{ab}^c t_c .$$

In a projective representation, let us examine the composition of two infinitesimal transformations of the form $I + \alpha t_a$ and $I + \beta t_b$. As $\zeta(I, g) = \zeta(g, I) = 0$, $\zeta(I + \alpha t_a, I + \beta t_b)$ is of order $\alpha \beta$

$$i\zeta(I + \alpha t_a, I + \beta t_b) = \alpha \beta z_{ab} .$$

(2-110)

The $t_a$ are represented by $T_a$, and by expanding to second order, we find

$$e^{-i\zeta(I + \alpha t_a, I + \beta t_b)} U(e^{\alpha t_a}) U(e^{\beta t_b}) = U(e^{\alpha t_a} e^{\beta t_b}) = U(e^{(\alpha t_a + \beta t_b) e^{1/2} \alpha \beta [t_a, t_b]})$$

and thus, with $U(e^{\alpha t_a}) = e^{\alpha T_a}$ etc,

$$\alpha \beta \left( -z_{ab} I + \frac{1}{2} [T_a, T_b] - \frac{1}{2} C_{ab}^c T_c \right) = 0$$

(which proves that $z_{ab}$ must be antisymmetric in $a, b$). One thus finds that the commutation relations of $T$ are modified by a central term (i.e. commuting with all the other generators)

$$[T_a, T_b] = C_{ab}^c T_c + 2z_{ab} I .$$

The existence of projective representations may thus imply the realization of a central extension of the Lie algebra. One calls that way the new Lie algebra generated by the $T_a$ and by one or several new generator(s) $C_{ab}$ commuting with all the $T_a$ (and among themselves)

$$[T_a, T_b] = C_{ab}^c T_c + C_{ab} \quad [C_{ab}, T_c] = 0 .$$

(2-111)

(In an irreducible representation of the algebra, Schur’s lemma ensures that $C_{ab} = c_{ab} I$.) The triviality (or non-triviality) of the cocycle $\zeta$ translates in infinitesimal form into the possibility (or impossibility) of getting rid of the central term by a redefinition of the $T$

$$T_a \rightarrow \tilde{T}_a = T_a + X_a \quad [\tilde{T}_a, \tilde{T}_b] = C_{ab}^c \tilde{T}_c ,$$

(2-112)

in a way consistent with the contraints on the $C_{ab}^c$ and $C_{ab}$ coming from the Jacobi identity.

Exercise. Write the constraint that the Jacobi identity puts on the constants $C_{ab}^c$ et $C_{ab}$. Show that $C_{ab} = C_{ab}^c D_c$ gives a solution and that a redefinition such as (2-112) is then possible.

One proves (Bargmann) that for a connected Lie group $G$, the cocycles are trivial if

1. there exists no non-trivial central extension of $\mathfrak{g}$;

2. $G$ is simply connected.

As for point 1), a theorem of Bargmann tells us that there is no non-trivial central extension for any semi-simple group, like the classical groups $\text{SU}(n)$, $\text{SO}(n)$, $\text{Sp}(2n)$. It is thus point 2) which is relevant.

If the group $G$ is not simply connected, one studies the (say unitary) representations of its universal covering $\tilde{G}$, which are representations up to a phase of $G$ (the group $\pi_1(G) = \tilde{G}/G$ is represented on $U(1)$). This is the case of the groups $\text{SO}(n)$ and their universal covering $\text{Spin}(n)$, (for example $\text{SO}(3)$), or of the Lorentz group $\text{O}(1,3)$, as recalled above.

*
A short bibliography (cont’d)

In addition to references already given in the Introduction and in Chap. 1,

General representation theory


For a proof of Peter-Weyl theorem, see for example


For a proof of Wigner theorem, see E. Wigner, [Wi], ou A. Messiah, [M] t. 2, p 774, ou S. Weinberg, [Wf] chap 2, app A.

On projective representations, see


 Appendix D. ‘Tensors, you said tensors?’

The word “tensor” covers several related but not quite identical concepts. The aim of this appendix is to (try to) clarify these matters...

D.1. Algebraic definition

Given two vector spaces $E$ et $F$, their tensor product is by definition the vector space $E \otimes F$ generated by the pairs $(x, y)$, $x \in E$, $y \in F$, denoted $x \otimes y$. An element of $E \otimes F$ thus reads

$$z = \sum_{\alpha} x^{(\alpha)} \otimes y^{(\alpha)}$$  \hspace{1cm} (D-1)

with a finite sum over vectors $x^{(\alpha)} \in E$, $y^{(\alpha)} \in F$ (a possible scalar coefficient $\lambda_\alpha$ has been absorbed into a redefinition of the vector $x^{(\alpha)}$).

If $A$, resp. $B$, is a linear operator acting in $E$, resp. $F$, $A \otimes B$ is the linear operator acting in $E \otimes F$ according to

$$A \otimes B(x \otimes y) = Ax \otimes By$$  \hspace{1cm} (D-2)

$$A \otimes B \sum_\alpha (x^{(\alpha)} \otimes y^{(\alpha)}) = \sum_\alpha Ax^{(\alpha)} \otimes By^{(\alpha)}$$  \hspace{1cm} (D-3)

In particular if $E$ et $F$ have two bases $e_i$ and $f_j$, $z = x \otimes y = \sum_{i,j} x^i y^j e_i f_j$, the basis $E \otimes F$ and the components of $z$ are labelled by pairs of indices $(i, j)$, and $A \otimes B$ is described in that basis by a matrix which is read off

$$(A \otimes B)z = \sum_{i,j,i',j'} A_{ii'} B_{jj'} x^{i'} y^{j'} e_i f_j =: (A \otimes B)_{i',j'} z^{i'} e_i \otimes f_j$$  \hspace{1cm} (D-4)
thus

\[(A \otimes B)_{ij'; j'} = A_{ii'}B_{jj'}, \tag{D-5}\]
a formula which is sometimes taken as a definition of tensor product of two matrices.

**D.2. Group action**

If a group \(G\) has representations \(D\) and \(D'\) in two vector spaces \(E\) and \(F\), \(x \in E \mapsto D(g)x = e_iD_{ij}x^j,\) and likewise for \(y \in F,\) the tensor product representation \(D \otimes D'\) in \(E \otimes F\) is defined by

\[D(g) \otimes D'(g)(x \otimes y) = D(g)x \otimes D'(g)y \tag{D-6}\]
in accord with (D-2). The matrix of \(D \otimes D'\) in a basis \(e_i \otimes f_j\) is \(D_{ii'}D'_{jj'}\).

Another way of saying it is: if \(x^i\) "transforms by the representation \(D^i\)" and \(y^j\) by \(D'^j\), under the action of \(g \in G, \) i.e. \(x' = D(g)x, y' = D'(g)y, \) \(x \otimes y \mapsto x' \otimes y',\) with

\[(x' \otimes y')^{ij} = x^i y^j = D_{ii'}D'_{jj'}x^i y^j, \tag{D-7}\]
another formula sometimes taken as a definition of a tensor (under the action of \(G\)).

The previous construction of rank 2 tensors \(z^{ij}\) may be iterated to make tensor products \(E_1 \otimes E_2 \otimes \cdots \otimes E_p\) and rank \(p\) tensors \(z^{i_1 \cdots i_p}\). This is what we did in Chap. 00, §3.3, in the construction of the representations of \(SU(2)\) by symmetrized tensor products of the spin \(\frac{1}{2}\) representation, or in §6.2 for those of \(SL(2, \mathbb{C})\), by symmetrized tensor products of the two representations with pointed or unpointed indices, \((0, \frac{1}{2})\) and \((\frac{1}{2}, 0)\).

**Appendix E. More on representation matrices of \(SU(2)\)**

We return to the representation matrices \(D^j\) of \(SU(2)\) defined and explicitly constructed in §3.2 of Chap. 00.

### 2.4.6 Orthogonality, completeness, characters

All unitary representations of \(SU(2)\) have been constructed in Chap. 00. Following the discussion of §3, the matrix elements of \(D^j\) satisfy orthogonality and completeness properties, which make use of the invariant measure on \(SU(2)\) introduced in Chap. 1 (§1.2.4 and App. C)

\[
(2j + 1) \int \frac{d\mu(U)}{2\pi^2} D^j_{mn}(U)D^{j*}_{mn'}(U) = \delta_{jj'}\delta_{mm'}\delta_{nn'} \tag{E-1}
\]

\[
\sum_{jmn}(2j + 1)D^j_{mn}(U)D^{j*}_{mn}(U') = 2\pi^2\delta(U, U').
\]

The "function" \(\delta(U, U')\) appearing in the rhs of (E-1) is the one adapted to the measure \(d\mu(U),\) such that \(\int d\mu(U')\delta(U, U')f(U') = f(U);\) in Euler angles parametrization \(\alpha, \beta, \gamma\) for example,

\[
\delta(U, U') = 8\delta(\alpha - \alpha')\delta(\cos \beta - \cos \beta')\delta(\gamma - \gamma'), \tag{E-2}
\]
2.4. PROJECTIVE REPRESENTATIONS. WIGNER THEOREM.

(see Appendix C of Chap. 1). The meaning of equation (E-1) is that functions \( D_{mn}^j(U) \) form a complete basis in the space of functions (continuous or square integrable) on the group SU(2) (Peter–Weyl theorem).

Characters of representations of SU(2) follow from the previous expressions

\[
\chi_j(U) = \chi_j(\psi) = \text{tr} D_j(n, \psi) = \sum_{m=-j}^j e^{im\psi} \quad \text{(E-3)}
\]

\[
= \frac{\sin \left( \frac{2j+1}{2} \psi \right)}{\sin \frac{\psi}{2}}.
\]

Note that these expressions are polynomials (so-called Chebyshev polynomials of 2nd kind) of the variable \( 2 \cos \frac{\psi}{2} \) (see exercise D at the end of this chapter). In particular

\[
\chi_0(\psi) = 1 \quad \chi_1(\psi) = 2 \cos \frac{\psi}{2} \quad \chi_1(\psi) = 1 + 2 \cos \psi \quad \text{etc}.
\]

One may then verify all the expected properties

unitarity and reality \[ \chi_j(U^{-1}) = \chi_j(U) \]

parity and periodicity \[ \chi_j(-U) = \chi_j(2\pi + \psi) = (-1)^{2j} \chi_j(U) \]

orthogonality \[
\chi_j(\psi) \chi'_j(\psi') = \frac{\pi}{2} \sin^2 \frac{\psi}{2} \delta(\psi - \psi') = \frac{\pi}{2} \sin^2 \frac{\psi}{2} \delta(\cos \frac{\psi}{2} - \cos \frac{\psi'}{2})
\]

completeness \[
\sum_{j=0,\frac{1}{2},...} \chi_j(\psi) \chi_j(\psi') = \frac{\pi}{\sin^2 \frac{\psi}{2}} \delta(\psi - \psi') = \frac{\pi}{2} \sin^2 \frac{\psi}{2} \delta(\cos \frac{\psi}{2} - \cos \frac{\psi'}{2})
\]

The latter expresses that characters form a complete basis of class functions, \textit{i.e.} of even \( 2\pi \)-periodic functions of \( \frac{1}{2} \psi \). This is a variant of the Fourier expansion.

Does the multiplicity formula (2-72) lead to the well known formulae (2-39)?

2.4.7 Special functions. Spherical harmonics

We are by now familiar with the idea that infinitesimal generators act in each representation as differential operators. This is true in particular in the present case of SU(2): the generators \( J_i \) appear as differential operators with respect to parameters of the rotation, compare with the case of a one-parameter subgroup \( \exp -iJ\psi \) for which \( J = i\partial/\partial \psi \). This gives rise to differential equations satisfied by the \( D_{mn}^j \) and exposes their relation with “special functions” of mathematical physics.

We also noticed that the construction of the \textit{Wigner D matrices} in Chap. 00, §3.3, applies not only to SU(2) matrices but also to arbitrary matrices \( A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \) in the linear group GL(2,\( \mathbb{C} \)). Equation (3.23) of Chap. 00 still holds true

\[
P_{jm}(\xi', \eta') = \sum_{m'} P_{jm'}(\xi, \eta) D_{m'm}^j(A) \quad \text{(00.3.23)}
\]

The combination \( (a\xi + c\eta)^{j+m}(b\xi + d\eta)^{j-m} \) clearly satisfies

\[
\left( \frac{\partial^2}{\partial a\partial a} - \frac{\partial^2}{\partial b\partial c} \right) (a\xi + c\eta)^{j+m}(b\xi + d\eta)^{j-m} = 0 \quad \text{(E-6)}
\]
and because of the independance of the $P_{jm}(\xi, \eta)$, the $D^j_{m'm'}(A)$ satisfy the same equation. If we now impose that $d = a^*$, $c = -b^*$, but $\rho^2 = |a|^2 + |b|^2$ is kept arbitrary, the matrices $A$ satisfy $AA^\dagger = \rho^2 I$, $\det A = \rho^2$, hence $A = \rho U$, $U \in SU(2)$, and (E-6) leads to

$$\Delta_4 D^j_{m'm'}(A) = 4 \left( \frac{\partial^2}{\partial a \partial a^*} + \frac{\partial^2}{\partial b \partial b^*} \right) D^j_{m'm'}(A) = 0 \quad (E-7)$$

where $\Delta_4$ is the Laplacian in the space $\mathbb{R}^4$ with variables $u_0, u$, and $a = u_0 + iu_3$, $b = u_1 + iu_2$. In polar coordinates

$$\Delta_4 = \frac{\partial^2}{\partial \rho^2} + \frac{3}{\rho} \frac{\partial}{\partial \rho} + \frac{1}{\rho^2} \Delta_{S^3} \quad (E-8)$$

where the last term $\Delta_{S^3}$, Laplacian on the unit sphere $S^3$, acts only on “angular variables” $U \in SU(2)$. The functions $D^j$ being homogeneous of degree 2 in $a, b, c, d$ hence in $\rho$, one finally gets

$$-\frac{1}{4} \Delta_{S^3} D^j_{m'm'}(U) = j(j + 1) D^j_{m'm'}(U) \quad . \quad (E-9)$$

For example, using the parametrization by Euler angles, one finds

$$\left\{ \frac{1}{\sin \beta} \frac{\partial}{\partial \beta} \sin \beta \frac{\partial}{\partial \beta} + \frac{1}{\sin^2 \beta} \left[ \frac{\partial^2}{\partial \alpha^2} + \frac{\partial^2}{\partial \gamma^2} - 2 \cos \beta \frac{\partial^2}{\partial \alpha \partial \gamma} \right] + j(j + 1) \right\} D^j(\alpha, \beta, \gamma)_{m'm} = 0 \quad . \quad (E-10)$$

For $m = 0$ (hence $j$ necessarily integer), the dependence on $\gamma$ disappears (see (00.3.14)). Choose for example $\gamma = 0$ and perform a change of notations $(j, m') \rightarrow (l, m)$ and $(\beta, \alpha) \rightarrow (\theta, \phi)$, so as to recover classical notations. The equation reduces to

$$\left[ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + l(l + 1) \right] D^l_{m0}(\phi, \theta, 0) = 0 \quad . \quad (E-11)$$

The differential operator made of the first two terms is the Laplacian $\Delta_{S^2}$ on the unit sphere $S^2$. Equation (E-11) thus defines spherical harmonics $Y^m_l(\theta, \phi)$ as eigenvectors of the Laplacian $\Delta_{S^2}$. The correct normalisation is

$$\left[ \frac{2l + 1}{4\pi} \right]^{\frac{1}{2}} D^l_{m0}(\phi, \theta, 0) = Y^m_l(\theta, \phi) \quad . \quad (E-12)$$

Introduce also the Legendre polynomials and functions $P_l(u)$ and $P^m_l(u)$, which are defined for integer $l$ and $u \in [-1, 1]$ by

$$P_l(u) = \frac{1}{2^l l!} \frac{d^l}{du^l} (u^2 - 1)^l \quad \quad (E-13)$$

$$P^m_l(u) = (1 - u^2)^{\frac{m}{2}} \frac{d^m}{du^m} P_l(u) \quad \quad \text{for } 0 \leq m \leq l \quad . \quad (E-14)$$

The Legendre polynomials $P_l(u)$ are orthogonal polynomials on the interval $[-1, 1]$ with the weight 1: $\int_{-1}^1 P_l(u) P_l'(u) du = \frac{2}{2l + 1} \delta_{ll'}$. The first $P_l$ read

$$P_0 = 1 \quad P_1 = u \quad P_2 = \frac{1}{2} (3u^2 - 1) \quad P_3 = \frac{1}{2} (5u^3 - 3u) \quad , \cdots \quad (E-15)$$
while $P_l^0 = P_l$, $P_l^1 = (1-u^2)^{1/2} P_l$, etc. The spherical harmonics are related to Legendre functions $P_l^m(\cos \theta)$ (for $m \geq 0$) by

$$ Y_l^m(\theta, \phi) = (-1)^m \left[ \frac{(2l + 1)(l - m)}{4\pi(l + m)} \right]^{1/2} P_l^m(\cos \theta) e^{im\phi} $$  \hspace{1cm} (E-16)

and thus

$$ D^l_{m0}(\theta, 0, 0) = d^l_{m0}(\theta) = (-1)^m \left[ \frac{(l - m)}{l + m} \right]^{1/2} P_l^m(\cos \theta) = \left( \frac{4\pi}{2l + 1} \right)^{1/2} Y_l^m(\theta, 0) . $$  \hspace{1cm} (E-17)

In particular, $d^l_{00}(\theta) = P_l(\cos \theta)$. In general, $d^l_{m'm}(\theta)$ is related to the Jacobi polynomial

$$ P_j^{(\alpha, \beta)}(u) = \frac{(-1)^l}{2^l l!} (1-u)^{-\alpha} (1+u)^{-\beta} \frac{d^l}{du^l} [(1-u)^{\alpha+l} (1+u)^{\beta+l}] $$  \hspace{1cm} (E-18)

by

$$ d^l_{m'm}(\theta) = \left[ \frac{(j + m')!(j - m')!}{(j + m)!(j - m)!} \right]^{1/2} \left( \cos \frac{\theta}{2} \right)^{m+m'} \left( \sin \frac{\theta}{2} \right)^{m-m'} P_j^{(m' - m, m + m')}(\cos \theta) . $$  \hspace{1cm} (E-19)

Jacobi and Legendre polynomials pertain to the general theory of orthogonal polynomials, for which we shows that they satisfy 3-term linear recursion relations, and also differential equations. For instance, Jacobi polynomials are orthogonal for the measure

$$ \frac{1}{\sin \theta} d\theta (1-u)^{\alpha}(1+u)^{\beta} \frac{d}{du} P_j^{(\alpha, \beta)}(u) = \delta_{jj'} \left[ \frac{2^{\alpha+\beta+1} \Gamma(l+\alpha+1) \Gamma(l+\beta+1)}{(2l+\alpha+\beta+1)! \Gamma(l+\alpha+\beta+1)} \right] P_{j'}^{(\alpha, \beta)}(u) $$  \hspace{1cm} (E-20)

and satisfy the recursion relation

$$ 2(l+1)(l+\alpha+\beta+1)(2l+\alpha+\beta)P_{l+1}^{(\alpha, \beta)}(u) = (2l+\alpha+\beta+1)(2l+\alpha+\beta+2)u + \alpha^2 - \beta^2 P_l^{(\alpha, \beta)}(u) - 2(l+\alpha)(l+\beta)(2l+\alpha+\beta+2)P_{l-1}^{(\alpha, \beta)} . $$  \hspace{1cm} (E-21)

The Jacobi polynomial $P_l^{(\alpha, \beta)}(u)$ is a solution of the differential equation

$$ \{ (1-u^2) \frac{d^2}{du^2} + [\beta - \alpha - (2 + \alpha + \beta)u] \frac{d}{du} + l(l+\alpha+\beta+1) \} P_l^{(\alpha, \beta)}(u) = 0 . $$

The Legendre polynomials correspond to the case $\alpha = \beta = 0$. These relations appear here as related to those of the $D^j$. This is a general feature: many “special functions” (Bessel, etc) are related to representation matrices of groups. Group theory thus gives a geometric perspective to results of classical analysis.

Return to spherical harmonics and their properties.

(i) They satisfy the differential equations

$$ (\Delta_{S^2} + l(l+1)) Y_l^m = 0 $$  \hspace{1cm} (E-23)

$$ J_l Y_l^m = -i \frac{\partial}{\partial \phi} Y_l^m = m Y_l^m $$  \hspace{1cm} (E-24)

and may be written as

$$ Y_l^m(\theta, \phi) = \frac{(-1)^l}{2^l l!} \sqrt{\frac{(2l + 1)(l + m)!}{4\pi(l - m)!}} e^{im\phi} \sin^{-m} \theta \left( \frac{d}{d \cos \theta} \right)^{l-m} \sin^2 \theta . $$  \hspace{1cm} (E-25)
(ii) They are normalized to 1 on the unit sphere and more generally satisfy orthogonality and completeness properties

\[
\int d\Omega Y_l^m \bar{Y}_{l'}^{m'} = \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta Y_l^m \bar{Y}_{l'}^{m'} = \delta_{ll'}\delta_{mm'} \quad (E-26)
\]

\[
\sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_l^m(\theta, \phi)Y_l^m(\theta', \phi') = \delta(\Omega - \Omega') = \frac{\delta(\theta - \theta')\delta(\phi - \phi')}{\sin \theta} \quad (E-27)
\]

\[
= \delta(\cos \theta - \cos \theta')\delta(\phi - \phi') \quad (E-28)
\]

(iii) One may consider \(Y_l^m(\theta, \phi)\) as a function of the unit vector \(\hat{n}\) with polar angles \(\theta, \phi\). If the vector \(\hat{n}\) is transformed into \(\hat{n}'\) by the rotation \(R\), one has

\[
Y_l^m(\hat{n}') = Y_l^m(\hat{n})D^l(R)_{m'm} \quad (E-29)
\]

which expresses that the \(Y_l^m\) transform as vectors of the spin \(l\) representation.

(iv) One checks on the above expression the symmetry properties in \(m\)

\[
Y_l^m(\theta, \phi) = (-1)^m Y_l^{-m}(\theta, \phi) \quad (E-30)
\]

and parity

\[
Y_l^m(\pi - \theta, \phi + \pi) = (-1)^l Y_l^m(\theta, \phi) . \quad (E-31)
\]

Note that for \(\theta = 0\), \(Y_l^m(0, \phi)\) vanishes except for \(m = 0\), see (E-13, E-16).

(v) Spherical harmonics satisfy also recursion formulae of two types: those coming from the action of \(J_\pm\), differential operators acting as in (00.3.10)

\[
e^{\pm i\phi} \left[ \pm \frac{\partial}{\partial \theta} + icotg \theta \frac{\partial}{\partial \phi} \right] Y_l^m = \sqrt{l(l+1) - m(m \pm 1)} Y_l^{m \pm 1} \quad (E-32)
\]

and those coming from the tensor product with the vector representation

\[
\sqrt{2l + 1} \cos \theta Y_l^m = \left( \frac{2l+2}{2l-1} \right)^{\frac{1}{2}} Y_{l-1}^m + \left( \frac{2l+2}{2l+3} \right)^{\frac{1}{2}} Y_{l+1}^m . \quad (E-33)
\]

More generally, one has a product formula

\[
Y_l^m(\theta, \phi)Y_{l'}^{m'}(\theta, \phi) = \sum_L \langle lm; l'm' | L, m + m' \rangle \left[ \frac{(2l + 2)(2l' + 2)}{4\pi(2L + 1)} \right]^{\frac{1}{2}} Y_{L+m+m'}^m(\theta, \phi) . \quad (E-34)
\]

(vi) Finally let us quote the very useful “addition theorem”

\[
\frac{2l + 1}{4\pi} P_l(\cos \theta) = \sum_{m=-l}^{l} Y_l^m(\hat{n})Y_l^{m*}(\hat{n}') \quad (E-35)
\]

where \(\theta\) denotes the angle between directions \(\hat{n}\) and \(\hat{n}'\). This may be proved by showing that the rhs satisfies the same differential equation as the \(P_l\) (see exercise 1 below).

Exercises.

1. Prove that the Legendre polynomial \(P_l\) verifies

\[
(\Delta_{S^2} + l(l + 1)) P_l(\hat{n}, \hat{n}') = 0
\]
2.4. PROJECTIVE REPRESENTATIONS. WIGNER THEOREM.

as a function of \( \hat{n} \) or of \( \hat{n}' \), as well as \((\mathbf{J} + \mathbf{J}') P_l = 0\) where \( \mathbf{J} \) and \( \mathbf{J}' \) are generators of rotations of \( \hat{n} \) and \( \hat{n}' \) respectively. Conclude that there exists an expansion on spherical harmonics given by the addition theorem of (E-35) (Remember that \( P_l(1) = 1 \)).

2. Prove that a generating function of Legendre polynomials is

\[
\frac{1}{\sqrt{1 - 2ut + t^2}} = \sum_{l=0}^{\infty} t^l P_l(u). \tag{E-36}
\]

Hint: show that the differential equation of the \( P_l \) (a particular case of (E-22) for \( \alpha = \beta = 0 \)) is indeed satisfied and that the \( P_l \) appearing in that formula are polynomials in \( u \). Derive from it the identity (assuming \( r' < r \),

\[
\frac{1}{|\vec{r} - \vec{r}'|} = \sum_{l=0}^{\infty} \frac{r'^l}{r^{l+1}} P_l(\cos \theta) = \sum_{l,m} \frac{4\pi}{2l + 1} \frac{r'^l}{r^{l+1}} Y^m_l(\hat{n}) Y^m_l(\hat{n}'). \tag{E-37}
\]

The expression of the first \( Y^m_l \) may be useful

\[
Y^0_0 = 1, \quad Y^0_1 = \sqrt{\frac{3}{4\pi}} \cos \theta, \quad Y^{\pm 1}_1 = \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi}, \tag{E-38}
\]

\[
Y^0_2 = \sqrt{\frac{5}{16\pi}} (3\cos^2 \theta - 1), \quad Y^{\pm 1}_2 = \mp \sqrt{\frac{15}{8\pi}} \cos \theta \sin \theta e^{\pm i\phi}, \quad Y^{\pm 2}_2 = \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi}.
\]

2.4.8 Physical applications

1. Multipole moments

Consider the electric potential created by a static charge distribution \( \rho(\vec{r}) \)

\[
\phi(\vec{r}) = \frac{1}{4\pi \epsilon_0} \int \frac{d^3r' \rho(\vec{r}')}{|\vec{r} - \vec{r}'|}
\]

and expand it on spherical harmonics following (E-37). One finds

\[
\phi(\vec{r}) = \frac{1}{\epsilon_0} \sum_{l,m} \frac{1}{2l + 1} \frac{Y^m_l(\hat{n})}{r^{l+1}} Q_{lm} \tag{E-39}
\]

where the \( Q_{lm} \), defined by

\[
Q_{lm} = \int d^3r' \rho(\vec{r}') r'^l Y^m_l(\hat{n}') \tag{E-40}
\]

are the multipole moments of the charge distribution \( \rho \). For example, if \( \rho(\vec{r}) = \rho(r) \) is invariant by rotation, only \( Q_{00} \) is non vanishing and is equal to the total charge (up to a factor \( 1/\sqrt{4\pi} \))

\[
Q_{00} = \frac{Q}{\sqrt{4\pi}} = \sqrt{4\pi} \int r^2 d\phi \rho(r) \quad \phi(r) = \frac{Q}{4\pi \epsilon_0 r}.
\]

For an arbitrary \( \rho(\vec{r}) \), the three components of \( Q_{lm} \) reconstruct the dipole moment \( \int d^3r' \rho(\vec{r}') \vec{r}' \). More generally, under rotations, the \( Q_{lm} \) are the components of a tensor operator transforming according to the spin \( l \) representation and (see. (E-31), of parity \((-1)^l\)).
In Quantum Mechanics, the $Q_{lm}$ become operators in the Hilbert space of the theory. One may apply the Wigner-Eckart theorem and conclude that

$$\langle j_1, m_1 | Q_{lm} | j_2, m_2 \rangle = \langle j_1 | Q_l | j_2 \rangle \langle j_1, m_1 | l; j_2, m_2 \rangle$$

with a reduced matrix element which is independent of the $m$. In particular, if $j_1 = j_2 = j$, the expectation value of $Q_l$ is non-zero only for $l \leq 2j$.

2. Eigenstates of the angular momentum in Quantum Mechanics

Spherical harmonics may be interpreted as wave functions in coordinates $\theta, \phi$ of the eigenstates of the angular momentum $\vec{L} = \hbar \vec{J} = \hbar \vec{r} \wedge \vec{\nabla}$

$$Y_l^m(\theta, \phi) = \langle \theta, \phi|l, m \rangle$$

in analogy with

$$\frac{1}{(2\pi)^{3/2}} e^{i\vec{x} \cdot \vec{p}} = \langle \vec{x}|\vec{p} \rangle.$$ 

(We take $\hbar = 1$.) In particular, suppose that in a scattering process described by a rotation invariant Hamiltonian, a state of initial momentum $\vec{p}_i$ along the $z$-axis, (i.e. $\theta = \phi = 0$), interacts with a scattering center and comes out in a state of momentum $\vec{p}_f$, with $|p_i| = |p_f| = p$, along the direction $\hat{n} = (\theta, \phi)$. One writes the scattering amplitude

$$\langle p, \theta, \phi | T | p, 0, 0 \rangle = \sum_{ll'mm'} Y_{l'}^{m'}(\theta, \phi) \langle p, l, m | T | p, l', m' \rangle Y_{l}^{m*}(0, 0)$$

$$= \sum_{lm} Y_l^m(\theta, \phi) \langle p, l, m | T | p, l, m \rangle Y_l^{m*}(0, 0)$$

$$= \sum_{l} \frac{2l + 1}{4\pi} T_l(p) P_l(\cos \theta)$$

(E-41)

using once again the addition formula and $\langle plm | T | pl'm' \rangle = \delta_{ll'}\delta_{mm'} T_l(p)$ expressing rotation invariance. This is the very useful partial wave expansion of the scattering amplitude.

*
Exercises for chapter 2

A. Unitary representations of a simple group

Let $G$ be a simple non abelian group, and $D$ be unitary representation of $G$.

1. Show that $\det D$ is a representation of dimension 1 of the group, and a homomorphism of the group into the group $U(1)$.

2. What can be said about the kernel $K$ of this homomorphism? Show that any “commutator” $g_1g_2g_1^{-1}g_2^{-1}$ belongs to $K$ and thus that $K$ cannot be trivial.

3. Conclude that the representation is unimodular (of determinant 1).

4. Can we apply that argument to $SO(3)$? to $SU(2)$?

B. Adjoint representation

1. Show that if the Lie algebra $\mathfrak{g}$ of a Lie group $G$ is simple, the adjoint representation of $G$ is irreducible.

2. Show that if $\mathfrak{g}$ is semi-simple, its adjoint representation is faithful: $\ker \text{ad} = 0$.

C. Tensor product $D \otimes D^*$

Let $G$ be a compact group and $D^{(\rho)}$ its irreducible representations. Denote $D^{(1)}$ the identity representation, $D^{(\bar{\rho})}$ the conjugate representation of $D^{(\rho)}$.

What is the multiplicity of $D^{(1)}$ in the decomposition of $D^{(\rho)} \otimes D^{(\bar{\rho})}$ into irreducible representations?

D. Chebyshev polynomials

Consider the expression

$$U_l = \frac{\sin(l + 1)\theta}{\sin\theta},$$

where $l$ is an integer $\geq 0$.

1. By an elementary trigonometric calculation, express $U_{l-1} + U_{l+1}$ in terms of $U_l$, with an $l$ independent coefficient.

2. Conclude that $U_l$ is a polynomial in $z = 2\cos\theta$ of degree $l$, which we denote $U_l(z)$.

3. What is the group theoretic interpretation of the result in 1.?

4. With the minimum of additional computations, what can be said about

$$\frac{2}{\pi} \int_{-1}^{1} dz (1 - z^2)^{\frac{1}{2}} U_l(z)U_{\nu}(z)$$

and

$$\frac{2}{\pi} \int_{-1}^{1} dz (1 - z^2)^{\frac{1}{2}} U_l(z)U_{\nu}(z)U_{\nu}(z)?$$

The $U_l(z)$ are the Chebyshev polynomials (Tchebichev in the French transcription) of 2nd kind. They are orthogonal (the first relation in 4.) and satisfy a 3-term recursion relation (question 1.), which are two general properties of orthogonal polynomials.
E. Spherical Harmonics

Show that the integral
\[
\int d\Omega Y_{l_1}^{m_1}(\theta, \phi)Y_{l_2}^{m_2}(\theta, \phi)Y_{l_3}^{m_3}(\theta, \phi)
\]
is proportional to the Clebsch-Gordan coefficient \((-1)^m \langle l_1, m_1; l_2, m_2|l_3, -m_3 \rangle\), with an \(m\) independent factor to be determined.

Problem I. Decomposition of an amplitude

Consider two real unitary representations \((\rho)\) and \((\sigma)\) of a simple compact Lie group \(G\) of dimension \(d\). Denote \(|\rho, \alpha\rangle\), resp. \(|\sigma, \beta\rangle\), two bases of these representations, and \(T^{(\rho)a}_{\alpha\alpha'}\), resp. \(T^{(\sigma)a}_{\beta\beta'}\), \(a = 1, \cdots d\), the representation matrices in a basis of the Lie algebra. Explain why this basis may be assumed to be orthonormal wrt the Killing form. These matrices are taken to be real skew-symmetric and thus satisfy \(\text{tr} T^a T^b = -\delta_{ab}\). Consider now the quantity
\[
X_{\alpha\beta;\alpha'\beta'} := \sum_{a=1}^d T^{(\rho)a}_{\alpha\alpha'} T^{(\sigma)a}_{\beta\beta'}.
\]  
(2-114)

To simplify things, we assume that all irreducible representations appearing in the tensor product of representations \((\rho)\) et \((\sigma)\) are real and with multiplicity 1. Let \(|\tau\gamma\rangle\) be a basis of such a representation. The (real) Clebsch-Gordan coefficients are written as matrices
\[
\left(M^{(\tau\gamma)}\right)_{\alpha\beta} = \langle \tau\gamma|\rho\alpha;\sigma\beta \rangle.
\]  
(2-115)

1. Recall why these coefficients satisfy orthogonality and completeness relations and write them.

2. Show that it follows that
\[
X_{\alpha\beta;\alpha'\beta'} = -\sum_{\tau\gamma} \left(M^{(\tau\gamma)}\right)_{\alpha\beta} (T^{(\rho)a}_{\alpha\alpha'}) (T^{(\sigma)a}_{\beta\beta'}).
\]  
(2-116)

3. Acting with the infinitesimal generator \(T^a\) on the two sides of the relation
\[
|\rho\alpha;\sigma\beta \rangle = \sum_{\tau\gamma} \left(M^{(\tau\gamma)}\right)_{\alpha\beta} |\tau\gamma\rangle
\]  
(2-117)

show that one gets
\[
\sum_{\gamma'} T^{(\tau)a}_{\gamma'\gamma} \left(M^{(\tau\gamma')}\right)_{\alpha\beta} = \sum_{\alpha'} \left(M^{(\tau\gamma')}\right)_{\alpha'\beta} (T^{(\rho)a}_{\alpha'\alpha}) + \sum_{m_1^2} \left(M^{(\tau\gamma')}\right)_{\alpha\beta'} (T^{(\sigma)a}_{\beta'\beta})
\]  
(2-118)

or, in terms of matrices of dimensions \(\dim(\rho) \times \dim(\sigma)\)
\[
\sum_{\gamma'} T^{(\tau)a}_{\gamma'\gamma} M^{(\tau\gamma')} = -T^{(\rho)a} M^{(\tau\gamma)} + M^{(\tau\gamma)} T^{(\sigma)a}.
\]  
(2-119)
4. Using repeatedly this relation (2-119) in (2-116), show that one finds

\[ X_{\alpha\beta;\alpha'\beta'} = \frac{1}{2} \sum_{\gamma} (C_{\rho} + C_{\sigma} - C_{\tau}) (M^{(\tau\gamma)})_{\alpha\beta} (M^{(\tau\gamma)})_{\alpha'\beta'} \]  

(2-120)

where the \( C \) are Casimir operators, for example

\[ C_{\rho} = -\sum_{a} (T^{(\rho)a})^{2} . \]  

(2-121)

5. Why can one say that "large representations" \( \tau \) tend to make the coefficient \( (C_{\rho}+C_{\sigma}-C_{\tau}) \) increasingly negative? (One may take the example of SU(2) with \( \rho \) and \( \sigma \) two spin \( j \) \( (j \in \mathbb{N}) \) representations.

6. Can you propose a field theory in which the coefficient \( X_{\alpha\beta;\alpha'\beta'} \) would appear in a two-body scattering amplitude (in the tree approximation)? What is the consequence of the property derived in 5) on that amplitude?

**Problem II. Tensor product in SU(2)**

1. Let \( R_{\frac{1}{2}} \) denote the spin \( \frac{1}{2} \) representation of SU(2); we want to compute the multiplicity \( n_{r} \) of the identity representation in the decomposition into irreducible representations of the tensor product of \( r \) copies of \( R_{\frac{1}{2}} \).

   (a) Interpret \( n_{r} \) in terms of the number of linearly independent invariants, multilinear in \( \xi_{1}, \cdots, \xi_{r} \), where the \( \xi_{i} \) are spinors transforming under the representation \( R_{\frac{1}{2}} \).

   (b) By convention \( n_{0} = 1 \). With no calculation, what are \( n_{1} \) and \( n_{2} \)?

   (c) Show that \( n_{r} \) may be expressed with an integral involving characters \( \chi_{j}(\psi) \) of SU(2). (Do not attempt to compute this integral explicitly for arbitrary \( r \).)

   (d) Check that this formula gives the values of \( n_{1} \) et \( n_{2} \) found in b).

   (e) We shall now show that the \( n_{r} \) may also be obtained by the following graphical and recursive method. On the graph of 2.2, attach \( n_{0} = 1 \) to the leftmost vertex, then to each vertex \( S \), attach the sum \( \alpha = \beta + \gamma \) of numbers on vertices immediately on the left of \( S \).

   i. Show that the \( n_{r} \) are the numbers located on the horizontal axis. What is the interpretation of the horizontal and vertical axes?

   ii. Compute with this method the value of \( n_{4} \) and \( n_{6} \).

2. One wants to repeat this computation for the spin 1 representation \( R_{1} \), and hence to determine the number \( N_{r} \) of times where the identity representation appears in the tensor product of \( r \) copies of \( R_{1} \).
(a) How should the graph of fig 2.2 be modified to yield the $N_r$?

(b) Compute $N_2$, $N_3$ et $N_4$ by this method.

(c) What do these numbers represent in terms of vectors $V_1, \cdots, V_r$ transforming under the representation $R_1$?

Problem III. Real, complex and quaternionic representations

Preliminary question

Given a vector space $E$ of dimension $d$, one denotes $E \otimes E$ or $E^{\otimes 2}$ the space of rank 2 tensors and $(E \otimes E)_S$, resp. $(E \otimes E)_A$, the space of symmetric, resp. antisymmetric, rank 2 tensors, also called (anti)symmetrized tensor product. What is the dimension of spaces $E \otimes E$, $(E \otimes E)_S$, $(E \otimes E)_A$?

A. Real and quaternionic representations

1. Consider a compact group $G$. If $D(g)$ is a representation of $G$, show that $D^{-1T}(g)$ is also a representation, called the contragredient representation.

2. Recall briefly why one may assume with no loss of generality that the representations of $G$ are unitary, which we assume in the following.

   Show that the contragredient representation is then identical to the complex conjugate one.

3. Suppose that the unitary representation $D$ is (unitarily) equivalent to its contragredient (or conjugate) representation. Show that there exists a unitary matrix $S$ such that

   \[ D = SD^{-1T}S^{-1} \]  

   \[ (2-122) \]

4. Show that (2-122) implies that the bilinear form $S$ is invariant.

   Is this form degenerate?
5. Using (2-122) show that
\[ DSS^{-1T} = SS^{-1T}D. \] (2-123)

6. Show that if \( D \) is irreducible, \( S = \lambda S^T \), with \( \lambda^2 = 1 \).

7. Conclude that the invariant form \( S \) is either symmetric or antisymmetric.
   
   In the former case (\( S \) symmetric), the representation is called \textit{real}, in the latter (\( S \) antisymmetric), it is called \textit{pseudoreal} (or \textit{quaternionic}). One may prove that in the former case, there exists a basis on \( \mathbb{R} \) in which the representation matrices are real, and such a basis does not exist in the latter case.

8. Do you know an example of the second case?

B. Frobenius–Schur indicator

1. Let \( G \) be a finite or compact Lie group. Its irreducible representations are labelled by an index \( \rho \) and one denotes \( \chi^{(\rho)}(g) \) their character. Let \( \chi(g) \) be the character of some arbitrary representation, reducible or not.

   (a) For any function \( F \) on the finite \( G \), one denotes \( \langle F \rangle \) its group average
   \[ \langle F \rangle = \frac{1}{|G|} \sum_{g \in G} F(g). \] (2-124)

   How to extend that definition to the case of a compact Lie group (and a continuous function \( F \))?

   (b) - Recall why \( \langle \chi \rangle \) is an integer and what it means.
       - If \( \bar{\rho} \) denotes the conjugate representation of the irreducible representation \( \rho \), recall why \( \langle \chi^{(\rho)} \chi^{(\bar{\rho})} \rangle = 1 \) and what it implies on the decomposition of \( \rho \otimes \bar{\rho} \) into irreducible representations.

   (c) Show that an irreducible representation \( \rho \) is equivalent to \( \bar{\rho} \) iff
   \[ \langle (\chi^{(\rho})(g))^2 \rangle = 1. \]

   Evaluate this expression if \( \rho \) is not equivalent to \( \bar{\rho} \).

2. We now consider a representation \( D^{(\rho)} \) acting in a space \( E \), and its tensor square \( D^{(\rho) \otimes 2} \equiv D^{(\rho)} \otimes D^{(\rho)} \), which acts on rank 2 tensors of \( E \otimes E \).

   (a) Write explicitly the action of \( D^{(\rho) \otimes 2} \) on a tensor \( t = \{ t^{ij} \} \),
   \[ t^{ij} \mapsto t^{ij} = \ldots \]
(b) Show that any rank 2 tensor, \( t = \{ t^{ij} \} \), is the sum of a symmetric tensor \( t_S \) and of an antisymmetric one \( t_A \), transforming under independent representations. Write explicitly the transformation matrices, paying due care to the symmetry properties of the tensors under consideration.

(c) Show that the characters of the representations of symmetric and antisymmetric tensors are respectively

\[
\chi^{(\rho \otimes \rho)_S}(g) = \frac{1}{2} \left( (\chi^{(\rho)}(g))^2 \pm \chi^{(\rho)}(g^2) \right) .
\] (2-125)

(d) What is the value of these characters for \( g = e \), the identity in the group? Could this result have been anticipated?

3. One then defines the Frobenius–Schur indicator of the irreducible representation \( \rho \) by

\[
\text{ind}(\rho) = \langle \chi^{(\rho)}(g^2) \rangle .
\] (2-126)

(a) Using the results of 2., show that one may write

\[
\text{ind}(\rho) = \langle \chi^{(\rho \otimes \rho)_S} \rangle - \langle \chi^{(\rho \otimes \rho)_A} \rangle .
\]

(b) Using the results of 1., show that

\[
\langle (\chi^{(\rho)}(g))^2 \rangle = \langle \chi^{(\rho \otimes \rho)_S} \rangle + \langle \chi^{(\rho \otimes \rho)_A} \rangle
\]

takes the value 0 or 1, depending on the case: discuss.

(c) - Show that \( \langle \chi^{(\rho \otimes \rho)_S} \rangle \) and \( \langle \chi^{(\rho \otimes \rho)_A} \rangle \) are non negative integers and give a certain multiplicity to be discussed.

- Finally show that the Frobenius–Schur indicator of (2-126) can take only the three values 0 et \( \pm 1 \) according to cases to be discussed.

(d) What is the relation between this discussion and that of part A?

4. \( \star \) We now restrict to the case of a finite group \( G \). For any \( h \in G \), we define \( Q(h) := \sum_{\rho} \text{ind}(\rho) \chi^{(\rho)}(h) \). Prove the

**Theorem** \( Q(h) = \# \{ g \in G | g^2 = h \} \)
Chapter 3

Simple Lie algebras. Classification and representations. Roots and weights

3.1 Cartan subalgebra. Roots. Canonical form of the algebra

We consider a semi-simple (i.e. with no abelian ideal) Lie algebra. We want to construct a canonical form of commutation relations modeled on the case of SU(2)

\[ [J_z, J_\pm] = \pm J_\pm \quad [J_+, J_-] = 2J_z. \]  

It will be important to consider the algebra over \( \mathbb{C} \), at the price of “complexifying” it if it was originally real. The adjoint representation will be used. As it is a faithful representation for a semi-simple algebra, (i.e. \( \text{ad} \ X = 0 \Rightarrow X = 0 \), see exercise B of Chap. 2), no information is lost.

It may also be useful to remember that the complex algebra has a real compact version, in which the real structure constants lead to a negative definite Killing form, and, as the representations can be taken unitary, the elements of the Lie algebra (the infinitesimal generators) may be taken as Hermitian (or antiHermitian, depending on our conventions).

3.1.1 Cartan subalgebra

We define first the notion of Cartan subalgebra. This is a maximal abelian subalgebra of \( \mathfrak{g} \) such that all its elements are diagonalisable (hence simultaneously diagonalisable) in the adjoint representation. That such an algebra exists is non trivial and must be established, but we shall admit it.

If we choose to work with the unitary form of the adjoint representation, the elements of \( \mathfrak{g} \) are Hermitian matrices, and assuming that the elements of \( \mathfrak{h} \) are commuting among themselves ensures that they are simultaneously diagonalizable.

This Cartan subalgebra is non unique, but one may prove that two distinct choices are related by an automorphism of the algebra \( \mathfrak{g} \).

For instance if \( g \) is the Lie algebra of a Lie group \( G \) and if \( h \) is a Cartan subalgebra of \( g \), any conjugate \( ghg^{-1} \) of \( h \) by an arbitrary element of \( G \) is another Cartan subalgebra.

Let \( h \) be a Cartan subalgebra, call \( \ell \) its dimension, it is independent of the choice of \( h \) and it is called the rank of \( g \). For \( su(2) \), this rank is 1, (the choice of \( J_z \) for example); for \( su(n) \), the rank is \( n - 1 \). Indeed for \( su(n) \), a Cartan algebra is generated \(^1\) by diagonal traceless matrices, a basis of which is given by the \( n - 1 \) matrices

\[
H_1 = \text{diag}(1, -1, 0, \cdots, 0), \quad H_2 = \text{diag}(0, 1, -1, 0, \cdots, 0), \cdots, \quad H_{n-1} = \text{diag}(0, \cdots, 0, 1, -1).
\]

An arbitrary matrix of the Lie algebra, (in that representation), (anti-)Hermitian and trace-less, is diagonalisable by a unitary transformation; its diagonal form is traceless and is thus expressed as a linear combination of the \( h_j \); the original matrix is thus conjugate by a unitary transformation of a linear combination of the \( h_j \). This is a general property, and one proves (Cartan, see [Bu], chap. 16) that

*If \( g \) is the Lie algebra of a group \( G \), any element of \( g \) is conjugate by \( G \) of an element of \( h \).*

Application. Canonical form of antisymmetric matrices. Using the previous statement, prove the

**Proposition** If \( A = A^* = -A^T \) is a real skew-symmetric matrix of dimension \( N \), one may find a real orthogonal matrix \( O \) such that \( A = ODO^T \) where \( D = \text{diag}(\begin{pmatrix} 0 & \mu_j \\ -\mu_j & 0 \end{pmatrix} \quad j = 1, \ldots, n) \) if \( N = 2n \) and \( D = \text{diag}(0, \begin{pmatrix} 0 & \mu_j \\ -\mu_j & 0 \end{pmatrix} \quad j = 1, \ldots, n-1) \) if \( N = 2n + 1 \), with real \( \mu_j \).

If one allows the complexification of orthogonal matrices, one may fully diagonalise the matrix \( A \) in the form \( D = \text{diag}(\begin{pmatrix} i\mu_j & 0 \\ 0 & -i\mu_j \end{pmatrix} \quad j = 1, \ldots, n) \) or \( D = \text{diag}(0, \begin{pmatrix} i\mu_j & 0 \\ 0 & -i\mu_j \end{pmatrix} \quad j = 1, \ldots, n) \). For a proof making only use of matrix theory, see for example [M.L. Mehta, Elements of Matrix Theory, p 41].

3.1.2 Canonical basis of the Lie algebra

Let \( H_i, i = 1, \cdots, \ell \) be a basis of \( h \). It is convenient to choose the \( H_i \) Hermitian. By definition \([H_i, H_j] = 0\), (abelian subalgebra) or more precisely, since we are in the adjoint representation,

\[
[\text{ad} H_i, \text{ad} H_j] = 0.
\]

We may thus diagonalise simultaneously these \( \text{ad} H_i \). We already know (some?) eigenvectors of vanishing eigenvalue since \( \forall i, j, \quad \text{ad} H_i H_j = 0 \), and we may complete them to make a basis by finding a set of eigenvectors \( E_\alpha \) linearly independent of the \( H_j \)

\[
\text{ad} H_i E_\alpha = \alpha(i) E_\alpha
\]

\[i.e. \quad \text{a set of elements of } g \text{ such that}
\[
[H_i; E_\alpha] = \alpha(i) E_\alpha
\]

\(^1\)We use momentarily the “representation of definition” (made of \( n \times n \) matrices) rather than the adjoint representation.
with the $\alpha_{(i)}$ not all vanishing (otherwise the subalgebra $\mathfrak{h}$ would not be maximal).

The space $\mathfrak{h}^*$. In these expressions, the $\alpha_{(i)}$ are eigenvalues of the operators $\text{ad} H_i$. Since we chose Hermitian $\text{ad} H_i$, their eigenvalues $\alpha_{(i)}$ are real. By linearity, for an arbitrary element of $\mathfrak{h}$ written as $H = \sum_i h^i H_i$,

$$\text{ad} (H) E_\alpha = \alpha(H) E_\alpha ,$$  \hfill (3-6)$$

and the eigenvalue of $\text{ad} (H)$ on $E_\alpha$ is $\alpha(H) := \sum_i h^i \alpha_{(i)}$, which is a linear form on $\mathfrak{h}$. In general linear forms on a vector space $E$ form a vector space $E^*$, called the dual space of $E$. One may thus consider the root $\alpha$, of components $\alpha_{(i)}$, as a vector of the dual space of $\mathfrak{h}$, hence $\alpha \in \mathfrak{h}^*$, the root space. Note that $\alpha(H_i) = \alpha_{(i)}$.

Roots enjoy the following properties ($*$)

1. if $\alpha$ is a root, $-\alpha$ in another root;

2. the eigenspace of the eigenvalue $\alpha$ is of dimension 1 (no multiplicity);

3. if $\alpha$ is a root, the only roots of the form $\lambda \alpha$ are $\pm \alpha$;

4. roots $\alpha$ generate all the dual space $\mathfrak{h}^*$.

For proofs of 1., 2., 3., see below, for 4. see exercise A.

Number of roots. Since the $H_i$ are diagonalisable, the total number of their eigenvectors $E_\alpha$ and $H_i$ must be equal to the dimension of the space, here the dimension $d$ of the adjoint representation, i.e. of the Lie algebra $\mathfrak{g}$. As any (non vanishing by definition) root comes along with its opposite, the number of roots $\alpha$ is even and equal to $d - \ell$ (with $\ell = \text{rank}(\mathfrak{g})$). We denote $\Delta$ the set of roots.

In the basis $\{H_i, E_\alpha\}$ of $\mathfrak{g}$, the Killing form takes a simple form

$$(H_i, E_\alpha) = 0 \quad (E_\alpha, E_\beta) = 0 \quad \text{unless} \quad \alpha + \beta = 0 .$$ \hfill (3-7)$$

To show that, we write $(H, [H', E_\alpha]) = \alpha(H')(H, E_\alpha)$, and also, using the definition of the Killing form and the cyclicity of the trace

$$(H, [H', E_\alpha]) = \text{tr} (\text{ad} H [\text{ad} H', \text{ad} E_\alpha]) = \text{tr} ([\text{ad} H, \text{ad} H'] \text{ad} E_\alpha) = 0 \quad (3-8)$$

since $[\text{ad} H, \text{ad} H'] = 0$. It follows that $\forall H, H' \in \mathfrak{h}$, $\alpha(H')(H, E_\alpha) = 0$, hence that $(H, E_\alpha) = 0$. Likewise

$$(E_\alpha, E_\beta) = \alpha(H)(E_\alpha, E_\beta) = -(E_\alpha, [H, E_\beta]) = -\beta(H)(E_\alpha, E_\beta) \quad (3-9)$$

again by the cyclicity of the trace, and thus $(E_\alpha, E_\beta) = 0$ if $\exists H : (\alpha + \beta)(H) \neq 0$, i.e. if $\alpha + \beta \neq 0$. Note that the point 1. in ($*$) above follows simply from (3-7): if $-\alpha$ were not a root, $E_\alpha$ would be orthogonal to all elements of the basis hence to any element of $\mathfrak{g}$, and the form would be degenerate, contrary to the hypothesis of semi-simplicity (and Cartan’s criterion). For an elegant proof of points 2. et 3. of ($*$), see [OR, p. 29].

The restriction of this form to the Cartan subalgebra is non-degenerate, since otherwise one would have $\exists H \in \mathfrak{h}, \forall H' \in \mathfrak{h} : (H, H') = 0$, but $(H, E_\alpha) = 0$, thus $\forall X \in \mathfrak{g}$, $(H, X) = 0$ and the form would be degenerate, contrary to the hypothesis of semi-simplicity (and Cartan’s criterion, Chap. 1, §4.4). The Killing form being non-degenerate on $\mathfrak{h}$, it induces an isomorphism between $\mathfrak{h}$ and $\mathfrak{h}^*$: to $\alpha \in \mathfrak{h}^*$ one associates the unique $H_\alpha \in \mathfrak{h}$ such that

$$\forall H \in \mathfrak{h} \quad \alpha(H) = (H_\alpha, H) .$$ \hfill (3-10)$$
(Or said differently, one solves the linear system $g_{ij} h_i^j = \alpha_{(i)}$ which is of Cramer type since $g_{ij} = (H_i, H_j)$ is invertible.) One has also a bilinear form on $\mathfrak{h}^*$ inherited from the Killing form

$$\langle \alpha, \beta \rangle = (H_\alpha, H_\beta),$$

which we are going to use in § 2 to study the geometry of the root system.

It remains to find the commutation relations of the $E_\alpha$ among themselves. Using the Jacobi identity, one finds that

$$\text{ad}_{H_i}[E_\alpha, E_\beta] = [H_i, [E_\alpha, E_\beta]] = [E_\alpha, [H_i, E_\beta]] - [E_\beta, [H_i, E_\alpha]] = (\alpha + \beta)_{(i)}[E_\alpha, E_\beta].$$

Invoking the trivial multiplicity (=1) of roots, one sees that three cases may occur. If $\alpha + \beta$ is a root, $[E_\alpha, E_\beta]$ is proportional to $E_{\alpha+\beta}$, with a proportionality coefficient $N_{\alpha\beta}$ which will be shown below to be non zero (see § 2.1 and exercise B). If $\alpha + \beta \neq 0$ is not a root, $[E_\alpha, E_\beta]$ must vanish. Finally if $\alpha + \beta = 0$, $[E_\alpha, E_{-\alpha}]$ is an eigenvector of all $\text{ad}_{H_i}$ with a vanishing eigenvalue, thus $[E_\alpha, E_{-\alpha}] = H \in \mathfrak{h}$. To determine that $H$, let us proceed like in (3-9)

$$(H_i, [E_\alpha, E_{-\alpha}]) = \text{tr}(\text{ad}_{H_i}[\text{ad}_{E_\alpha}, \text{ad}_{E_{-\alpha}}]) = \text{tr}([\text{ad}_{H_i}, \text{ad}_{E_\alpha}] \text{ad}_{E_{-\alpha}})$$

$$= \alpha_{(i)}(E_\alpha, E_{-\alpha}) = (H_i, H_\alpha)(E_\alpha, E_{-\alpha})$$

hence

$$[E_\alpha, E_{-\alpha}] = (E_\alpha, E_{-\alpha})H_\alpha.$$  

(3-14)

To recapitulate, we have constructed a canonical basis of the algebra $\mathfrak{g}$

$$
\begin{align*}
[H_i, H_j] &= 0 \\
[H_i, E_\alpha] &= \alpha_{(i)}E_\alpha \\
[E_\alpha, E_\beta] &= \begin{cases} 
N_{\alpha\beta}E_{\alpha+\beta} & \text{if } \alpha + \beta \text{ is a root} \\
(E_\alpha, E_{-\alpha})H_\alpha & \text{if } \alpha + \beta = 0 \\
0 & \text{otherwise}
\end{cases}
\end{align*}
$$

(3-15)

Up to that point, the normalisation of the vectors $H_i$ and $E_\alpha$ has not been fixed. It is common to choose, in accord with (3-7)

$$
(H_i, H_j) = \delta_{ij} \quad (E_\alpha, E_\beta) = \delta_{\alpha+\beta,0}.
$$

(3-15)

(Indeed, the restriction of the Killing form to $\mathfrak{h}$, after multiplication by $i$ to make the $\text{ad}_{H_i}$ Hermitian, is positive definite.) With that normalisation, $H_\alpha$ defined above by (3-10) satisfies also

$$H_\alpha = \alpha.H := \alpha_{(i)}H_i.$$  

(3-16)
Note that $E_{\alpha}$, $E_{-\alpha}$ and $H_{\alpha}$ form an su(2) subalgebra

$$[H_{\alpha}, E_{\pm\alpha}] = \pm \langle \alpha, \alpha \rangle E_{\pm\alpha} \quad [E_{\alpha}, E_{-\alpha}] = H_{\alpha}.$$ (3-17)

(This is in fact $H_{\alpha}/\langle \alpha, \alpha \rangle$ that we identify with $J$, and that observation will be used soon.) Any semi-simple algebra thus contains an su(2) algebra associated to each of its roots.

Note that with the normalisations of (3-15), the Killing metric reads in the basis $\{H_{i}, E_{\alpha}, E_{-\alpha}\}$

$$g_{ab} =
\begin{pmatrix}
\mathbb{I}_{\ell} & 0 \\
0 & 1 & 0 \\
1 & 0 & 0 \\
0 & \cdots & 0 & 1 \\
1 & 0 & 0 \\
\end{pmatrix}
$$ (3-18)

where the first block is an identity matrix of dimension $\ell \times \ell$.

### 3.2 Geometry of root systems

#### 3.2.1 Scalar products of roots. The Cartan matrix

As noticed in (3-11), the space of roots, i.e. the space (of dimension $\ell$, see point 4. in (∗) above) generated by the $d - \ell$ roots $\alpha$ inherits the Euclidean metric of $\mathfrak{h}$

$$\langle \alpha, \beta \rangle := (H_{\alpha}, H_{\beta}) = \alpha(H_{\beta}) = \beta(H_{\alpha}) = (\alpha.H, \beta.H) = \sum_{i} \alpha(i)\beta(i) ,$$ (3-19)

where the various expressions aim at making the reader familiar with the notations introduced above. (Only the last two expressions depend on the choice of normalisation (3-15).) We shall now show that the geometry –lengths and angles– of roots is strongly constrained. First it is good to remember the lessons of the su(2) algebra: in a representation of finite dimension, $J_{z}$ has integer or half-integer eigenvalues. Thus here, where each $H_{\alpha}/\langle \alpha, \alpha \rangle$ plays the role of a $J_{z}$ and has $E_{\beta}$ as eigenvectors, ad $H_{\alpha}E_{\beta} = \langle \alpha, \beta \rangle E_{\beta}$, i.e.

$$[H_{\alpha}, E_{\beta}] = \langle \alpha, \beta \rangle E_{\beta}$$ (3-20)

we may conclude that

$$2\frac{\langle \alpha, \beta \rangle}{\langle \alpha, \alpha \rangle} = m \in \mathbb{Z} .$$ (3-21)

#### Root chains

It is in fact useful to refine the previous discussion. Like in the case of su(2), the idea is to repeatedly apply the “raising” $E_{\alpha}$ and “lowering” $E_{-\alpha}$ operators (aka ladder operators) on a given eigenvector $E_{\beta}$. We saw that if $\alpha$ and $\beta$ are two distinct roots, with $\alpha + \beta \neq 0$, it may happen that $\beta \pm \alpha$ are also roots. Let $p \leq 0$ be the smallest integer such that $(\text{ad} E_{-\alpha})^{p} E_{\beta}$ is non
zero, i.e. that $\beta + p\alpha$ is a root, and let $q \geq 0$ be the largest integer such that $(\text{ad } E_\alpha)^q E_\beta$ is non zero, i.e. that $\beta + q\alpha$ is a root. We call the subset of roots $\{\beta + p\alpha, \beta + (p+1)\alpha, \cdots, \beta, \cdots, \beta + q\alpha\}$ the $\alpha$-chain through $\beta$. Note that the $E_\beta'$, when $\beta'$ runs along that chain, form a basis of a finite dimensional representation of the $\text{su}(2)$ algebra generated by $H_\alpha$ and $E_{\pm \alpha}$. According to what we know about these representations of $\text{su}(2)$, the lowest and highest eigenvalues of $H_\alpha$ are opposite

$$\langle \alpha, \beta + p\alpha \rangle = -\langle \alpha, \beta + q\alpha \rangle$$

or $2\langle \beta, \alpha \rangle = -(q+p)\langle \alpha, \alpha \rangle$, thus with the notation (3-21)

$$m = -p - q . \quad (3-22)$$

This construction also shows that $\beta - m\alpha = \beta + (p+q)\alpha$ is in the $\alpha$-chain through $\beta$, (since $p \leq -m \leq q$), hence that this is a root.

Remark. The discussion of § 3.1 left the coefficients $N_{\alpha\beta}$ undetermined. One shows (see Exercise B), using the commutation relations of the $E'$s along a chain that the coefficients $N_{\alpha\beta}$ satisfy non linear relations and that they are determined up to signs by the geometry of the root system according to

$$|N_{\alpha\beta}| = \sqrt{\frac{1}{2}(1-p)q\langle \alpha, \alpha \rangle} . \quad (3-23)$$

Note that, as stated before, $N_{\alpha\beta}$ vanishes only if $q = 0$, i.e. if $\alpha + \beta$ is not a root.

### Weyl group

For any vector $x$ in the root space $\mathfrak{h}^*$, define the linear transformation

$$w_\alpha(x) = x - 2\frac{\langle \alpha, x \rangle}{\langle \alpha, \alpha \rangle}\alpha . \quad (3-24)$$

This is a reflection in the hyperplane orthogonal to $\alpha$: $(w_\alpha)^2 = I$, $w_\alpha(\alpha) = -\alpha$, and $w_\alpha(x) = x$ if $x$ is orthogonal to $\alpha$. This is of course an isometry, since it preserves the scalar product: $\langle w_\alpha(x), w_\alpha(y) \rangle = \langle x, y \rangle$. Such a $w_\alpha$ is called a Weyl reflection. By definition the Weyl group $W$ is the group generated by the $w_\alpha$, i.e. the set of all possible products of $w_\alpha$ over roots $\alpha$.

Thanks to the remark following (3-22), if $\alpha$ and $\beta$ are two roots, $w_\alpha(\beta) = \beta - m\alpha$ is also a root. The set of roots is thus globally invariant under the action of the Weyl group. The group $W$ is completely determined by its action on roots, which is a permutation. $W$ is thus a subgroup of the permutation group of the finite set $\Delta$, hence a finite group.

Example: for the algebra $\text{su}(n)$, one finds that $W = \mathfrak{S}_n$, the permutation group of $n$ objects, see below in § 3.3.2.

Note that if $\beta_+ = \beta + q\alpha$ is the highest root in the $\alpha$-chain through $\beta$, and $\beta_- = \beta + p\alpha$ the lowest one, $w_\alpha(\beta_\pm) = \beta_{\mp}$ and more generally, the roots of the chain are swapped by pairs under the action of $w_\alpha$. The chain is thus invariant by $w_\alpha$. (This is a generalisation of the $m \leftrightarrow -m$ symmetry of the $\text{su}(2)$ “multiplets” $(-j, -j+1, \cdots, j-1, j)$.)

---

2This property is far from trivial: generically, when $m$ vectors are given in the Euclidean space $\mathbb{R}^m$, the group generated by reflections in the hyperplanes orthogonal to these vectors is infinite. You need very peculiar configurations of vectors to make the group finite. Finite reflection groups have been classified by Coxeter. Weyl groups of simple algebras form a subset of Coxeter groups.
3.2. GEOMETRY OF ROOT SYSTEMS

Positive roots, simple roots. Cartan matrix

Roots are not linearly independent in \( h^* \). One may show that one can partition their set \( \Delta \) into “positive” and “negative” roots, the opposite of a positive root being negative, and find a basis \( \alpha_i, i = 1, \cdots, \ell \) of \( \ell \) simple roots, such that any positive (resp. negative) root is a linear combination with non negative (resp non positive) integer coefficients of these simple roots. As a consequence, a simple root cannot be written as the sum of two positive roots (check!).

Neither the choice of a set of positive roots, nor that of a basis of simple roots is unique. One goes from a basis of simple roots to another one by some operation of the Weyl group.

If \( \alpha \) and \( \beta \) are simple roots, \( \alpha - \beta \) cannot be a root (why?). The integer \( p \) in the previous discussion thus vanishes and \( m = -q \leq 0 \). It follows that \( \langle \alpha, \beta \rangle \leq 0 \).

The scalar product of two simple roots is non positive. \((P)\)

We now define the Cartan matrix

\[
C_{ij} = 2 \frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle}.
\]

\((3-25)\)

Beware, that matrix is \textit{a priori} non symmetric.\(^3\) Its diagonal elements are 2, its off-diagonal elements are \( \leq 0 \) integers.

One must remember that the scalar product appearing in the numerator of \((3-25)\) is positive definite. According to the Schwarz inequality, \( \langle \alpha, \beta \rangle^2 \leq \langle \alpha, \alpha \rangle \langle \beta, \beta \rangle \) with equality only if \( \alpha \) and \( \beta \) are colinear. This property, together with the integrity properties of their elements, suffices to classify all possible Cartan matrices, as we shall now see.

Write \( \langle \alpha_i, \alpha_j \rangle = \| \alpha_i \| \| \alpha_j \| \cos \overrightarrow{\alpha_i, \alpha_j} \). Then by multiplying or dividing the two equations \((3-21)\) for the pair \( \{ \alpha_i, \alpha_j \}, i \neq j \), namely \( C_{ij} = m_i \leq 0 \) and \( C_{ji} = m_j \leq 0 \), where the property \((P)\) above has been taken into account, one finds that if \( i \neq j \),

\[
\begin{align*}
\cos \overrightarrow{\alpha_i, \alpha_j} &= -\frac{1}{2} \sqrt{m_i m_j} \\
\| \alpha_i \| \| \alpha_j \| &= \sqrt{\frac{m_i}{m_j}}
\end{align*}
\]

with \( m_i, m_j \in \mathbb{N} \), \((3-26)\)

and the value \(-1\) of the cosine is impossible, since \( \alpha_i \neq -\alpha_j \) by assumption, so that the only possible values of that cosine are \( 0, -\frac{1}{2}, -\sqrt{\frac{3}{2}}, -\sqrt{\frac{5}{2}} \), \( i.e. \) the only possible angles between simple roots are \( \frac{\pi}{2}, \frac{2\pi}{3}, \frac{3\pi}{4} \) or \( \frac{5\pi}{6} \), with ratios of lengths of roots respectively equal to \(?\) (undetermined), \( 1, \sqrt{2}, \sqrt{3} \).

There exists of course only one algebra of rank 1, viz the (complexified) su(2) algebra, \((3-1)\) or \((3-17)\). It will be called \( A_1 \) below. It is then easy to classify the possible algebras of rank 2. The four cases are depicted on Fig. 3.1, with their Cartan matrices reading

\[
A_2 : \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad B_2 : \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix} \quad G_2 : \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \quad D_2 : \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.
\]

\((3-27)\)

The nomenclature, \( A_2, B_2, G_2 \) and \( D_2 \), is conventional, and so is the numbering of roots. The latter case, \( D_2 \), which has \( \langle \alpha_1, \alpha_2 \rangle = 0 \), is mentioned here for completeness: it corresponds to

\(^3\)Also, beware that some authors call Cartan matrix the transpose of \((3-25)!\)
a semi-simple algebra, the direct sum of two $A_1$ algebras. (Nothing forces its two roots to be of equal length.)

In general, if the set of roots may be split into two mutually orthogonal subsets, one sees that the Lie algebra decomposes into a direct sum of two algebras, and vice versa. Recalling that any semi-simple algebra may be decomposed into the direct sum of simple subalgebras (see end of Chap. 1), in the following we consider only simple algebras.

Dynkin diagrams

For higher rank, i.e. for higher dimension of the root space, it becomes difficult to visualise the root system. Another representation is adopted, by encoding the Cartan matrix into a diagram in the following way: with each simple root is associated a vertex of the diagram; two vertices are linked by an edge iff $\langle \alpha_i, \alpha_j \rangle \neq 0$; the edge is simple if $C_{ij} = C_{ji} = -1$ (angle of $2\pi/3$, equal lengths); it is double (resp. triple) if $C_{ij} = -2$ (resp. $-3$) and $C_{ji} = -1$ (angle of $3\pi/4$ resp. $5\pi/6$, with a length ratio of $\sqrt{2}$, resp. $\sqrt{3}$) and then carries an arrow (or rather a sign $>\$) from $i$ to $j$ indicating which root is the longest. (Beware that some authors use the opposite convention for arrows!).

Figure 3.1: Root systems of rank 2. The two simple roots are drawn in thick lines. For the algebras $B_2$, $G_2$ et $D_2$, only positive roots have been labelled.

Figure 3.2: Dynkin diagrams
3.2. GEOMETRY OF ROOT SYSTEMS

3.2.2 Root systems of simple algebras. Cartan classification

The analysis of all possible cases led Cartan\(^4\) to a classification of simple complex Lie algebras, in terms of four infinite families and five exceptional cases. The traditional notation is the following

\[
A_\ell, \ B_\ell, \ C_\ell, \ D_\ell, \ E_6, \ E_7, \ E_8, \ F_4, \ G_2.
\]

(3-28)

In each case, the lower index gives the rank of the algebra. The geometry of the root system is encoded in the Dynkin diagrams of Fig. 3.2.

The four infinite families are identified with the (complexified) Lie algebras of classical groups

\[
A_\ell = \text{sl}(\ell + 1, \mathbb{C}), \quad B_\ell = \text{so}(2\ell + 1, \mathbb{C}), \quad C_\ell = \text{sp}(2\ell, \mathbb{C}), \quad D_\ell = \text{so}(2\ell, \mathbb{C}).
\]

(3-29)

or with their unique compact real form, respectively \(A_\ell = \text{su}(\ell + 1)\), \(B_\ell = \text{so}(2\ell + 1)\), \(C_\ell = \text{usp}(\ell)\), \(D_\ell = \text{so}(2\ell)\).

The “exceptional algebras” \(E_6, \ldots, G_2\) have respective dimensions 78, 133, 248, 52 and 14. Those are algebras of ... exceptional Lie groups! The group \(G_2\) is the group of automorphisms of octonions, \(F_4\) is itself an automorphism group of octonion matrices, etc.

Among these algebras, the algebras \(A, D, E\), whose roots have the same length, are called simply laced. A curious observation is that many problems, finite subgroups of \(\text{su}(2)\), “simple” singularities, “minimal conformal field theories”, etc, are classified by the same \(ADE\) scheme... but this is another story!

The real forms of these simple complex algebras have also been classified by Cartan. One finds 12 infinite series and 23 exceptional cases!

3.2.3 Chevalley basis

There exists another basis of the Lie algebra \(\mathfrak{g}\), called Chevalley basis, with brackets depending only on the Cartan matrix. Let \(h_i, e_i, f_i, i = 1, \ldots, \ell\), be generators attached to simple roots \(\alpha_i\) according to

\[
e_i = \left( \frac{2}{\langle \alpha_i, \alpha_i \rangle} \right)^{\frac{1}{2}} E_{\alpha_i}, \quad f_i = \left( \frac{2}{\langle \alpha_i, \alpha_i \rangle} \right)^{\frac{1}{2}} E_{-\alpha_i}, \quad h_i = \frac{2\alpha_i}{\langle \alpha_i, \alpha_i \rangle} .
\]

(3-30)

Their commutation relations read

\[
[h_i, h_j] = 0, \quad [h_i, e_j] = C_{ji} e_j, \quad [h_i, f_j] = -C_{ji} f_j, \quad [e_i, f_j] = \delta_{ij} h_j
\]

(3-31)

(check!). The algebra is generated by the \(e_i, f_i, h_i\) and all their commutators, constrained by (3-31) and by the “Serre relations”

\[
\text{ad}(e_i)^{1-C_{ij}} e_j = 0 \quad (3-32)
\]

\[
\text{ad}(f_i)^{1-C_{ij}} f_j = 0 . \quad (3-33)
\]

This proves that the whole algebra is indeed encoded in the data of the simple roots and of their geometry (Cartan matrix or Dynkin diagram).

Note also the remarkable and a priori not obvious property that in that basis, all the structure constants (coefficients of the commutation relations) are integers.

---

\(^4\)This classification work, undertaken by Killing, was corrected and completed by É. Cartan, and later simplified by van der Waerden, Dynkin, ...
3.2.4 Coroots. Highest root. Coxeter number and exponents

We give here some complements on notations and concepts that are encountered in the study of simple Lie algebras and of their root systems.

As the combination

$$\alpha_i^\vee := \frac{2}{\langle \alpha_i, \alpha_i \rangle} \alpha_i,$$  \hspace{1cm} (3-34)

for $\alpha_i$ a simple root, appears frequently, it is given the name of coroot. The Cartan matrix may be rewritten as

$$C_{ij} = \langle \alpha_i, \alpha_j^\vee \rangle.$$  \hspace{1cm} (3-35)

The highest root $\theta$ is the positive root with the property that the sum of its components in a basis of simple roots is maximal: one proves that this characterizes it uniquely. Its components in the basis of simple roots and in that of coroots

$$\theta = \sum_i a_i \alpha_i, \quad \frac{2}{\langle \theta, \theta \rangle} \theta = \sum_i a_i^\vee \alpha_i^\vee,$$  \hspace{1cm} (3-36)

called Kac labels, resp dual Kac labels, play also a role, in particular through their sums,

$$h = 1 + \sum_i a_i, \quad h^\vee = 1 + \sum_i a_i^\vee.$$  \hspace{1cm} (3-37)

The numbers $h$ and $h^\vee$ are respectively the Coxeter number and the dual Coxeter number. When a normalisation of roots has to be picked, which we have not done yet, one usually imposes that $\langle \theta, \theta \rangle = 2$.

Lastly the diagonalisation of the symmetrized Cartan matrix

$$\hat{C}_{ij} := 2 \frac{\langle \alpha_i, \alpha_j \rangle}{\sqrt{\langle \alpha_i, \alpha_i \rangle \langle \alpha_j, \alpha_j \rangle}}$$  \hspace{1cm} (3-38)

yields a spectrum of eigenvalues

$$\text{eigenvalues of } \hat{C} = \left\{ 4 \sin^2 \left( \frac{\pi}{2h} m_i \right) \right\}, \quad i = 1, \cdots, \ell,$$  \hspace{1cm} (3-39)

in which a new set of integers $m_i$ appears, the Coxeter exponents, satisfying $1 \leq m_i \leq h - 1$ with possible multiplicities. These numbers are relevant for various reasons. They contain useful information on the Weyl group. After addition of 1, (making them $\geq 2$), one gets the degrees of algebraically independent Casimir operators, or the degrees where the Lie group has a non trivial cohomology, etc etc.

Examples: for $A_{n-1}$ alias $\text{su}(n)$, roots and coroots coincide. The highest root is $\theta = \sum_i \alpha_i$, thus $h = h^\vee = n$, the Coxeter exponents are $1, 2, \cdots, n - 1$. For $D_n$ alias $\text{so}(2n)$, roots and coroots are again identical, $\theta = \alpha_1 + 2\alpha_2 + \cdots + 2\alpha_{n-2} + \alpha_{n-1} + \alpha_n$, $h = 2n - 2$, and the exponents are $1, 3, \cdots, 2n - 3, n - 1$, with $n - 1$ double if $n$ is even.

See Appendix F for Tables of data on the classical simple algebras.

3.3 Representations of semi-simple algebras

3.3.1 Weights. Weight lattice

We now turn our attention to representations of semi-simple algebras, with an approach parallel to that of previous sections. In what follows, “representation” means finite dimensional irreducible representation. We also assume these representations to be unitary: this is the case of interest for representations of compact groups. The elements of the Cartan subalgebra commute among themselves, they also commute in any representation. Denoting with “bras” and
"kets" the vectors of that representation, and writing simply $X$ (instead of $d(X)$) for the representative of the element $X \in \mathfrak{g}$, one may find a basis $|\lambda_a\rangle$ which diagonalises simultaneously the elements of the Cartan algebra

$$H|\lambda_a\rangle = \lambda(H)|\lambda_a\rangle$$

(3-40)
or equivalently

$$H_i|\lambda_a\rangle = \lambda_{(i)}|\lambda_a\rangle,$$

(3-41)

with an eigenvalue $\lambda$ which is again a linear form on the space $\mathfrak{h}$, hence an element of $\mathfrak{h}^*$, the root space. Such a vector $\lambda = (\lambda_{(i)})$ of $\mathfrak{h}^*$ is called a weight. Note that for a unitary representation, the $H$ are Hermitian, hence $\lambda$ is real-valued: the weights are real vectors of $\mathfrak{h}^*$.

As the eigenvalue $\lambda$ may occur with some multiplicity, we have appended the eigenvectors with a multiplicity index $a$. The set of weights of a given representation forms in the space $\mathfrak{h}^*$ the weight diagram of the representation, see Fig. 3.5 below for examples in the case of su(3).

The adjoint representation is a particular representation of the algebra whose weights are the roots. The roots studied in the previous sections thus belong to the set of weights in $\mathfrak{h}^*$.

The vectors $|\lambda_a\rangle$ forming a basis of the representation, their total number, including the multiplicity, equals the dimension of the representation space $E$. This space $E$ contains representation subspaces for each of the su(2) algebras that we identified in §3.2, generated by $\{H_\alpha, E_\alpha, E_{-\alpha}\}$. By the same argument as in §2, one now show that any weight $\lambda$ satisfies

$$\forall \alpha, \quad 2\frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} = m' \in \mathbb{Z},$$

(3-42)

and conversely, it may be shown that any $\lambda \in \mathfrak{h}^*$ satisfying (3-42) is the weight of some finite dimensional representation. One may thus use (3-42) as an alternative definition of weights. To convince oneself that the weights of any representation satisfy (3-42), one may, like in §3.2, define the maximal chain of weights through $\lambda, \lambda + p'\alpha, \cdots, \lambda + q'\alpha$ which form a representation of the su(2) subalgebra, and then show that $m' = -p' - q'$.

Let $p'$ be the smallest $\leq 0$ integer such that $(E_{-\alpha})^{p'}|\lambda_a\rangle \neq 0$, and $q'$ the largest $\geq 0$ integer such that $(E_\alpha)^{q'}|\lambda_a\rangle \neq 0$, $H_\alpha$ has respective eigenvalues $\langle \lambda, \alpha \rangle + p'\langle \alpha, \alpha \rangle$, and $\langle \lambda, \alpha \rangle + q'\langle \alpha, \alpha \rangle$ on these vectors. Expressing that the eigenvalues of $2H_\alpha/\langle \alpha, \alpha \rangle$ are opposite integers, one finds

$$2q' + 2\frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} = 2j$$

$$2p' + 2\frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} = -2j.$$

Subtracting these equations gives $q' - p' = 2j$, and the length of the chain is $2j + 1$ (dimension of the spin $j$ representation of su(2)), while adding them to get rid of $2j$, one has

$$2\frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} = -(q' + p') =: m', \quad \text{as announced in (3-42)}.$$

This chain is invariant under the action of the Weyl reflection $w_\alpha$. (This is a generalisation of the $\mathbb{Z}_2$ symmetry of su(2) “multiplets” $(-j, -j + 1, \cdots, j - 1, j)$.) More generally the set of weights is invariant under the Weyl group: if $\lambda$ is a weight of a representation, so is $w_\alpha(\lambda)$,
and one shows that they have the same multiplicity. The weight diagram of a representation is thus invariant under the action of $W$.

The set of weights is split by the Weyl group $W$ into “chambers”, whose number equals the order of $W$. The chamber associated with the element $w$ of $W$ is the cone

$$C_w = \{ \lambda | \langle w \lambda, \alpha_i \rangle \geq 0 , \ \forall i = 1, \ldots, \ell \} ,$$

(3-43)

where the $\alpha_i$ are the simple roots. (This is not quite a partition, as some weights belong to the “walls” between chambers.) The fundamental chamber is $C_1$, corresponding to the identity in $W$. The weights belonging to that fundamental chamber are called dominant weights. Any weight may be brought into $C_1$ by some operation of $W$: it is on the “orbit” (for the Weyl group) of a unique dominant weight. Among the weights of a representation, at least one belongs to $C_1$.

On the other hand, from $[H_i, E_\alpha] = \alpha(i) E_\alpha$ follows that

$$H_i E_\alpha \lambda_a \rangle = ( [H_i, E_\alpha] + E_\alpha H_i ) \lambda_a \rangle = (\alpha(i) + \lambda(i)) E_\alpha \lambda_a \rangle$$

hence that $E_\alpha \lambda_a \rangle$, if non vanishing, is an eigenvector of weight $\lambda + \alpha$. Now, in an irreducible representation, all vectors are obtained from one another by such actions of $E_\alpha$, and we conclude that

$\triangleright$ Two weights of the same (irreducible) representation differ by a integer-coefficient combination of roots,

(but this combination is in general not a root).

One then introduces a partial order on weights of the same representation: $\lambda' > \lambda$ if $\lambda' - \lambda = \sum_i n_i \alpha_i$, with non negative (integer) coefficients $n_i$. Among the weights of that representation, one proves there exists a unique highest weight $\Lambda$, which is shown to be of multiplicity 1. The highest weight vector will be denoted $|\Lambda\rangle$ (with no index $a$). It is such that for any positive root $E_\alpha |\Lambda\rangle = 0$, (otherwise, it would not be the highest), hence $q' = 0$ in the discussion above, and $\langle \Lambda, \alpha \rangle = \frac{1}{2} \langle \alpha, \alpha \rangle j > 0$, $\Lambda$ is thus a dominant weight.

$\triangleright$ The highest weight of a representation is a dominant weight, $\Lambda \in C_1$.

This highest weight vector characterises the irreducible representation. (In the case of $su(2)$, this would be a vector $|j, m = j\rangle$.) In other words, two representations are equivalent iff they have the same highest weight.

One then introduces the Dynkin labels of the weight $\lambda$ by

$$\lambda_i = 2 \frac{\langle \lambda, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} \in \mathbb{Z} \quad (3-44)$$

with $\alpha_i$ the simple roots. For a dominant weight, thus for any highest weight of a representation, these indices are non negative, i.e. in $\mathbb{N}$.

The fundamental weights $\Lambda_i$ satisfy by definition

$$2 \frac{\langle \Lambda_i, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = \delta_{ij} . \quad (3-45)$$
Their number equals the rank \( \ell \) of algebra, and they make a basis of \( \mathfrak{h}^* \). Each one is the highest weight of an irreducible representation called \textit{fundamental}; hence there are \( \ell \) fundamental representations. We have thus obtained

\( \triangleright \) Any irreducible representation irreducible is characterised by its highest weight, and with a little abuse of notation, we denote \((\Lambda)\) the irreducible representation of highest weight \( \Lambda \).

\( \triangleright \) Any highest weight decomposes on fundamental weights, and its components are its \textit{Dynkin labels} (3-44),

\[
\Lambda = \sum_{j=1}^{\ell} \lambda_j A_j , \quad \lambda_i \in \mathbb{N} .
\]  

(3-46)

and any \( \Lambda \) of the form (3-46) is the highest weight of an irreducible representation.

Stated differently, the knowledge of the fundamental weight suffices to construct all irreducible representations of the algebra.

Using the properties just stated, show that the highest weight of the adjoint representation is necessarily \( \theta \), defined in § 2.4.

**Weight and root lattices**

Generally speaking, given a basis of vectors \( e_1, \cdots e_p \) in a \( p \) dimensional space, the \textit{lattice} generated by these vectors is the set of vectors \( \sum_{i=1}^{p} z^i e_i \) with coefficients \( z^i \in \mathbb{Z} \). This lattice is also denoted \( \mathbb{Z}e_1 + \cdots + \mathbb{Z}e_p \).

The \textit{weight lattice} \( P \) is the lattice generated by the fundamental \( \ell \) weights \( \Lambda_i \). The \textit{root lattice} \( Q \) is the one generated by the \( \ell \) simple roots \( \alpha_i \). This is a sublattice of \( P \). Any weight of an irreducible representation belongs to \( P \).

One may consider the \textit{congruence classes} of the additive group \( P \) wrt its subgroup \( Q \), that are the classes for the equivalence relation \( \lambda \sim \lambda' \) iff \( \lambda - \lambda' \in Q \). The number \( |P/Q| \) of these classes turns out to be equal to the determinant of the Cartan matrix(!). In the case of \( \text{su}(n) \), there are \( n \) classes, we shall return to that point later.

One may also introduce the \textit{lattice} \( Q^\vee \) generated by the \( \ell \) coroots \( \alpha_i^\vee \) (cf §2.4). It is the “dual” of \( P \), in the sense that \( (\alpha_i^\vee, \Lambda_j) \in \mathbb{Z} \).

One also shows that the subgroups of the finite group \( P/Q \) are isomorphic to homotopy groups of groups \( G \) having \( \mathfrak{g} \) as a Lie algebra! For example for \( \text{su}(n) \), we find below that \( P/Q = \mathbb{Z}_n \), and these subgroups are characterised by a divisor \( d \) of \( n \). For each of them, \( \text{SU}(n)/\mathbb{Z}_d \) has the \( \text{su}(n) \) Lie algebra. The case \( n = 2 \), with \( \text{SU}(2) \) and \( \text{SO}(3) \), is quite familiar.

**Dimension and Casimir operator**

It may be useful to know the dimension of a representation with a given highest weight and the value of the quadratic Casimir operator in that representation. These expressions are given in terms of the \textit{Weyl vector} \( \rho \), defined by any of the two (non trivially!) equivalent formulas

\[
\rho = \frac{1}{2} \sum_{\alpha > 0} \alpha = \sum_{j} A_j .
\]

(3-47)
A remarkable formula, due to Weyl, gives the dimension of the representation of highest weight \( \Lambda \) as a product over positive roots

\[
\dim(\Lambda) = \prod_{\alpha > 0} \frac{\langle \Lambda + \rho, \alpha \rangle}{\langle \rho, \alpha \rangle}
\]  

(3-48)

while the eigenvalue of the quadratic Casimir reads

\[
C_2(\Lambda) = \frac{1}{2} \langle \Lambda, \Lambda + 2\rho \rangle.
\]

(3-49)

A related question is that of the trace of generators of \( g \) in the representation \( \Lambda \). Let \( t_a \) be the basis of \( g \) such that \( \text{tr} t_a t_b = T_A \delta_{ab} \), with a coefficient \( T_A \) ("A" like adjoint) whose sign depends on conventions (\( t \) Hermitian or antihermitian, see Chap. 1). In the representation of highest weight \( \Lambda \), one has

\[
\text{tr} d_\Lambda(t_a)d_\Lambda(t_b) = T_\Lambda \delta_{ab}.
\]

(3-50)

But in that basis, the quadratic Casimir reads \( C_2 = \sum_a (d_\Lambda(t_a))^2 \) hence, taking the trace,

\[
\text{tr} C_2 = \sum_a \text{tr} (d_\Lambda(t_a))^2 = T_\Lambda \sum_a 1 = T_\Lambda \dim g
\]

\[
= C_2(\Lambda) \text{tr} I_{\Lambda} = C_2(\Lambda) \dim(\Lambda)
\]

(3-51)

whence

\[
T_\Lambda = C_2(\Lambda) \frac{\dim(\Lambda)}{\dim g},
\]

(3-52)

a useful formula in calculations (gauge theories . . .). In the adjoint representation, \( \dim(A) = \dim g \), hence \( T_\Lambda = T_A = C_2(A) \).

There is a host of additional, sometimes intriguing, formulas relating various aspects of Lie algebras and representation theory. For example the Freudenthal–de Vries “strange formula”, which connects the norms of the vectors \( \rho \) and \( \theta \) to the dimension of the algebra and the Coxeter number:

\[
\langle \rho, \rho \rangle = \frac{h^2}{24} \langle \theta, \theta \rangle \dim g.
\]

There is also a formula (Freudenthal) giving the multiplicity of a weight \( \lambda \) within a representation of given highest weight \( \Lambda \). And last, as a related issue, a formula by Weyl giving the character \( \chi_\Lambda(e^H) \) of that representation evaluated on an element of the Cartan torus, an abelian subgroup resulting from the exponentiation of the Cartan algebra \( h \).

Conjugate representation

Given a representation of highest weight \( \Lambda \), its complex conjugate representation is generally non equivalent. On may characterize what is its highest weight \( \overline{\Lambda} \) thanks to the Weyl group. The non-equivalence of representations \( (\Lambda) \) and \( (\overline{\Lambda}) \) has to do with the symmetries of the Dynkin diagram. For the algebras of type \( B, C, E_7, E_8, F_4, G_2 \) for which there is no non trivial symmetry, the representations are self-conjugate. This is also the case of \( D_2r \).

For the others, conjugation corresponds to the following symmetry on Dynkin labels

\[
A_\ell = \text{su}(\ell + 1) \quad \lambda_i \leftrightarrow \lambda_{\ell+1-i} \quad \ell > 1
\]

\[
D_{2r+1} = \text{so}(4r + 2) \quad \lambda_i \leftrightarrow \lambda_{\ell-1} \quad \ell = 2r + 1
\]

\[
E_6 \quad \lambda_i \leftrightarrow \lambda_{6-i}, \quad i = 1, 2.
\]

(3-53)

3.3.2 Roots and weights of \( \text{su}(n) \)

Let us construct explicitly the weights and thus the irreducible representations of \( \text{su}(n) \).

We first pick a convenient parametrization of the space \( h^* \), which is of dimension \( n - 1 \). Let \( e_i, i = 1, \ldots, n \), be \( n \) vectors of \( h^* = \mathbb{R}^{n-1} \) (hence necessarily dependent), satisfying \( \sum_{i=1}^{n} e_i = 0 \).
3.3. REPRESENTATIONS OF SEMI-SIMPLE ALGEBRAS

They are obtained starting from an orthonormal basis \( \hat{e}_i \) of \( \mathbb{R}^n \) by projecting the \( \hat{e}_i \) on the hyperplane orthogonal to \( \hat{\rho} := \sum_{i=1}^{n} \hat{e}_i \), thus \( e_i = \hat{e}_i - \frac{1}{n} \hat{\rho} \). (In other words, one is considering the hyperplane \( \sum_{i}^{n} x_i = 1 \) in the space \( \mathbb{R}^n \)). These vectors have scalar products given by

\[
\langle e_i, e_j \rangle = \delta_{ij} - \frac{1}{n}.
\]  

(3-54)

In terms of these vectors, the positive roots of \( \text{su}(n) = A_{n-1} \), whose number equals \( |\Delta_+| = n(n-1)/2 \), are

\[
\alpha_{ij} = e_i - e_j, \quad 1 \leq i < j \leq n,
\]  

(3-55)

and the \( \ell = n - 1 \) simple roots are

\[
\alpha_i = \alpha_{i+1} = e_i - e_{i+1}, \quad 1 \leq i \leq n - 1.
\]  

(3-56)

These roots have been normalized by \( \langle \alpha, \alpha \rangle = 2 \). The sum of positive roots is easily computed to be

\[
2\rho = (n-1)e_1 + (n-3)e_2 + \cdots + (n-2i+1)e_i + \cdots - (n-1)e_n
\]

\[
= (n-1)\alpha_1 + 2(n-2)\alpha_2 + \cdots + i(n-i)\alpha_i + \cdots + (n-1)\alpha_{n-1}.
\]  

(3-57)

One checks that the Cartan matrix is

\[
C_{ij} = \langle \alpha_i, \alpha_j \rangle = \begin{cases} 2 & \text{if } i = j \\ -1 & \text{if } i = j \pm 1 \end{cases}
\]

in accord with the Dynkin diagram of type \( A_{n-1} \), thus justifying (3-56). The fundamental weights \( \Lambda_i, i = 1, \cdots, n-1 \) are then readily written

\[
\Lambda_i = \sum_{j=1}^{i} e_j,
\]  

(3-58)

\[
e_1 = \Lambda_1, \quad e_i = \Lambda_i - \Lambda_{i-1} \text{ for } i = 2, \cdots, n - 1, \quad e_n = -\Lambda_{n-1}
\]  

(3-59)

with scalar products

\[
\langle \Lambda_i, \Lambda_j \rangle = \frac{i(n-j)}{n}, \quad i \leq j.
\]  

(3-60)

The Weyl group \( W \equiv S_N \) acts on roots and on weights by permuting the \( e_i \):

\[
w \in W \leftrightarrow \bar{w} \in S_N : w(e_i) = e_{\bar{w}(i)}.
\]

**Dimension of the representation of weight \( \Lambda \)**

Combining formulas (3-48) and (3-55), prove the following expression

\[
\dim(\Lambda) = \prod_{1 \leq i < j \leq n} \frac{f_i - f_j + j - i}{j - i} \quad \text{où} \quad f_i := \sum_{k=i}^{n-1} \lambda_k, \quad f_n = 0.
\]  

(3-61)
Conjugate representations

If $\Lambda = (\lambda_1, \cdots, \lambda_{n-1})$ is the highest weight of an irreducible representation of $\mathfrak{su}(n)$, $\overline{\Lambda} = (\lambda_{n-1}, \cdots, \lambda_1)$ is that of the complex conjugate, generally inequivalent, representation (see above in § 3.3.1). Note that neither the dimension, nor the value of the Casimir operator distinguish the representations $\Lambda$ and $\overline{\Lambda}$.

"n-ality".

There are $n$ congruence classes of $P$ with respect to $Q$. They are distinguished by the value of

$$\nu(\lambda) := \lambda_1 + 2\lambda_2 + \cdots + (n-1)\lambda_{n-1} \mod n,$$

(3-62)

to which one may give the unimaginative name of "n-ality", by extension of the "triality" of $\mathfrak{su}(3)$, see below. The elements of the root lattice thus have $\nu(\lambda) = 0$.

Examples of $\mathfrak{su}(2)$ and $\mathfrak{su}(3)$.

In the case of $\mathfrak{su}(2)$, there is one fundamental weight $\Lambda = \Lambda_1$ and one positive root $\alpha$, normalised by $\langle \alpha, \alpha \rangle = 2$, hence $\langle \Lambda, \alpha \rangle = 1$, $\langle \Lambda, \Lambda \rangle = \frac{1}{2}$. Thus $\alpha = 2\Lambda$, $\Lambda$ corresponds to the spin $\frac{1}{2}$ representation, $\alpha$ to spin 1. The weight lattice and the root lattice are easy to draw, see Fig. 3.3. The Dynkin label $\lambda_1$ is identical to the integer $2j$, the two congruence classes of $P$ wrt $Q$ correspond to representations of integer and half-integer spin, the dimension $\dim(\Lambda) = \lambda_1 + 1 = 2j + 1$ and the Casimir operator $C_2(\Lambda) = \frac{1}{4}\lambda_1(\lambda_1 + 2) = j(j + 1)$, in accord with well known expressions.

For $\mathfrak{su}(3)$, the weight lattice is triangular, see Fig. (3.4), on which the triality $\tau(\lambda) := \lambda_1 + 2\lambda_2 \mod 3$ has been shown and the fundamental weights and the highest weights of the "low lying" representations have been displayed. Following the common use, representations are referred to by their dimension

$$\dim(\Lambda) = \frac{1}{2}(\lambda_1 + 1)(\lambda_2 + 1)(\lambda_1 + \lambda_2 + 2),$$

(3-63)

supplemented by a bar to distinguish a representation from its conjugate, whenever necessary. The conjugate of representation of highest weight $\Lambda = (\lambda_1, \lambda_2)$ has highest weight $\overline{\Lambda} = (\lambda_2, \lambda_1)$. Only the representations lying on the bisector of the Weyl chamber are thus self-conjugate.

Exercise. Compute the eigenvalue of the quadratic Casimir operator in terms of Dynkin labels $\lambda_1, \lambda_2$ using the formulas (3-49) and (3-57).

\footnote{\textit{which may be ambiguous; for example, identify on Fig. (3.4) the weight of another representation of dimension 15.}}
3.3. REPRESENTATIONS OF SEMI-SIMPLE ALGEBRAS

Figure 3.4: Weights of su(3). Only the first Weyl chamber \( C_1 \) has been detailed, with some highest weights. The weights of triality 0 (forming the root lattice) are represented by a wide disk, those of triality 1, resp. 2, by a full, resp. open disk.

Figure 3.5: The weight diagrams of low lying representations of su(3), denoted by their dimension. Note that a rotation of 30° of the weight lattice has been performed with respect to the previous figure. In each representation, the highest weight is marked by a small indentation. The small dots are weights of multiplicity 1, the wider and open dots have multiplicity 2.
The set of weights of low lying representations is displayed on Fig. (3.5), after a rotation of the axes of the previous figures. The horizontal axis, colinear to $\alpha_1$, and the vertical axis, colinear to $\Lambda_2$, will indeed acquire a physical meaning: that of axes of isospin and “hypercharge” coordinates, see next chapter.

Remark. The case of $\mathfrak{su}(n)$ has been detailed. Analogous formulas for roots, fundamental weights, etc of other simple algebras are of course known explicitly and tabulated in the literature. See for example Appendix F for the identity card of “classical algebras”, of type $A, B, C, D$, and Bourbaki, chap.6, for more details on the other algebras.

3.4 Tensor products of representations of $\mathfrak{su}(n)$

3.4.1 Littlewood-Richardson rules

Given two irreducible representations of $\mathfrak{su}(n)$ (or of any other Lie algebra), a frequently encountered problem is to decompose their tensor product into a direct sum of irreducible representations. If one is only interested in multiplicities and if one has a character table of the corresponding compact group, one may use the formulae proved in Chap. 2, § 2.3.2.

There exists also fairly complex rules giving that decomposition into irreducible representations of a product of two irreducible representations $(\Lambda)$ and $(\Lambda')$ of $\mathfrak{su}(n)$. Those are the Littlewood-Richardson rules, which appeal to the expression in terms of Young tableaux (see next §). But it is often simpler to proceed step by step, noticing that the irreducible representation $(\Lambda')$ is found in an adequate product of fundamental representations, and examining the successive products of representation $\Lambda$ by these fundamental representations. By the associativity of the tensor product, one brings the original problem back to that of the tensor product of $(\Lambda)$ by the various fundamental representations.

The latter operation is easy to describe on the weight lattice. Given the highest weight $\Lambda$ in the first Weyl chamber $C_1$, the tensor product of $(\Lambda)$ by the fundamental representation of highest weight $\Lambda_i$ is decomposed into irreducible representations in the following way: one adds in all possible ways the dim$(\Lambda_i)$ weights of the fundamental to the vector $\Lambda$ and one keeps as highest weights in the decomposition only the weights resulting from this addition that belong to $C_1$. 

![Figure 3.6: Tensor product of the 8 representation by the 3 representation, depicted on the weight diagram of $\mathfrak{su}(3)$.](image)
Let us illustrate that on the case of $\text{su}(3)$. Suppose that we want to determine the decomposition of $8 \otimes 8$. One knows that the 8 representation (adjoint) is to be found in the product of two fundamental 3 and $\bar{3}$. The weights of the fundamental representation “3” of highest weight $\Lambda_1 = e_1$ are $e_1, e_2, e_3$. Those of the fundamental representation “$\bar{3}$” are their opposites. With the previous rule, one finds

$$
\begin{align*}
3 \otimes 3 &= \bar{3} + 6 \\
3 \otimes 6 &= 8 + 10 \\
3 \otimes 8 &= 3 + \bar{6} + 15 \\
3 \otimes 15 &= 6 + \bar{15} + 24
\end{align*}
$$

etc, and their conjugates, see Fig. 3.6. In general one adds the three vectors $e_1 = (1, 0)$, $e_2 = (-1, 1)$ et $e_3 = (0, -1)$ (in the basis $\Lambda_1, \Lambda_2$) to $\Lambda = (\lambda_1, \lambda_2)$: the highest weights of the decomposition are thus $(\lambda_1 + 1, \lambda_2)$, $(\lambda_1 - 1, \lambda_2 + 1)$ and $(\lambda_1, \lambda_2 - 1)$, among which those having a negative Dynkin label are discarded. Note the consistency with triality: all the representations appearing in the rhs have the same triality, the sum (modulo 3) of trialities of those of the lhs. For example, $\tau(3) = 1$, $\tau(15) = 1$, $\tau(6) = 2$, $\tau(\bar{15}) = 2$, etc.

Iterating this procedure, one then computes

$$8 \otimes (1 + 8) = 8 \otimes 3 \otimes \bar{3} = (3 + \bar{6} + 15) \otimes \bar{3} = 1 + 8 + 8 + 8 + 10 + \bar{10} + 27$$

from which one derives the formula

$$8 \otimes 8 = 1 + 8_s + 8_a + 10 + \bar{10} + 27. \quad (3-65)$$

In the latter expression, one added a subscript $s$ or $a$ to distinguish the two copies of the 8 representation: one is symmetric, the other antisymmetric in the exchange of the two representations 8 of the left hand side. This relation will be very useful in the following chapter, in the study of the SU(3) “flavor” symmetry group.

Though a bit tedious, this procedure is simple and systematic. There exist also codes computing these decompositions.

### 3.4.2 Explicit tensor construction of representations of SU(2) and SU(3)

Consider a vector $V \in \mathbb{C}^n$ in the defining representation of SU($n$). Under the action of $U \in \text{SU}(n)$, $V \mapsto V' = UV$, or component-wise $v^i \mapsto v'^i = U^i_j v^j$, with indices $i, j = 1, \cdots n$. Let $W$ be a vector which transforms by the complex conjugate representation, (like $W = V^*$), hence $W \mapsto W' = U^* W$. It is natural to denote the components of $W$ with lower indices, since $U^* = (U^\dagger)^T$, and therefore $w'_i = w_j (U^\dagger)_j^i$. Note that $V.W := v^i w_i$ is invariant, by virtue of $U^\dagger . U = I$. In other words the mixed tensor $\delta'^i_j$ is invariant

$$\delta'^i_j = U^i_\nu U^\dagger^\nu_j \delta'^\nu_\nu = (U U^\dagger)^i_j = \delta^i_j .$$
Consider now tensors of rank \((p, m)\), with \(p\) upper indices and \(m\) lower ones, transforming as \(V^{\otimes p} \otimes W^{\otimes m}\), hence according to
\[
t^i_{[k_1 \cdots k_p]} U^i_{j_1} \cdots U^i_{j_p} U^i_{[k_1 \cdots k_m] l_1 \cdots l_m} U^i_{]k_1 \cdots k_m].
\]

- In the case of \(SU(2)\), we know that the representations \(U\) and \(U^*\) are equivalent. This results from the existence of a matrix \(C = i\sigma_2\), such that \(C U C^{-1} = U^*\), thus \(C V^*\) transforms like \(V\). Or, since \(C_{ij} = \epsilon_{ij}\) and \(\epsilon_{ij} \epsilon_{j'k} = \epsilon_{ij} \det U = \epsilon_{ij}\), the antisymmetric tensor \(\epsilon\), invariant and invertible \((\epsilon_{ij} = -\epsilon_{ji}\), \(\epsilon_{ij} \epsilon_{jk} = \delta_{ik}\)), may be used to raise or lower indices, \((v_i := \epsilon_{ij} v^j\), hence \(v_1 = v^2\), \(v_2 = -v^1\)); and therefore it suffices to consider only tensors of rank \(p\) with upper indices. For any pair of indices, say \(i_1\) and \(i_2\), such a tensor may be written as a sum of symmetric and antisymmetric components in these indices
\[
t^{i_1 i_2 \cdots i_p}_{[i_1, i_2 \cdots i_p]} = \frac{1}{2}(t^{i_1 i_2 \cdots i_p} - t^{i_2 i_1 \cdots i_p})
\]
with \(t^{i_1, i_2 \cdots i_p} := \frac{1}{2}(t^{i_1 i_2 \cdots i_p} - t^{i_2 i_1 \cdots i_p})\) and \(t^{i_1, i_2 \cdots i_p} := \frac{1}{2}(t^{i_1 i_2 \cdots i_p} + t^{i_2 i_1 \cdots i_p})\). The antisymmetric component may be recast as \(t^{i_1, i_2 \cdots i_p} = \epsilon_{i_1 i_2} t^{[3\cdots i_p}\) with \(t^{[3\cdots i_p} = -\frac{1}{2} \epsilon_{ab} t^{a b i_1 \cdots i_p}\), and its rank has thus been reduced\(^6\). Consequently only tensors that are completely symmetric in all their upper indices give irreducible representations, and one recovers once construction of all irreducible representations of \(SU(2)\) by symmetrized tensor products of the representation of dimension 2, see Chap 00, and the rank \(p\) identifies with \(2j\). One checks in particular that the number of independent components of a rank \(p\) completely symmetric tensor in the space \(\mathbb{C}^2\) is \(p + 1\), since these components have 0, 1, \(\cdots\) \(p\) indices equal to 1, the other being equal to 2.

A rank \(p\) completely symmetric tensor will be represented by a “Young diagram” with \(p\) boxes \(\begin{array}{ccc}\end{array}\). For the general definition of a Young diagram, see next section. Take \(p = 3\) for definiteness. The tensor product of such a rank 3 tensor by a rank 1 tensor will be depicted as
\[
\begin{array}{ccc}\end{array} \otimes \begin{array}{c} \end{array} = \begin{array}{cccc} \end{array} + \begin{array}{c} \end{array}
\]
which means, in terms of components,
\[
4t^{ijk} u^l = (t^{ijk} u^l + t^{jkl} u^i + t^{ikl} u^j + t^{ijl} u^k) + (t^{ijk} u^l - t^{jkl} u^i) + (t^{ijk} u^l - t^{ikl} u^j) + (t^{ijk} u^l - t^{ijl} u^k)
\]
where the first term is completely symmetric in its \(p + 1 = 4\) indices, and the following terms are antisymmetric in \((i, l)\), \((j, l)\) or \((k, l)\). According to the previous argument, the latter may be reduced to rank 2 tensors.
\[
(t^{ijk} u^l - t^{jkl} u^i) = \epsilon^{il}_{\quad j} \tilde{t}^{jk}, \quad \tilde{t}^{jk} = -\epsilon_{ab} t^{ajk} u^b
\]
which we represent by erasing the columns with two boxes. Therefore
\[
\begin{array}{ccc}\end{array} \otimes \begin{array}{c} \end{array} = \begin{array}{ccccc} \end{array} + \begin{array}{c} \end{array}
\]
where we recognize the familiar rule \(j \otimes \frac{1}{2} = (j + \frac{1}{2}) \oplus (j - \frac{1}{2})\).

---

\(^6\)It may be useful to recall the identities \(\epsilon_{ab} \epsilon_{cd} = \delta_{ac} \delta_{bd} - \delta_{ad} \delta_{bc}\) and hence \(\epsilon_{ab} \epsilon_{bc} = -\delta_{ac}\).
Exercise: reproduce with this method the decomposition rule of \( j \otimes j' \).

- In the case of SU(\( n \)), \( n > 2 \), one must consider tensors with two types of indices, upper and lower, and reduce them. But this is only in the case of SU(3) that this construction will provide us with all irreducible representations. For \( n > 3 \) one has to introduce other tensors transforming under fundamental representations of SU(\( n \)) other than the defining representation (of dimension \( n \)) and its conjugate.

We thus restrict the discussion of the end of this section to the case of SU(3). The tensors are of type \( t_{j_1 \cdots j_m}^{i_1 \cdots i_p} \), \((i, j) = 1, 2, 3\), transforming under the representation \( 3^p \otimes \bar{3}^m \). We still have an invariant tensor \( \epsilon \), but now of rank 3, \( \epsilon_{i'j'k'}U_i^{i'j'}U_j^{j'k'} = \epsilon_{ijk}\det U = \epsilon_{ijk} \), which allows us to trade any pair of upper antisymmetric indices for a lower one, or vice versa, and thus to reduce the rank. But a pair of one upper and one lower indices may also be contracted, according to a remark at the beginning of the section. Consequently one may consider only completely symmetric and traceless tensors of rank \((p, m)\). One may prove that such tensors form an irreducible representation, which is nothing other than the representation of highest weight \( p\Lambda_1 + m\Lambda_2 \) in the notations of § 3.3.2. We content ourselves with a check that the dimensions of these representations are in accord with those given in (3-63), see Exercise D. With this representation we associate again a Young diagram with two rows, the first with \( p + m \), the second with \( m \) boxes.

The rules of decomposition of tensor products, in particular by the fundamental representations 3 and \( \bar{3} \) (see §4.1), are also recovered in this language: the new box must be added in all possible ways to the diagram (while preserving the decreasing of lengths of rows), and any column of height 3 is erased, reflecting the property that \( \det U = 1 \). Exercise: study the reduction of \( \square \otimes \square \) and recover the graphical rule of § 3.4.1 in this language.

A particular case that we shall use repeatedly in the next chapter is the following: the adjoint representation is that of rank \((1,1)\) traceless tensors. This is no surprise: the adjoint representation is spanned by the su(3) Lie algebra, hence by (anti)Hermitian \( 3 \times 3 \) traceless matrices. A tensor of that representation transforms by \( t_{ij} \mapsto t'_{ij} = U_i^iU_j^jU'_{ij} \), or in a matrix form
\[
t' = U t U^\dagger ,
\]
which is also expected, compare with the definition of the adjoint representation in Chap. 2.

Which Young diagram is associated with the adjoint representation?

### 3.5 Young tableaux and representations of GL(\( n \)) and SU(\( n \))

The previous construction extends to su(\( n \)), in fact to the group GL(\( n \)), and involves symmetrization and antisymmetrization operations related to the symmetric group of permutations \( S_n \). We just give a few indications.

Let \( E = \mathbb{C}^n \) be the vector space of dimension \( n \). The group GL(\( n \)) is naturally represented in \( E \)
\[
g \in \text{GL}(E) \cong \text{GL}(n), \quad x \in E \mapsto x' = g x .
\]

\[
(3-68)
\]
Form the \( m \)-th tensor power of \( E \): \( F = E^{\otimes m} = E \otimes \cdots \otimes E \). In \( F \), the group \( \text{GL}(n) \) acts by a representation, the \( m \)-th tensor power of (3-68)

\[
g \in \text{GL}(n), \quad D(g)x^{(1)} \otimes \cdots \otimes x^{(m)} = g.x^{(1)} \otimes \cdots \otimes g.x^{(m)}
\]

which is in general reducible. But in \( F \), there is also the action of the symmetric group \( S_m \) according to

\[
\sigma \in S_m, \quad \hat{D}(\sigma)x^{(1)} \otimes \cdots \otimes x^{(m)} = x^{(\sigma^{-1}1)} \otimes \cdots \otimes x^{(\sigma^{-1}m)}. \tag{3-70}
\]

Choose a basis \( e_i \) in \( E \), and denote \( g_{ij} \) the matrix elements of \( g \in \text{GL}(n) \) in that basis. The representation of \( \text{GL}(n) \) in \( F \) has a matrix

\[
D(g)_{\{i_1 \cdots i_m\}\{j_1 \cdots j_m\}} = \prod_{k=1}^{m} g_{i_k j_k} \tag{3-71}
\]

and that of \( S_m \)

\[
\hat{D}(\sigma)_{\{i_1 \cdots i_m\}\{j_1 \cdots j_m\}} = \prod_{k=1}^{m} \delta_{i_k j_k}. \tag{3-72}
\]

A tensor \( t \), element of \( F \), has components \( t^{i_1} \) in that basis and transforms under the action of \( g \in \text{GL}(n) \), resp. of \( \sigma \in S_m \), into a tensor \( t' \) of components \( t'^{i_1} = D_{i_1,j_1} t^{j_1} \), resp. \( \hat{D}_{i_1,j_1} t^{j_1} \). These two sets of matrices commute

\[
\sum_{\{j\}} D\{i_1,\cdots,i_m\} \hat{D}\{j_1,\cdots,j_m\} = \prod_{l} g_{i_l j_l} \delta_{j_l k_{\sigma^{-1}1}} = \prod_{l} g_{i_l j_l} \delta_{i_l k_{l}} = \sum_{\{j\}} \hat{D}\{i_1,\cdots,i_m\} D\{j_1,\cdots,j_m\} \tag{3-73}
\]

Define then a Young diagram. A Young diagram is made of \( m \) boxes set in \( k \) rows of non increasing length: \( f_1 \geq f_2 \geq \cdots \geq f_k, \sum f_i = m \). Here is an example for \( m = 8 \), with \( f_1 = 4, f_2 = 2, f_3 = 2 \)

\[
\begin{array}{ccc}
\cdot & \cdot & \cdot \\
\cdot & \cdot & \\
\cdot & \\
\end{array}
\]

The \( m \) boxes of a Young diagram may then be filled with different integers ranging between 1 and \( m \), thus making a Young tableau. A standard tableau is a tableau in which the integers are increasing in each row from left to right, and in each column from top to bottom.

The number \( n_Y \) of standard tableaux obtained from a Young diagram \( Y \) is computed as follows. One defines the numbers \( \ell_i = f_i + k - i, i = 1, \cdots, k \). They form a strictly increasing sequence: \( \ell_1 > \ell_2 > \cdots > \ell_k \). Then one proves that

\[
n_Y = \frac{n!}{\prod_i \ell_i! \prod_{i<j} (\ell_i - \ell_j)}. \tag{3-74}
\]

where the product in the numerator is 1 if there is a single row.

The representation theory of the symmetric group \( S_m \) tells us that there is a bijection between irreducible representations and Young diagrams with \( m \) boxes. The dimension of that representation is given by the number of standard tableaux (3-74).

A tensor is said to be of (symmetry) type \( Y \) if it transforms by \( S_m \) under that representation. The commutation of matrices \( D(g) \) and \( \hat{D}(\sigma) \) then ensures that tensors of type \( Y \) form an invariant sub-space under the action of \( \text{GL}(n) \).

Example. Consider the cases of \( m = 2 \) and \( m = 3 \). In the first case, rank 2 tensors may be decomposed into their symmetric and antisymmetric parts which transform independently under the action of \( \text{GL}(n) \)

\[
t^{i_1 i_2} = \frac{1}{2} \left(t^{i_1 i_2} + t^{i_2 i_1}\right) + \frac{1}{2} \left(t^{i_1 i_2} - t^{i_2 i_1}\right).
\]
This decomposition corresponds to the two Young tableaux with 2 boxes, arranged horizontally or vertically. For rank 3, one writes in a similar way the tensors associated with the 4 standard Young tableaux

\[
\begin{align*}
A &= t^{i_1 i_2 i_3} + t^{i_1 i_3 i_2} + t^{i_2 i_1 i_3} + t^{i_2 i_3 i_1} + t^{i_3 i_1 i_2} + t^{i_3 i_2 i_1} \\
B &= t^{i_1 i_2 i_3} + t^{i_1 i_3 i_2} + t^{i_2 i_1 i_3} - t^{i_2 i_3 i_1} - t^{i_3 i_1 i_2} - t^{i_3 i_2 i_1} \\
C_1 &= t^{i_2 i_3} - t^{i_3 i_2} + t^{i_1 i_3} - t^{i_3 i_1} \\
D_1 &= t^{i_2 i_3} - t^{i_3 i_2} + t^{i_1 i_3} - t^{i_3 i_1} 
\end{align*}
\]

where, to make the notations lighter, the indices \(i_1, i_2, i_3\) on \(A, \cdots, D_1\) have been omitted. Any rank 3 tensor decomposes on that basis:

\[6t^{i_1 i_2 i_3} = A + B + 2(C_1 + D_1)\]

The labels 1 on \(C\) and \(D\) recall that under the action of the group \(S_3\), these objects mix with another combination \(C_2 = t^{i_1 i_2 i_3} - t^{i_3 i_1 i_2} + t^{i_2 i_3 i_1} - t^{i_3 i_2 i_1}\) (resp. \(D_2 = t^{i_2 i_3} + t^{i_2 i_1 i_3} - t^{i_1 i_2 i_3} - t^{i_3 i_2 i_1}\)) to make dimension 2 representations. On the contrary the action of the group \(GL(n)\) mixes the different components of tensor \(A\), those of tensor \(B\), etc. Tensors \(C\) et \(D\) transform by equivalent representations.

All Young tableaux, however, do not contribute for a given \(n\). It is clear that a tableau with \(k > n\) rows implies an antisymmetrization of \(k\) indices taking their values in \(\{1, \cdots, n\}\) and vanishes. On the other hand it is easy to see that any tableau with \(k \leq n\) rows gives rise to a representation. One proves, and we admit, that this representation of \(GL(n)\) is irreducible and that its dimension is

\[
dim_Y^{(n)} = \frac{\Delta(f_1 + n - 1, f_2 + n - 2, \cdots, f_n)}{\Delta(n-1, n-2, \cdots, 0)} \quad (3-79)
\]

where \(\Delta(a_1, a_2, \cdots, a_n) = \prod_{i<j} (a_i - a_j)\) is the Vandermonde determinant of the \(a\)'s and the \(f_i\) denote the lengths of rows of the tableau \(Y\). This is a polynomial of degree \(m = \sum f_i\) in \(n\). Compare with (3-61).

In the case of a one-row tableau, the formula results from a simple combinatorial argument. The dimension equals the number of components of the completely symmetric tensor \(t^{i_1 \cdots i_m}\) in which one may assume that \(1 \leq i_1 \leq i_2 \leq \cdots \leq i_m \leq n\). One has to arrange in all possible ways the \(m\) indices \(i_1, \cdots, i_m\) to mark the successive blocks de 1, 2, \(\cdots, n\). The sought dimension is thus the binomial coefficient \(\binom{n+m-1}{n}\) \(\equiv C_n^{m-1}\), in accord in this particular case with (3-79).

In the previous example with \(m = 3\), the last two tensors \(C_1\) and \(D_1\) transform according to equivalent representations. One thus says that \(E^{\otimes 3}\) decomposes as

\[
\begin{array}{ccc}
\hline
& & \\
\hline
\hline
\end{array} + \begin{array}{c}
\hline
& \\
\hline
\end{array} + 2 \begin{array}{c}
\hline
\end{array}
\]

where the third representation comes with a multiplicity two. As a general rule, the multiplicity in \(E^{\otimes m}\) of some representation of \(GL(n)\) labeled by a Young tableau equals the dimension of the corresponding representation of \(S_m\).

This remarkable relation between representations of \(S_m\) and of \(GL(n)\) is due to Frobenius and H. Weyl and is called Frobenius–Weyl duality.

One may extend these considerations to other groups of linear transformations, \(SL(n)\), \(O(n)\), \(U(n)\), \(\ldots\) Because of the additional conditions on the \(g\) matrices in these groups, a further reduction of the representations may occur. For example, we saw in sect. 2.2.2 that the tensor power \(E^{\otimes 2}\) of the 3-dimensional Euclidian space reduced under the action of \(SO(3)\) into three subspaces, corresponding to tensors with a definite symmetry and traceless, and to an invariant scalar.

**Relations between Young diagrams and weights of \(su(n)\)**

Let us finally give the relation between the two descriptions of irreducible representations obtained for \(SU(n)\) or its Lie algebra \(su(n)\). In that case, one may limit the number of rows of the Young tableau \(Y\) to \(k \leq n - 1\)
Figure 3.7: Correspondence between a Young diagram and a highest weight (or Dynkin labels). Here $Y \leftrightarrow \Lambda = (2, 2, 0, 0, 1, 0, 2)$

to obtain all irreducible representations. The $i$-th fundamental weight is represented by a Young diagram made of one column of height $i$, for example $\Lambda_3 = \begin{array}{c}
\end{array}$. And the correspondence between the highest weight $\Lambda$ with Dynkin labels $\lambda_i$ and the tableau $Y$ with rows of length $f_i$ is as follows

$$\Lambda = (\lambda_1, \cdots, \lambda_{n-1}) \leftrightarrow Y = (f_i = \sum_{j=1}^{n-1} \lambda_j) . \quad (3-80)$$

In other words, $\lambda_k$ is the number of columns of $Y$ of height $k$, see Fig. 3.7.

*  

**A short bibliography (cont’d)**

The construction of roots and weights is described in many references given above: (Bump; Bröcker et Dieck; Gilmore …) but also in J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Graduate Texts in Mathematics 9, Springer.

The “Big Yellow Book” of P. Di Francesco, P. Mathieu et D. Sénéchal, [DFMS], *Conformal Field Theory*, Springer, contains a wealth of information on simple Lie algebras, their representations, the tensor products of the latter…

For explicit expressions of the constants $N_{\alpha \beta}$, see [Gi], or Wybourne, *Classical groups for physicists*, John Wyley.


Appendix F. The classical algebras of type $A, B, C, D$

F.1 $\mathfrak{sl}(N) = A_{N-1}$

Rank $= l = N - 1$, dimension $N^2 - 1$, Coxeter number $h = N$, dual Coxeter number $h^\vee = N$.

$e_i$, $i = 1, \cdots, N$ a set of vectors in $\mathbb{R}^N$ such that $\sum_{i=1}^N e_i = 0$, $\{e_i, e_j\} = \delta_{ij} - \frac{1}{N}$.

Roots $\alpha_{ij} = e_i - e_j$, $i \neq j = 1, \cdots, N$; positive roots $\alpha_{ij} = e_i - e_j$, $i < j$; their number $|\Delta_+| = N(N - 1)/2$;

simple roots $\alpha_i := \alpha_{i+1} = e_i - e_{i+1}$, $i = 1, \cdots, N - 1$.

Highest root $\theta = \alpha_1 + \cdots + \alpha_{N-1} = 2e_1 + e_2 + \cdots + e_{N-1} = \Lambda_1 + \Lambda_{N-1} = (1, 0, \cdots, 0, 1)$.

Sum of positive roots

$$2\rho = (N-1)e_1 + (N-3)e_2 + \cdots + (N-2i+1)e_i + \cdots - (N-1)e_N$$
$$= (N-1)\alpha_1 + 2(N-2)\alpha_2 + \cdots + i(N-i)\alpha_i + \cdots + (N-1)\alpha_{N-1}. \quad (F-1)$$

Cartan matrix $(\alpha_i, \alpha_j) = \begin{cases} 2 & \text{if } i = j \\ -1 & \text{if } i = j \pm 1 \\ 0 & \text{otherwise} \end{cases}$

Fundamental weights $\Lambda_i i = 1, \cdots, N - 1$, $\Lambda_i = \sum_{j=1}^i e_j$, $e_1 = \Lambda_1$, $e_i = \Lambda_i - \Lambda_{i-1}$ for $i = 2, \cdots, N - 1$, $e_N = -\Lambda_{N-1}$.

$\langle \Lambda_i, \Lambda_j \rangle = \frac{i(N-j)}{N}$ for $i \leq j$.

Weyl group: $W \equiv S_N$ acts on the weights by permuting the $e_i$: $w \in W \leftrightarrow \bar{w} \in S_N$: $w(e_i) = e_{\bar{w}(i)}$.

Coxeter exponents $\{1, 2, \cdots, N-1\}$.

F.2 $\mathfrak{so}(2l + 1) = B_l$, $l \geq 2$

Rank $= l$, dimension $l(2l + 1)$, Coxeter number $h = 2l$, dual Coxeter number $h^\vee = 2l - 1$.

$e_i$, $i = 1, \cdots, l$, $(e_i, e_j) = \delta_{ij}$ a basis of $\mathbb{R}^l$.

Roots $\pm e_i$, $1 \leq i \leq l$ and $\pm e_i \pm e_j$, $1 \leq i < j \leq l$. Basis of simple roots $\alpha_i = e_i - e_{i+1}$, $i = 1, \cdots, l - 1$, and $\alpha_l = e_l$.

Positive roots

$e_i = \sum_{i \leq k \leq l} \alpha_k$, $1 \leq i \leq l$,

$e_i - e_j = \sum_{i \leq k < j} \alpha_k$, $1 \leq i < j \leq l$, \quad (F-2)

$e_i + e_j = \sum_{i \leq k < j} \alpha_k + 2 \sum_{j \leq k \leq l} \alpha_k$, $1 \leq i < j \leq l$,

their number is $|\Delta_+| = l^2$.

Highest root $\theta = e_1 + e_2 = \alpha_1 + 2\alpha_2 + \cdots + 2\alpha_l$.

Sum of positive roots

$$2\rho = (2l - 1)e_1 + (2l - 3)e_2 + \cdots + (2l - 2i + 1)e_i + \cdots + 3e_{l-1} + e_l$$
$$= (2l - 1)\alpha_1 + 2(2l - 2)\alpha_2 + \cdots + i(2l - i)\alpha_i + \cdots + l^2\alpha_l. \quad (F-3)$$

Cartan matrix $(\alpha_i, \alpha_j) = \begin{cases} 2 & \text{if } 1 \leq i = j \leq l \\ -1 & \text{if } 1 \leq i = (j \pm 1) \leq l - 1 \\ -2 & \text{if } i = l - 1, j = l \\ -1 & \text{if } i = l, j = l - 1 \\ 0 & \text{otherwise} \end{cases}$
Fundamental weights \( \Lambda_i = \sum_{j=1}^{i} e_j, \ i = 1, \ldots, l-1, \ \Lambda_l = \frac{1}{2} \sum_{j=1}^{l} e_j; \) hence \( e_1 = \Lambda_1 = (1,0,\cdots,0), \ e_i = \Lambda_i - \Lambda_{i-1} = (0,\cdots,-1,i,0,\cdots), \ i = 2,\cdots,l-1, \ e_l = 2\Lambda_l - \Lambda_{l-1} = (0,\cdots,0,-1,2). \)

Dynkin labels of the roots
\[ \alpha_1 = (2,-1,0,\cdots), \ \alpha_i = (0,\cdots,-1,2,-1,0,\cdots), \ i = 2,\cdots,l-2; \ \alpha_{l-1} = (0,\cdots,0,-1,2,-2); \ \alpha_l = (0,\cdots,0,-1,2) \]
and \( \theta = (0,1,0,\cdots,0) \)

Weyl group: \( W \cong S_l \ltimes (\mathbb{Z}_2)^l \) of order \( 2^l l! \), acts on the weights by permuting the \( e_i \) and \( e_i \mapsto (\pm 1) e_i \).

Coxeter exponents \( \{1,3,5,\cdots,2l-1\} \).

\[ \text{F.3.} \ \mathfrak{sp}(2l) = C_l, \ l \geq 2 \]

Rank = \( l \), dimension \( l(2l+1) \), Coxeter number \( h = 2l \), dual Coxeter number \( h^* = l+1 \)
\( e_i, \ i = 1,\cdots,l, \ (e_i,e_j) = \frac{1}{2} \delta_{ij} \) a basis of \( \mathbb{R}^l \) \( \text{(Beware! factor 2 to enforce the normalisation } \theta^2 = 2) \). Basis of simple roots \( \alpha_i = e_i - e_{i+1}, \ i = 1,\cdots,l-1, \) and \( \alpha_l = 2e_l \).

Roots \( \pm 2e_i, 1 \leq i \leq l \) and \( \pm e_i \pm e_j, \ 1 \leq i < j \leq l \).

Positive roots
\[
e_i - e_j = \sum_{i \leq k < j} \alpha_k, \quad 1 \leq i < j \leq l, \\
e_i + e_j = \sum_{i \leq k < j} \alpha_k + 2 \sum_{j \leq k < l} \alpha_k + \alpha_l, \quad 1 \leq i < j \leq l, \\
2e_i = 2 \sum_{i \leq k < l} \alpha_k + \alpha_l, \quad 1 \leq i \leq l,
\]
their number is \( |\Delta_+| = l^2 \).

Highest root \( \theta = 2e_1 = 2\alpha_1 + 2\alpha_2 + \cdots + 2\alpha_{l-1} + \alpha_l \).

Sum of positive roots
\[
2\rho = 2le_1 + (2l-2)e_2 + \cdots + (2l-2i+2)e_i + \cdots + 4e_{l-1} + 2e_l = 2\alpha_1 + 2(2l-1)\alpha_2 + \cdots + i(2l-i+1)\alpha_i + \cdots + (l-1)(l-2)\alpha_{l-1} + \frac{1}{2}l(l+1)\alpha_l.
\]

Cartan matrix \( (\alpha_i,\alpha_j^\vee) \)
\[
\begin{bmatrix}
2 & \text{if } 1 \leq i = j \leq l \\
-1 & \text{if } 1 \leq i = (j \pm 1) \leq l-1 \\
-1 & \text{if } i = l-1, j = l \\
-2 & \text{if } i = l, j = l-1 \\
0 & \text{otherwise}
\end{bmatrix}
\]

Fundamental weights \( \Lambda_i = \sum_{j=1}^{i} e_j, \ i = 1,\cdots,l, \) hence \( e_1 = \Lambda_1 = (1,0,\cdots,0), \ e_i = \Lambda_i - \Lambda_{i-1} = (0,\cdots,-1,i,0,\cdots), \ i = 2,\ldots,l \).

Dynkin labels of the roots
\( \alpha_1 = (2,-1,0,\cdots), \ \alpha_i = (0,\cdots,-1,2,-1,0,\cdots), \ i = 2,\cdots,l-1; \ \alpha_l = (0,\cdots,0,-2,2) \) and \( \theta = (2,0,\cdots,0) \)

Weyl group: \( W \cong S_l \ltimes (\mathbb{Z}_2)^l \) of order \( 2^l l! \), acts on the weights by permuting the \( e_i \) and \( e_i \mapsto (\pm 1) e_i \).

\[ \text{F.4.} \ \mathfrak{so}(2l) = D_l, \ l \geq 3 \]

Rank = \( l \), dimension \( l(2l-1) \), Coxeter number = dual Coxeter number \( h = 2l-2 = h^* \)
\( e_i, \ i = 1,\cdots,l, \ (e_i,e_j) = \delta_{ij} \) a basis of \( \mathbb{R}^l \).

Basis of simple roots \( \alpha_i = e_i - e_{i+1}, \ i = 1,\cdots,l-1, \) and \( \alpha_l = e_{l-1} + e_l \).
3.5. YOUNG TABLEAUX AND REPRESENTATIONS OF GL(N) AND SU(N)

Positive roots

\[ e_i - e_j = \sum_{i \leq k \leq j \leq l} \alpha_k, \quad 1 \leq i < j \leq l, \]
\[ e_i + e_j = \sum_{i \leq k \leq j < l} \alpha_k + \alpha_{l-1} + \alpha_l, \quad 1 \leq i < j \leq l - 1, \]
\[ e_i + e_l = \sum_{i \leq k \leq l-2} \alpha_k + \alpha_l, \quad 1 \leq i \leq l - 1, \]

their number is \( |\Delta_+| = l(l-1) \).

Highest root \( \theta = e_1 + e_2 = \alpha_1 + 2\alpha_2 + \cdots + 2\alpha_{l-2} + \alpha_{l-1} + \alpha_l \).

Sum of positive roots

\[ 2\rho = 2(l-1)e_1 + 2(l-2)e_2 + \cdots + 2e_{l-1} = 2(l-1)\alpha_1 + 2(2l-3)\alpha_2 + \cdots + i(2l-i-1)\alpha_i + \cdots + \frac{l(l-1)}{2}(\alpha_{l-1} + \alpha_l). \]

Weyl group: \( W \equiv S_l \rtimes (\mathbb{Z}_2)^{l-1} \), of order \( 2^{l-1}l! \), acts on the weights by permuting the \( e_i \) and \( e_i \mapsto (-1)^i e_i \), with \( \prod_i (-1)^i = 1 \).

Coxeter exponents \( \{1, 3, 5, \ldots, 2l-3, l-1\} \), with \( l-1 \) appearing twice if \( l \) is even.

Cartan matrix \( (\alpha_i, \alpha_j) = \)

\[
\begin{cases}
2 & \text{if } 1 \leq i = j \leq l \\
-1 & \text{if } 1 \leq i = (j \pm 1) \leq l - 2 \\
-1 & \text{if } (i, j) = (l - 2, l) \text{ or } (l, l - 2) \\
0 & \text{otherwise}
\end{cases}
\]

Fundamental weights \( \Lambda_i = \sum_{j=1}^i e_j = \alpha_1 + 2\alpha_2 + \cdots + (i-1)\alpha_{i-1} + i(\alpha_i + \cdots + \alpha_{l-2}) + \frac{i}{2}(\alpha_{l-1} + \alpha_l) \) for \( i = 1, \cdots, l-2 \); \( \Lambda_{l-1} = \frac{1}{2}(e_1 + \cdots + e_{l-1} - e_l) = \frac{1}{2}(\alpha_1 + 2\alpha_2 + \cdots + (l-2)\alpha_{l-2}) + \frac{1}{2}\alpha_{l-1} + \frac{l-2}{2}\alpha_l \); \( \Lambda_l = \frac{1}{2}(e_1 + \cdots + e_{l-1} + e_l) = \frac{1}{2}(\alpha_1 + 2\alpha_2 + \cdots + (l-2)\alpha_{l-2}) + \frac{l-2}{2}\alpha_{l-1} + \frac{l}{2}\alpha_l \).

For the exceptional algebras of types \( E, F, G \), see Bourbaki.
Exercises for chapter 3

A. Cartan algebra and roots

1. Show that any element $X$ of $\mathfrak{g}$ may be written as $X = \sum x^i H_i + \sum_{\alpha \in \Delta} x^\alpha E_\alpha$ with the notations of § 3.1.2.

For an arbitrary $H$ in the Cartan algebra, determine the action of $\text{ad}(H)$ on such a vector $X$; conclude that $\text{ad}(H)\text{ad}(H')X = \sum_{\alpha \in \Delta} x^\alpha(a(H)\alpha(H'))E_\alpha$ and taking into account that the eigenspace of each root $\alpha$ has dimension 1, cf point (∗) 2. of § 3.1.2, that the Killing form reads

$$(H, H') = \text{tr}(\text{ad}(H)\text{ad}(H')) = \sum_{\alpha \in \Delta} \alpha(H)\alpha(H').$$  \hspace{1cm} (3-81)

2. One wants to show that roots $\alpha$ defined by (3-5) or (3-6) generate all the dual space $\mathfrak{h}^\ast$ of the Cartan subalgebra $\mathfrak{h}$. Prove that if it were not so, there would exist an element $H$ of $\mathfrak{h}$ such that $\forall \alpha \in \Delta \alpha(H) = 0$.

Using (3-81) show that this would imply $\forall H' \in \mathfrak{h}$, $(H, H') = 0$. Why is that impossible in a semi-simple algebra? (see the discussion before equation (3-10)).

3. Variant of the previous argument: under the assumption of 2. and thus of (3-82), show that $H$ would commute with all $H_i$ and all the $E_\alpha$, thus would belong to the center of $\mathfrak{g}$. Prove that the center of an algebra is an abelian ideal. Conclude in the case of a semi-simple algebra.

B. Computation of the $N_{\alpha\beta}$

1. Show that the real constants $N_{\alpha\beta}$ satisfy $N_{\alpha\beta} = -N_{\beta\alpha}$ and, by complex conjugation of $[E_\alpha, E_\beta] = N_{\alpha\beta}E_{\alpha+\beta}$ that

$$N_{\alpha\beta} = -N_{-\alpha,-\beta}. \hspace{1cm} (3-83)$$

2. Consider three roots satisfying $\alpha + \beta + \gamma = 0$. Writing the Jacobi identity for the triplet $E_\alpha, E_\beta, E_\gamma$, show that $\alpha(\gamma)N_{\beta\gamma} + \text{cycl.} = 0$. Derive from it the relation

$$N_{\alpha\beta} = N_{\beta,-\alpha,-\beta} = N_{-\alpha,-\beta,\alpha}. \hspace{1cm} (3-84)$$

3. Considering the $\alpha$-chain through $\beta$ and the two integers $p$ et $q$ defined in § 3.2.1, write the Jacobi identity for $E_\alpha, E_\alpha$ and $E_{\beta+k\alpha}$, with $p \leq k \leq q$, and show that it implies

$$\langle \alpha, \beta + k\alpha \rangle = N_{-\alpha,\beta+k\alpha}N_{\alpha,\beta+(k-1)\alpha} + N_{\beta+k\alpha,\alpha}N_{-\alpha,\beta+(k+1)\alpha}.$$

Let $f(k) := N_{\alpha,\beta+k\alpha}N_{-\alpha,-\beta-k\alpha}$. Using the relations (3-84), show that the previous equation may be recast as

$$\langle \alpha, \beta + k\alpha \rangle = f(k) - f(k - 1). \hspace{1cm} (3-85)$$

4. What are $f(q)$ et $f(q - 1)$? Show that the recursion relation (3-85) is solved by

$$f(k) = -(N_{\alpha,\beta+k\alpha})^2 = (k - q)(\alpha, \beta + \frac{1}{2}(k + q + 1)\alpha). \hspace{1cm} (3-86)$$

What is $f(p - 1)$? Show that the expression (3-21) is recovered. Show that (3-86) is in accord with (3-23). The sign of the square root is still to be determined . . . , see [Gi].
3.5. YOUNG TABLEAUX AND REPRESENTATIONS OF GL(N) AND SU(N)

C. Study of the $B_l = so(2l + 1)$ and $G_2$ algebras

1. $so(2l + 1) = B_l$, $l \geq 2$
   a. What is the dimension of the group $SO(2l + 1)$ or of its Lie algebra $so(2l + 1)$?
   b. What is the rank of the algebra? (Hint: diagonalize a matrix of $so(2l + 1)$ on $\mathbb{C}$, or write it as a diagonal of real $2 \times 2$ blocks, see lecture notes, §3.1)
   c. How many roots does the algebra have? How many positive roots? How many simple?
   d. Let $e_i$, $i = 1, \cdots, l$ be a orthonormal basis in $\mathbb{R}^l$, $\langle e_i, e_j \rangle = \delta_{ij}$. Consider the set of vectors
      \[
      \Delta = \{ \pm e_i, \ 1 \leq i \leq l \} \cup \{ \pm e_i \pm e_j, \ 1 \leq i < j \leq l \}
      \]
      What is the cardinal of $\Delta$? $\Delta$ is the set of roots of the algebra $so(2l + 1)$.
   e. A basis of simple roots is given $\alpha_i = e_i - e_{i+1}$, $i = 1, \cdots, l-1$, et $\alpha_l = e_l$. Explain why
      the roots
      \[
      e_i = \sum_{i \leq k \leq l} \alpha_k, \ 1 \leq i \leq l,
      \]
      \[
      e_i - e_j = \sum_{i \leq k < j} \alpha_k, \ 1 \leq i < j \leq l,
      \]
      \[
      e_i + e_j = \sum_{i \leq k < j} \alpha_k + 2 \sum_{j \leq k \leq l} \alpha_k, \ 1 \leq i < j \leq l,
      \]
      qualify as positive roots. Check that assertion on the case of $B_2 = so(5)$.
   f. Compute the Cartan matrix and check that it agrees with the Dynkin diagram given in the notes.
   g. Compute the sum $\rho$ of positive roots.
   h. The Weyl is the (“semi-direct”) product $W \equiv S_l \ltimes (Z_2)^l$, which acts on the $e_i$ (hence on weights and roots) by permutation and by independant sign changes $e_i \mapsto (\pm 1)e_i$. What is its order? In the case of $B_2$, check that assertion and draw the first Weyl chamber.
   i. Show that the vectors $\Lambda_i = \sum_{j=1}^l e_j$, $i = 1, \cdots, l - 1$, $\Lambda_l = \frac{1}{2} \sum_{j=1}^l e_j$ are the fundamental weights.
   j. Using Weyl formula: $\dim(\Lambda) = \prod_{\alpha > 0} \frac{\langle \Lambda + \rho, \alpha \rangle}{\langle \rho, \alpha \rangle}$ compute the dimension of the two fundamental representations of $B_2$ and of that of highest weight $2\Lambda_2$. In view of these dimensions, what are these representations of $SO(5)$?
   k. Draw on the same figure the roots and the low lying weights of $so(5)$.

2. $G_2$

In the space $\mathbb{R}^2$, we consider three vectors $e_1, e_2, e_3$ of vanishing sum, $\langle e_i, e_j \rangle = \delta_{ij} - \frac{1}{3}$, and construct the 12 vectors
      \[
      \pm(e_1 - e_2), \ \pm(e_1 - e_3), \ \pm(e_2 - e_3), \ \pm(2e_1 - e_2 - e_3), \ \pm(2e_2 - e_1 - e_3), \ \pm(2e_3 - e_1 - e_2)
      \]
      They make the root system of $G_2$, as we shall check.
   a. What can be said on the dimension of the algebra $G_2$?
b. Show that \( \alpha_1 = \epsilon_1 - \epsilon_2 \) and \( \alpha_2 = -2\epsilon_1 + \epsilon_2 + \epsilon_3 \) are two simple roots, in accord with the Dynkin diagram of \( G_2 \) given in the notes. Compute the Cartan matrix.

c. What are the positive roots? Compute the vector \( \rho \), half-sum of positive roots.

d. What is the group of invariance of the root diagram? Show that it is of order 12 and that it is the Weyl group of \( G_2 \). Draw the first Weyl chamber.

e. Check that the fundamental weights are

\[
\Lambda_1 = 2\alpha_1 + \alpha_2 \quad \Lambda_2 = 3\alpha_1 + 2\alpha_2
\]

f. What are the dimensions of the fundament representations?

g. In the two cases of \( B_2 \) and \( G_2 \), one observes that the highest weight of the adjoint representation is given by the highest root. Explain why this is true in general.

3. A little touch of particle physics

Why were the groups \( SO(5) \) or \( G_2 \) inappropriate as symmetry groups extending the \( SU(2) \) isospin group, knowing that several “octets” of particles had been observed?

D. Dimensions of \( SU(3) \) representations

We admit that the construction of \( \S \) 3.4.2, of completely symmetric traceless rank \((p, m)\) tensors in \( \mathbb{C}^3 \), does give the irreducible representations of \( SU(3) \) of highest weight \((p, m)\). Then we want to determine the dimension \( d(p, m) \) of the space of these tensors.

1. Show, by studying of the product of two tensors of rank \((p, 0)\) and \((0, m)\) and separating the trace terms (those containing a \( \delta^i_j \) between one lower and one upper index) that \((p, 0) \otimes (0, m) = ((p - 1, 0) \otimes (0, m - 1)) \oplus (p, m)\) and thus that

\[
d(p, m) = d(p, 0)d(0, m) - d(p - 1)d(0, m - 1) .
\]

2. Show by a computation analogous to that of \( SU(2) \) that

\[
d(p, 0) = d(0, p) = \frac{1}{2}(p + 1)(p + 2) .
\]

3. Derive from it the expression of \( d(p, m) \) and compare with (3-63).

*
Chapter 4

Global symmetries in particle physics

Particle physics offers a wonderful playground to illustrate the various manifestations of symmetries in physics. We will be only concerned in this chapter and the following one with “internal symmetries”, excluding space-time symmetries.

We shall examine in turn various types of symmetries and their realizations, as exact symmetries, or broken explicitly, spontaneously or by quantum anomalies.

4.1 Global exact or broken symmetries. Spontaneous breaking

4.1.1 Overview. Exact or broken symmetries

Transformations that concern us in this chapter are *global symmetries* and we discuss them in the framework of (classical or quantum) field theory. A group \( G \) acts on degrees of freedom of each field \( \phi(x) \) in the same way at all points \( x \) of space-time. For example, \( G \) acts on \( \phi \) by a linear representation, and to each element \( g \) of the group corresponds a matrix or operator \( D(g) \), independent of the point \( x \).

\[
\phi(x) \mapsto D(g)\phi(x) .
\] (4-1)

In a quantum theory, according to Wigner theorem, one assumes that this transformation is also realized on vector-states of the Hilbert space of the theory by a unitary operator \( U(g) \), and, as an operator, \( \phi(x) \mapsto U(g)\phi(x)U^\dagger(x) \).

This transformation may be a *symmetry* of dynamics, in which case \( U(g) \) commutes with the Hamiltonian of the system, or in the Lagrangian picture, it leaves the Lagrangian invariant and gives rise to Noether currents \( j_\mu^i \) of vanishing divergence (see chap.0) and to conserved charges \( Q_i = \int d^4x j^0_i(x,t) \), \( i = 1, \ldots, \dim G \). These charges act on fields as infinitesimal generators, classically in the sense of the Poisson bracket, \( \{ Q_i, \phi(x) \} \delta \alpha^i = \delta \phi(x) \), (see Chapter 0), and *if everything goes right in the quantum theory*, as operators in the Hilbert space with commutation
relations with the fields \([Q_i, \phi(x)]\delta \alpha^i = -i\hbar \delta \phi(x)\) and between themselves \([Q_i, Q_j] = iC_{ij}^k Q_k\).

An important question will be indeed to know if a symmetry which is manifest at the classical level, say on the Lagrangian, is actually realized in the quantum theory.

- An example of exact symmetry is provided by the \(U(1)\) invariance associated with electric charge conservation. A field carrying an electric charge \(q\) (times \(|e|\)) is a complex field, it transforms under the action of the group \(U(1)\) according to the irreducible representation labelled by the integer \(q\)

\[
\phi(x) \mapsto e^{i q \alpha} \phi(x) ; \quad \phi^\dagger(x) \mapsto e^{-i q \alpha} \phi^\dagger(x),
\]

and there is invariance (of the Lagrangian) if all fields transform that way, with a Noether current \(j^\mu(x)\), sum of contributions of the different charged fields, being divergence-less, \(\partial_\mu j^\mu(x) = 0\), and the associated charge \(Q = \sum q_i\) is conserved. The quantum theory is quantum electrodynamics, and there one proves that the classical \(U(1)\) symmetry, the current conservation (and gauge invariance) are preserved by quantization and in particular by renormalization, for example that all electric charges renormalize in the same way, see the course of Quantum Field Theory.

Other invariances and conservation laws of a similar nature are those associated with baryonic or leptonic charges, which are conserved (until further notice . . . ).

- A symmetry may also be broken explicitly. For example the Lagrangian contains terms that are non invariant under the action of \(G\). In that case, the Noether currents are non conserved, but their divergence reads

\[
\partial_\mu j^\mu_i(x) = \frac{\partial L(x)}{\partial \phi_i},
\]

see Chapitre 0, § 4.2. We will see below with flavor \(SU(3)\) an example of a broken (or “approximate”) symmetry.

Certain types of breakings, called “soft”, are such that the symmetry is restored at short distance or high energy. This is for example the case of scale invariance (by space dilatations), broken by the presence of any mass scale in the theory, but restored—in a fairly subtle way—at short distance, see the study of the Renormalization Group in the courses of quantum or statistical field theory.

- A more subtle mechanism of symmetry breaking is that of spontaneous symmetry breaking. This refers to situations where the ground state of the system does not have a symmetry apparent on the Lagrangian or on the equations of motion. The simplest illustration of this phenomenon is provided by a classical system with one degree of freedom, described by the “double well potential” of Fig. 4.1(a). Although the potential exhibits a manifest \(Z_2\) symmetry under \(x \rightarrow -x\), the system chooses a ground state in one of the two minima of the potential, which breaks symmetry. This mechanism plays a fundamental role in physics, with diverse manifestations ranging from condensed matter—ferromagnetism, superfluidity, supraconductivity. . . —to particle physics—chiral symmetry, Higgs phenomenon— and cosmology.

Example. Spontaneous breaking in the \(O(n)\) model

The Lagrangian of the bosonic (and Minkovskian, here) “\(O(n)\) model” for a real \(n\)-component field \(\phi = \{\phi^i\}\)

\[
\mathcal{L} = \frac{1}{2} (\partial \phi)^2 - \frac{1}{2} m^2 \phi^2 - \frac{\lambda}{4} (\phi^2)^2
\]
4.1. GLOBAL EXACT OR BROKEN SYMMETRIES. SPONTANEOUS BREAKING

Figure 4.1: Potentials (a) with a “double well”; (b) “mexican hat”

is invariant under the $O(n)$ rotation group. The Noether current $j_\mu^a = \partial_\mu \phi^i (T^a)_{ij} \delta^j$ (with $T^a$ real antisymmetric) has a vanishing divergence, which implies the conservation of a “charge” etc. The minimum of the potential corresponds to the ground state, alias the vacuum, of the theory. If the parameter $m^2$ is taken negative, the minimum of the potential $V = \frac{1}{2} m^2 \phi^2 + \lambda \phi^4$ is no longer at $\phi^2 = 0$ but at some value $v^2$ of $\phi^2$ such that $-m^2 = \lambda v^2$, see Fig. 4.1(b). The field $\phi$ “chooses” spontaneously a direction $\hat{n} (\hat{n}^2 = 1)$ in the internal space, in which its vacuum expectation value (“vev” in the jargon) is non vanishing $\langle 0 | \phi | 0 \rangle = v \hat{n}$. (4-4)

This “vev” breaks the initial invariance group $G = O(n)$ down to its subgroup $H$ that leaves invariant the vector $\langle 0 | \phi | 0 \rangle = v \hat{n}$, hence a group isomorphic to $O(n - 1)$. The fact that a vacuum expectation value of a non invariant field be non zero, $\langle 0 | \phi | 0 \rangle \neq 0$, signals that the vacuum is not invariant: this is a case of spontaneous symmetry breaking. This is the mechanism at work in a low temperature ferromagnet, for example, in which the non zero magnetization signals the spontaneous breaking of the space rotation symmetry.

Exercise (see F. David's course): Set $\phi = (v + \sigma) \hat{n} + \pi$, where $\pi$ denote the $n - 1$ components of the field $\phi$ orthogonal to $\langle \phi \rangle = v \hat{n}$, determine the terms of $V(\sigma, \pi)$ that are linear and quadratic in the fields $\sigma$ and $\pi$; check that the linear term in $\sigma$ vanishes (minimum of the potential), that $\sigma$ has a non-zero mass term, but that the $\pi$ are massless, they are the Nambu–Goldstone bosons of the spontaneously broken symmetry. This is a general phenomenon: any continuous spontaneously broken symmetry is accompanied by the appearance of massless excitations whose number equals that of the generators of the broken symmetry (Goldstone theorem). More precisely when a group $G$ is spontaneously broken into a subgroup $H$ (group of residual symmetry, invariance group of the ground state), a number $d(G) - d(H)$ of massless Goldstone bosons appears. In the previous example, $G = O(n)$, $H = O(n - 1)$, $d(G) - d(H) = n - 1$.

Let us give a simple proof of that theorem in the case of a Lagrangian field theory. We write $L = \frac{1}{2} (\partial \phi)^2 - V(\phi)$ with quite generic notations, $\phi$ denotes a set of fields $\{ \phi_i \}$ on which acts a continuous transformation group $G$. The potential $V$ is assumed to be invariant under the action infinitesimal transformations $\delta^a \phi_i$, $a = 1, \cdots, \text{dim} G$. For example for linear transformations: $\delta^a \phi_i = T^a_{ij} \phi_j$. We thus have

$$\frac{\partial V(\phi(x))}{\partial \phi_i(x)} \delta^a \phi_i(x) = 0.$$
Differentiate this equation with respect to \( \phi_j(x) \) (omitting everywhere the argument \( x \))

\[
\frac{\partial V}{\partial \phi_i} \frac{\partial \delta \phi_i}{\partial \phi_j} + \frac{\partial^2 V}{\partial \phi_i \partial \phi_j} \delta^a \phi_i = 0
\]

and evaluate it at \( \phi(x) = v \), a (constant, \( x \)-independent) minimum of the potential: the first term vanishes, the second tells us

\[
\left. \frac{\partial^2 V}{\partial \phi_i \partial \phi_j} \right|_{\phi = v} \delta^a v_i = 0 ,
\]

where we write (with a little abuse of notation) \( \delta^a v_i = \delta^a \phi_i |_{\phi = v} \). On the other hand, the theory is quantized near that minimum \( v \) ("vacuum" of the theory) by writing \( \phi(x) = v + \varphi(x) \) and by expanding

\[
V(\phi) = V(v) + \frac{1}{2} \frac{\partial^2 V}{\partial \phi_i \partial \phi_j} \big|_{\phi = v} \varphi_i \varphi_j + \cdots
\]

and the masses of the fields \( \varphi \) are then read off the quadratic form. But (4-5) tells us that the "mass matrix"

\[
\frac{\partial^2 V}{\partial \phi_i \partial \phi_j} |_{\phi = v}
\]

has as many "zero modes" (eigenvalues of vanishing eigenvalue) as there are independent variations \( \delta^a v_i \neq 0 \). If \( H \) is the invariance group of \( v \), \( \delta^a v_i \neq 0 \) for the generators of \( G \) that are not generators of \( H \), and there are indeed \( \dim G - \dim H \) massless modes, qed.

### 4.1.2 Chiral symmetry breaking

Consider a Lagrangian that involves massless fermions

\[
\mathcal{L} = \bar{\psi}i\gamma^\mu \psi + g(\bar{\psi} \gamma^\mu \psi)(\bar{\psi} \gamma_5 \psi) ,
\]

where \( \psi = \{ \psi_\alpha \}_{\alpha = 1, \ldots, N} \) is a \( N \)-component vector. That Lagrangian is invariant under the action of two types of infinitesimal transformations

\[
\delta \psi(x) = \delta A \psi(x) \quad (4-7)
\]

\[
\delta \psi(x) = \delta B \gamma_5 \psi(x) ,
\]

where the matrices \( A \) and \( B \) are infinitesimal \( N \times N \) antihermitian. The conserved Noether currents are respectively

\[
J_{\mu}^a = \bar{\psi} T^a \gamma_\mu \psi \quad J_{\mu}^{a(5)} = \bar{\psi} T^a \gamma_5 \gamma_\mu \psi ,
\]

with \( T^a \) infinitesimal generators of the unitary group \( U(N) \). The transformations of the first line are called "vector", those of the second, which involve \( \gamma_5 \), are "axial". One may also rephrase it in terms of independent transformations of \( \psi_L := \frac{1}{2}(I - \gamma_5) \psi \) and of \( \psi_R := \frac{1}{2}(I + \gamma_5) \psi \); one recalls that \( (\gamma_5)^2 = I \), that \( \frac{1}{2}(I \pm \gamma_5) \) are thus projectors and that \( \gamma_5 = \gamma_5^\dagger \) anticommutes with the \( \gamma_\mu \); one has thus \( \bar{\psi}_L = \psi_L^\dagger \gamma_0 = \frac{1}{2} \bar{\psi}(I + \gamma_5) \), etc, and

\[
\mathcal{L} = \bar{\psi}_L i\gamma^\mu \psi_L + \bar{\psi}_R i\gamma^\mu \psi_R + (\bar{\psi}_L \gamma_\mu \psi_L + \bar{\psi}_R \gamma_\mu \psi_R)(\bar{\psi}_L \gamma^\mu \psi_L + \bar{\psi}_R \gamma^\mu \psi_R)
\]

which is clearly invariant under the finite unitary transformations \( \psi_L \rightarrow U_1 \psi_L , \psi_R \rightarrow U_2 \psi_R \), with \( U_1 , U_2 \in U(N) \). The group of chiral symmetry is thus \( U(N) \times U(N) \).

If we now introduce a mass term \( \delta \mathcal{L} = -m \bar{\psi} \psi \) (which "couples" the left and right components \( \psi_L \) and \( \psi_R \): \( \delta \mathcal{L} = -m(\bar{\psi}_R \psi_L + \bar{\psi}_L \psi_R) \)), the "vector" symmetry is preserved, but the axial one is not and gives rise to a divergence

\[
\partial^\mu J_{\mu}^{a(5)}(x) \propto m \bar{\psi} T^a \gamma_5 \psi .
\]
The residual symmetry group is $U(N)$, “diagonal” subgroup of $U(N) \times U(N)$ (diagonal in the sense that one takes $U_1 = U_2$ in the transformations of $\psi_{L,R}$.)

The axial symmetry may also be spontaneously broken. Let us start from a Lagrangian, sum of terms of the type (4-6) with $N = 2$ and (4-3) for $n = 4$, with a coupling term between the fermions and the four bosons, traditionally denoted $\sigma$ and $\pi$

$$\mathcal{L} = \bar{\psi}(i\partial + g(\sigma + i\pi \tau\gamma_5))\psi + \frac{1}{2}((\partial\sigma)^2 + (\partial\sigma)^2) - \frac{1}{2}m^2(\sigma^2 + \pi^2) - \frac{\lambda}{4}(\sigma^2 + \pi^2)^2,$$ (4-10)

in which the Pauli matrices have been exceptionally denoted by $\tau$ not to confuse them with the field $\sigma$. The symmetry group is $U(2) \times U(2)$, with fields $\psi_L, \psi_R$ and $\sigma + i\pi \tau$ transforming respectively by the representations $\left(\frac{1}{2}, 0\right), \left(0, \frac{1}{2}\right)$ and $\left(\frac{1}{2}, \frac{1}{2}\right)$ of $SU(2) \times SU(2)$ (see exercise B). If $m^2 < 0$, the field $\phi = (\sigma, \pi)$ develops a non-zero vev, that may be oriented in the direction $\sigma$ if one has initially introduced a small explicit breaking term $\delta L = c\sigma$, the analogue of a small magnetic field, which is then turned off. The vev is given as above by $v = -m^2/\lambda$, and, rewriting $\sigma(x) = \sigma'(x) + v$, where the field $\sigma'$ has now a vanishing vev, one sees that the fermions have acquired a mass $m_\psi = -g v$, whereas the $\pi$ are massless. This Lagrangian, the $\sigma$-model of Gell-Mann–Lévy, has been proposed as a model explaining the chiral symmetry breaking and the low mass of the $\pi$ mesons, regarded as “quasi Nambu–Goldstone quasi-bosons” (“quasi” in the sense that the chiral symmetry is only approximate before being spontaneously broken). Some elements of that model will reappear in the standard model.

### 4.1.3 Quantum symmetry breaking. Anomalies

Another mode of symmetry breaking, of purely quantum nature, manifests itself in anomalies of quantum field theories. A symmetry, which is apparent at the classical level of the Lagrangian, is broken by the effect of “quantum corrections”. This is for instance what takes place with some chiral symmetries of the type just studied: an axial current which is classically divergenceless may acquire by a “one-loop effect” a divergence $\partial_\mu J_\mu^a \neq 0$. If the “anomalous” current is the Noether current of an internal classical symmetry, that symmetry is broken by the quantum anomaly, which may cause interesting physical effects (see discussion of the decay $\pi^0 \to \gamma\gamma$, for example in [IZ] chap 11). But in a theory like a gauge theory where the conservation of the axial current is crucial to ensure consistency –renormalizability, unitarity–, the anomaly constitutes a potential threat that must be controlled. This is what happens in the standard model, and we return to it in Chap. 5. Another example is provided by dilatation (scale) invariance of a massless theory, see the study of the renormalization group in F. David’s course.

### 4.2 The SU(3) flavor symmetry and the quark model.

An important approximate symmetry is the “flavor” SU(3) symmetry, to which we devote the rest of this chapter.

#### 4.2.1 Why SU(3)?

We saw (Chap. 00) that if the weak and electromagnetic interactions are neglected, hadrons, i.e. particles subject to strong interactions such as proton and neutron, $\pi$ mesons etc, fall into
“multiplets” of a SU(2) group of isospin. Or said differently, the Hamiltonian (or Lagrangian) of strong interactions is invariant under the action of that SU(2) group and consequently, the SU(2) group is represented in the space of hadronic states by unitary representations. Proton and neutron belong to a representation of dimension 2 and of isospin $\frac{1}{2}$, the three pions $\pi^\pm, \pi^0$ form a representation of dimension 3 and isospin 1, etc. The electric charge $Q$ of each of these particles is related to the eigenvalue of the third component $I_z$ of isospin by

$$Q = \frac{1}{2}B + I_z \quad \text{[for SU(2)]}$$

(4-11)

where a new quantum number $B$ appears, the baryonic charge, supposed to be (additively) conserved in all interactions (until further notice). $B$ is 0 for $\pi$ mesons, 1 for “baryons” as proton or neutron, $-1$ for their antiparticles, 4 for an $\alpha$ particle (Helium nucleus), etc.

This relation between $Q$ and $I_z$ must be amended for a new family of mesons ($K^\pm, K^0, \bar{K}^0, \cdots$) and baryons $\Lambda^0, \Sigma, \Xi, \ldots$ discovered at the end of the fifties. One assigns them a new quantum number $S$, strangeness. This strangeness is assumed to be additively conserved in strong interactions. Thus, if $S$ is $-1$ for the $\Lambda^0$ and $+1$ for the $K^+$ and the $K^0$, the reaction $p + \pi^- \rightarrow \Lambda^0 + K^0$ conserves strangeness, whereas the observed decay $\Lambda^0 \rightarrow p + \pi^-$ violates that conservation law, as it proceeds through weak interactions. Relation (4-11) must be modified into the Gell-Mann–Nishima relation

$$Q = \frac{1}{2}B + \frac{1}{2}S + I_z = \frac{1}{2}Y + I_z,$$

(4-12)

where we introduced the hypercharge $Y$, which, at this stage, equals $Y = B + S$.

These conservation laws and different properties of mesons and baryons discovered then, in particular their organisation into “octets”, led at the beginning of the sixties M. Gell-Mann and Y. Ne’eman to postulate the existence of a group SU(3) of approximate symmetry of strong interactions. The quantum numbers $I_z$ and $Y$ that are conserved and simultaneously measurable are interpreted as eigenvalues of two commuting charges, hence of two elements of a Cartan algebra of rank 2, and the algebra of SU(3) which is the natural candidate, as it possesses an 8-dimensional representation (see exercise C of Chap. 3). In the defining representation 3 of SU(3), one constructs a basis of the Lie algebra $su(3)$, made of 8 Hermitian matrices $\lambda_a$ that play the role of Pauli matrices $\sigma_i$ for $su(2)$. These matrices are normalised by

$$\text{tr} \lambda_a \lambda_b = 2\delta_{ab} .$$

(4-13)

$\lambda_1$ and $\lambda_2$, $\lambda_4$ and $\lambda_5$, $\lambda_6$ and $\lambda_7$ have the same matrix elements as $\sigma_1$ and $\sigma_2$ at the $\ast$ locations

$$\begin{pmatrix} \ast & \ast & \ast \\ \ast & \ast & \ast \\ \ast & \ast & \ast \end{pmatrix}, \begin{pmatrix} \ast & \ast & \ast \\ \ast & \ast & \ast \\ \ast & \ast & \ast \end{pmatrix}, \begin{pmatrix} \ast & \ast & \ast \\ \ast & \ast & \ast \\ \ast & \ast & \ast \end{pmatrix}$$

respectively, where dots stand for zeros. The two generators of the Cartan algebra are

$$\lambda_3 = \begin{pmatrix} 1 & & \\ & -1 & \\ & & 1 \end{pmatrix}, \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & & \\ & 1 & \\ & & -2 \end{pmatrix} .$$

(4-14)

The charges $I_z$ and $Y$ are then representatives in the representation under study of $\frac{1}{2}\lambda_3$ and $\frac{1}{\sqrt{3}}\lambda_8$. 
4.2. THE SU(3) FLAVOR SYMMETRY AND THE QUARK MODEL.

Figure 4.2: Octets of pseudoscalar \( (J^P = 0^-) \) and of vector mesons \( (J^P = 1^-) \)

Figure 4.3: Baryon octet \( (J^P = \frac{1}{2}^+) \) and decuplet \( (J^P = \frac{3}{2}^+) \)
The matrices $\lambda_a$ satisfy commutation relations
\[ [\lambda_a, \lambda_b] = 2i f_{abc} \lambda_c \] (4-15)
with structure constants (real and completely antisymmetric) $f_{abc}$ of the su(3) Lie algebra. It is useful to also consider the anticommutators
\[ \{\lambda_a, \lambda_b\} = \frac{4}{3} \delta_{ab} + 2 d_{abc} \lambda_c . \] (4-16)
Thanks to (4-13), (4-15) and (4-16) may be rewritten as $\text{tr}(\lambda_a, \lambda_b) = 4$ if $abc$ and $\text{tr}(\{\lambda_a, \lambda_b\}) = 4 d_{abc}$. These numbers $f$ and $d$ are tabulated in the literature ... but they are easily computable! Beware that in contrast with (4-15), relation (4-16) and the (real, completely symmetric) constants $d_{abc}$ are proper to the 3-dimensional representation.

Hadrons are then organized in SU(3) representations. Each multiplet gathers particles with the same spin $J$ and parity $P$. For instance two octets of mesons with $J^P$ equal to $0^-$ or $1^-$ and one octet and one “decuplet” of baryons of baryonic charge $B = 1$ are easily identified. Contrary to isospin symmetry, the SU(3) symmetry \footnote{said to be “of flavour”, according to the modern terminology, moderne, but called “unitary symmetry” or “eightfold way” at the time of Gell-Mann and Ne’eman...} is not an exact symmetry of strong interactions. The conservation laws and selection rules that follow are only approximate.

At this stage one may wonder about the absence of other representations of zero triality, such as the 27, or of those of non zero triality, like the 3 and the 3. We return to that point in § 4.2.5.

### 4.2.2 Consequences of the SU(3) symmetry

#### The octets of fields

Let us look more closely at the two octets of baryons $\mathcal{N} = (N, \Sigma, \Xi, \Lambda)$ and of pseudoscalar mesons $\mathcal{P} = (\pi, K, \eta)$. Recalling what was said in Chap. 3, § 4.2, namely that the adjoint representation is made of traceless tensors of rank $(1, 1)$, it is natural to group the 8 fields associated to these particles in the form of traceless matrices
\[ \Phi = \begin{pmatrix} \frac{1}{\sqrt{2}} \pi^0 - \frac{1}{\sqrt{6}} \eta & \pi^+ & K^+ \\ \pi^- & -\frac{1}{\sqrt{2}} \pi^0 - \frac{1}{\sqrt{6}} \eta & K^0 \\ K^- & \bar{K}^0 & \frac{\sqrt{2}}{\sqrt{3}} \eta \end{pmatrix}, \] (4-17)
and
\[ \Psi = \begin{pmatrix} \frac{1}{\sqrt{2}} \Sigma^0 - \frac{1}{\sqrt{6}} \Lambda & \Sigma^+ & p \\ \Sigma^- & -\frac{1}{\sqrt{2}} \Sigma^0 - \frac{1}{\sqrt{6}} \Lambda & n \\ \Xi^- & \Xi^0 & \frac{2}{\sqrt{3}} \Lambda \end{pmatrix}. \] (4-18)

To make sure that the assignments of fields/particles to the different matrix elements are correct, it suffices to check their charge and hypercharge. The generators of charge $Q$ and hypercharge $Y$
\[ Q = I_z + \frac{1}{2} Y = \frac{1}{3} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad Y = \frac{1}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \] (4-19)
4.2. THE SU(3) FLAVOR SYMMETRY AND THE QUARK MODEL.

act in the adjoint representation by commutation and one has indeed

\[
[Q, \Phi] = \begin{pmatrix} 0 & \pi_+ & K_+ \\ -\pi_- & 0 & 0 \\ -K_- & 0 & 0 \end{pmatrix} \quad [Y, \Phi] = \begin{pmatrix} 0 & 0 & K_+ \\ 0 & 0 & K_0 \\ -K_- & -K_0 & 0 \end{pmatrix}.
\]

Exercises: (i) with no further calculation, what is \([I_z, \Phi]\) ? Check.
(ii) Compute \(\text{tr} \Phi^2\), and explain why the result justifies the choice of normalization of the matrix elements in (4-17). See also Problem 2.c.

Tensor products in SU(3) and invariant couplings

Recall that in SU(3), with notations of Chap. 3,

\[
8 \otimes 8 = 1 \oplus 8 \oplus 8 \oplus 10 \oplus \bar{10} \oplus 27.
\]

(4-20)

Let us show that this has immediate implications on the number of invariant couplings between fields.

- We want to write an SU(3) invariant Lagrangian involving the previous octets of fields \(\Phi\) and \(\Psi\). What is the number of independent “Yukawa couplings”, i.e. of the form \(\bar{\Psi}\Phi\Psi\), that are invariant under SU(3)? In other words, what is the number of (linearly independent) invariants in \(8 \otimes 8 \otimes 8\)? According to a reasoning already done in Chap 2. § 3.2, this number equals the number of times the representation 8 appears in \(8 \otimes 8\), hence, according to (4-20), 2. There are thus two independent invariant Yukawa couplings. If the two octets \(\Psi\) and \(\Phi\) are written as traceless 3 \(\times\) 3 matrices as in (4-17) and (4-18) \(\Psi = \{\psi^i_j\}\) and \(\Phi = \{\phi^i_k\}\), these two couplings read

\[
\text{tr} \bar{\Psi}\Phi\Psi = \bar{\psi}^j_i \psi^k_j \phi^i_k \quad \text{and} \quad \text{tr} \bar{\Psi}\Phi\Psi = \bar{\psi}^j_i \phi^i_k \psi^j_k
\]

(4-21)

(this compact writing omits indices of Dirac spinors, a possible \(\gamma_5\) matrix, etc). An often preferred expression uses the sum and difference of these two terms, hence \(\text{tr} \bar{\Psi}\{\Phi, \Psi\}\) and \(\text{tr} \bar{\Psi}\{\Phi, \Psi\}\), traditionally called \(f\) term and \(d\) term, by reference to (4-15) and (4-16).

- Another question of the same nature is: what is \(a\ priori\) the number of SU(3) invariant amplitudes in the scattering of two particles belonging to the octets \(N\) and \(P\) : \(N_i + P_i \rightarrow N_f + P_f\)? (One takes only SU(3) invariance into account and does not consider possible discrete symmetries.) The issue is thus the number of invariants in the fourth tensor power of representation 8. Or equivalently, the number of times the same representation appears in the two products \(8 \otimes 8\) and \(8 \otimes 8\). If \(m_i\) are the multiplicities appearing in \(8 \otimes 8\), namely \(m_1 = 1, m_8 = 2\), etc, see (4-20), this number is \(\sum_i m_i^2 = 8\). There are thus eight invariant amplitudes. In other words, one may write \(a\ priori\) the scattering amplitude in the form

\[
\langle N_f P_f | T | N_i P_i \rangle = \sum_{r=1}^8 A_r(s,t) \left( (I, I, I, Y)_{(N_f)}, (I, I, I, Y)_{(P_f)} | r, (I, I, I, Y)_{(r)} \right) \langle r, (I, I, I, Y)_{(r)} | (I, I, I, Y)_{(N_i)}, (I, I, I, Y)_{(P_i)} \rangle
\]

(4-22)
(with $s$ and $t$ the usual relativistic invariants $s = (p_1 + p_2)^2$, $t = (p_1 - p_3)^2$), and all the dependence in the nature of the scattered particles, identified by the values of their isospin and hypercharge, is contained in SU(3) Clebsch-Gordan coefficients.

- Let $\Phi_i, i = 1, 2, 3, 4$ be four distinct octet fields. How many quartic (degree 4) SU(3) invariant couplings may be formed with these four fields? On the one hand, the previous argument gives 8 couplings; on the other hand, it is clear that terms $\text{tr}(\Phi P_1 \Phi P_2 \Phi P_3 \Phi P_4)$ and $\text{tr}(\Phi P_1 \Phi P_2) \text{tr}(\Phi P_3 \Phi P_4)$ are invariant under all permutations $P$. A quick counting leads to 9 different terms, in contradiction with the previous argument. Where is the catch? For more, go to the Problem 1 at the end of this chapter...

4.2.3 Electromagnetic breaking of the SU(3) symmetry

The SU(3) symmetry is broken, as we said, by strong interactions. Of course, just like the isospin SU(2) symmetry, it is also broken by electromagnetic and by weak interactions. We won’t examine the latter but describe now two consequences of the strong and electromagnetic breakings.

The interaction Lagrangian of a particle of charge $q$ with the electromagnetic field $A$ reads

$$\mathcal{L}_{em} = -q j^\mu A_\mu$$

where $j$ is the electric current. The field $A$ is invariant under SU(3) transformations, but how does $j$ transform? One knows the transformation of its charge $Q = \int d^4x j_0(x, t)$, since following (4-12), $Q$ is a linear combination of two generators $Y$ and $I_z$. $Q$ thus transforms according to the adjoint representation (8, alias (1, 1) in terms of Dynkin labels). And it is natural to assume that the current $j$ also transforms in the same way. This is indeed what is found when the current $j^\mu$ is regarded as the Noether current of the U(1) symmetry (exercise, check it).

Magnetic moments

The electromagnetic form factors of the baryon octet are defined as

$$\langle B|j_\mu(x)|B'\rangle = e^{ikx}\tilde{u}(F^B_{ee}(k^2)\gamma_\mu + F^B_{em}(k^2)\sigma_\mu\nu k^\nu)u'$$

where $\tilde{u}$ and $u'$ are Dirac spinors which describe respectively the baryons $B$ and $B'$; $k$ is the four-momentum transfer from $B'$ to $B$. $F_e$ is the electric form factor, if $B = B'$, $F_e(0) = q_B$, the electric charge of $B$, whereas $F_m$ is the magnetic form factor and $F^B_{mm}(0)$ gives the magnetic moment of baryon $B$. One wants to compute these form factors to first order in the electromagnetic coupling and to zeroth order in the other terms that might break the SU(3) symmetry.

From a group theoretical point of view, the matrix element $\langle B|j_\mu(x)|B'\rangle$ comes under the Wigner-Eckart theorem: there are two ways to project $8 \times 8$ on $8$ (see (4-20)), (or also, there are two ways to construct an invariant with $8 \otimes 8 \otimes 8$). There are thus two “reduced matrix elements”, hence two independent amplitudes for each of the two form factors, dressed with SU(3) Clebsch-Gordan coefficients. By an argument similar to (4-21), one finds that one may write

$$F^B_{ee,m}(k^2) = F^{(1)}_{e,m}(k^2) \text{tr} BQB' + F^{(2)}_{e,m}(k^2) \text{tr} BB'Q$$
where $Q$ is the matrix of (4-19)

$$Q = \begin{pmatrix}
\frac{2}{3} & 0 & 0 \\
0 & -\frac{1}{3} & 0 \\
0 & 0 & -\frac{1}{3}
\end{pmatrix},$$

and $\text{tr} \bar{B}QB'$ means the coefficient of $\bar{B}B'$ in the matrix trace $\text{tr} \bar{Ψ}QΨ$, and likewise for $\text{tr} \bar{B}QB'$. For example, the magnetic moment of the neutron $\mu(n)$ is proportional to the magnetic term in $\bar{nn}$, namely $-\frac{1}{3} (F_m^{(1)} + F_m^{(2)})$. The four functions $F_{e,m}^{(1,2)}$ are unknown (their computation would involve the theory of strong interactions) but one may eliminate them and find relations

$$\mu(n) = 2\mu(\Xi^0) = 2\mu(\Lambda) = -2\mu(\Sigma^0), \quad \mu(\Sigma^+) = \mu(p)$$

$$\mu(\Xi^-) = \mu(\Sigma^-) = -(\mu(p) + \mu(n)), \quad \mu(\Sigma^0 \rightarrow \Lambda) = \frac{\sqrt{3}}{2} \mu(n),$$

where the last quantity is the transition magnetic moment $\Sigma^0 \rightarrow \Lambda$. These relations are in qualitative agreement with experimental data.

The magnetic moments of “hyperons” (baryons of higher mass than the nucleons) are measured by their spin precession in a magnetic field or in transitions within “exotic atoms” (i.e. in the nucleus of which a nucleon has been substituted for a hyperon). The transition magnetic moment $\Sigma^0 \rightarrow \Lambda$ is determined from the cross-section $\Lambda \rightarrow \Sigma^0$ in the Coulomb field of a heavy nucleus. One reads in tables

$$\mu(p) = 2.792847351 \pm 0.00000028 \mu_N, \quad \mu(n) = -1.9130427 \pm 0.0000005 \mu_N$$

$$\mu(\Lambda) = -0.613 \pm 0.004 \mu_N, \quad |\mu(\Sigma^0 \rightarrow \Lambda)| = 1.61 \pm 0.08 \mu_N$$

$$\mu(\Sigma^+) = 2.458 \pm 0.010 \mu_N, \quad \mu(\Sigma^-) = -1.160 \pm 0.025 \mu_N$$

$$\mu(\Xi^0) = -1.250 \pm 0.014 \mu_N, \quad \mu(\Xi^-) = -0.6507 \pm 0.0025 \mu_N$$

where $\mu_N$ is the nuclear magneton, $\mu_N = \frac{e\hbar}{2m_p} = 3.152 \times 10^{-14}$ MeV T$^{-1}$.

**Electromagnetic mass splittings**

With similar assumptions and methods, one may also find relations between mass splittings of particles with same hypercharge and isospin $I$ but different charge, due to electromagnetic interactions, see Problem 3.

### 4.2.4 “Strong” mass splittings. Gell-Mann–Okubo mass formula

In view of the discrepancies between masses within a SU(3) multiplet, the mass term in the Lagrangian (or Hamiltonian) cannot be an invariant of SU(3). Gell-Mann and Okubo made the assumption that the non invariant term $\Delta M$ transforms under the representation 8, more precisely, since it must have vanishing isospin and hypercharge, that it transforms like the $\eta$ or $\Lambda$ component of octets. One is thus led to consider matrix elements $\langle H|\Delta M|H \rangle$ for the hadrons $H$ of a multiplet, and to appeal once more to Wigner–Eckart theorem. According to the decomposition rules of tensor products given in Chap. 3, the representation 8 appears at most twice in the product of an irreducible representation of SU(3) by its conjugate, (check it, recalling that $8 = 3 \otimes \bar{3} \oplus 1$); there are at most two independent amplitudes describing mass splittings within the multiplet, which leads to relations between these mass splittings.
An elegant argument enables one to avoid the computation of Clebsch–Gordan coefficients and to find these two amplitudes in any representation. As the eight infinitesimal generators transform themselves according to the representation 8 (adjoint representation), they may be set as before into a matrix

\[
G = \begin{pmatrix}
\frac{1}{2}Y + I_z & \sqrt{2}I_+ & * \\
\sqrt{2}I_- & \frac{1}{2}Y - I_z & * \\
* & * & -Y
\end{pmatrix}
\]

where the * stand for strangeness-changing generators that are of no concern to us here. (Note that \(G_{11} = \frac{1}{2}Y + I_z = Q\), the electric charge, is invariant under the action (by commutation with \(G\)) of generators

\[
X = \begin{pmatrix}
0 & 0 & 0 \\
0 & * & * \\
0 & * & *
\end{pmatrix}
\]

which preserve the electric charge.) One seeks two combinations of the generators \(I_z\) and \(Y\) transforming like the element \((3, 3)\) of that matrix. One is of course \(Y\) itself, the other is given by the element \((3, 3)\) of the cofactor of \(G\),

\[
\text{cof}G_{33} = \frac{1}{4}Y^2 - I_z^2 - 2I_+I_- = \frac{1}{4}Y^2 - \vec{I}^2.
\]

One gets in that way a mass formula for any representation (any multiplet)

\[
M = m_1 + m_2Y + m_3(I(I + 1) - \frac{1}{4}Y^2) \tag{4-27}
\]

which leaves three undetermined constants (that depend on the multiplet). For example for the baryon octet, one has four experimental masses, which leads to a sum rule

\[
\frac{M_\Xi + M_N}{2} = \frac{3M_\Lambda + M_\Sigma}{4} \tag{4-28}
\]

which is experimentally well verified: one finds 1128.5 MeV/c^2 in the left hand side, 1136 MeV/c^2 in the rhs^2. For the decuplet, show that the same formula gives equal mass differences between the four particles \(\Delta, \Sigma^*, \Xi^*\) and \(\Omega^-\). The latter result led to an accurate prediction of the existence and mass of the \(\Omega^-\) particle, which was regarded as one of the major achievements of SU(3). For the octet of pseudoscalar mesons, the mass formula is empirically better verified in terms of the square masses

\[
m_K^2 = \frac{3m_\eta^2 + m_\Xi^2}{4}.
\]

### 4.2.5 Quarks

The representations 3 and \(\bar{3}\) are so far absent from the scene: among the observed particles, no “triplet” seems to show up. The Gell-Mann–Zweig model makes the assumption that a triplet (representation 3) of quarks \((u, d, s)\) (“up”, “down” and “strange”) and its conjugate representation \(\bar{3}\) of antiquarks \((\bar{u}, \bar{d}, \bar{s})\) encompass the elementary constituents of all hadrons (known at the time). Their charges and hypercharges are respectively

---

^2The observed masses of these hadrons are \(M_N \approx 939\) MeV/c^2, \(M_\Lambda = 1116\) MeV/c^2, \(M_{\Sigma^0} \approx 1195\) MeV/c^2, \(M_{\Xi^0} \approx 1318\) MeV/c^2; those of pseudoscalar mesons \(m_\pi \approx 137\) MeV/c^2, \(m_K \approx 496\) MeV/c^2 and \(m_\eta = 548\) MeV/c^2. For the decuplet, \(M_\Delta \approx 1232\) MeV/c^2, \(M_{\Sigma^*} \approx 1385\) MeV/c^2, \(M_{\Xi^*} \approx 1530\) MeV/c^2, \(M_\Omega \approx 1672\) MeV/c^2.
4.2. THE SU(3) FLAVOR SYMMETRY AND THE QUARK MODEL.

Fig. 4.4: The triplets of quarks and antiquarks.

<table>
<thead>
<tr>
<th>Quarks</th>
<th>u</th>
<th>d</th>
<th>s</th>
<th>(\bar{u})</th>
<th>(\bar{d})</th>
<th>(\bar{s})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isospin (I_z)</td>
<td>(\frac{1}{2})</td>
<td>(-\frac{1}{2})</td>
<td>0</td>
<td>(-\frac{1}{2})</td>
<td>(\frac{1}{2})</td>
<td>0</td>
</tr>
<tr>
<td>Baryonic charge (B)</td>
<td>(\frac{1}{3})</td>
<td>(\frac{1}{3})</td>
<td>(\frac{1}{3})</td>
<td>(-\frac{1}{3})</td>
<td>(-\frac{1}{3})</td>
<td>(-\frac{1}{3})</td>
</tr>
<tr>
<td>Strangeness (S)</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Hypercharge (Y)</td>
<td>(\frac{1}{3})</td>
<td>(\frac{1}{3})</td>
<td>(-\frac{2}{3})</td>
<td>(-\frac{1}{3})</td>
<td>(-\frac{1}{3})</td>
<td>(\frac{2}{3})</td>
</tr>
<tr>
<td>Electric charge (Q)</td>
<td>(\frac{2}{3})</td>
<td>(-\frac{1}{3})</td>
<td>(-\frac{1}{3})</td>
<td>(-\frac{2}{3})</td>
<td>(\frac{1}{3})</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1. Quantum numbers of quarks \(u\), \(d\), \(s\)

One recalls (Chap. 3 §4) that any irreducible representation of SU(3) appears in the decomposition of iterated tensor products of representations 3 and \(\bar{3}\); in particular, \(3 \otimes \bar{3} = 1 \oplus 8\) and \(3 \otimes 3 \otimes 3 = 1 \oplus 8 \oplus 8 \oplus 10\). Mesons and baryons observed in Nature and classified as above in representations 8 and 10 of SU(3) are bound states of pairs \(q\bar{q}\) or \(qqq\), respectively. More generally, one assumes that only representations of zero triality may give rise to observable particles. Thus

\[ p = uud, \, n = udd, \, \Omega^- = sss, \]
\[ \pi^+ = ud, \, \pi^0 = \frac{(u\bar{u} - d\bar{d})}{\sqrt{2}}, \, \pi^- = d\bar{u}, \]
\[ \Delta^{++} = uuu, \, \cdots, \, \Delta^- = ddd, \]
\[ \eta_8 = \frac{(u\bar{u} + d\bar{d} - 2s\bar{s})}{\sqrt{6}}, \, K^+ = u\bar{s}, \, K^0 = d\bar{s} \quad \text{etc.} \] (4-29)

The quark model interprets the singlet that appears in the product \(3 \times 3\) as a bound state \(\eta_1 = \frac{(u\bar{u} + d\bar{d} + s\bar{s})}{\sqrt{3}}\).

The physically observed particles \(\eta\) (mass 548 MeV) and \(\eta'\) (958 MeV) result from a “mixing” (i.e. a linear combination) due to SU(3) breaking interactions of these \(\eta_1\) and \(\eta_8\). Exercise: complete on Fig. 4.3 the interpretations of baryons as bound states of quarks, making use of the knowledge of their charges and other quantum numbers.

4.2.6 Hadronic currents and weak interactions

The weak interactions are phenomenologically well described by an effective “current-current” Lagrangian (Fermi)

\[ \mathcal{L}_{\text{Fermi}} = -\frac{G}{\sqrt{2}} J^\rho(x) J^\dagger_\rho(x) \] (4-30)
where $G$ is the Fermi constant, whose value (in units where $\hbar = c = 1$) is

$$G = (1.026 \pm 0.001) \times 10^{-5} m_p^{-2}. \quad (4-31)$$

(But this interaction Lagrangian has the major flaw of being non renormalisable, a flaw which will be corrected by the gauge theory of the Standard Model. At low energy, however, $L_{\text{Fermi}}$ offers a good description of physics, whence the name “effective”.) The current $J_{\rho}$ is the sum of a leptonic and a hadronic contributions

$$J_{\rho}(x) = l_{\rho}(x) + h_{\rho}(x) \quad (4-32)$$

The leptonic current

$$l_{\rho}(x) = \bar{\psi}_e(x) \gamma_\rho (1 - \gamma_5) \psi_{\nu_e} + \bar{\psi}_\mu(x) \gamma_\rho (1 - \gamma_5) \psi_{\nu_\mu} \quad (4-33)$$

is the sum of contributions of the lepton families (or generations), $e, \mu$ (and $\tau$ that we omit in this first approach). The hadronic current, if one restricts to the first two generations, reads

$$h_{\rho} = \cos \theta_C h_{\rho}^{(\Delta S=0)} + \sin \theta_C h_{\rho}^{(\Delta S=1)} \quad (4-34)$$

i.e. a combination of strangeness-conserving and non-conserving currents, weighted by the Cabibbo angle $\theta_C \approx 0.25$. (This “mixing” extends to the introduction of the third generation, see next Chapter.) Finally each of these currents $h_{\rho}^{(\Delta S=0)}, h_{\rho}^{(\Delta S=1)}$ has the “$V - A$” form, following an idea of Feynman and Gell-Mann, i.e. is a combination of vector and axial currents,

$$h_{\rho}^{(\Delta S=0)} = (V_{\rho}^1 - iV_{\rho}^2) - (A_{\rho}^1 - iA_{\rho}^2) \quad (4-35)$$

The vector currents $V_{\rho}^{1,2,3}$ are the Noether currents of isospin, the other components of $V_{\rho}$ are those of the SU(3) symmetry. One shows that their conservation (exact for isospin, approximate for the others) implies that in the matrix element $G\langle p | h_{\rho}^{(\Delta S=0)} | n \rangle = \bar{u}_{\rho}(0) \gamma_\rho (G_V(q^2) - G_A(q^2) \gamma_5) u_n$ measured in beta decay at quasi-vanishing momentum transfer, the vector form factor $G_V(0) = G$. On the contrary, the axial currents are non conserved and $G_A(0)$ is “renormalized” (that is, dressed) by strong interactions, $G_A/G_V \approx 1.22$. The electromagnetic current is nothing other that the combination $j_{\rho} = V_{\rho}^3 + \frac{1}{\sqrt{3}} V_{\rho}^8$. In the quark model, these hadronic currents have the form

$$V_{\rho}^a(x) = \bar{q}(x) \frac{\lambda^a}{2} \gamma_\rho q(x) \quad (4-36)$$

$$A_{\rho}^a(x) = \bar{q}(x) \frac{\lambda^a}{2} \gamma_\rho \gamma_5 q(x).$$

We will meet them again in the Standard Model.

### 4.3 From SU(3) to SU(4) to six flavors

#### 4.3.1 New flavors

The discovery in the mid 70’s of particles of a new type revived the game: these particles carry another quantum number, “charm” (whose existence had been postulated beforehand
by Glashow, Iliopoulos and Maiani and by Kobayashi and Maskawa for different reasons). This introduces a third direction in the space of internal symmetries, on top of isospin and strangeness (or hypercharge). The relevant group is SU(4), which more severely broken than SU(3). Particles fall into representations of that SU(4), etc. A fourth flavor, charm, is thus added, and a fourth charmed quark $c$ constitutes with $u,d,s$ the representation 4 of SU(4), as inobservable as the 3 of SU(3), according to the same principle.

As for today, one believes there are in total six flavors, the last two being beauty or bottomness and truth (or topness ??), hence two additional quarks $b$ and $t$. $B$ mesons, which are bound states $u\bar{b}$, $d\bar{b}$ etc, are observed in everyday experiments, in particular at LHCb, whereas the experimental evidence for the existence of the $t$ quark is more indirect. The hypothetical flavor group SU(6) is very strongly broken, as attested by masses of the 6 quarks

$$
\begin{align*}
m_u &\approx 1.5 - 4 \text{ MeV} , \\
m_d &\approx 4 - 8 \text{ MeV} , \\
m_s &\approx 80 - 130 \text{ MeV} \\
m_c &\approx 1.15 - 1.35 \text{ GeV} , \\
m_b &\approx 4 - 5 \text{ GeV} , \\
m_t &\approx 175 \text{ GeV}
\end{align*}
$$

and this limits its usefulness. One may however rewrite (4-12) in the form

$$
Q = \frac{1}{2} Y + I_z \\
Y = B + S + C + B + T
$$

with different quantum numbers contributing additively to hypercharge. The convention is that the flavor $S,C,B,T$ of a quark vanishes or is of the same sign as its electric charge $Q$. Thus $C(c) = 1$, $B(b) = -1$ etc. Table 1 must now be extended as follows

---

3One should of course make precise what is meant by mass of an invisible particle, and this may be done in an indirect way and with several definitions, whence the range of given values.
4.3.2 Introduction of color

Various problems with the original quark model have led to the hypothesis (Han–Nambu) that each flavor comes with a multiplicity 3, which reflects the existence of a group SU(3), different from the previous one, the color group SU(3)\(_c\).

Considerations leading to that triplicating hypothesis are first the study of the \(\Delta^{++}\) particle, with spin \(3/2\), made of 3 quarks \(u\). This system of 3 quarks has a spin \(3/2\) and an orbital angular momentum \(L = 0\), which give it a symmetric wave function, in contradiction with the fermionic character of quarks. The additional color degree of freedom allows an extra antisymmetrization, (which leads to a singlet state of color), and thus removes the problem. On the other hand, the decay amplitude of \(\pi^0 \rightarrow 2\gamma\) is proportional to the sum \(\sum Q^2 I_z\) over the set of fermionic constituents of the \(\pi^0\). The proton, with its charge \(Q = 1\) and \(I_z = \frac{1}{2}\), gives a value in agreement with experiment. Quarks \((u,d,s)\) with \(Q = (\frac{2}{3}, \frac{1}{3}, -\frac{1}{3})\) and \(I_z = (\frac{1}{2}, -\frac{1}{2}, 0)\) lead to a result three times too small, and color multiplicity corrects it to the right value.

According to the confinement hypothesis, only states of the representation 1 of SU(3)\(_c\) are observable. The other states, which are said to be “colored”, are bound in a permanent way inside hadrons. This applies to quarks, but also to gluons, which are vector particles (spin 1) transforming by the representation 8 of SU(3)\(_c\), whose existence is required by the construction of the gauge theory of strong interactions, Quantum Chromodynamics (QCD), see Chap. 5.

To be more precise, the confinement hypothesis applies to zero or low temperature, and quark or gluon deconfinement may occur in hadronic matter at high temperature or high density (within the “quark gluon plasma”).

The quark model with its color group SU(3)\(_c\) is now regarded as part of quantum chromodynamics. The six flavors of quarks are grouped into three “generations”, \((u,d), (c,s), (t,b)\), which are in correspondence with three generations of leptons, \((e^-, \nu_e), (\mu^-, \nu_\mu), (\tau^-, \nu_\tau)\). That correspondence is important for the consistency of the Standard Model (anomaly cancellation), see next chapter.
Further references for Chapter 4

On flavor SU(3), the standard reference containing all historical papers is M. Gell-Mann and Y. Ne'eman, *The Eightfold Way*, Benjamin 1964.

In particular one finds there tables of SU(3) Clebsch-Gordan coefficients by J.J. de Swart.


All the properties of particles mentioned in this chapter may be found in the tables of the Particle Data Group, on line on the site [http://pdg.lbl.gov/2010/tables/contents_tables.html](http://pdg.lbl.gov/2010/tables/contents_tables.html)

Exercises for chapter 4

A. Sigma model and chiral symmetry breaking

Consider the Lagrangian (4-10) and define $W = \sigma + i\pi$.

1. Compute $\det W$. Show that one may write $\mathcal{L}$ in terms of $\psi_{L,R}$ and $W$ as

$$\mathcal{L} = \tilde{\psi}_R i\partial_\mu \psi_R + \tilde{\psi}_L i\partial_\mu \psi_L + g(\tilde{\psi}_L W \psi_R + \tilde{\psi}_R W^\dagger \psi_L) + \mathcal{L}_K - \frac{1}{2} m^2 \det W - \frac{\lambda}{4} (\det W)^2$$

where $\mathcal{L}_K$ is the kinetic term of the fields $(\sigma, \pi)$. One may also give that term the form $\mathcal{L}_K = \frac{1}{2} (\det \partial_\mu W - \sum_{i=1}^3 \det \partial_i W)$ (which looks a bit odd, but which is indeed Lorentz invariant!).

2. Show that $\mathcal{L}$ is invariant under transformations of SU(2) x SU(2) with $\psi_L \rightarrow U \psi_L$, $\psi_R \rightarrow V \psi_R$, provided $W$ transforms in a way to be specified. Justify the assertion made in §4.1.2: $\psi_L$, $\psi_R$ and $W$ transform respectively under the representations $(\frac{1}{2}, 0)$, $(0, \frac{1}{2})$ and $(\frac{1}{2}, \frac{1}{2})$.

3. If the field $W$ acquires a vev $v$, for example along the direction of $\sigma$, $\langle \sigma \rangle = v$, show that the field $\psi$ acquires a mass $M = -gv$.

B. Changes of basis in SU(3)

In SU(3), write the change of basis which transforms the weights $\Lambda_1$, $\Lambda_2$ of Chap. 3 into the axes used in figures 2, 3 and 4. Derive the transformation of the coordinates $(\lambda_1, \lambda_2)$ (Dynkin labels) into the physical coordinates $(I_z, Y)$. What is the dimension of the representation of SU(3) expressed in terms of the isospin and hypercharge of its highest weight?

C. Gell-Mann–Okubo formula

Complete and justify all the arguments sketched in § 4.2.2, 4.2.3 and 4.2.4. In particular check that the formula (4-27) does lead for the baryon octet to the rule (4-28), and for the decuplet, to constant mass splittings.

D. Counting amplitudes

How many independent amplitudes are necessary to describe the scattering $BD \rightarrow BD$, where $B$ and $D$ refer to the baryonic octet and decuplet?
Problems

1. SU(3) invariant four-field couplings

Consider a Hermitian, 3 × 3 and traceless matrix $A$.

a. Show that its characteristic equation

$$A^3 - (\text{tr } A) A^2 + \frac{1}{2} ((\text{tr } A)^2 - \text{tr } A^2) A - \det A = 0$$

implies a relation between \(\text{tr } A^4\) and \((\text{tr } A^2)^2\).

b. If the group SU(3) acts on $A$ by $A \rightarrow U A U^\dagger$, show that any sum of products of traces of powers of $A$ is invariant. We call such a sum an “invariant polynomial in $A$”. How many linearly independent such invariant polynomials in $A$ of degree 4 are there?

c. One then “polarises” the identity found in a., which means one writes $A = \sum_{i=1}^4 x_i A_i$ with 4 matrices $A_i$ of the previous type and 4 arbitrary coefficients $x_i$, and one identifies the coefficient of $x_1 x_2 x_3 x_4$. Show that this gives an identity of the form (Burgoyne’s identity)

$$\sum_P \text{tr } (A_{P1} A_{P2} A_{P3} A_{P4}) = a \sum_P \text{tr } (A_{P1} A_{P2}) \text{tr } (A_{P3} A_{P4})$$

(4-38)

with sums over permutations $P$ of 4 elements and a coefficient $a$ to be determined. How many distinct terms appear in each side of that identity?

d. How many polynomials of degree 4, quadrilinear in $A_1, \ldots, A_4$, invariant under the action of SU(3) $A_i \rightarrow U A_i U^\dagger$ and linearly independent, can one write? Why is the identity (4-38) useful?

2. Hidden invariance of a bosonic Lagrangian

One wants to write a Lagrangian for the field $\Phi$ of the pseudoscalar meson octet, see (4-17).

a. Why is it natural to impose that this Lagrangian be even in the field $\Phi$?

b. Using the results of Problem 1., write the most general form of a SU(3) which is invariant Lagrangian of degree less or equal to 4 (for renormalizability) and even in $\Phi$.

c. One then writes each complex field by making explicit its real and imaginary parts, for example $K^+ = \frac{1}{\sqrt{2}} (K_1 - iK_2)$, $K^- = \frac{1}{\sqrt{2}} (K_1 + iK_2)$, and likewise with $K^0, \bar{K}^0$ and with $\pi^\pm$. Compute $\text{tr } \Phi^2$ with that parametrization and show that one gets a simple quadratic form in the 8 real components. What is the invariance group $G$ of that quadratic form? Is $G$ a subgroup of SU(3)?

d. Conclude that any Lagrangian of degree 4 in $\Phi$ which is invariant under SU(3) is in fact invariant by this group $G$.

3. Electromagnetic mass splittings in an SU(3) octet

Preliminary question.

Given a vector space $E$ of dimension $d$, we denote $E \otimes E$ the space of rank 2 tensors and $(E \otimes E)_S$, resp. $(E \otimes E)_A$, the spaces of symmetric, resp. antisymmetric, rank 2 tensors, also called (anti-)symmetrized tensor product. What is the dimension of spaces $E \otimes E$, $(E \otimes E)_S$, $(E \otimes E)_A$?
One assumes that SU(3) is an exact symmetry of strong interactions, and one wants to study mass splittings due to electromagnetic effects.

a. How many independent mass differences between baryons with the same quantum numbers $I$ and $Y$ but different charges $Q$ (or $I_z$ component), are there in the baryon octet $J^P = \frac{1}{2}^+$? We admit that these electromagnetic effects result from second order perturbations in the Lagrangian $\mathcal{L}_{em}(x) = -q_j \mu^j(x) A_\mu(x)$. If $|B\rangle$ is a baryon state, one should thus compute

$$\delta M_B = \langle B|\left(\int d^4x \mathcal{L}_{em}\right)^2|B\rangle.$$  \hspace{1cm} (4-39)

For lack of a good way of computing that matrix element, one wants to determine the number of independent amplitudes that contribute.

b. Why does this calculation amounts to counting the number of invariants appearing in the tensor product of four representations $8$? In view of the calculations done in sect. 4.2.2, what should be that number?

c. But caution! the product of the two Lagrangians is symmetric. As for the product $\int \mathcal{L}_{em} \int \mathcal{L}_{em}$, one must decompose into irreducible representations the symmetrized tensor product $(8 \otimes 8)_S$. Use the result of the Preliminary question to calculate the number of independent symmetric rank 2 tensors in the representation 8. Show that this number is consistent with the decomposition that we admit

$$(8 \otimes 8)_S = 1 \oplus 8 \oplus 27.$$ \hspace{1cm} (4-40)

d. i) What is then the number of invariant amplitudes contributing to $\delta M_B$?

d. ii) What is the number of invariant amplitudes contributing to $\delta M_B - \delta M_{B'}$ for hadrons $B$ and $B'$ with the same quantum numbers, as discussed in a.?

d. iii) In the spirit of what is done in §4.2.3 for magnetic moments, write a basis on invariants in terms of the matrices $\Psi$, $\bar{\Psi}$ et $Q$?

e. i) Show that a priori the number of amplitudes determined in question d. ii) implies one relation between electromagnetic mass splittings within the baryon octet.

e. ii) Calculate $\Delta_{em} M = \alpha tr \bar{\Psi}Q^2\Psi + \beta tr \bar{\Psi}\Psi Q^2 + \gamma tr \bar{\Psi}Q\Psi Q$, (the use of Maple or of Mathematica may be helpful...), identify in that expression the coefficients $\Delta_{em} M_p$ of $\bar{p}p$, $\Delta_{em} M_n$ of $\bar{nn}$, etc, and check the relation

$$M_{\Xi^-} - M_{\Xi^0} = M_{\Sigma^-} - M_{\Sigma^0} + M_p - M_n.$$ \hspace{1cm} (4-41)

The experimental values are $M_n = 939, 56$ MeV/$c^2$, $M_p = 938, 27$ MeV/$c^2$, $M_{\Xi^-} = 1321, 71$ MeV/$c^2$, $M_{\Xi^0} = 1314, 86$ MeV/$c^2$, $M_{\Sigma^-} = 1197, 45$ MeV/$c^2$, $M_{\Sigma^0} = 1192, 64$ MeV/$c^2$, $M_{\Sigma^+} = 1189, 37$ MeV/$c^2$.

Calculate the values of the two sides of relation (4-41). Comment.

f. Octet of pseudoscalar mesons. Could one do a similar reasoning for pseudoscalar mesons?

g. What about the electromagnetic mass splittings within the $(\frac{3}{2})^+$ decuplet?

*
Chapter 5

Gauge theories. Standard model

Transformations considered so far were global, space-time independent, transformations. Another type of symmetry, which is restricting the dynamics of the system in a much more stringent way, considers local transformations. At each point of space–time, acts a distinct copy of the transformation group. Such a symmetry, called gauge symmetry, is familiar in electrodynamics. Its extension by Yang and Mills to non-abelian transformation groups turned out to be one of the most fruitful theoretical ideas of the second half of the XXth century. A full course should be devoted to it. More modestly, the present chapter gives an elementary introduction and overview.

5.1 Gauge invariance. Minimal coupling. Yang–Mills Lagrangian

5.1.1 Gauge invariance

The study of electrodynamics has introduced the notion of local invariance. The Lagrangian

\[ \mathcal{L} = \bar{\psi}(i\not{D} - e\not{A} - m)\psi - \frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) \]  

is invariant under infinitesimal gauge transformations

\[ \delta A_\mu(x) = -\partial_\mu \delta \alpha(x) \]
\[ \delta \psi(x) = ie\delta \alpha(x)\psi(x) \]

since the electromagnetic field tensor

\[ F_{\mu\nu} = (\partial_\mu A_\nu - \partial_\nu A_\mu) \]

is invariant, and the combination

\[ i\not{D} \psi(x) := (i\not{D} - e\not{A})\psi(x) \]
also transforms as $\psi$. The finite form of these transformations is readily written

$$
A_\mu(x) \mapsto A_\mu(x) - \partial_\mu \alpha(x) \\
\psi(x) \mapsto e^{i\epsilon \alpha(x)} \psi(x),
$$

(5-3)

which shows that the transformations are a local (i.e. $x$ dependent) version of those of the group $U(1)$ or $\mathbb{R}$ (depending whether one identifies or not $\alpha$ and $\alpha + 2\pi/e$). The corresponding global transformations are those leading to a conserved Noether current, which implies the conservation of electric charge. The Lagrangian displays the “minimal coupling” of the field $\psi$ to the electromagnetic field\(^1\). Any other charged field of charge $q$ couples to the electromagnetic field through a term involving the “covariant derivative” $i\partial_\mu - qA_\mu(x)$.

This is for example the case of a charged, hence complex, boson field $\phi$, whose contribution to the Lagrangian reads

$$
\delta L = (\partial_\mu - iq A_\mu)\phi^*(\partial^\mu + iq A_\mu)\phi - V(\phi^*\phi)
$$

(5-4)

which is indeed invariant under $\phi(x) \mapsto e^{iq\alpha(x)}\phi(x)$, $A_\mu(x) \mapsto A_\mu(x) - \partial_\mu \alpha(x)$.

### 5.1.2 Non abelian Yang–Mills extension

Following the brilliant observation of Yang and Mills (1954), this construction may be transposed to the case of a non-abelian Lie group $G$, with however a few interesting modifications... Let $\psi$ be a field (which we denote as a fermion field, but this is irrelevant) transforming under $G$ by some representation $D$. Let $T_a$ be the infinitesimal generators in that representation, which we assume antihermitian: $[T_a, T_b] = C_{ab}^c T_c$; the infinitesimal transformation thus reads

$$
\delta \psi(x) = T_a \delta \alpha^a \psi(x).
$$

(5-5)

(In this chapter, we denote $t_a$ the corresponding matrices in the adjoint representation.) To extend the notion of local transformation, we need a gauge field $A_\mu$, which allows to construct a covariant derivative $D_\mu \psi$. It is natural to consider that $A_\mu$ lives in the Lie algebra of $G$, and hence carries indices of the adjoint representation

$$
A_\mu(x) = \{ A^a_\mu(x) \}
$$

(5-6)

or equivalently, is represented in any representation by the antihermitian matrix\(^2\)

$$
A_\mu(x) = T_a A^a_\mu(x).
$$

(5-7)

The covariant derivative reads

$$
D_\mu \psi(x) := (\partial_\mu - A_\mu(x))\psi(x),
$$

(5-8)

---

\(^1\)An additional term in the Lagrangian like $\bar{\psi}[\gamma_\mu, \gamma_\nu] \psi F^{\mu\nu}$ would be gauge invariant but non minimal.

\(^2\)Caution! this convention implies that some expressions differ by a factor $i$ from the abelian case.
5.1. GAUGE INVARIANCE. MINIMAL COUPLING. YANG–MILLS LAGRANGIAN

or, componentwise

\[
D_\mu \psi_A(x) := \left( \partial_\mu \delta_{AB} - A_\mu^B(x) (T_a)_A^B \right) \psi_B(x) .
\]  

(5-9)

That covariant derivative does transform as \( \psi \), just like in the abelian case, provided one imposes that \( A_\mu \) transforms according to

\[
\delta A_\mu^a(x) = \partial_\mu \delta \alpha^a(x) + C_{bc}^a \delta \alpha^b(x) A_\mu^c(x) .
\]  

(5-10)

The term \( \partial_\mu \delta \alpha^a(x) \) notwithstanding, one sees that \( \{ A_\mu^a \} \) transforms as the adjoint representation (whose matrices are \((t_a)_c^b = -C_{ab}^c\)). Lastly a field tensor transforming in a covariant way \( (i.e. \) without any inhomogeneous term in \( \partial \delta \alpha^a(x) \)) may be constructed

\[
F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - [A_\mu, A_\nu] \]  

(5-11)

or in components

\[
F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - C_{bc}^a A_\mu^b A_\nu^c .
\]  

(5-12)

One proves, after some algebra and using the Jacobi identity, that

\[
\delta F_{\mu\nu}^a(x) = C_{bc}^a \delta \alpha^b(x) F_{\mu\nu}^c(x) ,
\]  

(5-13)

which is indeed an infinitesimal transformation in the adjoint representation.

It is in fact profitable, and more enlightening, to look at the effect of a finite local transformation \( g(x) \) of the group \( G \),

\[
\psi(x) \mapsto D(g(x))\psi(x) \\
A_\mu = A_\mu^a T_a \mapsto D(g(x))(-\partial_\mu + A_\mu(x))D(g^{-1}(x)) ,
\]  

(5-14)

and for the covariant derivative acting on \( \psi \),

\[
D_\mu \psi(x) \mapsto D(g(x))D_\mu \psi(x) \\
or equivalently\(^3\)

\[
D_\mu \mapsto D(g(x))D_\mu D(g^{-1}(x)) .
\]  

(5-16)

Now one verifies easily that in a given representation

\[
[D_\mu, D_\nu] = -F_{\mu\nu} := -F_{\mu\nu}^a T_a
\]  

(5-17)

from which follows that \( F_{\mu\nu}(x) \mapsto D(g(x))F_{\mu\nu} D(g^{-1}(x)) \), and in particular, in the adjoint representation, the finite transformation of \( F_{\mu\nu} = F_{\mu\nu}^a t_a \) is

\[
F_{\mu\nu}(x) \mapsto g(x)F_{\mu\nu}(x)g^{-1}(x) ,
\]  

(5-18)

of which (5-13) is the infinitesimal version.

\(^3\)Beware the notations! In that equation (5-16), which deals with a differential operator, the derivative \( \partial_\mu \) contained in \( D_\mu \) acts on everything sitting on its right, whereas in the second equation (5-14), it acts only on \( D(g^{-1}(x)) \).
Pure gauge

If the tensor $F_{\mu\nu}$ vanishes in the neighbourhood of a point $x_0$, one may write locally (i.e. in that neighbourhood) $A_\mu(x)$ as a “pure gauge”, i.e.

$$F_{\mu\nu} = 0 \iff A_\mu(x) = (\partial_\mu g(x)) g^{-1}(x) . \quad (5-19)$$

(The naming “pure gauge” is justified by the fact that such an $A_\mu(x) = (\partial_\mu g(x)) g^{-1}(x)$ is the gauge transform of a vanishing gauge field! Proving $\iff$ is a trivial calculation, as for $\Rightarrow$, see a few lines below…) We insist on the local character of that property.

Parallel transport along a curve

Another interesting object is the group element attached to a curve $C$ going from $x_0$ to $x$

$$\gamma(C) := P \exp \left( \int_C dx^\mu A_\mu(x) \right) \quad (5-20)$$

where the symbol $P$ means that a parametrization $x(s)$ of the curve being chosen, and terms in the expansion of the exponential are ordered from right to left with increasing $s$ (compare with the $T$-product in quantum field theory). One shows that under the gauge transformation (5-14)

$$\gamma(C) \mapsto g(x)\gamma(C)g^{-1}(x_0) . \quad (5-21)$$

More generally, for any representation $\mathcal{D}$ and with $A = A^a T_a$, (5-20) defines a $\gamma_\mathcal{D}(C)$ in the representation $\mathcal{D}$ that transforms as $\mathcal{D}(g(x))\gamma_\mathcal{D}(C)\mathcal{D}(g^{-1}(x_0))$. 

Exercise. Prove that statement by first considering an infinitesimal path from $x$ to $x + dx$, hence $\gamma(C) \approx 1 + A_\mu(x) dx^\mu$, and by performing a finite gauge transformation $A_\mu \rightarrow g(x)(-\partial_\mu + A_\mu(x))g^{-1}(x)$, show that $\gamma(C) \rightarrow g(x + dx)\gamma(C)g^{-1}(x)$. The result for a finite curve follows by recombining these infinitesimal elements.

Given an objet, like the field $\psi$, transforming by some representation $\mathcal{D}$, the role of $\gamma_\mathcal{D}(C)$ is to “transport” $\psi(x_0)$ into an object denoted $^t \psi(x)$ transforming like $\psi(x)$. Show that for an infinitesimal curve $(x, x + dx)$ the difference $^t \psi(x + dx) - \psi(x + dx)$ is expressed in a natural way in terms of the covariant derivative.

Consider then the case where $x = x_0$ in (5-20). From (5-21), it follows that for a closed loop $C$, $\gamma(C)$ transforms in a covariant way, $\gamma(C) \mapsto g(x_0)\gamma(C)g^{-1}(x_0)$. Let us examine again the case of an infinitesimal closed loop. One finds that then

$$\gamma(C) \approx \exp \frac{1}{2} \int_S dx^\mu \wedge dx^\nu F_{\mu\nu} , \quad (5-22)$$

where the integration is carried out on an infinitesimal surface $S$ of boundary $C$.

Exercise: Prove that statement by considering an elementary square circuit extending from $x$ along the coordinate axes $\mu$ and $\nu$: $(x \rightarrow x + dx^\mu \rightarrow x + dx^\mu + dx^\nu \rightarrow x + dx^\nu \rightarrow x)$, and expand to second order in $dx$ to find $\gamma(C) \approx 1 + dx^\mu dx^\nu F_{\mu\nu}$ (with no summation over $\mu, \nu$). Hint: the use of the commutator formula of Chap. 1, (1.23), simplifies the computation a great deal!
This has an immediate consequence. If \( F = 0 \), any \( \gamma(C) \) of the form (5-20) is insensitive to small variations of the curve \( C \) with fixed end points \( x_0 \) and \( x \), hence depends only on these end-points \( x_0 \) and \( x \). The element \( g(x, x_0) := g(C) \) that follows satisfies \( (\partial_\mu - A_\mu)g(x, x_0) = 0 \), (check!), thus completing the proof of (5-19).

**Wilson loop**

Return to the case of a closed loop \( C \) with \( x = x_0 \) in (5-20). As just mentionned, \( \gamma(C) \) transforms in a covariant way, \( \gamma(C) \mapsto g(x_0)\gamma(C)g^{-1}(x_0) \). Its trace

\[
W(C) = \text{tr} \, \gamma(C) = \text{tr} \, P \exp \oint dx^\mu A_\mu(x)
\]

(5-23) is thus invariant. We postulate that any physical quantity in a gauge theory must be “gauge invariant”, i.e. invariant under a gauge transformation. This is the case of \( \text{tr} \, F_{\mu\nu}F^{\mu\nu} \), \( \bar{\psi}(i\partial - \vec{A})\psi \) etc. The interest of \( W(C) \) is that it is a *non local* invariant quantity, which depends on the contour \( C \). Note that it depends on the representation in which \( A = A^aT_a \) is evaluated.

This *Wilson loop* was proposed by Wilson and Polyakov as a way to measure the interaction potential between particles propagating along \( C \), and as a good indicator of confinement. See below § 3.1 and the Problem at the end of this Chapter for a discrete version of that quantity.

### 5.1.3 Geometry of gauge fields

The previous considerations show that the theory of gauge fields has a strong geometric flavor. The appropriate language to discuss these matters is indeed the *theory of fiber bundles*, principal fiber bundle for the gauge group itself, vector bundle for each matter field like \( \psi \), above the base space which is space–time. The gauge field is a connection on the fiber bundle, which permits to define a parallel transport from point to point. The tensor \( F_{\mu\nu} \) is its curvature, as expressed by (5-17). All these notions are defined locally, in a system of local coordinates (*a chart*), and changes of chart imply transformations of the form (5-14). This language becomes particularly useful when one looks at topological (instantons etc) or global (“Gribov problem”) properties of gauge theories. For a mere introduction to properties of local symmetry and the *perturbative* construction of the standard model, we won’t need it.

### 5.1.4 Yang–Mills Lagrangian

The Lagrangian describing a gauge field coupled to a matter field like \( \psi \) via the minimal coupling reads

\[
\mathcal{L} = \frac{1}{2g^2} \text{tr} \left( F_{\mu\nu}F^{\mu\nu} \right) + \bar{\psi}(i\partial - \vec{A}) - m \psi ,
\]

(2-24) with a parameter, the coupling constant \( g \). The value of that coupling depends of course on the normalization of the matrices \( T_a \) that appear in \( F_{\mu\nu} = F_\mu^a T_a \). One proves (see Exercise B at the end of this chapter) that for any simple Lie algebra one may choose a basis such that in any representation \( R \), \( \text{tr} \, T_a T_b = -T_R \delta_{ab} \), with \( T_R \) a real positive coefficient that depends on the group and on the representation. We will choose in \( F_{\mu\nu} \) the fundamental representation of lowest dimension, (the defining representation of dimension \( N \) in the case of \( SU(N) \)) with a normalization \( T_f = \frac{1}{2} \), hence \( \text{tr} \, T_a T_b = -\frac{1}{2} \delta_{ab} \). To the Lagrangian \( \mathcal{L} \), one may add the
contribution of other fermion fields or of boson fields. Note that the representations “carried” by the fermions and other matter fields, that appear in their covariant derivatives \( D_\mu = \partial_\mu - A^a_\mu T_a \), may differ from the fundamental representation.

As such, \( \mathcal{L} \) of (2-24) resembles very much the Lagrangian of the abelian case (5-1), after a change \( iA \rightarrow gA \) has been carried out.

Let us review the most salient features of that construction:

- like in the abelian case, the gauge invariance principle implies a coupling of a universal type, namely through the covariant derivative;

- contrary to the abelian case where each charge is independent and unquantized (at least if the gauge group is \( \mathbb{R} \) rather than \( U(1) \)), the coupling constant \( g \) of all fields to the gauge fields is the same, within each simple component of the gauge group; (for example, the standard model, based on the group \( U(1) \times SU(2) \times SU(3) \) possesses three independent couplings, see below.)

- like in the abelian case, the gauge field comes naturally without a mass term: a mass term \( \frac{1}{2} M^2 A_\mu A^\mu \) does break gauge invariance. This looks most embarrassing for physical applications, since the massless vector fields (of spin 1) are quite exceptional in Nature (the electromagnetic field and its photonic excitations being the basic counter-example); this will lead us to introduce “soft” mechanisms of (spontaneous) breaking of gauge invariance to remedy it;

- contrary to the abelian case, the gauge field itself carries a charge of the group: we saw that for global (\( x \) independent) transformations of the group \( G \), \( A_\mu \) transforms by the adjoint representation. The property of the gauge field to be charged has important implications in many phenomena, from the infrared effects (confinement), to the ultraviolet ones (sign of the \( \beta \) function), as we shall see below.

### 5.1.5 Quantization. Renormalizability

The quantization of the Yang–Mills theory requires to overcome serious difficulties that we only briefly evoke. As in electrodynamics, the quadratic form in the gauge field in \( \mathcal{L} \), namely

\[
(\partial_\mu A_\nu - \partial_\nu A_\mu)^2 \quad \text{or in Fourier space} \quad A_\mu(-k)(k^\nu k^\nu - k^2 g^{\mu\nu})A_\nu(k)
\]

is degenerate, thus non invertible, which reflects gauge invariance. Consequently the propagator of the field \( A_\mu \) is \( a \) priori undefined. One must first “fix the gauge”, by imposing a non-invariant “gauge condition” (like the Coulomb gauge in QED), and the Faddeev and Popov procedure, justified by their general study of constrained systems leads to the introduction of auxiliary fields and to explicit Feynman rules, (see for example [IZ, chap. 12] and the courses of the second semester).

One then proves, and that was a decisive step in the building of the Standard Model\(^4\), that the theory so quantized is renormalizable: all ultraviolet divergences appearing in Feynman

---

\(^4\) G. ’t Hooft and M. Veltman, Nobel prize 1999
diagrams may, at any finite order of perturbation theory, be absorbed into a redefinition of parameters—couplings, field normalization, masses—of the Lagrangian. Thus, to the one loop order, diagrams of Fig. 5.1 have divergences that may be absorbed into a change of normalization of the $A$ field ("wave function renormalization") and a renormalization of the coupling constant $g$

$$g \rightarrow g_0 = \left(1 - \frac{g^2}{(4\pi)^2} \left(\frac{11}{3}C_2 - \frac{4}{3}T_f \right) \log \frac{\Lambda}{\mu}\right) g,$$

(2.25)

where $\Lambda$ is a scale of ultraviolet "cutoff" and $\mu$ a mass scale which must be introduced for a definition of the renormalisation procedure.

5.2 Massive gauge fields

5.2.1 Weak interactions and intermediate bosons

We saw in Chap. 4 (equ. (4-30)) that the Fermi Lagrangian

$$\mathcal{L}_{\text{Fermi}} = -\frac{G}{\sqrt{2}} J^\rho(x) J^\rho_\mu(x)$$

(2.26)

gives a good description of the low energy physics of weak interactions: leptonic processes like $\bar{\nu}_e e^- \rightarrow \bar{\nu}_e e^-$ or $\bar{\nu}_\mu \mu^-$, semi-leptonic ones like $\pi^+ \rightarrow \mu^+ \nu_\mu$ or the $\beta$ decay $n \rightarrow p e^- \bar{\nu}_e$, or non-leptonic ones : $\Lambda \rightarrow p \pi^-$, $K^0 \rightarrow \pi \pi$, etc. But this Lagrangian is theoretically unsatisfactory, since it leads to a non renormalizable theory, making impossible any calculation beyond the "Born term", the first order of perturbation theory, which violates unitarity.

The violation of unitarity appears in the calculation of the total cross section $\sigma$ of any process, to first order of the perturbation series. A simple dimensional argument gives at high energy

$$\sigma \sim \text{const. } G^2 s$$

where $s$ is the square center-of-mass energy. But that behavior contradicts general results based on unitarity that predict that $\sigma$ must decrease in each partial wave like $1/s$. A violation of unitarity by the Born term is thus expected at an energy of the order of $\sqrt{s} \sim G^{-\frac{1}{2}} \sim 300$ GeV. And the non-renormalisability of the theory precludes an improvement of that Born term.
The idea is thus to regard $\mathcal{L}_\text{Fermi}$ as an approximation of a theory where the charged current $J^\rho$ is coupled to a vector field $W$ of mass $M$, in the large mass limit\footnote{The inverse mass $M^{-1}$ represents the range of weak interactions, which is known to be short, and the mass $M$ is thus high (of the order of 100 GeV, as we see below).} Consider the new Lagrangian

$$\mathcal{L}_\text{int.boson} = g J^\rho(x) W_{\rho\mu}^\dagger(x) + h.c. - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} + M^2 W_{\rho\mu}^\dagger W^\mu. \quad (2-27)$$

In the large mass $M$ limit, one may neglect the kinetic term $-\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$ with respect to the mass term, the field $W$ becomes a simple auxiliary field with no dynamics, that one may integrate out by “completing the square”, and one recovers $\mathcal{L}_\text{Fermi}$ provided

$$\frac{G}{\sqrt{2}} = \frac{g^2}{M^2} \quad (2-28)$$

which relates the new coupling $g$ to the Fermi constant $G$. Is the theory (2-27) with its “intermediate boson” $W$, vector of weak interactions, a good theory of weak interactions? In fact the propagator of the massive $W$ field reads

$$-\frac{ig_{\mu\nu} - \frac{k_{\mu}k_{\nu}}{M^2}}{k^2 - M^2} \quad (2-29)$$

which has a bad behavior as $k \gg M$ and makes again the theory non-renormalizable: the problem has just been shifted! The solution stems from a soft and subtle (!) way to introduce the mass of the $W$ field, via a spontaneous breaking of gauge symmetry.

### 5.2.2 Spontaneous breaking of gauge symmetry. Brout–Englert–Higgs mechanism

Let us return to the abelian case described by (5-1), (5-4) and suppose now that the potential $V$ has a minimum localized at a non–zero value of $\phi^* \phi$. Consequently, the field $\phi$ acquires a vev $\langle \phi \rangle = v/\sqrt{2} \neq 0$. Reparametrizing the field $\phi$ according to

$$\phi(x) = e^{iq\theta(x)/v} \frac{v + \varphi(x)}{\sqrt{2}} \quad (2-30)$$

with $v$ real and $\varphi$ hermitian, and accompanying it by a U(1) gauge transformation

$$\phi(x) \mapsto \phi'(x) = e^{-iq\theta(x)/v} \phi(x) = \frac{v + \varphi(x)}{\sqrt{2}}$$

$$A_\mu(x) \mapsto A'_\mu(x) = A_\mu(x) + \frac{1}{v} \partial_\mu \theta(x) \quad (2-31)$$

together with the corresponding transformation of other charged fields ($\psi \ldots$), one sees that the Lagrangian $\delta \mathcal{L}$ of (5-4) reads

$$\delta \mathcal{L} = (\partial_\mu - iqA'_\mu) \phi' (\partial^\mu + iqA'_\mu) \phi' - V(\phi^2) - \frac{1}{2} |(\partial_\mu - iqA'_\mu) \varphi|^2 + \frac{1}{2} q^2 v^2 A'_\mu A^\mu - V \left( \frac{1}{2} (v + \varphi)^2 \right) \quad (2-32)$$
Finally, one sees that the spontaneous breaking of the $U(1)$ symmetry by the boson field $\phi$ leads to the appearance of a mass term for the gauge field $A'_\mu$! One also notes that the field $\theta$ which, in the absence of a gauge field, would be the Goldstone field, has completely disappeared, being “swallowed” by the new massive (“longitudinal”) mode of the vector field $A_\mu$; the total number of degrees of freedom of these fields is thus not modified: we started with 2 transverse modes of the massless electromagnetic field + 2 modes of the charged field (its real and its imaginary parts, say) and we end up with $3 + 1$. This is the Brout–Englert–Higgs mechanism, in its abelian version. If the boson $\phi$ is coupled to a fermion field $\psi$ by a term of the type $\bar{\psi}\phi\psi$, the appearance of its “vev” gives rise to a mass term $\frac{G}{\sqrt{2}}\bar{\psi}\psi$ for the $\psi$.

This Brout–Englert–Higgs mechanism extends to a non-abelian group. The details depend on the scheme of breaking and on the choice of representation for the boson field. (See the courses of second semester for a detailed analysis.) In general, if the group $G$ is broken into a subgroup $H$, the $r = \dim G - \dim H$ would-be Goldstone bosons, that are in correspondence with generators of the “coset” $G/H$, become longitudinal massive modes of $r$ vectors. It remains $\dim H$ massless vector fields. Example: in the electro-weak standard model of § 5.3.2 below: $G = SU(2) \times U(1)$, $H = U(1)$ (not the $U(1)$ factor of $G$!), three gauge fields become massive, one remains massless.

A crucial step in the construction of the standard model was to understand that this spontaneous symmetry breaking in a gauge theory, that we just described at the classical level, is compatible with the quantization of the theory. Renormalizability in 4 dimensions of the gauge theory is not affected by that breaking, and the resulting theory is unitary. Only physical states (massive gauge fields, remaining bosons after the symmetry breaking, etc) contribute to the sum over intermediate states in the unitarity relation.

### 5.3 The standard model

What is presently called the standard model of particle physics is a gauge theory based on a non simple gauge group: $U(1 \times SU(2) \times SU(3)$, in which the different factors play distinct roles. As the group has three simple factors, the theory depends a priori on three independent coupling constants, with gauge fields for each, that are coupled to matter fields, quarks and leptons, as well as to boson fields that play an auxiliary but crucial role!

#### 5.3.1 The strong sector

The group $SU(3) \equiv SU(3)_c$ is the gauge group of color (see Chap. 4, § 3.2.). The gauge fields $A_\mu$ have indices of the adjoint representation (of dimension 8). The associated particles, called gluons, with spin 1 and zero mass, have never been directly observed so far. The gluon fields are coupled to color degrees of freedom of fermionic quark fields, $\psi_{Ai}$, which carry an index $A$ of the representation 3 (or $\bar{3}$ for the $\bar{\psi}$) (and also a flavor index $i = u, d, s, c, b, t$, on which the color group $SU(3)_c$ does not act). The theory so defined is Quantum Chromodynamics (QCD in short). It describes the physics of all strong interactions. Its Lagrangian is of the type (2-24),
with fermionic mass terms depending on flavor, generated by the electroweak sector.

### Asymptotic freedom

Knowing the coupling constant renormalization, (2-25), one may compute the corresponding beta function. One finds

\[ \beta(g) = -\frac{\Lambda}{\partial \Lambda} g(\Lambda) \bigg|_{g_0} = -\frac{g^3}{(4\pi)^2} \left( \frac{11}{3} C_2 - \frac{4}{3} T_f \right) + O(g^5) \]  

(2-33)

It thus appears that this beta function is negative in the vicinity of \( g = 0 \), as long as the coefficient \( \frac{11}{3} C_2 - \frac{4}{3} T_f > 0 \) (not too many matter fields!), in other words that \( g = 0 \) is an attractive ultraviolet fixed point of the renormalization group: \( \frac{d g^2(\lambda)}{d \log \lambda} < 0 \Rightarrow g^2(\lambda) \sim (2b \log \lambda)^{-1} \to 0 \) as \( \lambda \to \infty \), with \( b = \) coefficient of the term \(-g^3\) in (2-33). This is **asymptotic freedom**, a fundamental property of strong interactions. Exercise : how many triplets of quarks are compatible with asymptotic freedom of QCD ?

This non-abelian gauge theory is the only local and renormalizable theory in 4 dimensions that possesses that property of asymptotic freedom. As such, it is the only one consistent with results of deep-inelastic scattering experiments of leptons off hadrons, that reveal the internal structure of the latter as made, at very short distances, of quasi-free constituents.

This SU(3)\(_c\) gauge group is not broken, either explicitly, or spontaneously. This is essential for the consistency of the scenario imagined to account for the quark confinement of quarks and gluons (see Chap. 4, § 3.2.) : non singlet particles of the gauge group are supposed inobservable, as bound to one another inside singlet states and being submitted to forces of growing intensity as one attempts to pull them apart.

This “infrared slavery” (infrared = large distance) is the reciprocal of asymptotic freedom. It shows that confinement is a strong coupling phenomenon, which is by essence non perturbative, namely inaccessible to perturbative calculations.

A non-perturbative approach that has provided many qualitative and quantitative results is the discretization of QCD into a **lattice gauge theory**. This opened the possibility to use methods borrowed to Statistical Mechanics of lattice models, either analytical (strong coupling or high temperature calculations, mean field, . . . ) or numerical (Monte-Carlo). The confinement scenario seems confirmed in that approach by the study of the expectation value of the Wilson loop defined above (§ 5.1.2). Following the idea of Wilson and Polyakov, for a rectangular loop \( C \) of dimensions \( T \times R, T \gg R \), and carrying the representation \( \sigma \) of the gauge group, \( W^{(\sigma)}(C) \) describes the evolution during time \( T \) of a pair of static particles (of very high mass), belonging to representation \( \sigma \), and “frozen” at a relative distance \( R \). One wants to compute the potential between these static charges

\[ V_\sigma(R) = -\lim_{T \to \infty} \frac{1}{T} \log W^{(\sigma)}(C). \]

If the Wilson loop has an “area law”, \( \log W(C) \sim -\kappa RT \), the potential between static charges grows linearly at large distance, \( V \sim \kappa R \), which is accord with the idea of confinement. This is what happens in general in a lattice gauge theory at strong coupling, see the Problem at the end of this chapter. The Monte-Carlo computations confirm that this behavior persists at weak couplings, which are relevant for making contact with the continuous theory (the coupling of the lattice theory must be thought of as the effective coupling at the scale of the lattice spacing \( a \), thus, according to asymptotic freedom, \( g^2 = g^2(\Lambda = 1/a) \to 0 \). These Monte Carlo computations allow to determine numerically the coefficient \( \kappa \) in \( V \), or string tension.

---

6David J. Gross, H. David Politzer, Frank Wilczek, Nobel prize 2004
5.3. THE STANDARD MODEL

QCD is still a very active research field. Strong interactions are indeed ubiquitous in particle physics and the observation of any other interaction, of any other effect, assumes a knowledge as precise as possible of the strong contribution. At the time LHC has started to yield data, precise calculations of QCD contributions are of fundamental importance.

5.3.2 The electro-weak sector, a sketch

The group $\text{U}(1) \times \text{SU}(2)$ describes the electro-weak interactions (Glashow–Salam–Weinberg model\(^7\)). Generators of these groups $\text{SU}(2)$ and $\text{U}(1)$ are referred to as weak isospin and weak hypercharge. We present only the main lines of that construction, without explaining the details nor the reasons that led to choices of groups, of representations etc.

Call $A^a_\mu$, $W^i_\mu$ and $B_\mu$ the gauge fields of $\text{SU}(3)$, $\text{SU}(2)$ and $\text{U}(1)$ respectively. The left-handed, $\psi_L := \frac{1}{2}(1 - \gamma_5)\psi$, and right-handed, $\psi_R := \frac{1}{2}(1 + \gamma_5)\psi$, quarks and leptons are coupled to fields $W_\mu$ and $B_\mu$ in a different way. One writes the covariant derivative of one of these fields as

$$D_\mu \psi = (\partial_\mu - g_3 A^a_\mu T^a - g_2 W^j_\mu t^j - ig_1 y B_\mu) \psi$$

where the representations assigned to each field, either lepton or quark, left or right, are the triplet representation of $\text{SU}(3)_c$ for quarks and the trivial one for leptons, of course, and for the electroweak part, are given in the following Table

<table>
<thead>
<tr>
<th>Quarks &amp; Leptons</th>
<th>$\nu^e_L, e_L$</th>
<th>$\nu^e_R$</th>
<th>$e_R$</th>
<th>$(u_L, d_L)$</th>
<th>$u_R$</th>
<th>$d_R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak isospin $t_z$</td>
<td>$(\frac{1}{2}, -\frac{1}{2})$</td>
<td>$0$</td>
<td>$0$</td>
<td>$(\frac{1}{2}, -\frac{1}{2})$</td>
<td>$0$</td>
<td>$0$</td>
</tr>
<tr>
<td>Weak hypercharge $y$</td>
<td>$(-1, -1)$</td>
<td>$0$</td>
<td>$2$</td>
<td>$(\frac{1}{3}, \frac{1}{3})$</td>
<td>$\frac{2}{3}$</td>
<td>$-\frac{2}{3}$</td>
</tr>
<tr>
<td>Electric charge $Q = \frac{1}{2}y + t_z$</td>
<td>$(0, -1)$</td>
<td>$0$</td>
<td>$-1$</td>
<td>$(\frac{2}{3}, -\frac{1}{3})$</td>
<td>$\frac{2}{3}$</td>
<td>$-\frac{1}{3}$</td>
</tr>
</tbody>
</table>

Table 1. Weak quantum numbers of leptons $\nu^e$ and $e$ and of quarks $u, d$.

Things repeat themselves identically for the next generations.

A remarkable consequence of the use of $\text{SU}(2)$ as a symmetry group of weak interactions is that, beside the two charged currents $J^{1,2}_\mu$ (or $J^\pm_\mu$) of Fermi theory, a third component $J^3_\mu$ appears. This neutral current, which is not the electromagnetic current and which is coupled to the gauge field $W^0_3$, is necessarily present and contributes for example to the $e^-\nu_e \rightarrow e^-\nu_e$ scattering which is forbidden in Fermi theory. The experimental discovery of these neutral currents (1973)\(^8\) was the first confirmation of the validity of the Standard Model.

The group $\text{U}(1)_{\text{em}}$ of electromagnetism is identified thanks to the charges of the fields. There is a “mixing” of the initial $\text{U}(1)$ factor and of a $\text{U}(1)$ subgroup of $\text{SU}(2)$. This mixing is characterized by an angle $\theta_W$, called Weinberg angle: if the $\text{U}(1)$ and $\text{SU}(2)$ gauge fields are denoted $B_\mu$ and $W_\mu$ respectively, the electromagnetic field is $A^\text{em}_\mu = \cos \theta_W B_\mu + \sin \theta_W W^0_\mu$, with the orthogonal combination corresponding to another neutral vector field called $Z^0$.

\(^7\)S. Glashow, A. Salam, S. Weinberg, Nobel prize 1979

\(^8\)The history of that discovery may be read in http://cerncourier.com/cws/article/cern/29168
Let us examine the “neutral current” terms that couple for example the electron and its neutrino to neutral boson vectors $W^0$ and $B$. They are read off the covariant derivatives (2-34) with quantum numbers of Table 1

$$\frac{1}{2} i \left[ \bar{e}_L (-g_2 W^0 - g_1 B) e_L + \bar{e}_R (-2g_1 B) e_R + \bar{\nu}_e (g_2 W^0 - g_1 B) \nu_e \right]$$

The rotation $W^0 = \cos \theta_W Z^0 + \sin \theta_W A$, $B = -\sin \theta_W Z^0 + \cos \theta_W A$ must be such that the electric charge $e$ (coupling to $A$) is the same for $e_L$ and $e_R$ and zero for $\nu_e$. One finds

$$2e = g_2 \sin \theta_W + g_1 \cos \theta_W = 2g_1 \cos \theta_W \quad \text{et} \quad g_2 \sin \theta_W - g_1 \cos \theta_W = 0$$

which are indeed compatible and give

$$\tan \theta_W = \frac{g_1}{g_2} \quad e = g_1 \cos \theta_W = g_2 \sin \theta_W.$$

The result of this calculation does not of course depend on the representation in which it is carried. At this stage we have just made a change of parameters, $(g_1, g_2) \mapsto (e, \theta_W)$ but the latter are physically observable.

The Lagrangian contains also a coupling to a boson field of spin 0, assumed to be a complex doublet of SU(2) $\Phi = \left( \begin{array}{c} \phi^+ \\ \phi^0 \end{array} \right)$ (weak isospin $\frac{1}{2}$, weak hypercharge $y = +1$), $|D\Phi|^2 \equiv D_\mu \Phi = (\partial_\mu - g_2 W_\mu^i \frac{\tau^i}{2} - \frac{1}{2} g_1 B_\mu) \Phi$. The field $\Phi$ is endowed with a potential $V(\Phi)$ with a “mexican hat” shape, which is responsible of the spontaneous breaking of $U(1) \times SU(2)$ into $U(1)_{em.}$ and hence of the generation of the masses of vector fields according to the mechanism described in § 5.2.2, and also of those of fermions. This Higgs field (2 complex components, hence 4 Hermitian ones) has three of its components that disappear, traded for longitudinal modes of massive gauge fields. Only one of these four components remains, and this is that component $\varphi$ (and the “Higgs particle” it creates) that is actively searched in current experiments ATLAS and CMS at LHC. In parallel, three of the four gauge fields, the $W^\pm$ and the $Z^0$, become massive, whereas the fourth, the electromagnetic field $A$ remains massless.

The symmetry breaking of $U(1) \times SU(2)$ by the field $\Phi$ occurs in a direction that preserves $U(1)_{em.}$ (Or more exactly the direction of that breaking determines what is called $U(1)_{em.}$.) One writes, by generalizing (2-30) to the SU(2) group of generators $i\tau^j_\tau$ ($\tau^j$ = Pauli matrices)

$$\Phi(x) = e^{i \xi_\tau(x)} \frac{\tau^j}{2} \left( \begin{array}{c} 0 \\ \sqrt{\xi_\tau(x)^2 + \frac{g_1^2}{g_2^2}} \end{array} \right),$$

which is accompanied by a gauge transformation, which causes the fields $\xi_\tau$ to disappear and gives the fields $W$ and $B$ the quadratic form (mass terms)

$$\mathcal{L}_{(2)} = \frac{1}{8} \varepsilon^{ij} \left[ (g_1 B - g_2 W^3)^2 + g_2^2 (|W|^2 + (W^2)^2) \right]$$

As expected the component $Z^0 = (g_1 B - g_2 W^3)/\sqrt{g_1^2 + g_2^2}$ becomes massive, and so do $W^{1,2}$, whereas the orthogonal combination $A = (g_2 B + g_1 W^3)/\sqrt{g_1^2 + g_2^2}$ remains massless. One finds

$$M_{W^\pm} = \frac{1}{2} g_2 \sqrt{g_1^2 + g_2^2} \quad M_{Z^0} = \frac{1}{2} g_1 \sqrt{g_1^2 + g_2^2}$$

(2-36)

and using (2-35), the relation $G = \frac{g_2^2}{3M_W}$ read off the Lagrangian and the experimental values of $e$ and of $G = 10^{-5} m_e^2$ one computes

$$M_{W^\pm} = \frac{38}{\sin \theta_W} \text{GeV} \quad M_{Z^0} = \frac{M_W}{\cos \theta_W} = \frac{38}{\sin \theta_W \cos \theta_W} \text{GeV}.$$
These expressions then undergo small perturbative corrections. Lastly the mass of the famous and so eagerly expected Higgs boson $\phi$ is not predicted by the theory but should lie in the range 100–200 GeV. The latest results from LHC seem to exclude most of that range, leaving only a “window” between 120 and 140 GeV for its mass in the above simplest scenario.

The “intermediate bosons” associated with the massive vector fields $W^\pm$ and $Z^0$ have been discovered experimentally at the end of the seventies$^9$; their masses $M_{W^\pm} = 80.4\text{ GeV}$ and $M_{Z^0} = 91.2\text{ GeV}$ are compatible with the following value of the Weinberg angle

$$\sin^2 \theta_W \approx 0.23,$$

which is also compatible with all the other experimental results.

To summarize, the Lagrangian that describes all interactions but gravitation has a remarkably simple and compact form

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \sum \bar{\psi}_i \gamma^\mu D_\mu \psi_i + |D\Phi|^2 - V(\Phi) + \text{Higgs - fermions couplings},$$

where $F_{\mu\nu}$ denote the gauge field tensors of $A$, $W$ and $B$. Note that the $SU(2) \times U(1)$ invariance forbids couplings between left and right fermions (which transform under different representations), and thus forbids fermionic mass terms. The only mass scale lies in $V(\Phi)$, and it is the Higgs mechanism and the coupling of $\Phi$ to fermions –leptons and quarks– which give rise to the masses of fermions and of (some of) vector bosons. This coupling, called after Yukawa, has the general form (written for quarks),

$$\mathcal{L}_Y = -Y^d_{ij} \bar{\psi}_{Li}^d \Phi d_R j - Y^u_{ij} \bar{\psi}_{Li}^u \bar{\Phi}^i u_R j + \text{h.c.},$$

with a priori arbitrary matrices $Y^d_{ij}$, $Y^u_{ij}$. $i, j = 1, 2, 3$ are generation indices, the dot denotes the scalar product of isospin doublets $\Phi$ and $\bar{\Phi}^i = \left( \begin{array}{c} \phi^0_1 \\ -\phi^+_{-1} \end{array} \right)$ with quark doublets

$$\psi_{Li} = \left( \begin{array}{c} u_i \\ d_i \end{array} \right)_L = \left( \begin{array}{c} u \\ d \end{array} \right)_L \left( \begin{array}{c} c \\ s \end{array} \right)_L \left( \begin{array}{c} t \\ b \end{array} \right)_L.$$

Couplings of the same type appear between leptons and scalar fields.

The vev $v/\sqrt{2}$ of $\phi^0$ then gives rise to a “mass matrix”. A complication of the theory described by (2-38) is that the diagonalization of that quark mass matrix involves a unitary rotation of $(u_L, c_L, t_L)$ and of $(d_L, s_L, b_L)$ with respect to the basis coupled to gauge fields in (2-38): if $(u_L, c_L, t_L)$ and $(d_L, s_L, b_L)$ stand for the mass eigenstates, the charged hadronic current coupled to the field $W^+$ is

$$J_\mu = (\bar{u} \gamma^\mu \ell)L \gamma_\mu M \left( \begin{array}{c} d \\ s \\ b \end{array} \right)_L,$$
CHAPTER 5. GAUGE THEORIES. STANDARD MODEL

with $M$ the unitary Cabibbo-Kobayashi-Maskawa matrix\(^\text{10}\). This mechanism generalizes to 3 generations the mixing with the Cabibbo angle encountered in Chap. 4, (equ. (4-33)) in the case of 2 generations. The matrix $M$ is written as

$$M = \begin{pmatrix}
V_{ud} & V_{us} & V_{ub} \\
V_{cd} & V_{cs} & V_{cb} \\
V_{td} & V_{ts} & V_{tb}
\end{pmatrix} =
\begin{pmatrix}
c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\
-s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\
s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13}
\end{pmatrix}$$

with 4 angles $\delta$ and $\theta_{ij}$, ($c_{ij} = \cos \theta_{ij}$ and $s_{ij} = \sin \theta_{ij}$), and $\theta_{12} = \theta_C = $ Cabibbo angle. Experimentally, $0 \ll \theta_{13} \ll \theta_{23} \ll \theta_{12} \ll \pi/2$. The accurate measure of the matrix elements of $M$ is presently the object of an intense experimental activity, in connection with the study of violations of the $CP$ symmetry (due to a large extent to the phase $e^{i\delta}$) and of “flavor oscillations”.

For a much more comprehensive discussion of details and achievements of the standard model, see the courses of the 2nd semester.

5.4 Complements

5.4.1 Standard Model and beyond

The Standard Model is both remarkably well verified and not very satisfactory. Beside massive neutrinos, whose existence is now beyond any doubt, and which request little amendments to the Lagrangian (2-38), no significative disagreement has been found to this day between experimental results and predictions of the model. Still, the non satisfactory aspects of the standard model are numerous: the excessively large number (about twenty) of free parameters in the model, the lack of “naturalness” in the way certain terms have to be adjusted in a very fine way; the question of the Higgs boson whose discovery is expected at LHC, but that many physicists regard as an ad hoc construction; etc.

Attempts at improving the standard model by fusing the three gauge groups within a larger group, in a “grand-unified” (GUT) theory should be mentionned. The next subsection is devoted to that issue.

The currently most popular extensions of the standard model are those based on supersymmetry. The “MSSM”, (“Maximally Supersymmetric (extension of the) Standard Model”), or the “NMSSM” (“Next-to . . .”), resolve the hierarchy problem, predict a convergence of electro-weak and strong couplings and also the existence of supersymmetric partners for all known particles. For confirmation of this scheme, people eagerly await results from LHC . . .

5.4.2 Grand-unified theories or GUTs

An empirical observation is that the three coupling constants $g_1, g_2, g_3$, starting from their value experimentally values measured at current energies, seem to converge under the renormalization group flow to a common value at some energy of the order of $10^{15}$ or $16$GeV. This was a very strong incentive towards a grand-unification, see

\(^{10}\)M. Kobayashi, T. Maskawa, Nobel prize 2008, avec Y. Nambu
5.4. COMPLEMENTS

Fig. 5.2. The resulting grand-unified theory should not only be a gauge theory with a single coupling if the unification group \( G \) is simple, but also be capable of predicting the matter field and particle content according to the representations of \( SU(3) \times SU(2) \times U(1) \) from some representations of the group \( G \). For various reasons, the group \( SU(5) \) turns out to be the best candidate. This GUT possesses \( \text{dim} \, SU(5) = 24 \) gauge fields.

The main reason of that choice of \( SU(5) \) comes from the number of chiral fermions by generation. Each generation of the standard model contains two quark flavors coming each in 3 colors, plus one lepton, and each of these 6+1 fields may have two chiralities, plus a neutrino assumed to be massless and chiral. In total there are 15 chiral fermions per generation. (Remember that the antiparticle of a right fermion is left-handed: it is thus sufficient to consider left fermions.) One thus seeks a simple group \( G \) possessing a representation (reducible or irreducible) of dimension 15, that may accommodate all left-handed fermions of each generation. The only candidate is finally the group \( SU(5) \) which has representations of dimension 15: the symmetric tensor representation, and representations sums of 5 (or 5) and 10 (or 10).

The group \( SU(5) \) of unitary 5×5 matrices contains a \( SU(3) \) subgroup (3×3 submatrices of the upper left corner, say) and a \( SU(2) \) subgroup (2×2 blocks of the lower right corner), which give the corresponding generators of \( SU(3) \times SU(2) \); the \( U(1) \) subgroup is generated by the diagonal traceless matrix diag \((-\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}, \frac{1}{2}, \frac{1}{2})\). It is clear that these three subgroups commute with one another.

One must then decompose all fields (in representations 5, 10, 15 and 24) into representations of \( SU(3) \times SU(2) \). This exercise shows that representation 15 must be discarded and that the reducible representation \( \bar{5} \oplus 10 \) is the appropriate one for fermion fields: the 5 decomposes into representations \((3, 1) \oplus (1, 2)\) and contains antiquarks \( \bar{d}_L \) and left leptons \( e_L \); the 10 decomposes into \((1, 1) \oplus (3, 2) \oplus (3, 1)\) containing the left lepton \( e^+_L \), singlet of \( SU(2) \) and of \( SU(3) \), the two left quarks \( u_L, d_L \) which form a doublet of \( SU(2) \) and the antiquarks \( \bar{u}_L \).

Likewise, the 24 gauge fields include the 8 gluon fields, the 3+1 vectors of the electroweak sector, plus 12 supplementary fields, which acquire a very large mass at the expected breakdown of \( SU(5) \rightarrow SU(3) \times SU(2) \times U(1) \).

The breakdown \( SU(5) \rightarrow SU(3) \times SU(2) \times U(1) \) should take place at a grand-unification energy of the order of \( 10^{15} \) or \( 10^{16} \) GeV, an energy where couplings \( g_3, g_2, g_1 \) of \( SU(3), SU(2) \) and \( U(1) \) seem to converge (Fig. 5.2). The infinitesimal generators being now rigidly bound within the simple group \( SU(5) \), one may related the couplings to the \( U(1) \) and \( SU(2) \) gauge fields and predict the Weinberg angle: one finds \( \sin^2 \theta = \frac{3}{8}, \ldots \) but this calculation applies to the unification energy! The angle is renormalized between that energy and energies of
current experiments.

A striking consequence of the quarks–leptons unification within SU(5) multiplets is a violation of separate conservations of lepton and baryon numbers. In particular the existence of interaction terms, for example \( X^\rho (\bar{d} \gamma^\rho e^+ + \bar{u} \gamma^\rho u) \), with one of the new gauge fields (the matrices of the generators have been omitted), allows proton decay \( p = d u u \rightarrow d e^+ + \pi^0 e^+ \), and by other channels as well. The decay rate must be carefully computed to see if it is consistent with experimental data on proton lifetime (present bound \( 10^{32 \pm 1} \) years),... which is not the case !

One should also show to which representation the Higgs boson fields belong to permit a two-step breaking \( SU(5) \rightarrow SU(3) \times SU(2) \times U(1) \rightarrow SU(3) \times U(1) \) with two very different breakings...

Finally, the SU(5) GUT

- incorporates by construction the structure of fermion generations;
- puts leptons and quarks in the same representation and explains the commensurability of their electric charges and the cancellation of anomalies (see discussion below);
- reduces the number of parameters in the standard model and predicts the value of the Weinberg angle (at the unification scale);
- does not explain the why of the three observed generations;
- does not elucidate the question of “naturalness” (just evoked above), nor the related issue of “hierarchy” (why is the ratio \( M_{GUT}/M_W \) so large?) ;
- last but not least, fatal disease, predicts effects such as the proton decay at rates seemingly inconsistent with observation.

This is the latter point that led to abandon this unification scheme and to favor supersymmetric routes to unification.

### 5.4.3 Anomalies

We mentioned in Chap. 4 the existence of chiral anomalies, that affect the axial current \( J^{(5)}_\mu \) of the classical \( U(1) \) symmetry. In the gauge theory of the Standard Model, the electroweak gauge fields are coupled differently to left-handed and right-handed fermions, more precisely, they are coupled to axial currents, see the Lagrangian

\[
\mathcal{L} = i \bar{\psi} (\bar{\theta} - A) \left( \frac{1 - \gamma_5}{2} \right) \psi
\]

which contains a term \( A_\mu J^\mu \) with \( J^\mu = \bar{\psi} T_a \left( \frac{1 - \gamma_5}{2} \right) \psi \). Classically that current \( J^\mu \) should have a vanishing covariant derivative (in the adjoint representation) if the fermions are massless. One may again compute the (covariant) divergence of that current to the one-loop order, and one finds that

\[
D_\mu J^\mu = \frac{i}{24\pi^2} \partial_\mu \epsilon^{\mu\nu\rho\sigma} \text{tr} T_a (A_\nu \partial_\rho A_\sigma + \frac{1}{2} A_\nu A_\rho A_\sigma) .
\]

Curiously the right hand side is not gauge invariant, but its forms is not arbitrary and is dictated by geometric considerations (“descent equations”) that are beyond the scope of the present discussion. The anomaly of this “non-singlet” current (\( i.e. \) carrying a non-trivial representation of the gauge group) thus breaks gauge invariance. As such, it jeopardizes all the consistency, renormalisability and unitarity, of the theory. One conceives that controlling this anomaly is crucial for the construction of a physically sensible theory.

Then one observes that the “group theoretical coefficient” of the anomaly is proportional to

\[
d_{abc} = \text{tr} \left( T_a \{ T_b, T_c \} \right)
\]

where \( \{ T_b, T_c \} \) is the anticommutator of infinitesimal generators, see Exercise B.3.

In practice one ensures the anomaly cancellation in two cases:
• a) Suppose that the fermions all belong to real or pseudoreal representations. One recalls (see Chap 2) that this refers to situations where the representation is (unitarily) equivalent to its complex conjugate representation, \( T_a^* = C T_a C^{-1} \). In unitary representations the \( T_a \) are antihermitian,\( T_a = -T_a^\dagger = -T_a^{T^*} \). One then verifies (see Exercise B.3) that the group theoretical factor \( d_{abc} = -d_{abc} = 0 \) vanishes and so does the anomaly. Thus (4-dimensional) theories with gauge group \( SU(2) \) (in which all representations are real or pseudoreal) have no anomaly.

• b) Another situation is that there is cancellation of anomalies coming from different fermion representations. This is what takes place in the standard model. According to the argument of a), there is no anomaly associated with the weak isospin currents, coupled to an \( SU(2) \) gauge field. But there may a priori be some with weak hypercharge currents (\( U(1) \) group), as well as mixed anomalies, for example one \( U(1) \) current and two \( SU(2) \) etc. One must thus check that for all choices of three generators labelled by \( a, b, c \), the constant \( d_{abc} \) vanishes when one sums over all fermion representations. Finally one shows that it reduces to the vanishing of \( \text{tr}(t_2^3 Q) \) for each generation, which is indeed satisfied in the Standard Model. This is also what happens for the \( SU(5) \) theory discussed in the previous section: one shows that for each generation, contributions of representations 5 and 10 cancel one another.

Further references for Chapter 5


On gauge theories, Yang-Mills, the standard model, etc, one may consult any book of quantum field theory later than 1975, for example [IZ], [PS], [Wf], [Z-J].

On group theoretical aspects of gauge theories, voir L. O’Raifeartaigh, op. cit..

A very good review of grand-unification is given in *Introduction to unified theories of weak, electromagnetic and strong interactions - SU(5)*, A. Billoire and A. Morel, rapport Saclay DPh-T/80/068.


Exercises for chapter 5

A. Non abelian gauge field
1. Complete the proofs of (5-21) et (5-22).
2. Let \( A \) be a non abelian gauge field and \( F \) its field tensor. Show that the covariant derivative of \( F \) is such that
   \[
   D_{\mu a} F_{\nu b} k^a = [D_\mu, F_{\nu \rho}] = \partial_\mu F_{\nu \rho} - [A_\mu, F_{\nu \rho}].
   \]
   Prove the identity
   \[
   [D_\mu, F_{\nu \rho}] + [D_\nu, F_{\rho \mu}] + [D_\rho, F_{\mu \nu}] = 0.
   \]
   Recall what is the abelian version of that identity and its interpretation.
3. Consider the operator \( \slashed{D} = \partial - A \) acting on Dirac fermions in a representation \( R \). One wants to compute \( \slashed{D}^2 \). Writing \( D_\mu D_\nu \gamma^\mu \gamma^\nu = \frac{1}{2} D_\mu D_\nu \{\gamma^\mu, \gamma^\nu\} + \frac{1}{2} [D_\mu, D_\nu] \gamma^\mu \gamma^\nu \), show that one may write \( \slashed{D}^2 \) as a sum of \( D^2 = D_\mu D^\mu \) and of a term of the form \( a F_{\mu \nu} \sigma^{\mu \nu} \), where \( \sigma^{\mu \nu} = \frac{i}{2} [\gamma^\mu, \gamma^\nu] \). Compute \( a \).

B. Group theoretical factors…
1. Casimir operators
   Let \( G \) be a simple compact Lie group of dimension \( d \), \( R \) one of its representations, that one may assume irreducible and unitary. Let \( t_a \) be a basis of the Lie algebra \( g \) of \( G \), \( T_a \) its representatives in representation \( R \). The \( t_a \) and \( T_a \) are chosen antihermitian. One then considers the bilinear form on the Lie algebra defined by
   \[
   (X, Y)^{(R)} = \text{tr} (T_a T_b) x^a y^b
   \]
   if \( X = x^a t_a \) and \( Y = y^b t_b \in g \) (with summation over repeated indices).
   \begin{itemize}
   \item Prove that this form is invariant in the sense that
     \[
     \forall Z \in g \quad ([X, Z], Y)^{(R)} + (X, [Y, Z])^{(R)} = 0.
     \]
   \end{itemize}
   One recalls that any invariant bilinear form on a simple Lie algebra is a multiple of the Killing form.
   \begin{itemize}
   \item b) Prove that one may choose a basis of \( t_a \) and hence of \( T_a \) such that
     \[
     \text{tr} (T_a T_b) = -T_R \delta_{ab}
     \]
     with \( T_R \) a coefficient that depends on the representation.
   \item c) What is the sign of \( T_R \) ?
   \item d) Consider then the quadratic Casimir operator
     \[
     C_2^{(R)} = - \sum_a (T_a)^2.
     \]
     On how many values of \( a \) does one sum in that expression?
5.4. COMPLEMENTS

• e) Recall why \( C_2^{(R)} \) is a multiple of the identity in the representation space of \( R \)

\[
C_2^{(R)} = c_2(R) \mathbb{1}
\]

• f) Why are the assumptions of simplicity of \( G \) and of irreductibility of \( R \) important for that result?

• g) What is the sign of \( c_2(R) \)? Justify.

• h) Show that \( T_R \) is related to the value \( c_2(R) \) of the quadratic Casimir operator. For that purpose, one may compute in two different ways the quantity

\[
\text{tr} \sum_a (T_a)^2.
\]

• i) To what does this relation boil down for the adjoint representation \( \text{adj} \) of \( G \)?

• j) Normalize the (antihermitian) generators of \( SU(N) \) in such a way that in the defining representation, \( \text{tr} T_a T_b = -\frac{1}{2} \delta_{ab} \), thus \( T_f = \frac{1}{2} \). Is this verified by infinitesimal generators \( i \frac{\sigma}{2} \) of \( SU(2) \)? What is then the value of \( c_2 \) in that defining representation?

2. Computation of traces and of Casimir operators in representations of \( SU(N) \)

• a) Show that the expression (3-49) of Chap. 3 \( c_2(\Lambda) = \frac{1}{2} (\text{tr} (\Lambda^2 + 2 \rho \Lambda + \rho^2)) \) may be rewritten as \( c_2(\Lambda) = \frac{1}{2} ((\Lambda^2 + 2 \rho \Lambda + \rho^2) - \rho) \), thus for \( SU(N) \), using expressions (3-47) and (3-60) of the same chapter

\[
c_2(\Lambda) = \frac{1}{2N} \sum_{i=1}^{N-1} \left( [(\lambda_i + 1)^2 - 1]i(N - i) + 2 \sum_{j=i+1}^{N-1} [(\lambda_i + 1)(\lambda_j + 1) - 1]i(N - j) \right).
\]

• b) Compute that expression for the defining representation. Does the result agree with that found in question 1.j) above?

• c) Recall why the highest weight of the adjoint representation is the highest root (denoted \( \theta \) in the Appendix F of Chap 3). Is the expression \( \theta = \Lambda_1 + \Lambda_{N-1} \) in accord with what is known on the adjoint representation?

• d) Calculate the value of \( c_2(\Lambda) \) for the adjoint representation.

• e) Check this value for \( SU(2) \) by a direct calculation of \( c_2(\text{adj}) \).

• f) What is the value of \( T_{\text{adj}} \) in \( SU(N) \), that follows as a consequence of question 1.i) ?

3. Anomaly coefficients

We keep the same notations and conventions as above.
• a) In the computation of some Feynman diagrams in a gauge theory of group $G$, one encounters the coefficient

$$d_{\alpha\beta\gamma} = \text{tr} \left( T_\alpha (T_\beta T_\gamma + T_\gamma T_\beta) \right).$$

Show that $d_{\alpha\beta\gamma}$ is completely symmetric in its three indices.

• b) We recall that a representation is said to be real or pseudoreal if it is (unitarily) equivalent to its complex conjugate, hence if in a basis where the $T_\alpha$ are antihermitian, one may find a unitary matrix $U$ such that the complex conjugate of each $T_\alpha$ verifies

$$(T_\alpha)^* = U T_\alpha U^{-1}.$$ 

Show that if that condition is satisfied, $d_{\alpha\beta\gamma}$ vanishes identically. That condition is important to ensure the consistency of gauge theory, this is the condition of anomaly cancellation.

• c) Is the spin $\frac{1}{2}$ representation of SU(2) real or pseudoreal? That of spin $j$? Justify your answer.

• d) Give two examples of (non necessarily irreducible) non trivial representations of SU(3) that are real or pseudoreal, and two that are not.

• e) What is the coefficient $d$ for the U(1) groupe and a representation of charge $q$?

C. Spontaneous breaking of SU(2) gauge symmetry
Consider an SU(2) gauge theory coupled to a boson field $\vec{\Phi}$ of spin 1, considered as a vector of dimension 3. The potential of that field is denoted $V(\vec{\Phi}^2)$.

1. Write the Lagrangian and the gauge transformations of the fields $\vec{A}_\mu$ and $\vec{\Phi}$.

2. We suppose that the symmetry is spontaneously broken: the field $\Phi$ acquires a vev $v$ along some direction, say 3 : $\langle \vec{\Phi} \rangle = \begin{pmatrix} 0 \\ 0 \\ v \end{pmatrix}$. What is the residual group of symmetry? What will be the effect of the field $A_\mu$? Give a description of the fields and physical particles after symmetry breaking.

Problem. Lattice gauge theory
In the following, $G$ denotes a compact Lie group, $\chi^{(\rho)}$ the character of its irreducible unitary representation $\rho$.

1. Show that the orthogonality relations of $\mathcal{D}^{(\rho)}$ imply the following formulas:

$$\int_G \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g.g_1.g^{-1}.g_2) = \frac{1}{n_\rho} \chi^{(\rho)}(g_1)\chi^{(\rho)}(g_2),$$

(2-41)
and
\[ \int_G \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g.g_1)\chi^{(\sigma)}(g^{-1}.g_2) = \frac{\delta_{\rho,\sigma}}{n_{\rho}} \chi^{(\rho)}(g_1.g_2). \]  
(2-42)

Recall why a representation of \( G \) may always be regarded as unitary and show that then
\[ \chi^{(\rho)}(g^{-1}) = \chi^{(\bar{\rho})}(g) = (\chi^{(\rho)}(g))^*, \]
(2-43)

where \( \bar{\rho} \) is the complex conjugate representation of \( \rho \).

We make a frequent use of these three relations in the following.

2. Let \( \chi \) be the character of a real representation \( r \) (not necessarily irreducible) of \( G \), \( \beta \) a real parameter.

a) Show that one may expand \( \exp \beta \chi(g) \) on characters of irreducible representations of \( G \) according to
\[ e^{\beta \chi(g)} = \sum_{\rho} n_{\rho} b_{\rho}(\beta) \chi^{(\rho)}(g), \]
with functions \( b_{\rho}(\beta) \).

Express the function \( b_{\rho}(\beta) \) in terms of a group integral.

Using (2-43), show that the functions \( b_{\rho}(\beta) \) are real, \( b_{\rho}(\beta) = (b_{\rho}(\beta))^* = b_{\bar{\rho}}(\beta) \).

b) Show that \( b_{\rho} \) is non vanishing provided representation \( \rho \) appears in some tensor power \( r \otimes n \).

c) For \( G = SU(2) \) and \( r = (j = \frac{1}{2}) \), the representation of spin \( \frac{1}{2} \), is condition b) satisfied for any \( \rho \)? Why?

If \( r = (j = 1) \), what are the representations for which \( b_{\rho} \) is a priori zero?

d) For \( G = SU(3) \) and \( \chi = \chi^{(3)} + \chi^{(\bar{3})} \), show that \( b_{\rho} \) is non zero for all \( \rho \).

For \( \beta \rightarrow 0 \), what is the leading behaviour of \( b_{\alpha}(\beta) \) when \( \beta \rightarrow 0 \) if \( \alpha \) denotes the adjoint representation of \( SU(3) \)? More generally what is the leading behaviour of \( b_{\rho}(\beta) \) where \( \rho \) is the representation of highest weight \( \Lambda = (\lambda_1, \lambda_2) \)?

3. One defines a model of statistical mechanics in \( d \) dimensions in the following way. On a hypercubic lattice of dimension \( d \) and of lattice spacing \( a \), the degrees of freedom are attached to links (edges) between neighbouring sites and take their value in the compact group \( G \). With each oriented link \( \ell = \vec{ij} \) one associates the element of \( G \) denoted \( g_\ell = g_{ij} \), with \( -\ell = \vec{ji} \), one associates \( g_{ji} = g_\ell^{-1} \). With each elementary square (alias “plaquette”) \( p = ijkl \), one associates the product of the link elements:
\[ g_p = g_{ij}.g_{jk}.g_{kl}.g_{ti} \]

and the “energy” of a configuration of these variables is given by
\[ E = -\sum_{\text{plaquettes } p} \chi(g_p) \]
(2-44)

where \( \chi \) is, like in question 2, the character of some real representation \( r \) of the group. The Boltzmann weight is thus
\[ e^{-\beta E} = \prod_p e^{\beta \chi(g_p)} \], \[ \beta = \frac{1}{kT} \]
and the partition function reads

\[ Z = \prod_{\text{links } \ell} \int_G \frac{d\mu(g_\ell)}{v(G)} \prod_{\text{plaquettes } p} e^{\beta \chi(g_p)}. \quad (2-45) \]

\( \text{a) } \) Show that the energy \( E \) is invariant by redefinition of \( g_{ij} \) as \( g_{ij} \mapsto g_i g_{ij} g_{ji}^{-1}, \) where \( g_i \in G, \) (this is a \textit{local} invariance, the analogue in that discrete formalism of the gauge invariance studied in this chapter), and that \( E \) does not depend on the orientation of plaquettes.

\( \text{b) } \) One wants to understand the relation with the formalism of § 5.1. The degrees of freedom \( g_{ij} \) represent the path-variables defined in (5-20), \( g_{ij} = g(j,i) \) along the edge from \( i \) to site \( j \)

\( g_{ij} \equiv P \exp \int_{l=i}^{j} A_\mu dx^\mu \)

- For a small lattice spacing \( a, \) show, by using for example the BCH formula and by expanding to the first non vanishing order that

\[ g_p = \exp \left( a^2 F_{\mu\nu} + o(a^2) \right) \]

where \( \mu \) and \( \nu \) denote the directions of the edges of plaquette \( p. \) (One is here interested in an Euclidean version of gauge theory, and position of indices \( \mu, \nu \) is irrelevant.) Show then that the energy \( E_p \) (2-44) reads

\[ E_p \sim \text{const. } a^4(F_{\mu\nu})^2 + \text{const.'} \]

where the first constant will be determined as a function of the representation \( r \) chosen for \( \chi. \)

- Explain why the parameter \( \beta \) identifies (up to a factor) with the inverse of coupling \( g^2 \) in the continuous gauge theory. In fact this is rather the “bare (or unrenormalized) coupling constant”, why?

One first restricts oneself for simplicity to \( d = 2 \) dimensions. For a finite lattice of \( N \) plaquettes, for example a rectangle of size \( L_1 \times L_2 \) (see Fig.5.3), one wants to calculate \( Z. \) One chooses “free boundary conditions”, in other words variables \( g_\ell \) on the boundary of the rectangle are independent. One is also interested in the expectation value \( W^{(\sigma)}(C) \) of \( \chi^{(\sigma)}(g_C) \) where \( g_C \) is the ordered product of \( g_\ell \) along a closed oriented curve \( C \) for some irreducible representation \( \sigma \) of \( G \)

\[ W^{(\sigma)}(C) := \langle \chi^{(\sigma)}(g_C) \rangle = \frac{1}{Z} \prod_{\text{links } \ell} \int_G \frac{d\mu(g_\ell)}{v(G)} \chi^{(\sigma)}(\prod_{\ell \in C} g_\ell) \prod_p e^{\beta \chi(g_p)}. \quad (2-46) \]

\( \text{d) } \) Using the results of question 2 show that one may expand each \( \exp \beta \chi(g_p) \) on characters of irreducible representations of \( G \) according to

\[ e^{\beta \chi(g_p)} = \sum_{\rho} n_{\rho} b_{\rho} \chi^{(\rho)}(g_p). \quad (2-47) \]
e) One inserts in (2-45) or (2-46) the expansion (2-47) for each plaquette. Show that if two plaquettes share one link $\ell$, formulas of part 1 permit an integration over the variable $g_\ell$ of that link and that the two representations carried by the two adjacent plaquettes are then identical. Using repeatedly these formulas of part 1, show that one may integrate over all variables $g_\ell$ and that

$$ Z = b_1^N \quad W^{(\sigma)}(C) = n_\sigma \left( \frac{b_\sigma}{b_1} \right)^A $$

where $A$ is the area of the curve $C$, i.e. the number of plaquettes it encompasses, and the index 1 refers to the identity representation.

f) One now consider the case of dimension $d = 3$. Variables $g_\ell$ are attached to links of a cubic lattice. Energy is again given (2-44), where the sum runs over all plaquettes of this 3-dimensional lattice. As before, $W^{(\sigma)}(C) = \langle \chi^{(\sigma)}(g_C) \rangle$ receives contributions from plaquette configurations that form a surface bounded by $C$.

Let us show that contributions to the Wilson loop $W^{(\sigma)}(C)$ may also come from plaquette configurations forming a tube resting on the contour $C$ (Fig. 5.4).

- Show that for such a configuration, the repeated application of formulas (2-41) and (2-42) on all variables $g_\ell$ leads to the following expression

$$ W^{(\sigma)}(C) \bigg|_{\text{tube}} = \sum_\rho \left( \frac{b_\rho}{b_1} \right)^P \int_G \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g) \chi^{(\rho)}(g^{-1}) \chi^{(\sigma)}(g) $$

where $P$ is the number of plaquettes making the tube.

- Under which condition $\mathcal{C}$ on representation $\sigma$ of the loop $C$ is the contribution of representation $\rho$ to the right hand side of (2-49) non vanishing?

- Give an example for G=SU(2) of representations $\sigma$ for which this condition $\mathcal{C}$ is never satisfied for any $\rho$, and hence these tabular configurations do not contribute.

- Inversely give an example (again for SU(2)) of a possible choice of $\sigma$ which satisfies it.
We admit that at high temperature, (small $\beta$), the dominant contribution to $W^{(\sigma)}(C)$ is of type (2-49) if condition $C$ may be satisfied, and of type (2-48) in the opposite case.

4. The evaluation of the expectation value of the Wilson loop $W^{(\sigma)}(C)$ in the limit of a large loop $C$ which is a rectangle $R \times T$ allows to compute the potential $V_{\sigma}(R)$ between two static “charged” particles separated by distance $R$, one carrying representation $\sigma$ of the group and the other one being its antiparticle. More precisely we admit that

$$V_{\sigma}(R) = -\lim_{T \to \infty} \frac{1}{T} \log W^{(\sigma)}(C).$$

Evaluate the dependence of $V_{\sigma}(R)$ in $R$ which follows either from (2-48), or from the contribution to (2-49) due to representation $\rho$. What do you conclude on the interaction between the two particles in those two situations?

Physically, this kind of considerations gives a discrete (lattice) and simplified (2 or 3 dimensions) model of QCD. One may repeat these calculations in higher dimension, where the above result appear as the leading term in a small $\beta$ (“high temperature”) expansion. The fact that $W^{(\sigma)}(C)$ decays like $x^A$ ($x = b_\sigma/b_1 < 1$ for $\beta$ small enough) for large areas is a signal of quark confinement in that theory, that is of the impossibility to separate a pair quark-antiquark at large distance . . .

*
Index

α-chain, 76

action of a group on a set, 29
adjoint map, 14
adjoint representation, 43, 44, 65, 71, 81, 83, 88, 91, 100
algebra, 11
alternating group, 4
anomalies, 105
anomaly cancellation, 137, 140
asymptotic freedom, 130

Baker-Campbell-Hausdorff formula, 14
baryonic charge, 106
beauty, 115
bottom quark, 115
Brout–Englert–Higgs mechanism, 128
Cabibbo angle, 114
Cabibbo-Kobayashi-Maskawa matrix, 134
Cartan criteria, 20
Cartan subalgebra, 71
Cartan torus, 84
Casimir operator, 21, 30, 67, 80, 83, 138
center, 3
central extension, 16, 56
centralizer, 3, 4
chain of roots, 76
chain of weights, 81
character, 34
character, of SU(2), 59
charm, 114
Chebyshev polynomials, 59, 65
Chevalley basis, 79
chiral symmetry, 104
class function, 35, 47

Clebsch-Gordan coefficients, 41
Clebsch-Gordan decomposition, 39
cocycle, 51
color, 116, 129
commutant, 3
commutator in a group, 13
compact, 25
compact group, 9
compact Lie algebra, 17
compact, locally, 25
complex representation, 68
complexified Lie algebra, 17
confinement of color, 116, 130, 144
conformal transformation, 30
conjugacy classes, 3
connected group, 6
conservation laws, 53
coroot, 80
coset, 4
coupling constant renormalization, 127
covariant derivative, 122
Coxeter exponents, 80
Coxeter number, 80
cyclic group, 1
diffeomorphisms of the circle, 16
Dynkin diagram, 78
Dynkin labels, 82
electromagnetic form factors, 110
electromagnetic mass splittings, 111, 118
exponential map, 13, 26
Fermi constant, 114
Fermi Lagrangian, 114, 127
flavor, 114, 116
<table>
<thead>
<tr>
<th>Index Term</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frobenius–Schur indicator</td>
<td>69</td>
</tr>
<tr>
<td>Frobenius–Weyl duality</td>
<td>93</td>
</tr>
<tr>
<td>fundamental group</td>
<td>7</td>
</tr>
<tr>
<td>Galilean group</td>
<td>2</td>
</tr>
<tr>
<td>gauge invariance</td>
<td>121</td>
</tr>
<tr>
<td>gauge invariant</td>
<td>125</td>
</tr>
<tr>
<td>gauge symmetry</td>
<td>121</td>
</tr>
<tr>
<td>Gell-Mann–Nishima relation</td>
<td>106</td>
</tr>
<tr>
<td>Gell-Mann–Okubo mass formula</td>
<td>111</td>
</tr>
<tr>
<td>generations</td>
<td>116</td>
</tr>
<tr>
<td>Goldstone theorem</td>
<td>103</td>
</tr>
<tr>
<td>grand-unified theory</td>
<td>134</td>
</tr>
<tr>
<td>GUT</td>
<td>134</td>
</tr>
<tr>
<td>Haar measure</td>
<td>10, 26</td>
</tr>
<tr>
<td>Higgs field</td>
<td>132</td>
</tr>
<tr>
<td>homeomorphism</td>
<td>7</td>
</tr>
<tr>
<td>homogeneous space</td>
<td>29</td>
</tr>
<tr>
<td>homomorphism</td>
<td>4</td>
</tr>
<tr>
<td>homotopy</td>
<td>6</td>
</tr>
<tr>
<td>homotopy class</td>
<td>6</td>
</tr>
<tr>
<td>homotopy group</td>
<td>7</td>
</tr>
<tr>
<td>infrared slavery</td>
<td>130</td>
</tr>
<tr>
<td>intermediate boson</td>
<td>128</td>
</tr>
<tr>
<td>intertwiner</td>
<td>34, 38</td>
</tr>
<tr>
<td>invariances of a quantum system</td>
<td>53</td>
</tr>
<tr>
<td>invariant subgroup</td>
<td>5</td>
</tr>
<tr>
<td>isotropy group</td>
<td>29</td>
</tr>
<tr>
<td>Jacobi identity</td>
<td>11</td>
</tr>
<tr>
<td>Jacobi polynomials</td>
<td>61</td>
</tr>
<tr>
<td>kernel</td>
<td>4, 5</td>
</tr>
<tr>
<td>Killing form</td>
<td>19</td>
</tr>
<tr>
<td>Kobayashi-Maskawa matrix</td>
<td>134</td>
</tr>
<tr>
<td>ladder operators</td>
<td>75</td>
</tr>
<tr>
<td>lattice gauge theory</td>
<td>130</td>
</tr>
<tr>
<td>Legendre polynomials and functions</td>
<td>60</td>
</tr>
<tr>
<td>Lie algebra</td>
<td>11</td>
</tr>
<tr>
<td>Lie algebras, dimension 3</td>
<td>29</td>
</tr>
<tr>
<td>Lie group</td>
<td>10</td>
</tr>
<tr>
<td>Littlewood-Richardson rules</td>
<td>88</td>
</tr>
<tr>
<td>loop</td>
<td>6</td>
</tr>
<tr>
<td>magnetic moments</td>
<td>110</td>
</tr>
<tr>
<td>manifold</td>
<td>25</td>
</tr>
<tr>
<td>mass matrix</td>
<td>133</td>
</tr>
<tr>
<td>minimal coupling</td>
<td>122</td>
</tr>
<tr>
<td>multipolar expansion</td>
<td>63</td>
</tr>
<tr>
<td>multipole moments</td>
<td>63</td>
</tr>
<tr>
<td>Nambu–Goldstone bosons</td>
<td>103</td>
</tr>
<tr>
<td>neighborhood</td>
<td>24</td>
</tr>
<tr>
<td>non compact group, representations</td>
<td>46</td>
</tr>
<tr>
<td>normal subgroup</td>
<td>5</td>
</tr>
<tr>
<td>orbit</td>
<td>29</td>
</tr>
<tr>
<td>orthogonality and completeness of $D$ matrices</td>
<td>44</td>
</tr>
<tr>
<td>orthogonality and completeness of characters</td>
<td>45</td>
</tr>
<tr>
<td>partial wave expansion</td>
<td>64</td>
</tr>
<tr>
<td>Peter–Weyl theorem</td>
<td>47</td>
</tr>
<tr>
<td>projective representation</td>
<td>50</td>
</tr>
<tr>
<td>pseudoreal representation</td>
<td>36, 68, 137, 140</td>
</tr>
<tr>
<td>pure gauge</td>
<td>124</td>
</tr>
<tr>
<td>QCD</td>
<td>116</td>
</tr>
<tr>
<td>quaternionic representation</td>
<td>68</td>
</tr>
<tr>
<td>quaternions</td>
<td>23</td>
</tr>
<tr>
<td>quotient group</td>
<td>5</td>
</tr>
<tr>
<td>rank of a Lie algebra</td>
<td>72</td>
</tr>
<tr>
<td>ray</td>
<td>51</td>
</tr>
<tr>
<td>real representation</td>
<td>36, 68, 137, 140</td>
</tr>
<tr>
<td>representation</td>
<td>33</td>
</tr>
<tr>
<td>representation, completely reducible</td>
<td>35</td>
</tr>
<tr>
<td>representation, complex</td>
<td>68</td>
</tr>
<tr>
<td>representation, conjugate</td>
<td>36</td>
</tr>
<tr>
<td>representation, contragredient</td>
<td>36</td>
</tr>
<tr>
<td>representation, equivalent</td>
<td>34</td>
</tr>
<tr>
<td>representation, faithful</td>
<td>33</td>
</tr>
<tr>
<td>representation, indecomposable</td>
<td>35</td>
</tr>
<tr>
<td>representation, irreducible</td>
<td>35</td>
</tr>
</tbody>
</table>
representation, of compact Lie group, 44
representation, of finite groups, 48
representation, of Lie algebra, 42
representation, projective, 50
representation, pseudoreal, 36
representation, quaternionic, 68
representation, real, 36, 68
representation, reducible, 35
representation, unitarisable, 37
representation, unitary, 36
representations of $\text{SU}(2)$ and $\text{SU}(2)$, 58
root, 73
root lattice, 83
root space, 73
root, highest, 80
root, simple, 77

Schur lemma, 38
semi-simple group, 5
semi-simple Lie algebra, 17
Serre relations, 79
simple group, 5
simple Lie algebra, 17
simple Lie group, 17
$\text{SL}(2,\mathbb{R})$, 29
$\text{SO}(2,1)$, 29
$\text{SO}(4)$ Lie algebra, 17
$\text{SO}(n)$ Lie algebra, 16
spherical harmonics, 60, 61, 66
spontaneous breaking, 102, 128
stabilizer, 29
standard model, 129
standard Young tableau, 92
strangeness, 106
structure constants, 18
$\text{SU}(1,1)$, 29
$\text{SU}(3)$ flavor group, 105
$\text{SU}(3)_c$ color group, 116, 129
$\text{SU}(4)$, 115
$\text{SU}(5)$ grand-unification, 135
symplectic group, 2, 24
tangent space, 12, 25
tensor product of representations, 39
tensor products in $\text{SU}(2)$, 67
tensors, 57
tetrahedron group, 49
top quark, 115
topological group, 6
truth, 115
universal covering space, 7

V-A, 114
Vandermonde determinant, 28, 93
Virasoro algebra, 16
weak hypercharge, 131
weak isospin, 131
weight, 81
weight diagram of a representation, 81
weight lattice, 83
weight, dominant, 82
weight, fundamental, 82
weight, highest, 82
Weinberg angle, 131, 136
Weyl group, 76
Weyl vector, 83
Wigner $D$ matrices, 59
Wigner theorem, 51
Wigner–Eckart theorem, 54
Wilson loop, 125, 130, 143
Witt algebra, 16

Yang–Mills, 122
Yang–Mills Lagrangian, 125
Young diagram, 90, 92
Young tableau, 92