M2/International Centre for Fundamental Physics
Parcours of Physique Théorique

Invariances in Physics and Group Theory

Jean-Bernard Zuber

Figure 1: Some of the major actors of group theory mentionned in the first part of these notes.

Foreword

The following notes cover the content of the course "Invariances in Physique and Group Theory" given in the fall 2012. Additional lectures were given during the week of "prérentrée" on the SO(3), SU(2), SL(2,$\mathbb{C}$) groups, see below Chap. 0. More material on Special Relativity, classical field theory and the Dirac equation is contained in the chapters 0 and 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

- [BC] N.N. Bogolioubov et D.V. Chirkov, *Introduction à la théorie quantique des champs*, Dunod.
- [BDm] J.D. Bjorken and S. Drell: *Relativistic Quantum Mechanics*, McGraw Hill.
- [BDf] J.D. Bjorken and S. Drell: *Relativistic Quantum Fields*, McGraw Hill.
- [Bo] N. Bourbaki, *Groupes et Algèbres de Lie*, Chap. 1-9, Hermann 1960-1983.
- [Bu] D. Bump, *Lie groups*, Series "Graduate Texts in Mathematics", vol. **225**, Springer 2004.
- [DFMS] P. Di Francesco, P. Mathieu et D. Sénéchal, *Conformal Field Theory*, Springer,
- [DNF] B. Doubrovine, S. Novikov et A. Fomenko, *Géométrie contemporaine*, 3 volumes, Éditions de Moscou 1982.
- [FH] W. Fulton and J. Harris, *Representation Theory*, Springer.
- [Gi] R. Gilmore, *Lie groups, Lie algebras and some of their applications*, Wiley.
- [Ha] M. Hamermesh, *Group theory and its applications to physical problems*, Addison-Wesley
- [IZ] C. Itzykson et J.-B. Zuber, *Quantum Field Theory*, McGraw Hill 1980; Dover 2006.
- [Ki] A.A. Kirillov, *Elements of the theory of representations*, Springer.
- [LL] L. Landau et E. Lifschitz, *Théorie du Champ*, Editions Mir, Moscou ou *The Classical Theory of Fields*, Pergamon Pr.
- [M] A. Messiah, *Mécanique Quantique*, 2 tomes, Dunod.

- [OR] L. O' Raifeartaigh, *Group structure of gauge theories*, Cambridge Univ. Pr. 1986.
- [PS] M. Peskin and D.V. Schroeder, *An Introduction to Quantum Field Theory*, Addison Wesley.
- [Po] L.S. Pontryagin, *Topological Groups*, Gordon and Breach, 1966.
- [W] H. Weyl, *Classical groups*, Princeton University Press.
- [Wf] S. Weinberg, *The Quantum Theory of Fields*, vol. 1, 2 and 3, Cambridge University Press.
- [Wg] S. Weinberg, *Gravitation and Cosmology*, John Wiley & Sons.
- [Wi] E. Wigner, *Group Theory and its Applications to Quantum Mechanics.* Academ. Pr. 1959.
- [Z-J] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, Oxford Univ. Pr.

Contents

Chapter 0

Some basic elements on the groups SO(3), SU(2) and SL(2,$\mathbb{C}$)

0.1 Rotations of $\mathbb{R}^3$, the groups SO(3) and SU(2)

0.1.1 The group SO(3), a 3-parameter group

Let us consider the rotation group in three-dimensional Euclidean space. These rotations leave invariant the squared norm of any vector $\mathbf{OM}$, $\mathbf{OM}^2 = x_1^2 + x_2^2 + x_3^2 = x^2 + y^2 + z^2$ [1] and are represented in an orthonormal bases by 3×3 orthogonal real matrices, of determinant 1 : they form the "special orthogonal" group SO(3).

Olinde Rodrigues formula

Any rotation of SO(3) is a rotation by some angle ψ around an axis colinear to a unit vector $\mathbf{n}$, and the rotations associated to $(\mathbf{n}, \psi)$ and $(-\mathbf{n}, -\psi)$ are identical. We denote $R_{\mathbf{n}}(\psi)$ this rotation. In a very explicit way, one writes $\mathbf{x} = \mathbf{x}_\parallel + \mathbf{x}_\perp = (\mathbf{x}.\mathbf{n})\mathbf{n} + (\mathbf{x} - (\mathbf{x}.\mathbf{n})\mathbf{n})$ and $\mathbf{x}' = \mathbf{x}_\parallel + \cos\psi\, \mathbf{x}_\perp + \sin\psi\, \mathbf{n} \times \mathbf{x}_\perp$, whence Rodrigues formula

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

As any unit vector $\mathbf{n}$ in $\mathbb{R}^3$ depends on two parameters, for example the angle θ it makes with the Oz axis and the angle ϕ of its projection in the Ox, Oy plane with the Ox axis (see Fig. 1) an element of SO(3) is parametrized by 3 continuous variables. One takes

$$0 \le \theta \le \pi, \quad 0 \le \phi < 2\pi, \quad 0 \le \psi \le \pi\ . \tag{0.2}$$

But there remains an innocent-looking redundancy, $R_{\mathbf{n}}(\pi) = R_{-\mathbf{n}}(\pi)$, the consequences of which we see later ...

[1] In this chapter, we use alternately the notations (x, y, z) or (x_1, x_2, x_3) to denote coordinates in an orthonormal frame.

SO(3) is thus a dimension 3 manifold. For the rotation of axis $\mathbf{n}$ colinear to the Oz axis, we have the matrix

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

whereas around the Ox and Oy axes

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

Conjugation of $R_{\mathbf{n}}(\psi)$ by another rotation
A relation that we are going to use frequently reads

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

where $\mathbf{n}'$ is the transform of $\mathbf{n}$ by rotation R, $\mathbf{n}' = R\mathbf{n}$ (check it!). Conversely any rotation of angle ψ around a vector $\mathbf{n}'$ can be cast under the form (0.5) : we'll say later that the "conjugation classes" of the group SO(3) are characterized by the angle ψ.

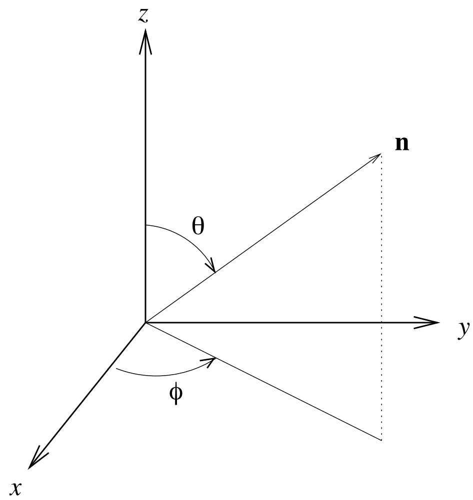


Fig. 1

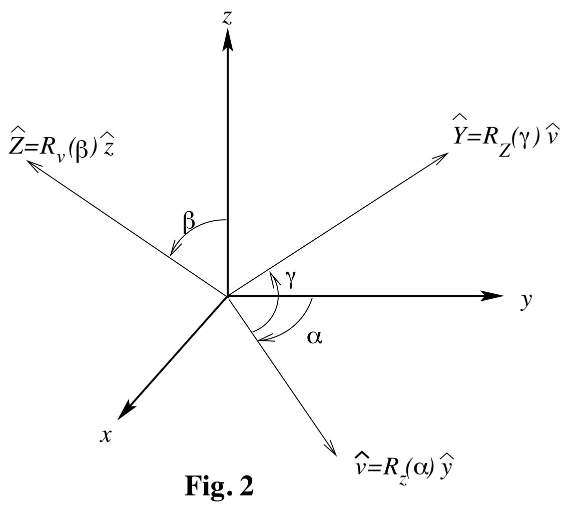


Fig. 2

Euler angles
Another description makes use of *Euler angles* : given an orthonormal frame (Ox, Oy, Oz), any rotation around O that maps it onto another frame (OX, OY, OZ) may be regarded as resulting from the composition of a rotation of angle α around Oz, which brings the frame onto (Ou, Ov, Oz), followed by a rotation of angle β around Ov bringing it on (Ou', Ov, OZ), and lastly, by a rotation of angle γ around OZ bringing the frame onto (OX, OY, OZ), (see Fig. 2). One thus takes $0 \le \alpha < 2\pi,\ 0 \le \beta \le \pi,\ 0 \le \gamma < 2\pi$ and one writes

$$R(\alpha, \beta, \gamma) = R_Z(\gamma)R_v(\beta)R_z(\alpha) \tag{0.6}$$

but according to (0.5)

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

thus, by inserting into (0.6)

$$R(\alpha,\beta,\gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma)\ . \tag{0.7}$$

where one used the fact that $R_z(\alpha)R_z(\gamma)R_z^{-1}(\alpha) = R_z(\gamma)$ since rotations around a given axis commute (they form an abelian subgroup, isomorphic to $SO(2)$).

Exercise : using (0.5), write the expression of a matrix R which maps the unit vector $\mathbf{z}$ colinear to Oz to the unit vector $\mathbf{n}$, in terms of $R_z(\phi)$ and $R_y(\theta)$; then write the expression of $R_{\mathbf{n}}(\psi)$ in terms of R_y and R_z. Write the explicit expression of that matrix and of (0.7) and deduce the relations between θ,ϕ,ψ and Euler angles. (See also below, equ. (0.66).)

0.1.2 From SO(3) to SU(2)

Consider another parametrization of rotations. To the rotation $R_{\mathbf{n}}(\psi)$, we associate the unitary 4-vector u : $(u_0 = \cos\frac{\psi}{2}, \mathbf{u} = \mathbf{n}\sin\frac{\psi}{2})$; we have $u^2 = u_0^2 + \mathbf{u}^2 = 1$, and u belongs to the unit sphere S^3 in the space $\mathbb{R}^4$. Changing the determination of ψ by an odd multiple of 2π changes u into $-u$. There is thus a bijection between $R_{\mathbf{n}}(\psi)$ and the pair $(u,-u)$, *i.e.* between SO(3) and $S^3/\mathbb{Z}_2$, the sphere in which diametrically opposed points are identified. We shall say that the sphere S^3 is a "covering group" of SO(3). In which sense is this sphere a group? To answer that question, introduce Pauli matrices σ_i, $i = 1,2,3$.

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\ . \tag{0.8}$$

Together with the identity matrix $\mathbf{I}$, they form a basis of the vector space of 2×2 Hermitian matrices. They satisfy the identity

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

with ϵ_{ijk} the completely antisymmetric tensor, $\epsilon_{123} = +1$, $\epsilon_{ijk} =$ (the signature of permutation (ijk)).

From u a real unit 4-vector unitary (*i.e.* a point of S^3), we form the matrix

$$U = u_0\mathbf{I} - i\mathbf{u}.\boldsymbol{\sigma} \tag{0.10}$$

which is unitary and of determinant 1 (check it and also show the converse: any unimodular (= of determinant 1) unitary 2×2 matrix is of the form (0.10), with $u^2 = 1$). These matrices form the *special unitary* group SU(2) which is thus isomorphic to S^3. By a power expansion of the exponential and making use of (0.9), one may verify that

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

It is then suggested that the multiplication of matrices

$$U_n(\psi) = e^{-i\frac{\psi}{2}\mathbf{n}.\boldsymbol{\sigma}} = \cos\frac{\psi}{2} - i\sin\frac{\psi}{2}\mathbf{n}.\boldsymbol{\sigma}, \qquad 0\le\psi\le 2\pi, \quad \mathbf{n}\in S^2 \tag{0.12}$$

gives the desired group law in S^3. Let us show indeed that to a matrix of SU(2) one may associate a rotation of SO(3) and that to the product of two matrices of SU(2) corresponds the product of the SO(3) rotations (this is the *homomorphism* property). To the point x of $\mathbb{R}^3$ of coordinates x_1, x_2, x_3, we associate the Hermitian matrix

$$X = \mathbf{x}.\boldsymbol{\sigma} = \begin{pmatrix} x_3 & x_1 - ix_2 \\ x_1 + ix_2 & -x_3 \end{pmatrix} , \tag{0.13}$$

with conversely $x_i = \frac{1}{2}\mathrm{tr}\,(X\sigma_i)$, and let SU(2) act on that matrix according to

$$X \mapsto X' = UXU^\dagger , \tag{0.14}$$

which defines a linear transform $x \mapsto x' = \mathcal{T}x$. One readily computes that

$$\det X = -(x_1^2 + x_2^2 + x_3^2) \tag{0.15}$$

and as $\det X = \det X'$, the linear transform $x \mapsto x' = \mathcal{T}x$ is an isometry, hence $\det \mathcal{T} = 1$ or -1. To convince oneself that this is indeed a rotation, *i.e.* that the transformation has a determinant 1, it suffices to compute that determinant for $U = \mathbf{I}$ where $\mathcal{T}=$ the identity, hence $\det \mathcal{T} = 1$, and then to invoke the connexity of the manifold $SU(2)(\cong S^3)$ to conclude that the continuous function $\det \mathcal{T}(U)$ cannot jump to the value -1. In fact, using identity (0.9), the explicit calculation of X' leads, after some algebra, to

$$\begin{aligned} X' &= (\cos\frac{\psi}{2} - i\mathbf{n}.\boldsymbol{\sigma}\sin\frac{\psi}{2})X(\cos\frac{\psi}{2} + i\mathbf{n}.\boldsymbol{\sigma}\sin\frac{\psi}{2}) \\ &= \big(\cos\psi\,\mathbf{x} + (1-\cos\psi)(\mathbf{x}.\mathbf{n})\,\mathbf{n} + \sin\psi\,(\mathbf{n}\times\mathbf{x})\big).\boldsymbol{\sigma} \end{aligned} \tag{0.16}$$

which is nothing else than the Rodrigues formula (0.1). We thus conclude that the transformation $x \mapsto x'$ performed by the matrices of SU(2) in (0.14) is indeed the rotation of angle ψ around $\mathbf{n}$. To the product $U_{n'}(\psi')U_n(\psi)$ in SU(2) corresponds in SO(3) the composition of the two rotations $R_{\mathbf{n}'}(\psi')R_{\mathbf{n}}(\psi)$ of SO(3). There is thus a "homomorphism" of the group SU(2) into SO(3). This homomorphism maps the two matrices U and $-U$ onto one and the same rotation of SO(3).

Let us summarize what we have learnt in this section. The group SU(2) is a covering group (of order 2) of the group SO(3) (the precise topological meaning of which will be given in Chap. 1), and the 2-to-1 homomorphism from SU(2) to SO(3) is given by equations (0.12)-(0.14).

0.2 Infinitesimal generators. The $su(2)$ Lie algebra

0.2.1 Infinitesimal generators of SO(3)

Rotations $R_{\mathbf{n}}(\psi)$ around a given axis $\mathbf{n}$ form a *one-parameter subgroup*, isomorphic to $SO(2)$. In this chapter, we follow the common use (among physicists) and write the infinitesimal generators of rotations as *Hermitian* operators $J = J^\dagger$. Thus

$$R_{\mathbf{n}}(\mathrm{d}\psi) = (I - i\mathrm{d}\psi J_{\mathbf{n}}) \tag{0.17}$$

where $J_{\mathbf{n}}$ is the "generator" of these rotations, a Hermitian 3×3 matrix. Let us show that we may reconstruct the finite rotations from these infinitesimal generators. By the group property,

$$R_{\mathbf{n}}(\psi + \mathrm{d}\psi) = R_{\mathbf{n}}(\mathrm{d}\psi) R_{\mathbf{n}}(\psi) = (I - i\mathrm{d}\psi J_{\mathbf{n}}) R_{\mathbf{n}}(\psi) \;, \tag{0.18}$$

or equivalently

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

which, on account of $R(0) = I$, may be integrated into

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

To be more explicit, introduce the three basic J_1, J_2 and J_3 describing the infinitesimal rotations around the corresponding axes[2]. From the infinitesimal version of (0.3) it follows that

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \qquad J_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \qquad J_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{0.21}$$

which may be expressed by a unique formula

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

with the completely antisymmetric tensor ϵ_{ijk}.

We now show that matrices (0.21) form a basis of infinitesimal generators and that $J_{\mathbf{n}}$ is simply expressed as

$$J_{\mathbf{n}} = \sum_k J_k n_k \tag{0.23}$$

which allows us to rewrite (0.20) in the form

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

The expression (0.23) follows simply from the infinitesimal form of Rodrigues formula, $R_{\mathbf{n}}(\mathrm{d}\psi) = (\mathbf{I} + \mathrm{d}\psi \mathbf{n}\times)$ hence $-iJ_{\mathbf{n}} = \mathbf{n}\times$ or alternatively $-i(J_{\mathbf{n}})_{ij} = \epsilon_{ikj} n_k = n_k(-iJ_k)_{ij}$, q.e.d. (Here and in the following, we make use of the convention of summation over repeated indices: $\epsilon_{ikj} n_k \equiv \sum_k \epsilon_{ikj} n_k$, etc.)

A comment about (0.24): it is obviously wrong to write in general $R_{\mathbf{n}}(\psi) = e^{-i\psi \sum_k n_k J_k} \stackrel{?}{=} \prod_{k=1}^{3} e^{-i\psi n_k J_k}$ because of the non commutativity of the J's. On the other hand, formula (0.7) shows that any rotation of SO(3) may be written under the form

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

The three matrices J_i, $i = 1, 2, 3$ satisfy the very important commutation relations

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

[2] Do not confuse $J_{\mathbf{n}}$ labelled the unit vector $\mathbf{n}$ with J_k, k-th component of $\mathbf{J}$. The relation between the two will be explained shortly.

which follow from the identity verified by the tensor ϵ

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

Exercise: note the structure of this identity (i is fixed, b summed over, cyclic permutation over the three others) and check that it implies (0.26).

In view of the importance of relations (0.23–0.26), it may be useful to recover them by another route. Note first that equation (0.5) implies that for any R

$$Re^{-i\psi J_{\mathbf{n}}}R^{-1} = e^{-i\psi RJ_{\mathbf{n}}R^{-1}} = e^{-i\psi J_{\mathbf{n}'}} \tag{0.28}$$

with $\mathbf{n}' = R\mathbf{n}$, whence

$$RJ_{\mathbf{n}}R^{-1} = J_{\mathbf{n}'} \ , \tag{0.29}$$

i.e. $J_{\mathbf{n}}$ transform like the vector $\mathbf{n}$.

The tensor ϵ_{ijk} is invariant under rotations

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

since the matrix $\mathcal{R}$ is of determinant 1. That matrix being also orthogonal, one may push one $\mathcal{R}$ to the right-hand side

$$\epsilon_{lmn}\mathcal{R}_{jm}\mathcal{R}_{kn} = \epsilon_{ijk}\mathcal{R}_{il} \tag{0.31}$$

which thanks to (0.22) expresses that

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

i.e. for any R and its matrix $\mathcal{R}$,

$$RJ_lR^{-1} = J_i\mathcal{R}_{il} \ . \tag{0.33}$$

Let R be a rotation which maps the unit vector $\mathbf{z}$ colinear to Oz on the vector $\mathbf{n}$, thus $n_k = \mathcal{R}_{k3}$ and

$$J_{\mathbf{n}} \overset{(0.29)}{=} RJ_3R^{-1} \overset{(0.33)}{=} J_k\mathcal{R}_{k3} = J_kn_k \ , \tag{0.34}$$

which is just (0.23). Note that equations (0.33) and (0.34) are compatible with (0.29)

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

As we shall see later in a more systematic way, the commutation relation (0.26) of infinitesimal generators J encodes an infinitesimal version of the group law. Consider for example a rotation of infinitesimal angle $d\psi$ around Oy acting on J_1

$$R_2(d\psi)J_1R_2^{-1}(d\psi) \overset{(0.33)}{=} J_k[\mathcal{R}_2(d\psi)]_{k1} \tag{0.35}$$

but to first order, $R_2(d\psi) = \mathbf{I} - id\psi J_2$, and thus the left hand side of (0.35) equals $J_1 - id\psi[J_2, J_1]$ while in the right hand side, $[\mathcal{R}_2(d\psi)]_{k1} = \delta_{k1} - id\psi(J_2)_{k1} = \delta_{k1} - d\psi\delta_{k3}$ by (0.22), whence $i[J_1, J_2] = -J_3$, which is one of the relations (0.26).

0.2.2 Infinitesimal generators in SU(2)

Let us examine now things from the point of view of SU(2). Any unitary matrix U (here 2×2) may be diagonalized by a unitary change of basis $U = V\exp\{i\,\mathrm{diag}\,(\lambda_k)\}V^\dagger$, V unitary, and hence written as

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

with H Hermitian, $H = V\,\mathrm{diag}\,(\lambda_k)\,V^\dagger$. The sum converges (for the norm $||M||^2 = \mathrm{tr}\,MM^\dagger$). The unimodularity condition $1 = \det U = \exp i\mathrm{tr}\,H$ is ensured if $\mathrm{tr}\,H = 0$. The set of such Hermitian traceless matrices forms a vector space $\mathcal{V}$ of dimension 3 over $\mathbb{R}$, with a basis given by the three Pauli matrices

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

which may be inserted back into (0.36). (In fact we already observed that any unitary 2×2 matrix may be written in the form (0.11)). Comparing that form with (0.24), or else comparing its infinitesimal version $U_n(\mathrm{d}\psi) = (I - i\,\mathrm{d}\psi\mathbf{n}.\frac{\boldsymbol{\sigma}}{2})$ with (0.17), we see that matrices $\frac{1}{2}\sigma_j$ play in SU(2) the role played by infinitesimal generators J_j in SO(3). But these matrices $\frac{1}{2}\sigma_.$ verify the same commutation relations

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

with the same *structure constants* ϵ_{ijk} as in (0.26). In other words, we have just discovered that infinitesimal generators J_i (eq. (0.21) of SO(3) and $\frac{1}{2}\sigma_i$ of SU(2) satisfy the same commutation relations (we shall say later that they are the bases of two different representations of the same Lie algebra $su(2) = so(3)$). This has the consequence that calculations carried out with the $\frac{1}{2}\vec{\sigma}$ and *making only use of commutation relations* are also valid with the $\vec{J}$, and vice versa. For instance, from (0.33), for example $R_2(\beta)J_kR_2^{-1}(\beta) = J_l\mathcal{R}_y(\beta)_{lk}$, it follows immediately, with no further calculation, that for Pauli matrices, we have

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

where the matrix elements $\mathcal{R}_y$ are read off (0.3'). Indeed there is a general identity stating that $e^ABe^{-A} = B + \sum_{n=1}^{\infty}\frac{1}{n!}\underbrace{[A[A,[\cdots,[A,B]\cdots]]]}_{n\ \text{commutators}}$, see Chap. 1, eq. (1.29), and that computation thus involves only commutators. On the other hand, the relation

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

(which does *not* involve only commutators) is specific to the dimension 2 representation of the su(2) algebra.

0.2.3 Lie algebra $su(2)$

Let us recapitulate: we have just introduced the commutation algebra (or Lie algebra) of infinitesimal generators of the group SU(2) (or SO(3)), denoted $su(2)$ or $so(3)$. It is defined by relations (0.26), that we write once again

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

We shall also make frequent use of the three combinations

$$J_z \equiv J_3, \quad J_+ = J_1 + iJ_2, \quad J_- = J_1 - iJ_2\ . \tag{0.40}$$

It is then immediate to compute

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

One also verifies that the *Casimir operator* defined as

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

commutes with all the J's

$$[\mathbf{J}^2, J_.] = 0 \;, \tag{0.43}$$

which means that it is invariant under rotations.

Anticipating a little on the following, we shall be mostly interested in "unitary representations", where the generators J_i, $i = 1,2,3$ are Hermitian, hence

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

Let us finally mention an interpretation of the J_i as differential operators acting on differentiable functions of coordinates in the space $\mathbb{R}^3$. In that space $\mathbb{R}^3$, an infinitesimal rotation acting on the vector $\mathbf{x}$ changes it into

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

hence a *scalar* function of $\mathbf{x}$, $f(\mathbf{x})$, is changed into $f'(\mathbf{x}') = f(\mathbf{x})$ or

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

We thus identify

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

which allows us to compute it in arbitrary coordinates, for example spherical, see Appendix 0. (Compare also (0.46) with the expression of (orbital) angular momentum in Quantum Mechanics $L_i = \frac{\hbar}{i}\epsilon_{ijk} x_j \frac{\partial}{\partial x_k}$). Exercise : check that these differential operators do satisfy the commutation relations (0.26).

Among the combinations of J that one may construct, there is one that must play a particular role, namely the Laplacian on the sphere S^2, a second order differential operator which is invariant under changes of coordinates (see Appendix 0). It is in particular rotation invariant, of degree 2 in the $J_.$, this may only be the Casimir operator $\mathbf{J}^2$ (up to a factor). In fact the Laplacian in $\mathbb{R}^3$ reads in spherical coordinates

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

For the sake of simplicity we have restricted this discussion to scalar functions, but one might more generally consider the transformation of a collection of functions "forming a representation" of SO(3), *i.e.* transforming linearly among themselves under the action of that group

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

or else

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

for example a vector field transforming as

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

What are now the infinitesimal generators for such objects ?

0.3 Representations of SU(2)

0.3.1 Representations of the groups SO(3) and SU(2)

We are familiar with the notions of vectors or tensors in the geometry of the space $\mathbb{R}^3$. They are objects that transform linearly under rotations

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

More generally we call *representation* of a group G in a vector space E a homomorphism of G into the group $\mathrm{GL}(E)$ of linear transformations of E (see Chap. 2). Thus, as we just saw, the group SO(3) admits a representation in the space $\mathbb{R}^3$ (the vectors V of the above example), another representation in the space of rank 2 tensors, etc. We now want to build the general representations of SO(3) and SU(2). For the needs of physics, in particular of quantum mechanics, we are mostly interested in unitary representations, in which the representation matrices are unitary. In fact, as we'll see, it is enough to study the representations of SU(2) to also get those of SO(3), and even better, it is enough to study the way the group elements close to the identity are represented, *i.e.* to find the representations of the infinitesimal generators of SU(2) (and SO(3)).

To summarize : to find all the unitary representations of the group SU(2), it is thus sufficient to find the representations by Hermitian matrices of its Lie algebra $su(2)$, that is, Hermitian operators satisfying the commutation relations (0.26).

0.3.2 Representations of the algebra $su(2)$

We now proceed to the classical construction of representations of the algebra $su(2)$. As above, $J_\pm$ and J_z denote the representatives of infinitesimal generators in a certain representation. They thus satisfy the commutation relations (0.41) and hermiticity (0.44). Commutation of operators J_z and $\mathbf{J}^2$ ensures that one may find common eigenvectors. The eigenvalues of these Hermitian operators are real, and moreover, $\mathbf{J}^2$ being semi-definite positive, one may always write its eigenvalues in the form $j(j+1)$, j real non negative (*i.e.* $j \geq 0$), and one thus considers a common eigenvector $|j\,m\,\rangle$

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

with m a real number, a priori arbitrary at this stage. By a small abuse of language, we call $|jm\rangle$ an "eigenvector of eigenvalues (j,m)".
(i) Act with J_+ and $J_- = J_+^\dagger$ on $|j\,m\rangle$. Using the relation $J_\pm J_\mp = \mathbf{J}^2 - J_z^2 \pm J_z$ (a consequence of (0.41)), the squared norm of $J_\pm|j\,m\rangle$ is computed to be:

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

These squared norms cannot be negative and thus

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

which implies

$$-j\leq m\leq j\ . \tag{0.51}$$

Moreover $J_+|j\,m\rangle = 0$ iff $m=j$ and $J_-|j\,m\rangle=0$ iff $m=-j$

$$J_+|j\,j\rangle = 0 \qquad\qquad J_-|j\ -j\rangle = 0\ . \tag{0.52}$$

(ii) If $m\neq j$, $J_+|j\,m\rangle$ is non vanishing, hence is an eigenvector of eigenvalues $(j,m+1)$. Indeed

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

Likewise if $m\neq -j$, $J_-|j\,m\rangle$ is a (non vanishing) eigenvector of eigenvalues $(j,m-1)$.
(iii) Consider now the sequence of vectors

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

If non vanishing, they are eigenvectors of J_z of eigenvalues $m, m-1, m-2,\cdots,m-p\cdots$. As the allowed eigenvalues of J_z are bound by (0.51), this sequence must stop after a finite number of steps. Let p be the integer such that $J_-^p|j\,m\rangle\neq 0$, $J_-^{p+1}|j\,m\rangle=0$. By (0.52), $J_-^p|j\,m\rangle$ is an eigenvector of eigenvalues $(j,-j)$ hence $m-p=-j$, *i.e.*

$$(j+m)\ \text{is a non negative integer}\,. \tag{0.54}$$

Acting likewise with $J_+, J_+^2,\cdots$ sur $|j\,m\rangle$, we are led to the conclusion that

$$(j-m)\ \text{is a non negative integer}\,. \tag{0.55}$$

and thus j *and* m *are simultaneously integers or half-integers.* For each value of j

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

m may take the $2j+1$ values [3]

$$m = -j, -j+1, \cdots, j-1, j \ . \tag{0.56}$$

Starting from the vector $|j\, m = j\,\rangle$, ("highest weight vector"), now chosen of norm 1, we construct the orthonormal basis $|j\, m\,\rangle$ by iterated application of J_- and we have

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

These $2j+1$ vectors form a basis of the "spin j representation" of the $su(2)$ algebra.

In fact this representation of the *algebra* $su(2)$ extends to a representation of the *group* SU(2), as we now show.

Remark. The previous discussion has given a central role to the unitarity of the representation and hence to the hermiticity of infinitesimal generators, hence to positivity: $||J_\pm|j\,m\,\rangle||^2 \geq 0 \Longrightarrow -j \leq m \leq j$, etc, which allowed us to conclude that the representation is necessarily of finite dimension. Conversely one may insist on the latter condition, and show that it suffices to ensure the previous conditions on j and m. Starting from an eigenvector $|\psi\,\rangle$ of J_z, the sequence $J_+^p|\psi\,\rangle$ yields eigenvectors of J_z of increasing eigenvalue, hence linearly independent, as long as they do not vanish. If by hypothesis the representation is of finite dimension, this sequence is finite, and there exists a vector denoted $|j\,\rangle$ such that $J_+|j\,\rangle = 0$, $J_z|j\,\rangle = j|j\,\rangle$. By the relation $\mathbf{J}^2 = J_-J_+ + J_z(J_z+1)$, it is also an eigenvector of eigenvalue $j(j+1)$ of $\mathbf{J}^2$. It thus identifies with the highest weight vector denoted previously $|j\,j\,\rangle$, a notation that we thus adopt in the rest of this discussion. Starting from this vector, the $J_-^p|j\,j\,\rangle$ form a sequence that must also be finite

$$\exists q \quad J_-^{q-1}|j\,j\,\rangle \neq 0 \qquad J_-^q|j\,j\,\rangle = 0 \ . \tag{0.58}$$

One easily shows by induction that

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

hence $q = 2j+1$. The number j is thus integer or half-integer, the vectors of the representation built in that way are eigenvectors of $\mathbf{J}^2$ of eigenvalue $j(j+1)$ and of J_z of eigenvalue m satisfying (0.56). We have recovered all the previous results. In this form, the construction of these "highest weight representations" generalizes to other Lie algebras, (even of infinite dimension, such as the Virasoro algebra, see Chap. 1, § 1.3.6).

The matrices $\mathcal{D}^j$ of the spin j representation are such that under the action of the rotation $U \in SU(2)$

$$|j\,m\,\rangle \mapsto D^j(U)|j\,m\,\rangle = |j\,m'\,\rangle \mathcal{D}^j_{m'm}(U) \ . \tag{0.60}$$

Depending on the parametrization $((\mathbf{n},\psi)$, angles d'Euler, ...), we write $\mathcal{D}^j_{m'm}(\mathbf{n},\psi)$, $\mathcal{D}^j_{m'm}(\alpha,\beta,\gamma)$, etc. By (0.7), we thus have

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

[3]In fact, we have just found a necessary condition on the j, m. That all these j give indeed rise to representations will be verified in the next subsection.

where the Wigner matrix d^j is defined by

$$d^j_{m'm}(\beta) = \langle\, j\, m'|e^{-i\beta J_y}|j\, m\,\rangle\ . \tag{0.62}$$

An explicit formula for d^j will be given in the next subsection. We also have

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

Exercise : Compute $\mathcal{D}^j(\mathbf{x},\psi)$. (Hint : use (0.5).)

One notices that $\mathcal{D}^j(\mathbf{z},2\pi) = (-1)^{2j}I$, since $(-1)^{2m} = (-1)^{2j}$ using (0.55), and this holds true for any axis $\mathbf{n}$ by the conjugation (0.5)

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

This shows that a 2π rotation in SO(3) is represented by $-I$ in a half-integer-spin representation of SU(2). Half-integer-spin representations of SU(2) are said to be "projective", (*i.e.* here, up to a sign), representations of SO(3); we return in Chap. 2 to this notion of projective representation.

We also verify the *unimodularity* of matrices $\mathcal{D}^j$ (or equivalently, the fact that representatives of infinitesimal generators are traceless). If $\mathbf{n} = R\mathbf{z}$, $\mathcal{D}(\mathbf{n},\psi) = \mathcal{D}(R)\mathcal{D}(\mathbf{z},\psi)\mathcal{D}^{-1}(R)$, hence

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

It may be useful to write explicitly these matrices in the cases $j=\frac{1}{2}$ and $j=1$. The case of $j=\frac{1}{2}$ is very simple, since

$$\begin{aligned}\mathcal{D}^{\frac12}(U) &= U = e^{-i\frac12\psi\mathbf{n}.\boldsymbol{\sigma}} = \begin{pmatrix}\cos\frac{\psi}{2} - i\cos\theta\sin\frac{\psi}{2} & -i\sin\frac{\psi}{2}\sin\theta\, e^{-i\phi}\\ -i\sin\frac{\psi}{2}\sin\theta\, e^{i\phi} & \cos\frac{\psi}{2} + i\cos\theta\sin\frac{\psi}{2}\end{pmatrix}\\ &= e^{-i\frac{\alpha}{2}\sigma_3}e^{-i\frac{\beta}{2}\sigma_2}e^{-i\frac{\gamma}{2}\sigma_3} = \begin{pmatrix}\cos\frac{\beta}{2}e^{-\frac{i}{2}(\alpha+\gamma)} & -\sin\frac{\beta}{2}e^{-\frac{i}{2}(\alpha-\gamma)}\\ \sin\frac{\beta}{2}e^{\frac{i}{2}(\alpha-\gamma)} & \cos\frac{\beta}{2}e^{\frac{i}{2}(\alpha+\gamma)}\end{pmatrix}\end{aligned} \tag{0.66}$$

an expected result since the matrices U of the group form obviously a representation. (As a by-product, we have derived relations between the two parametrizations, $(\mathbf{n},\psi) = (\theta,\phi,\psi)$ and Euler angles (α,β,γ).) For $j=1$, in the basis $|1,1\,\rangle$, $|1,0\,\rangle$ and $|1,-1\,\rangle$ where J_z is diagonal (which *is not* the basis (0.21) !)

$$J_z = \begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}\qquad J_+ = \sqrt{2}\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}\qquad J_- = \sqrt{2}\begin{pmatrix}0&0&0\\1&0&0\\0&1&0\end{pmatrix} \tag{0.67}$$

whence

$$d^1(\beta) = e^{-i\beta J_y} = \begin{pmatrix}\frac{1+\cos\beta}{2} & -\frac{\sin\beta}{\sqrt2} & \frac{1-\cos\beta}{2}\\ \frac{\sin\beta}{\sqrt2} & \cos\beta & -\frac{\sin\beta}{\sqrt2}\\ \frac{1-\cos\beta}{2} & \frac{\sin\beta}{\sqrt2} & \frac{1+\cos\beta}{2}\end{pmatrix} \tag{0.68}$$

as the reader may check.

In the following subsection, we write more explicitly these representation matrices of the group SU(2), and in Appendix B of Chap. 2, give more details on the differential equations they satisfy and on their relations with "special functions", orthogonal polynomials and spherical harmonics...

Irreductibility

A central notion in the study of representations is that of irreducibility. A representation is irreducible if it has no invariant subspace. Let us show that the spin j representation of SU(2) that we have just built is irreducible. We show below in Chap. 2 that, as the representation is unitary, it is either irreducible or "completely reducible" (there exists an invariant subspace and its supplementary space is also invariant) ; in the latter case, there would exist block-diagonal operators, different from the identity and commuting the matrices of the representation, in particulier with the generators J_i. But in the basis (0.5) any matrix M that commutes with J_z is diagonal, $M_{mm'} = \mu_m \delta_{mm'}$, (check it !), and commutation with J_+ forces all μ_m to be equal: the matrix M is a multiple of the identity and the representation is indeed irreducible.

One may also wonder why the study of finite dimensional representations that we just carried out suffices to the physicist's needs, for instance in quantum mechanics, where the scene usually takes place in an infinite dimensional Hilbert space. We show below (Chap. 2) that *Any representation of SU(2) or SO(3) in a Hilbert space is equivalent to a unitary representation, and is completely reducible into a (finite or infinite) sum of finite dimensional irreducible representations.*

0.3.3 Explicit construction

Let ξ and η be two complex variables on which matrices $U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ of SU(2) act according to $\xi' = a\xi + c\eta$, $\eta' = b\xi + d\eta$. In other terms, ξ and η are the basis vectors of the representation of dimension 2 (representation of spin $\frac{1}{2}$) of SU(2). An explicit construction of the previous representations is then obtained by considering homogenous polynomials of degree $2j$ in the two variables ξ and η, a basis of which is given by the $2j+1$ polynomials

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

(In fact, the following considerations also apply if U is an arbitrary matrix of the group GL(2,$\mathbb{C}$) and provide a representation of that group.) Under the action of U on ξ and η, the $P_{jm}(\xi,\eta)$ transform into $P_{jm}(\xi',\eta')$, also homogenous of degree $2j$ in ξ and η, which may thus be expanded on the $P_{jm}(\xi,\eta)$. The latter thus span a dimension $2j+1$ representation of SU(2) (or GL(2,$\mathbb{C}$)), which is nothing else than the previous spin j representation. This enables us to write quite explicit formulae for the $\mathcal{D}^j$

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

We find explicitly

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

For $U = -\mathbf{I}$, one may check once again that $\mathcal{D}^j(-\mathbf{I}) = (-1)^{2j}\mathbf{I}$. In the particular case of $U = e^{-i\psi\frac{\sigma_2}{2}} = \cos\frac{\psi}{2}\mathbf{I} - i\sin\frac{\psi}{2}\sigma_2$, we thus have

$$d^j_{m'm}(\psi) = \big((j+m)!(j-m)!(j+m')!(j-m')!\big)^{\frac{1}{2}} \sum_{k\geq 0} \frac{(-1)^{k+j-m}\cos\frac{\psi}{2}^{2k+m+m'}\sin\frac{\psi}{2}^{2j-2k-m-m'}}{(m+m'+k)!(j-m-k)!(j-m'-k)!k!} \ . \tag{0.72}$$

The expression of the infinitesimal generators acting on polynomials P_{jm} is obtained by considering U close to the identity. One finds

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

on which it is easy to check commutation relations as well as the action on the P_{jm} in accordance with (0.57). This completes the identification of (0.69) with the spin j representation.

Remarks

1. Repeat the proof of irreducibility of the spin j representation in that new form.
2. Notice that the space of the homogenous polynomials of degree $2j$ in the variables ξ and η is nothing else than the symmetrized $2j$-th tensor power of the representation of dimension 2.

0.4 Direct product of representations of SU(2)

0.4.1 Direct product of representations and the "addition of angular momenta"

Consider the direct (or tensor) product of two representations of spin j_1 and j_2 and to their decomposition on vectors of given total spin ("decomposition into irreducible representations"). We start with the product representation spanned by the vectors

$$|j_1\, m_1\,\rangle \otimes |j_2\, m_2\,\rangle \equiv |j_1\, m_1; j_2\, m_2\,\rangle \quad \text{written in short as} \quad |m_1\, m_2\,\rangle \tag{0.74}$$

on which the infinitesimal generators act as

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

The upper index indicates on which space the operators act. By an abuse of notation, one frequently writes, instead of (0.75)

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

and (in Quantum Mechanics) one talks about the "addition of angular momenta" $\mathbf{J}^{(1)}$ and $\mathbf{J}^{(2)}$. The problem is thus to decompose the vectors (0.74) onto a basis of eigenvectors of $\mathbf{J}$ and

J_z. As $\mathbf{J}^{(1)2}$ and $\mathbf{J}^{(2)2}$ commute with one another and with $\mathbf{J}^2$ and J_z, one may seek common eigenvectors that we denote

$$|(j_1\, j_2)\, J\, M\,\rangle \qquad \text{or more simply} \qquad |J\, M\,\rangle \tag{0.76}$$

where it is understood that the value of j_1 and j_2 is fixed. The question is thus twofold: which values can J and M take; and what is the matrix of the change of basis $|m_1\, m_2\,\rangle \to |J\, M\,\rangle$? In other words, what is the (Clebsch-Gordan) decomposition and what are the Clebsch-Gordan coefficients?

The possible values of M, eigenvalue of $J_z = J_z^{(1)} + J_z^{(2)}$, are readily found

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

and the only value of M such that $\langle\, m_1\, m_2 | J\, M\,\rangle \neq 0$ is thus

$$M = m_1 + m_2\ . \tag{0.78}$$

For j_1, j_2 and M fixed, there are as many independent vectors with that eigenvalue of M as there are couples (m_1, m_2) satisfying (0.78), thus

$$n(M) = \begin{cases} 0 & \text{if } |M| > j_1 + j_2 \\ j_1 + j_2 + 1 - |M| & \text{if } |j_1 - j_2| \le |M| \le j_1 + j_2 \\ 2\mathrm{inf}(j_1, j_2) + 1 & \text{if } 0 \le |M| \le |j_1 - j_2| \end{cases} \tag{0.79}$$

(see the left Fig. 3 in which $j_1 = 5/2$ and $j_2 = 1$). Let N_J be the number of times the representation of spin J appears in the decomposition of the representations of spin j_1 et j_2. The $n(M)$ vectors of eigenvalue M for J_z may also be regarded as coming from the N_J vectors $|J\, M\,\rangle$ for the different values of J compatible with that value of M

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

hence, by subtracting two such relations

$$\begin{aligned} N_J &= n(J) - n(J+1) \\ &= 1 \qquad \text{iff si} \quad |j_1 - j_2| \le J \le j_1 + j_2 \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{0.81}$$

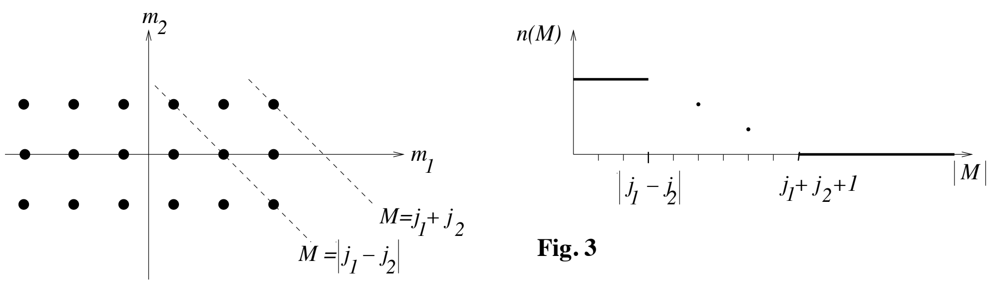


Fig. 3

To summarize, we have just shown that the $(2j_1+1)(2j_2+1)$ vectors (0.74) (with j_1 and j_2 fixed) may be reexpressed in terms of vectors $|J\,M\,\rangle$ with

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

Note that multiplicities N_J take the value 0 or 1 ; it is a pecularity of SU(2) that multiplicities larger than 1 do not occur in the decomposition of the tensor product of irreducible representations, *i.e.* here of fixed spin.

0.4.2 Clebsch-Gordan coefficients, 3-j and 6-j symbols...

The change of orthonormal basis $|j_1\,m_1; j_2\,m_2\,\rangle \to |(j_1\,j_2)\,J\,M\,\rangle$ is carried out by the Clebsch-Gordan coefficients (C.G.) $\langle\,(j_1\,j_2); J\,M|j_1\,m_1; j_2\,m_2\,\rangle$ which form a unitary matrix

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

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

Note that in the first line, M is fixed in terms of m_1 and m_2; and that in the second one, m_2 is fixed in terms of m_1, for given M. Each relation thus implies only one summation. The value of these C.G. depends in fact on a choice of a relative phase between vectors (0.74) and (0.76); the usual convention is that for each value of J, one chooses

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

The other vectors are then unambiguously defined by (0.57) and we shall now show that all C.G. are real. C.G. satisfy recursion relations that are consequences of (0.57). Applying indeed $J_\pm$ to the two sides of (0.83), one gets

$$\begin{aligned} \sqrt{J(J+1)-M(M\pm1)} \qquad & \langle\,(j_1\,j_2)\,J\,M|j_1\,m_1; j_2\,m_2\,\rangle \\ = \ & \sqrt{j_1(j_1+1)-m_1(m_1\pm1)}\langle\,(j_1\,j_2)\,J\,M\pm1|j_1\,m_1\pm1; j_2\,m_2\,\rangle \\ + \ & \sqrt{j_2(j_2+1)-m_2(m_2\pm1)}\langle\,(j_1\,j_2)\,J\,M\pm1|j_1\,m_1; j_2\,m_2\pm1\,\rangle \end{aligned} \tag{0.86}$$

which, together with the normalization $\sum_{m_1,m_2}|\langle\,j_1\,m_1; j_2\,m_2|(j_1\,j_2)\,J\,M\,\rangle|^2 = 1$ and the convention (0.85), allows one to determine all the C.G. As stated before, they are clearly all real.

The C.G. of the group SU(2), which describe a change of orthonormal basis, form a unitary matrix and thus satisfy orthogonality and completeness properties

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

$$\sum_{J=|j_1-j_2|}^{j_1+j_2} \langle\,j_1\,m_1; j_2\,m_2|(j_1\,j_2)\,J\,M\,\rangle\langle\,j_1\,m_1'; j_2\,m_2'|(j_1\,j_2)\,J\,M\,\rangle = \delta_{m_1m_1'}\delta_{m_2m_2'} \quad \text{if } |m_1|\le j_1,\ |m_2|\le j_2\ . \tag{0.87}$$

Once again, each relation implies only one non trivial summation.

Rather than the C.G. coefficients, one may consider another set of equivalent coefficients, called *3-j symbols*. They are defined through

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

and they enjoy simple symmetry properties:

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

is invariant under cyclic permutation of its three columns, and changes by the sign $(-1)^{j_1+j_2+j_3}$ when two columns are interchanged *or* when the signs of m_1, m_2 and m_3 are reversed. The reader will find a multitude of tables and explicit formulas of the C.G. and $3j$ coefficients in the literature.

Let us just give some values of C.G. for low spins

$$\frac{1}{2}\otimes\frac{1}{2}\ :\quad \begin{array}{rcl} |(\frac12,\frac12)1,1\,\rangle &=& |\frac12,\frac12;\frac12,\frac12\,\rangle \\ |(\frac12,\frac12)1,0\,\rangle &=& \frac{1}{\sqrt2}\left(|\frac12,\frac12;\frac12,-\frac12\,\rangle + |\frac12,-\frac12;\frac12,\frac12\,\rangle\right) \\ |(\frac12,\frac12)0,0\,\rangle &=& \frac{1}{\sqrt2}\left(|\frac12,\frac12;\frac12,-\frac12\,\rangle - |\frac12,-\frac12;\frac12,\frac12\,\rangle\right) \\ |(\frac12,\frac12)1,-1\,\rangle &=& |\frac12,-\frac12;\frac12,-\frac12\,\rangle \end{array} \tag{0.89}$$

and

$$\frac{1}{2}\otimes 1\ :\quad \begin{array}{rcl} |(\frac12,1)\frac32,\frac32\,\rangle &=& |\frac12,\frac12;1,1\,\rangle \\ |(\frac12,1)\frac32,\frac12\,\rangle &=& \frac{1}{\sqrt3}\left(\sqrt2|\frac12,\frac12;1,0\,\rangle + |\frac12,-\frac12;1,1\,\rangle\right) \\ |(\frac12,1)\frac32,-\frac12\,\rangle &=& \frac{1}{\sqrt3}\left(|\frac12,\frac12;1,-1\,\rangle + \sqrt2|\frac12,-\frac12;1,0\,\rangle\right) \\ |(\frac12,1)\frac32,-\frac32\,\rangle &=& |\frac12,-\frac12;1,-1\,\rangle \\ |(\frac12,1)\frac12,\frac12\,\rangle &=& \frac{1}{\sqrt3}\left(-|\frac12,\frac12;1,0\,\rangle + \sqrt2|\frac12,-\frac12;1,1\,\rangle\right) \\ |(\frac12,1)\frac12,-\frac12\,\rangle &=& \frac{1}{\sqrt3}\left(-\sqrt2|\frac12,\frac12;1,-1\,\rangle + |\frac12,-\frac12;1,0\,\rangle\right) \end{array} \tag{0.90}$$

One notices on the case $\frac12\otimes\frac12$ the property that vectors of total spin $j=1$ are symmetric under the exchange of the two spins, while those of spin 0 are antisymmetric. This is a general property: in the decomposition of the tensor product of two representations of spin $j_1 = j_2$, vectors of spin $j = 2j_1, 2j_1-2, \cdots$ are symmetric, those of spin $2j_1-1, 2j_1-3, \cdots$ are antisymmetric.

This is apparent on the expression (0.88) above, given the announced properties of the 3-j symbols.

In the same circle of ideas, consider the completely antisymmetric product of $2j+1$ copies of a spin j representation. One may show that this representation is of spin 0 (following exercise). (This has consequences in atomic physics, in the filling of electronic orbitals: a complete shell has a total orbital momentum and a total spin that are both vanishing, hence also a vanishing total angular momentum.)

Exercise. Consider the completely antisymmetric tensor product of $N=2j+1$ representations of spin j. Show that this representation is spanned by the vector $\epsilon_{m_1m_2\cdots m_N}|j\,m_1, j\,m_2,\cdots,j\,m_N\,\rangle$, that it is invariant under the action of SU(2) and thus that the corresponding representation has spin $J=0$.

One also introduces the *6-j symbols* that describe the two possible recombinations of 3 representations of spins j_1, j_2 and j_3

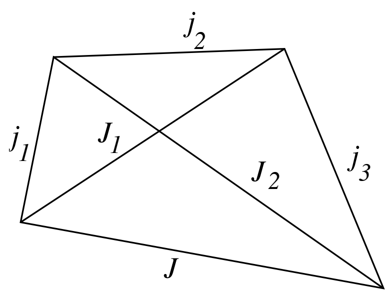


Fig. 4

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

depending on whether one composes first j_1 and j_2 into J_1 and then J_1 and j_3 into J, or first j_2 and j_3 into J_2 and then j_1 and J_2 into J'. The matrix of the change of basis is denoted

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

and the $\{\quad\}$ are the 6-j symbols. One may visualise this operation of "addition" of the three spins by a tetrahedron (see Fig. 4) the edges of which carry j_1, j_2, j_3, J_1, J_2 and J and the symbol is such that two spins carried by a pair of opposed edges lie in the same column. These symbols are tabulated in the literature.

0.5 A physical application: isospin

The group SU(2) appears in physics in several contexts, not only as related to the rotation group of the 3-dimensional Euclidian space. We shall now illustrate another of its avatars by the isospin symmetry.

There exists in nature elementary particles subject to nuclear forces, or more precisely to "strong interactions", and thus called *hadrons*. Some of those particles present similar properties but have different electric charges. This is the case with the two "nucleons", *i.e.* the proton and the neutron, of respective masses M_p =938,28 MeV/c^2 and M_n =939,57 MeV/c^2, and also with the "triplet" of pi mesons, π^0 (masse 134,96 MeV/c^2) and $\pi^\pm$ (139,57 MeV/c^2), with K mesons etc. According to a great idea of Heisenberg these similarities are the manifestation of a symmetry broken by electromagnetic interactions. In the absence of electromagnetic interaction proton and neutron on the one hand, the three π mesons on the other, etc, would have the same mass, differing only by an "internal" quantum number, in the same way as the two spin states of an electron in the absence of a magnetic field. In fact the group behind that symmetry is also SU(2), but a SU(2) group acting in an abstract space differing from the usual space. One gave the name *isotopic spin* or in short, *isospin*, to the corresponding quantum number. To summarize, the idea is that there exists a SU(2) group of symmetry of strong interactions, and that different particles subject to these strong interactions (hadrons) form representations of SU(2) : representation of isospin $I = \frac{1}{2}$ for the nucleon (proton $I_z = +\frac{1}{2}$, neutron $I_z = -\frac{1}{2}$), isospin $I = 1$ pour the pions ($\pi^\pm : I_z = \pm 1$, $\pi^0 : I_z = 0$) etc. The isospin is thus a "good quantum number", conserved in these interactions. Thus the "off-shell" process $N \to N + \pi$,

(N for nucleon) important in nuclear physics, is consistent with addition rules of isospins ($\frac{1}{2}\otimes 1$ "contains" $\frac{1}{2}$). The different scattering reactions $N+\pi \to N+\pi$ allowed by conservation of electric charge

$$\begin{array}{rcll} p+\pi^+ & \to & p+\pi^+ & I_z=\frac{3}{2} \\ p+\pi^0 & \to & p+\pi^0 & I_z=\frac{1}{2} \\ & \to & n+\pi^+ & \quad '' \\ p+\pi^- & \to & p+\pi^- & I_z=-\frac{1}{2} \\ & \to & n+\pi^0 & \quad '' \\ n+\pi^- & \to & n+\pi^- & I_z=-\frac{3}{2} \end{array}$$

also conserve total isospin I and its I_z component but the hypothesis of SU(2) isospin invariance tells us more. The matrix elements of the transition operator responsible for the two reactions in the channel $I_z=\frac{1}{2}$, for example, must be related by addition rules of isospin. Inverting the relations (0.90), one gets

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

whereas for $I_z=3/2$

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

Isospin invariance implies that $\langle\, I\, I_z\,|\mathcal{T}|I'\,I_z'\,\rangle=\mathcal{T}_I\delta_{II'}\delta_{I_z\,I_z'}$, as we shall see later (Chap. 2): not only are I and I_z conserved, but the resulting amplitude depends only on I, not I_z. Calculating then the matrix elements of the transition operator $\mathcal{T}$ between the different states,

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

one finds that amplitudes satisfy a relation

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

a non trivial consequence of isospin invariance, which implies triangular inequalities between squared modules of these amplitudes and hence between cross-sections of the reactions

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

which are experimentally well verified. Even better, one finds experimentally that at a certain energy of about 180 MeV, cross sections (proportional to squares of amplitudes) are in the ratios

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

which indicates that at that energy, scattering in the channel $I = 3/2$ is dominant. In fact, this signals the existence of an intermediate πN state, a very unstable particle called "resonance", denoted Δ, of isospin 3/2 and hence with four states of charge

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

the contribution of which dominates the scattering amplitude. This particle has a spin 3/2 and a mass $M(\Delta) \approx 1230$ MeV/c^2.

In some cases one may obtain more precise predictions. This is for instance the case with the reactions

$$^2\mathrm{H}\, p \to {}^3\mathrm{He}\, \pi^0 \quad \text{and} \quad {}^2\mathrm{H}\, p \to {}^3\mathrm{H}\, \pi^+$$

which involve nuclei of deuterium ^{2}H, of tritium ^{3}H and of helium ^{3}He. To these nuclei too, one may assign an isospin, 0 to the deuteron which is made of a proton and a neutron in an antisymmetric state of their isospins (so that the wave function of these two *fermions*, symmetric in space and in spin, be antisymmetric), $I_z = -\frac{1}{2}$ to ^{3}H and $I_z = \frac{1}{2}$ to ^{3}He which form an isospin $\frac{1}{2}$ representation. Notice that in all cases, the electric charge is related to the I_z component of isospin by the relation $Q = \frac{1}{2}\mathcal{B} + I_z$, with $\mathcal{B}$ the baryonic charge, equal here to the number of nucleons (protons or neutrons).
Exercise: show that the ratio of cross-sections $\sigma(^2\mathrm{H}\, p \to {}^3\mathrm{He}\, \pi^0)/\sigma(^2\mathrm{H}\, p \to {}^3\mathrm{H}\, \pi^+)$ is $\frac{1}{2}$.

Remark : invariance under isospin SU(2) that we just discussed is a symmetry of *strong interactions.* There exists also in the framework of the Standard Model a notion of "weak isospin", a symmetry of *electroweak interactions*, to which we return in Chap. 5.

0.6 Representations of SO(3,1) and SL(2,$\mathbb{C}$)

0.6.1 A short reminder on the Lorentz group

Minkowki space is a $\mathbb{R}^4$ space endowed with a pseudo-Euclidean metric of signature $(+,-,-,-)$. In an orthonormal basis with coordinates $(x^0 = ct, x^1, x^2, x^3)$, the metric is diagonal

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

and thus the squared norm of a 4-vector reads

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

The isometry group of that quadratic form, called O(1,3) or the Lorentz group $\mathcal{L}$, is such that

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

i.e.

$$\Lambda^\mu{}_\rho g_{\mu\nu} \Lambda^\nu{}_\sigma = g_{\rho\sigma} \quad \text{or} \quad \Lambda^T g \Lambda = g \,. \tag{0.93}$$

These pseudo-orthogonal matrices satisfy $(\det\Lambda)^2 = 1$ and (by taking the 00 matrix element of (0.93)) $(\Lambda^0{}_0)^2 = 1 + \sum_{i=1}^3 (\Lambda^0{}_i)^2 \geq 1$ and thus $\mathcal{L} \equiv$ O(1,3) has four connected components (or "sheets") depending on whether $\det\Lambda = \pm 1$ and $\Lambda^0{}_0 \geq 1$ or ≤ -1. The subgroup of *proper orthochronous* transformations satisfying $\det\Lambda = 1$ and $\Lambda^0{}_0 \geq 1$ is denoted $\mathcal{L}_+^\uparrow$. Any transformation of $\mathcal{L}_+^\uparrow$ may be written as the product of an "ordinary" rotation of SO(3) and a "special Lorentz transformation" or "boost".

A major difference between the SO(3) and the $\mathcal{L}_+^\uparrow$ groups is that the former is *compact* (the range of parameters is bounded and closed, see (0.2)), whereas the latter is not : in a boost along the 1 direction, say, $x'_1 = \gamma(x_1 - vx_0/c), x'_0 = \gamma(x_0 - vx_1/c)$, with $\gamma = (1 - v^2/c^2)^{-\frac{1}{2}}$, the velocity $|v| < c$ does not belong to a compact domain (or alternatively, the "rapidity" variable β, defined by $\cosh\beta = \gamma$ can run to infinity). This compactness/non-compactness has very important implications on the nature and properties of representations, as we shall see.

The Poincaré group, or inhomogeneous Lorentz group, is generated by Lorentz transformations $\Lambda \in \mathcal{L}$ and space-time translations; generic elements denoted (a, Λ) have an action on a vector x and a composition law given by

$$\begin{aligned} (a,\Lambda) \quad : \quad x \mapsto x' &= \Lambda x + a \\ (a',\Lambda')(a,\Lambda) &= (a' + \Lambda' a, \Lambda'\Lambda) \ ; \end{aligned} \tag{0.94}$$

the inverse of (a, Λ) is $(-\Lambda^{-1}a, \Lambda^{-1})$ (check it !).

0.6.2 Lie algebra of the Lorentz and Poincaré groups

An infinitesimal Poincaré transformation reads $(\alpha^\mu, \Lambda^\mu{}_\nu = \delta^\mu_\nu + \omega^\mu{}_\nu)$. By taking the infinitesimal form of (0.93), one easily sees that the tensor $\omega_{\rho\nu} = \omega^\mu{}_\nu g_{\rho\mu}$ has to be antisymmetric: $\omega_{\nu\rho} + \omega_{\rho\nu} = 0$. This leaves 6 real parameters: the Lorentz group is a 6-dimensional group, and the Poincaré group is 10-dimensional.

To find the Lie algebra of the generators, let us proceed like in § 0.2.3: look at the Lie algebra generated by differential operators acting on functions of space-time coordinates; if $x'^\lambda = x^\lambda + \delta x^\lambda = x^\lambda + \alpha^\lambda + \omega^{\lambda\nu}x_\nu$, $\delta f(x) = f(x^\mu - \alpha^\lambda - \omega^{\lambda\nu}x_\nu) - f(x) = (I - i\alpha^\mu P_\mu - \frac{i}{2}\omega^{\mu\nu}J_{\mu\nu})f(x)$, (cp (0.45)), thus

$$J_{\mu\nu} = i(x_\mu\partial_\nu - x_\nu\partial_\mu) \qquad\qquad P_\mu = -i\partial_\mu \tag{0.95}$$

the commutators of which are then easily computed

$$\begin{aligned} [J_{\mu\nu}, P_\rho] &= i\,(g_{\nu\rho}P_\mu - g_{\mu\rho}P_\nu) \\ [J_{\mu\nu}, J_{\rho\sigma}] &= i\,(g_{\nu\rho}J_{\mu\sigma} - g_{\mu\rho}J_{\nu\sigma} + g_{\mu\sigma}J_{\nu\rho} - g_{\nu\sigma}J_{\mu\rho}) \\ [P_\mu, P_\nu] &= 0\,. \end{aligned} \tag{0.96}$$

Note the structure of these relations: antisymmetry in $\mu \leftrightarrow \nu$ of the first one, in $\mu \leftrightarrow \nu$, in $\rho \leftrightarrow \sigma$ and in $(\mu,\nu) \leftrightarrow (\rho,\sigma)$ of the second one; the first one shows how a vector (here P_ρ) transforms under the infinitesimal transformation by $J_{\mu\nu}$, and the second then has the same pattern in the indices ρ and σ, expressing that $J_{\rho\sigma}$ is a 2-tensor.

Generators that commute with P_0 (which is the generator of time translations, hence the Hamiltonian) are the P_μ and the J_{ij} but not the J_{0j} : $i[P_0, J_{0j}] = P_j$.

Set

$$J_{ij} = \epsilon_{ijk}J^k \qquad K^i = J_{0i}\ . \tag{0.97}$$

Then

$$\begin{aligned}
[J^i, J^j] &= i\epsilon_{ijk} J^k \\
[J^i, K^j] &= i\epsilon_{ijk} K^k \\
[K^i, K^j] &= -i\epsilon_{ijk} J^k
\end{aligned} \tag{0.98}$$

and also

$$\begin{aligned}
[J^i, P^j] &= i\epsilon_{ijk} P^k \qquad & [K^i, P^j] &= iP^0 \delta_{ij} \\
[J^i, P^0] &= 0 \qquad & [K^i, P^0] &= iP^i \ .
\end{aligned} \tag{0.99}$$

Remark. The first two relations (0.98) and the first one of (0.99) express that, as expected, $\mathbf{J} = \{J^i\}$, $\mathbf{K} = \{K^i\}$ and $\mathbf{P} = \{P^i\}$ transform like vectors under rotations of $\mathbb{R}^3$. Now form the combinations

$$M^j = \frac{1}{2}(J^j + iK^j) \qquad N^j = \frac{1}{2}(J^j - iK^j) \tag{0.100}$$

which have the following commutation relations

$$\begin{aligned}
[M^i, M^j] &= i\epsilon_{ijk} M^k \\
[N^i, N^j] &= i\epsilon_{ijk} N^k \\
[M^i, N^j] &= 0 \ .
\end{aligned} \tag{0.101}$$

By considering the complex combinations M and N of its generators, one thus sees that the Lie algebra of $\mathcal{L} = O(1,3)$ is isomorphic to $su(2) \oplus su(2)$. The introduction of $\pm i$, however, implies that *unitary* representations of $\mathcal{L}$ do not follow in a simple way from those of SU(2)$\times$ SU(2). On the other hand, representations of finite dimension of $\mathcal{L}$, which are *non unitary* , are labelled by a pair (j_1, j_2) of integers or half-integers.

0.6.3 Covering groups of $\mathcal{L}_+^\uparrow$ and $\mathcal{P}_+^\uparrow$

We have seen that the study of SO(3) led us to SU(2), its “covering group” (the deep reasons of which will be explained in Chap. 1 and 2). Likewise in the case of the Lorentz group one is finds that its “covering group” turns out to be SL(2,ℂ).

There is a simple way to see how SL(2,ℂ) and $\mathcal{L}_+^\uparrow$ are related, which is a 4-dimensional extension of the method followed in § 0.1.2. One considers matrices σ_μ made of $\sigma_0 = \mathbf{I}$ and of the three familiar Pauli matrices. Note that

$$\mathrm{tr}\,\sigma_\mu\sigma_\nu = 2\delta_{\mu\nu} \qquad \sigma_\mu^2 = \mathbf{I} \quad \text{with no summation over the index } \mu \ .$$

With any real vector $x \in \mathbb{R}^4$, associate the Hermitian matrix

$$X = x^\mu \sigma_\mu \qquad x^\mu = \frac{1}{2}\mathrm{tr}\,(X\sigma_\mu) \qquad \det X = x^2 = (x^0)^2 - \mathbf{x}^2 \ . \tag{0.102}$$

A matrix $A \in \mathrm{SL}(2,\mathbb{C})$ acts on X according to

$$X \mapsto X' = AXA^\dagger \tag{0.103}$$

which is indeed Hermitian and thus defines a real $x'^{\mu} = \frac{1}{2}\mathrm{tr}\,(X'\sigma_\mu)$, with $\det X' = \det X$, hence $x^2 = x'^2$. This is a linear transformation of $\mathbb{R}^4$ that preserves the Minkowski norm x^2, and thus a Lorentz transformation, and one checks by an argument of continuity that it belongs to $\mathcal{L}_+^\uparrow$ and that $A \to \Lambda$ is a homomorphism of $\mathrm{SL}(2,\mathbb{C})$ into $\mathcal{L}_+^\uparrow$. In the following we denote $x' = A.x$ if $X' = AXA^\dagger$.

As is familiar from the case of SU(2), the transformations A and $-A \in \mathrm{SL}(2,\mathbb{C})$ give the same transformation of $\mathcal{L}_+^\uparrow$: $\mathrm{SL}(2,\mathbb{C})$ is a covering of order 2 of $\mathcal{L}_+^\uparrow$. For the Poincaré group, likewise, its covering is the ("semi-direct") product of the translation group by $\mathrm{SL}(2,\mathbb{C})$. If one denotes $\underline{a} := a^\mu\sigma_\mu$, then

$$(\underline{a}, A)(\underline{a}', A') = (\underline{a} + A\underline{a}'A^\dagger, AA') \tag{0.104}$$

and one sometimes refers to it as the "inhomogeneous $\mathrm{SL}(2,\mathbb{C})$ group", or $\mathrm{ISL}(2,\mathbb{C})$).

0.6.4 Irreducible finite-dimensional representations of $\mathrm{SL}(2,\mathbb{C})$

The construction of § 0.3.3 yields an explicit representation of GL(2,ℂ) and hence of SL(2,ℂ). For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{C})$, (0.71) gives the following expression for $\mathcal{D}^j_{mm'}(A)$:

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

Note that $\mathcal{D}^T(A) = \mathcal{D}(A^T)$ (since exchanging $m \leftrightarrow m'$ amounts to $n_2 \leftrightarrow n_3$, hence to $b \leftrightarrow c$) and $(\mathcal{D}(A))^* = \mathcal{D}(A^*)$ (since the numerical coefficients in (0.71) are real) thus $\mathcal{D}^\dagger(A) = \mathcal{D}(A^\dagger)$.

This representation is called $(j, 0)$, it is of dimension $2j + 1$. There exists another one of dimension $2j+1$, which is non equivalent, denoted $(0, j)$, this is the "contragredient conjugate" representation (in the sense of Chap 2. §2.1.3) $\mathcal{D}^j(A^{\dagger\,-1})$. Replacing A by $A^{\dagger\,-1}$ may be interpreted in the construction of § 0.6.3 if instead of associating $X = x^\mu\sigma_\mu$ with x, one associates $\widetilde{X} = x^0\sigma_0 - \mathbf{x}.\boldsymbol{\sigma}$. Notice that $\sigma_2(\sigma_i)^T\sigma_2 = -\sigma_i$ for $i = 1, 2, 3$ hence $\widetilde{X} = \sigma_2 X^T\sigma_2$. For the transformation A : $X \mapsto X' = AXA^\dagger$, we have

$$\widetilde{X}' = \sigma_2(X')^T\sigma_2 = \sigma_2(AXA^\dagger)^T\sigma_2 = (\sigma_2A^T\sigma_2)^\dagger\widetilde{X}(\sigma_2A^T\sigma_2)\ .$$

Any matrix A of $\mathrm{SL}(2,\mathbb{C})$ may itself be written as $A = a^\mu\sigma_\mu$, with $(a^\mu) \in \mathbb{C}^4$, and as $\det A = (a^0)^2 - \mathbf{a}^2 = 1$ (the "S" of $\mathrm{SL}(2,\mathbb{C})$), one verifies immediately that $A^{-1} = a^0\sigma_0 - \mathbf{a}.\boldsymbol{\sigma}$, donc

$$\sigma_2A^T\sigma_2 = A^{-1}\ . \tag{0.105}$$

Finally

$$X' = AXA^\dagger \quad \Longleftrightarrow \quad \widetilde{X}' = (A^{-1})^\dagger\widetilde{X}A^{-1}\ . \tag{0.106}$$

Remark. The two representations $(j, 0)$ and $(0, j)$ are inequivalent on $\mathrm{SL}(2,\mathbb{C})$, but equivalent on SU(2). Indeed in SU(2), $A = U = (U^\dagger)^{-1}$.

Finally, one proves that any finite-dimensional representation of $\mathrm{SL}(2,\mathbb{C})$ is completely reducible and may be written as a direct sum of irreducible representations. The most general finite-dimensional irreducible representation of $\mathrm{SL}(2,\mathbb{C})$ is denoted (j_1, j_2), with j_1 and $j_2 \geq 0$ integers or half-integers; it is defined by

$$(j_1, j_2) = (j_1, 0) \otimes (0, j_2) \; . \tag{0.107}$$

All these representations may be obtained from the representations $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ by tensoring: $(j_1, 0)$ and $(0, j_2)$ are obtained by symmetrized tensor product of representations $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$, respectively, as was done for SU(2). Only representations (j_1, j_2) with j_1 and j_2 simultaneously integers or half-integers are true representations of $\mathcal{L}_\uparrow^+$. The others are representations up to a sign.

Exercise : show that the representation $(0, j)$ is "equivalent" (equal up to a change of basis) to the complex conjugate of representation $(j, 0)$. (Hint: show it first for $j = \frac{1}{2}$ by recalling that $(A^{-1})^\dagger = \sigma_2 A^* \sigma_2$, then for representations of arbitrary j obtained by $2j$-th tensor power of $j = \frac{1}{2}$.)

Spinor representations

Return to the "spinor representations" $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$. Those are representations of dimension 2 (two-component spinors). It is traditional to note the indices of components with "pointed" or "unpointed" indices, for representation $(0, \frac{1}{2})$ and $(\frac{1}{2}, 0)$, respectively. With $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in$ $\mathrm{SL}(2,\mathbb{C})$, we thus have

$$\begin{array}{lll} (\frac{1}{2}, 0) & \quad \xi = (\xi^\alpha) \mapsto \xi' = A\xi = \begin{pmatrix} a\xi^1 + b\xi^2 \\ c\xi^1 + d\xi^2 \end{pmatrix} \\ (0, \frac{1}{2}) & \quad \xi = (\xi^{\dot\alpha}) \mapsto \xi' = A^*\xi = \begin{pmatrix} a^*\xi^{\dot 1} + b^*\xi^{\dot 2} \\ c^*\xi^{\dot 1} + d^*\xi^{\dot 2} \end{pmatrix} \end{array} \tag{0.108}$$

Note that the alternating (=antisymmetric) form $(\xi, \eta) = \xi^1\eta^2 - \xi^2\eta^1 = \xi^T(i\sigma_2)\eta$ is invariant in $(\frac{1}{2}, 0)$ (and also in $(0, \frac{1}{2})$), which follows once again from (0.105)

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

One may thus use that form to lower indices α (or $\dot\alpha$). Thus

$$\begin{array}{lllll} \text{in } (\frac{1}{2}, 0) & : & (\xi, \eta) = \xi_\alpha \eta^\alpha & \quad \xi_2 = \xi^1 & \quad \xi_1 = -\xi^2 \\ \text{in } (0, \frac{1}{2}) & : & (\xi, \eta) = \xi_{\dot\alpha} \eta^{\dot\alpha} & \quad \xi_{\dot 2} = \xi^{\dot 1} & \quad \xi_{\dot 1} = -\xi^{\dot 2} \end{array} \tag{0.109}$$

(j_1, j_2) representation

Tensors $\{\xi^{\alpha_1\alpha_2\cdots\alpha_{2j_1}\dot\beta_1\dot\beta_2\cdots\dot\beta_{2j_2}}\}$ symmetric in $\alpha_1, \alpha_2, \cdots, \alpha_{2j_1}$ and in $\dot\beta_1, \dot\beta_2, \cdots, \dot\beta_{2j_2}$, form the irreducible representation (j_1, j_2). (One cannot lower the rank by taking traces, since the only invariant tensor is the previous *alternating* form). The dimension of that representation is $(2j_1 + 1)(2j_2 + 1)$. The most usual representations encountered in field theory are $(0, 0)$, $(\frac{1}{2}, 0)$

and $(0,\frac{1}{2})$, $(\frac{1}{2},\frac{1}{2})$. The reducible representation $(\frac{1}{2},0)\oplus(0,\frac{1}{2})$ describes the (4-component) Dirac fermion; the $(\frac{1}{2},\frac{1}{2})$ corresponds to 4-vectors, as seen above:

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

i.e.

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

which shows that X transforms indeed according to the $(\frac{1}{2},\frac{1}{2})$ representation .

Exercise. Show that representations $(1,0)$ and $(0,1)$, of dimension 3, describe rank 2 tensors $F^{\mu\nu}$ that are "self-dual" ou "anti-self-dual", *i.e.* satisfy

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

0.6.5 Irreducible unitary representations of the Poincaré group. One particle states.

According to a theorem of Wigner which will be discussed in Chap, 2, the action of proper orthochronous transformations of the Lorentz or Poincaré groups on state vectors of a quantum theory is described by means of *unitary* representations of these groups, or rather of their "universal covers" SL(2,ℂ) and ISL(2,ℂ). As will be seen below (Chap. 2), unitary representations (of class L^2) of the non compact group $\mathrm{SL}(2,\mathbb{C})$ are necessarily of infinite dimension (with the possible exception of the trivial representation $(0,0)$, which describes a state invariant by rotation and by boosts, *i.e.* the vacuum !,..., and which is in fact not of class L^2!).

Returning to commutation relations of the Lie algebra (0.96), one seeks a maximal set of commuting operators. The four P_μ commute. Let (p_μ) be an eigenvalue for a common eigenvector of P_μ, describing a "one-particle state". We assume that the eigenvector denoted $|p\rangle$ is labelled only by p^μ and by discrete indices: (this is indeed the meaning of "one-particle state", in contrast with a two-particle state that would depend on a relative momentum, a continuous variable)

$$P_\mu|p\rangle = p_\mu|p\rangle \ . \tag{0.110}$$

One also considers the *Pauli-Lubanski tensor*

$$W^\lambda = \frac{1}{2}\epsilon^{\lambda\mu\nu\rho}J_{\mu\nu}P_\rho \tag{0.111}$$

and one verifies (exercise !) that (0.96) implies

$$\begin{aligned}
[W_\mu, P_\nu] &= 0 \\
[W^\mu, W^\nu] &= -i\epsilon^{\mu\nu\rho\sigma}W_\rho P_\sigma \\
[J_{\mu\nu}, W_\lambda] &= i(g_{\nu\lambda}W_\mu - g_{\mu\lambda}W_\nu) \ .
\end{aligned} \tag{0.112}$$

The latter relation means that W is a Lorentz 4-vector (compare with (0.96)). One also notes that $W.P = 0$ because of the antisymmetry of tensor ϵ. One finally shows (check it!) that $P^2 = P_\mu P^\mu$ and $W^2 = W_\mu W^\mu$ commute with all generators P and J : those are the Casimir operators of the algebra. According to a lemma by Schur, (see below Chap. 2, § 2.1.4), these Casimir operators are in any irreducible representation proportional to the identity, in other words, their eigenvalues may be used to label the irreducible representations.

In physics, one encounters only two types of representations for these one-particle states[4]: representations with $P^2 > 0$ and those with $P^2 = 0$, $W^2 = 0$. Their detailed study will be done in Adel Bilal's course.

[4]which does not mean that there are no other irreducible representations; for example "unphysical" representations where $P^2 = -M^2 < 0$

Bibliography

The historical reference for the physicist is the book by E. Wigner [Wi].
For a detailed discussion of the rotation group, with many formulas and tables, see:
J.-M. Normand, *A Lie group : Rotations in Quantum Mechanics*, North-Holland.
For a deep and detailed study of physical representations of Lorentz and Poincaré groups, see P. Moussa and R. Stora, *Angular analysis of elementary particle reactions*, in *Analysis of scattering and decay*, edited by M. Nikolic, Gordon and Breach 1968.

Problem

One considers two spin $\frac{1}{2}$ representations of the group SU(2) and their direct (or tensor) product. One denotes $\mathbf{J}^{(1)}$ and $\mathbf{J}^{(2)}$ the infinitesimal generators acting in each representation, and $\mathbf{J} = \mathbf{J}^{(1)} + \mathbf{J}^{(2)}$ those acting in their direct product, see (0.75), (0.75').

- What can be said about the operators $\mathbf{J}^{(1)\,2}$, $\mathbf{J}^{(2)\,2}$ and $\mathbf{J}^2$ and their eigenvalues ?
- Show that $\mathbf{J}^{(1)}.\mathbf{J}^{(2)}$ may be expressed in terms of these operators and that operators

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

  are projectors on spaces to be identified.
- Taking into account the symmetries of the vectors under exchange, what can you say about the operator

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

Appendix 0. The laplacian on the S^2 and S^3 spheres.

Consider a Riemannian manifold, *i.e.* a manifold endowed with a metric :

$$ds^2 = g_{\alpha\beta}\, d\xi^\alpha d\xi^\beta \tag{0.113}$$

with a metric tensor g and (local) coordinates ξ^α, $\alpha = 1, \cdots, n$; $n =$ the dimension of the manifold. This ds^2 must be invariant under changes of coordinates, $\xi \to \xi'$, which dictates the change of the tensor g

$$\xi \mapsto \xi' \ , \quad g \mapsto g' \qquad : \quad g'_{\alpha'\beta'} = \frac{\partial \xi^\alpha}{\partial \xi'^{\alpha'}} \frac{\partial \xi^\beta}{\partial \xi'^{\beta'}} g_{\alpha\beta} \, , \tag{0.114}$$

meaning that g is a covariant rank-2 tensor. The metric tensor is assumed to be non singular, *i.e.* invertible, and its inverse tensor is denoted with upper indices

$$g_{\alpha\beta}\, g^{\beta\gamma} = \delta_\alpha{}^\gamma \, . \tag{0.115}$$

Also, its determinant is traditionnally denoted g

$$g = \det(g_{\alpha\beta})\,. \tag{0.116}$$

There is then a general method to construct a volume element on the manifold (*i.e.* an integration measure) and a Laplacian, both invariant under changes of coordinates

$$\begin{aligned} \mathrm{d}\mu(\xi) &= \sqrt{g}\prod_{\alpha=1}^{n}\mathrm{d}\xi^\alpha \\ \Delta &= \frac{1}{\sqrt{g}}\partial_\alpha\sqrt{g}g^{\alpha\beta}\partial_\beta \end{aligned} \tag{0.117}$$

where ∂_α is a shorthand notation for the differential operator $\frac{\partial}{\partial\xi^\alpha}$.
Exercise: check that $\mathrm{d}\mu(\xi)$ and Δ are invariant under a change of coordinates $\xi \mapsto \xi'$.

This may be applied in many contexts, and will be used in Chap. 1 to define an integration measure on compact Lie groups.

Let us apply it here to the n-dimensional Euclidean space $\mathbb{R}^n$. In spherical coordinates, one writes

$$\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\Omega^2$$

where $\mathrm{d}\Omega$ is a generic notation that collects all the angular variables. The metric tensor is thus of the general form $\begin{pmatrix}1 & 0\\ 0 & r^2A\end{pmatrix}$ with a $(n-1)\times(n-1)$ matrix A which is r-independent and depends only on angular variables. The latter give rise to the Laplacian on the unit sphere S^{n-1}, denoted $\Delta_{S^{n-1}}$; $\sqrt{g} = r^{n-1}\sqrt{\det A}$; and (0.117) tells us that the Laplacian on $\mathbb{R}^n$ takes the general form

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

Let us write more explicit formulae for the S^2 and S^3 unit spheres. Consider first the unit sphere S^2 with angular coordinates $0 \le \theta \le \pi$, $0 \le \phi \le 2\pi$ (Fig. 1). We thus have

$$\begin{aligned} ds^2 &= d\theta^2 + \sin^2\theta\, d\phi^2 \\ \sqrt{g} &= \sin\theta \\ d\mu(x) &= \sin\theta\, d\theta\, d\phi \\ \Delta_{S^2} &= \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\,. \end{aligned} \tag{0.118}$$

The generators J_i read (see (0.46))

$$\begin{aligned} J_3 &= -i\frac{\partial}{\partial\phi} \\ J_1 &= -i\left[-\cos\phi\,\mathrm{cotg}\,\theta\frac{\partial}{\partial\phi} - \sin\phi\frac{\partial}{\partial\theta}\right] \\ J_2 &= -i\left[-\sin\phi\,\mathrm{cotg}\,\theta\frac{\partial}{\partial\phi} + \cos\phi\frac{\partial}{\partial\theta}\right] \end{aligned} \tag{0.119}$$

and one verifies that $-\Delta_{S^2} = \vec{J}^2 = J_1^2 + J_2^2 + J_3^2$.

For the unit sphere S^3 one finds similar formulas. In the parametrization (0.12), one takes for example

$$\mathrm{d}s^2 = \frac{1}{2}\mathrm{tr}\,\mathrm{d}U\mathrm{d}U^\dagger = \left(\mathrm{d}\frac{\psi}{2}\right)^2 + \sin^2\frac{\psi}{2}\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right) \tag{0.120}$$

invariant under $U \to UV$, $U \to VU$ or $U \to U^{-1}$, whence a measure invariant under the same transformations

$$\mathrm{d}\mu(U) = \frac{1}{2}\left(\sin\frac{\psi}{2}\right)^2 \sin\theta\,\mathrm{d}\psi\,\mathrm{d}\theta\,\mathrm{d}\phi\ . \tag{0.121}$$

In the Euler angles parametrization,

$$U = e^{-i\alpha\frac{\sigma_3}{2}}e^{-i\beta\frac{\sigma_2}{2}}e^{-i\gamma\frac{\sigma_3}{2}} \tag{0.122}$$

thus

$$\mathrm{d}s^2 = \frac{1}{2}\mathrm{tr}\,\mathrm{d}U\mathrm{d}U^\dagger = \frac{1}{4}\left(\mathrm{d}\alpha^2 + 2\mathrm{d}\alpha\mathrm{d}\gamma\cos\beta + \mathrm{d}\gamma^2 + \mathrm{d}\beta^2\right) \tag{0.123}$$

and with $\sqrt{g} = \sin\beta$ one computes

$$d\mu(U) = \frac{1}{8}\sin\beta\,d\alpha\,d\beta\,d\gamma \tag{0.124}$$

$$\Delta_{S^3} = \frac{4}{\sin^2\beta}\left[\frac{\partial^2}{\partial\alpha^2} + \frac{\partial^2}{\partial\gamma^2} + \frac{\partial^2}{\partial\alpha\partial\gamma}\right] + \frac{4}{\sin\beta}\frac{\partial}{\partial\beta}\sin\beta\frac{\partial}{\partial\sin\beta}\ . \tag{0.125}$$

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 $.$, $\times$ 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 groupe 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, $\{e^{2i\pi q/p}\}$, $q = 0, 1, \cdots, p-1$, or as the additive group of integers modulo p;
  - the groups of rotation invariance and the groups of rotation *and* reflexion 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 reflexions 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 ω_1 and ω_2 of non real ratio (why ?), this group describes the changes of basis $(\omega_1, \omega_2)^T \to (\omega'_1, \omega'_2)^T = A(\omega_1, \omega_2)^T$ that leave invariant the area of the elementary cell ($\Im(\omega_2\omega_1^*) = \Im(\omega'_2\omega'^*_1)$) and their effect on $\tau = \omega_2/\omega_1$: $\tau \to (a\tau + b)/(c\tau + d)$. This 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 $\mathrm{GL}(n,\mathbb{R})$ ou $\mathrm{GL}(n,\mathbb{C})$, for some n. In particular
  - $\mathrm{U}(n)$, the group of complex unitary matrices, $UU^\dagger = I$, which is the invariance group of the sesquilinear form $(x, y) = \sum x^{*i}y^i$;
  - $\mathrm{SU}(n)$ its unimodular subgroup, of unitary matrices of determinant $\det U = 1$;
  - $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ are orthogonal groups of invariance of the symmetric bilinear form $\sum_{i=1}^n x_i y_i$. Matrices of $\mathrm{SO}(n)$ have determinant 1 ;
  - $\mathrm{U}(p, q)$, $\mathrm{SU}(p, q)$, resp. $\mathrm{O}(p, q)$, $\mathrm{SO}(p, q)$, 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 $\mathrm{O}(n,\mathbb{R})$, $\mathrm{SO}(n,\mathbb{R})$ of matrices with real coefficients, but groups $\mathrm{O}(n,\mathbb{C})$, $\mathrm{SO}(n,\mathbb{C})$ of invariance of the same bilinear form over the complex numbers may also play a role.
  - $\mathrm{Sp}(2n,\mathbb{R})$: Let Z be the matrix $2n \times 2n$ made of a diagonal of n blocks $i\sigma_2$: $Z = \mathrm{diag}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, and consider the bilinear *skew-symmetric* form

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

    The symplectic group $\mathrm{Sp}(2n,\mathbb{R})$ 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 $\omega = \sum_{i=1}^n dp_i \wedge dq_i = \frac{1}{2} Z_{ij} d\xi_i \wedge d\xi_j$ in the coordinates $\xi = (p_1, q_1, p_2, \cdots, q_n)$; ω is invariant by action of $\mathrm{Sp}(2n,\mathbb{R})$ on ξ. For $n = 1$, verify that Sp(2,$\mathbb{R}$)=SL(2,$\mathbb{R}$).

    One may also consider the complex symplectic group $\mathrm{Sp}(2n,\mathbb{C})$. A related group, often denoted $\mathrm{Sp}(n)$ but that I shall denote $\mathrm{USp}(n)$ to avoid confusion with the previous ones, the *unitary symplectic group*, is the invariance group of a Hermitian quaternionic form, $\mathrm{USp}(n)$=$\mathrm{U}(2n) \cap \mathrm{Sp}(2n,\mathbb{C})$ 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 $\mathbf{x}' = R\mathbf{x} + \mathbf{v}t + \mathbf{x}_0$, $t' = t + t_0$,
  - the Poincaré group, in which translations are adjoined to the Lorentz group O(3,1),
  - etc etc.

1.1.2 Conjugacy classes of a group

In a group G we define the following equivalence relation:

$$a \sim b \;\text{ iff }\; \exists\, g \in G \;:\quad a = g.b.g^{-1} \tag{1.2}$$

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 a 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 (in Chap. 0) that in the rotation group $SO(3)$, a conjugacy class is characterized by the rotation angle ψ (around some unitary vector $\mathbf{n}$). But this notion is also familiar in the group $\mathrm{U}(n)$, 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 σ 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 σ is made of p_1 cycles of length 1, p_2 of length 2, etc, with $\sum i p_i = n$, 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:

  $$Z = \{a \mid \forall g \in G, a.g = g.a\} \tag{1.3}$$

  Z is a subgroup G, and is proper if G is nonabelian. Examples: the center of the group $GL(2, \mathbb{R})$ of regular 2×2 matrices is the set of matrices multiple of I; the center of $\mathrm{SU}(2)$ 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.

  $$Z_a = \{g \in G | a.g = g.a\} \tag{1.4}$$

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 $a\mathbf{I} + b\sigma_1$, $a^2 - b^2 \neq 0$.

- 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

$$Z(S) = \{y : \forall s \in S \quad y.s = s.y\} \tag{1.5}$$
$$N(S) = \{x : x^{-1}.S.x = S\} . \tag{1.6}$$

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 ρ of G into G' which respects the composition law:

$$\forall g, h \in G, \qquad \rho(g.h) = \rho(g).\rho(h) \tag{1.7}$$

In particular, ρ 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 :

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

which may also be rewritten as

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

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

$$g \sim_{\mathrm{L}} g' \iff g^{-1}.g' \in H \Leftrightarrow g \in g'.H. \tag{1.10}$$

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 (=cardinality) 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 $\forall g \in G, \quad gHg^{-1} = H$.

Exercise: check the equivalence of these three assertions.

The important property to remember is the following:

• *If $H$is an invariant subgroup 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 $g_2 \sim g_2'$, $\exists h_1, h_2 \in H \quad : \quad g_1 = h_1.g_1'$, $g_2 = g_2'.h_2$, 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 $g_1^{-1} = {g'}_1^{-1}.h_1^{-1} \sim {g'}_1^{-1}$. 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 ρ 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 ρ.

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?).

Direct, semi-direct product

Consider two groups G_1 and G_2 and their *direct product* $G = G_1 \times G_2$: it is the set of pairs (g_1, g_2) endowed with the natural product $(g'_1, g'_2).(g_1, g_2) = (g'_1 g_1, g'_2 g_2)$. Obviously its subgroups $\{(g_1, e)\} \simeq G_1$ and $\{(e, g_2)\} \simeq G_2$ are invariant subgroups, and G is not simple.

A more subtle construction appeals to the automorphism group of G_1 denoted $\mathrm{Aut}(G_1)$: this is the group of bijections β of G_1 into itself that respect its product (group homomorphism): $\beta(g'_1 g_1) = \beta(g'_1)\beta(g_1)$. Suppose there is a group homomorphism φ from another group G_2 into $\mathrm{Aut}(G_1)$: $\forall g_2 \in G_2, \varphi(g_2) \in \mathrm{Aut}(G_1)$ We now define on pairs (g_1, g_2) the following product

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

Exercise: show that this defines on these pairs a group structure. This is the semi-direct product of G_1 and G_2 (for a given φ) and is denoted $G_1 \rtimes_\varphi G_2$. Check that the subgroup $\{(g_1, e)\} \simeq G_1$ is an invariant subgroup of G.

Examples : the group of (orientation preserving) displacements, generated by translations and rotations in Euclidean $\mathbb{R}^n$, is the semi-direct product of $\mathbb{R}^n \rtimes \mathrm{SO}(n)$, with $(\vec{a}', R')(\vec{a}, R) = (\vec{a}' + R'.\vec{a}, R'R)$. Likewise the Poincaré group in Minkowski space is the semi-direct product $\mathbb{R}^4 \rtimes \mathcal{L}$.

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 $(g, h) \mapsto g.h$ 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 $g'.h'$ is nearby $g.h$ 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, or of gauge transformations in gauge theories.

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 topological space E is connected if its only subspaces that are both open and closed are E and $\emptyset$), 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 $\mathcal{L}$=O(3,1), the connected component of I is its proper orthochronous subgroup $\mathcal{L}_+^\uparrow$, (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]See Appendix B for a reminder of some points of vocabulary...

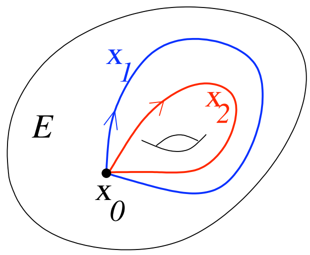


Figure 1.1: The paths x_1 et x_2 are homotopic. But none of them is homotopic to the "trivial" path that stays at x_0. The space is not simply connected.

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 $x_1(.)$ et $x_2(.)$ 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

$$\begin{aligned} \forall \xi \in [0,1] \qquad f(0,\xi) &= f(1,\xi) = x_0 & &: \text{ closed paths} \\ \forall t \in [0,1] \qquad f(t,0) &= x_1(t) \qquad f(t,1) = x_2(t) & &: \text{ interpolation .} \end{aligned} \tag{1.11}$$

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.

One may also *compose* paths: If $x_1(.)$ and $x_2(.)$ 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 $x_1(.)$ 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 in a connected space, homotopy groups relative to different end-points x_0 are isomorphic; if E is a connected group, see below, one may take for example the base point x_0 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

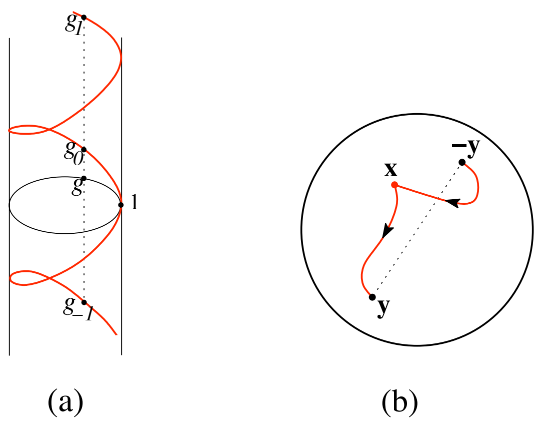


Figure 1.2: (a) The group U(1), identified with the circle and its universal covering group $\mathbb{R}$, identified with the helix. An element $g \in$ U(1) is lifted to points $\cdots, g_{-1}, g_0, g_1, \cdots$ on the helix. (b) In the ball B^3 representing SO(3), the points $\mathbf{y}$ and $-\mathbf{y}$ of the surface are identified. A path going from $\mathbf{x}$ to $\mathbf{x}$ via $\mathbf{y}$ and $-\mathbf{y}$ is thus closed but non contractible: SO(3) is non simply connected.

a neighborhood V_x and that $V_x \mapsto p(V_x)$ is a homeomorphism, *i.e.* a bicontinuous bijection [2]. 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 $\widetilde{G}$ is also a group, the *universal covering group*, and moreover, that the map p is a homomorphism of $\widetilde{G}$ into G. Its kernel which is an invariant subgroup of $\widetilde{G}$, is proved to be isomorphic to the homotopy group $\pi_1(G)$ ([Po], sect. 51). The quotient group is isomorphic to G

$$\widetilde{G}/\pi_1(G) \simeq G\ , \tag{1.12}$$

(according to a general property of the quotient group by the kernel of an homomorphism, cf. sect. 1.1.6).

One may construct the universal covering group $\widetilde{G}$ by considering paths that join the identity e to a point g, and their equivalence classes under continuous deformation **with fixed ends**. $\widetilde{G}$ 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 and gives $\widetilde{G}$ a group structure and one shows that $\widetilde{G}$ is simply connected (cf. [Po] sect. 51). The projection p of $\widetilde{G}$ 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 $\pi_1(\mathrm{U}(1)) = \mathbb{Z}$. The group $\widetilde{G}$ 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

[2] "bicontinuous" means that the map and its inverse are both continuous.

number modulo an integer multiple of 2π. 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. 0. To see this fact, represent the rotation $R_{\mathbf{n}}(\psi)$ by the point $\mathbf{x} = \tan\frac{\psi}{4}\mathbf{n}$ 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 π on the sphere $S^2 = \partial B^3$, but because of $R_{\mathbf{n}}(\pi) = R_{-\mathbf{n}}(\pi)$, (see Chap 0, 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 $\mathbf{x}$ to $\mathbf{x}$ 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 $\mathbf{x}$ to $\mathbf{x}$ and the group SO(3) is doubly connected. Its homotopy group is $\pi_1(\mathrm{SO}(3)) = \mathbb{Z}_2$. 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 $\pm U_{\mathbf{n}}(\psi) = \pm(\cos\frac{\psi}{2} - i\sin\frac{\psi}{2}\sigma.\mathbf{n}) \mapsto R_{\mathbf{n}}(\psi)$, see Chap. 0, sect. 1.2.

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 $R = R_{\mathbf{n}}(\psi)$ 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 $R_{\mathbf{n}}(\psi)$ and points that are arbitrarily close to its inverse $R_{-\mathbf{n}}(\psi)$ 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 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 $\pi_1(\mathrm{U}(n)) = \mathbb{Z}$;

• for the group SO(2)$\cong$ U(1), we have seen that π_1(SO(2))$= \mathbb{Z}$;

• for any $n > 2$, SO(n) is doubly connected, π_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 loops, *i.e.* to maps of S^1 into a manifold $\mathcal{V}$ (a group G here), may be extended to maps of a sphere S^n into $\mathcal{V}$. 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 $\pi_n(\mathcal{V})$. For example $\pi_n(S^n) = \mathbb{Z}$. See [DNF] for more details and the determination of these groups π_n. This notion is important in physics to describe topological defects, solitons, instantons, monopoles, etc. See Fig. 1.3 for *vortex* and *anti-vortex* configurations of unit vectors in 2 dimensions, of respective winding number (or *vorticity*) ± 1.

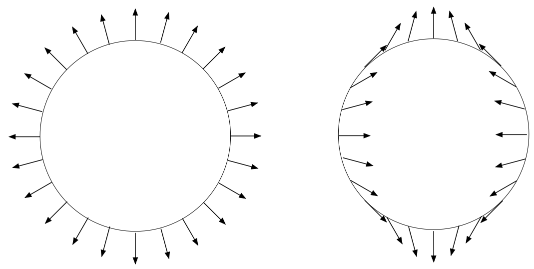


Figure 1.3: Two configurations of unit vectors realizing homotopically non trivial mapping $S^1 \to S^1$. Those are respectively the *vortex* and *anti-vortex* of the XY model of statistical mechanics, see for example `http://www.ibiblio.org/e-notes/Perc/xy.htm` for more details and nice figures.

1.2.3 Compact and non compact groups

If the domain $\mathcal{D}$ 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 space 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 $\mathcal{D}$ of $\mathbb{R}^d$, being compact is equivalent to being closed and bounded.

Examples. The unitary groups $\mathrm{U}(n)$ and their subgroups $\mathrm{SU}(n)$, $\mathrm{O}(n)$, $\mathrm{SO}(n)$, $\mathrm{USp}(n/2)$ (n even), are compact. The groups $\mathrm{SL}(n,\mathbb{R})$ or $\mathrm{SL}(n,\mathbb{C})$, $\mathrm{Sp}(n,\mathbb{R})$ or $\mathrm{Sp}(n,\mathbb{C})$, 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

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

(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, and invariance under inversion:

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

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

Conversely if the group is non compact, left and right measures may still exist, they may even coincide, (for 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 $d\mu_L(g) = y^{-2}dxdy\,, \quad d\mu_R(g) = y^{-1}dxdy$ 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 $G \times G \to G$ and $G \to G$ 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. There exist, however, infinite dimensional Lie groups.

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 $(\xi^1, \xi^2, \cdots \xi^d) \in \mathcal{D} \subset \mathbb{R}^d$, and in the group operations $g(\xi'') = g(\xi').g(\xi)$, and $g(\xi)^{-1} = g(\xi'')$, the ξ''^i are continuous (in fact analytic) functions of the ξ^j (and ξ'^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 $\mathcal{D}_j$, 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 $\mathrm{U}(n)$ is n^2, that of $\mathrm{SU}(n)$ is n^2-1, that $\mathrm{O}(n)$ or $\mathrm{SO}(n)$ is $n(n-1)/2$. What is the dimension of $\mathrm{Sp}(2n, \mathbb{R})$? of the Galilean group in $\mathbb{R}^3$? of the Lorentz and Poincaré groups ?

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

[3] For 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 !

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

$$(\lambda_1 X_1 + \lambda_2 X_2) * Y \quad = \lambda_1 X_1 * Y + \lambda_2 X_2 * Y \tag{1.13}$$

$$X * (\mu_1 Y_1 + \mu_2 Y_2) \quad = \mu_1 X * Y_1 + \mu_2 X * Y_2 \ . \tag{1.14}$$

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

$$[X, Y] = -[Y, X] \tag{1.15}$$

$$[X_1, [X_2, X_3]] + [X_2, [X_3, X_1]] + [X_3, [X_1, X_2]] = 0 \ . \tag{1.16}$$

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, $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u}.\mathbf{w})\mathbf{v} - (\mathbf{u}.\mathbf{v})\mathbf{w}$.

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)

$$g(t_1)g(t_2) = g(t_1 + t_2) \qquad g^{-1}(t) = g(-t) \ . \tag{1.17}$$

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

$$g(t+\delta t) = g(t)g(\delta t) \quad \Leftrightarrow \quad g^{-1}(t)g(t+\delta t) = g(\delta t) \ . \tag{1.18}$$

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

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

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

$$X = \frac{d}{dt} g(t)\Big|_{t=0} \ , \tag{1.19}$$

this is the velocity at $t = 0$ (or at $g = e$) along the curve. Equation (1.18) thus reads

$$g'(t) = g(t)X \ . \tag{1.20}$$

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 ξ^α 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 $G \subset \mathrm{GL}(n, \mathbb{R})$ 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 (1.20) as

$$g(t) = \exp tX = \sum_{n=0} \frac{t^n}{n!} X^n \ , \tag{1.21}$$

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 from $\mathfrak{g}$ to G, 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 $\mathrm{GL}(n, \mathbb{R})$, $\mathfrak{g}$ is the algebra of real $n \times n$ matrices, denoted $M(n, \mathbb{R})$. If G is the group of unitary matrices $\mathrm{U}(n)$, $\mathfrak{g}$ is the space of anti-Hermitian $n \times n$ matrices. Moreover they are traceless if $G = \mathrm{SU}(n)$. Likewise, for the orthogonal $\mathrm{O}(n)$ group, $\mathfrak{g}$ is made of skew-symmetric, hence traceless, matrices.
For the symplectic group $G = \mathrm{USp}(n)$, $\mathfrak{g}$ is generated by "anti-selfdual" quaternionic matrices, see Appendix A.
For each of these cases, check that the characteristic property (anti-Hermitian, skew-symmetric, traceless, ...) is preserved by the commutator, thus making $\mathfrak{g}$ a Lie algebra.

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 =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 SU(2) which is simply connected, if $\mathbf{n}$ is a unit vector, $e^{i\pi \mathbf{n}.\boldsymbol{\sigma}} = -I$, hence all elements $\pi i \mathbf{n}.\boldsymbol{\sigma}$ de $\mathfrak{g}$ =su(2) have the same image!

⋆ Example of a non-compact group for which the exp map is non surjective: G=SL(2,$\mathbb{R}$), for which $\mathfrak{g}$=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 $\mathrm{tr}\ A^{2n+1} = 0$, $\mathrm{tr}\ A^{2n} = 2(-\det A)^n$, hence $\mathrm{tr}\ e^A = 2\cosh\sqrt{-\det A} \ge -2$. There exist in G, however, matrices of trace < -2, for instance $\mathrm{diag}\,(-2, -\frac{1}{2})$.

⋆ 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 $\det e^X = e^{\mathrm{tr}\, X}$, 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

$$\begin{aligned} e^{tX}e^{uY}e^{-tX}e^{-uY} &= (I + tX + \tfrac{1}{2}t^2X^2)(I + uY + \tfrac{1}{2}u^2Y^2)(I - tX + \tfrac{1}{2}t^2X^2)(I - uY + \tfrac{1}{2}u^2Y^2) \\ &= I + (XY - YX)tu + O(t^3) \ . \end{aligned} \tag{1.22}$$

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, $XY - YX$, *i.e.* the Lie bracket of matrices X and Y. In general, for an arbitrary Lie group, the bracket is defined by

$$e^{tX}e^{uY}e^{-tX}e^{-uY} = e^Z \quad , \quad Z = tu[X,Y] + O(t^3) \tag{1.23}$$

and one proves that this bracket has the properties (1.15) 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 $\mathrm{ad}\, X$ be the linear operator in the Lie algebra defined by

$$Y \mapsto (\mathrm{ad}\, X)Y := [X,Y] \ , \tag{1.24}$$

hence

$$(\mathrm{ad}\,^p X)Y = [X, [X, \cdots [X, Y] \cdots]]$$

with p brackets (commutators).

Given two elements X and Y in $\mathfrak{g}$, and e^X and e^Y the elements they generate in G, does there exist a $Z \in \mathfrak{g}$ such that $e^X e^Y = e^Z$? 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 *Baker-Campbell-Hausdorff formula*, that we admit, gives an explicit expression of Z.

$$\begin{aligned} e^X e^Y &= e^Z \\ Z &= X + \int_0^1 dt \psi(\exp \mathrm{ad}\, X \ \exp t\, \mathrm{ad}\, Y) Y \end{aligned} \tag{1.25}$$

where $\psi(.)$ is the function

$$\psi(u) = \frac{u \ln u}{u-1} = 1 + \frac{1}{2}(u-1) - \frac{1}{6}(u-1)^2 + \cdots , \tag{1.26}$$

which is regular at $u = 1$. The first terms in the expansion in powers of X and Y read explicitly

$$Z = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}\Big([X, [X, Y]] + [Y, [Y, X]]\Big) + \cdots \tag{1.27}$$

This complicated formula has some useful particular cases. Hence if X et Y commute with $[X, Y]$, (1.25) simplifies to

$$e^X e^Y = e^{X+Y+\frac{1}{2}[X,Y]} = e^{X+Y} e^{\frac{1}{2}[X,Y]} , \tag{1.28}$$

a formula that one may prove directly using the general identity

$$e^X Y e^{-X} = \sum_0^\infty \frac{1}{n!} \mathrm{ad}\,^n X\, Y \tag{1.29}$$

(which is nothing else than the Taylor expansion at $t = 0$ of $e^{tX} Y e^{-tX}$ evaluated at $t = 1$), and writing and solving the differential equation satisfied by $f(t) = e^{tX} e^{tY}$, $f(0) = 1$

$$f'(t) = (X + e^{tX} Y e^{-tX}) f(t) \tag{1.30}$$

$$= (X + Y + t[X, Y]) f(t) . \tag{1.31}$$

On the other hand, to first order in Y, one may replace the argument of ψ in (1.25) by $\exp \mathrm{ad}\, X$ and then

$$Z = X + \sum_{n=0}^{\infty} \frac{B_n}{n!} (-1)^n \,(\mathrm{ad}\, X)^n\ Y + \mathrm{O}(Y^2) \tag{1.32}$$

where the B_n are the Bernoulli numbers: $\frac{t}{e^t - 1} = \sum_0 B_n \frac{t^n}{n!}$, $B_0 = 1, B_2 = \frac{1}{6}, B_4 = -\frac{1}{30}$ and, beside $B_1 = -\frac{1}{2}$, all B of odd index vanish. Still to first order in Y, one has also

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

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 et 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 $\mathrm{GL}(d, \mathbb{R})$. 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. 0, 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 $x \to x' = Ox$, 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 $\delta x^i = \omega^i{}_j x^j$ may also be written

$$\delta x^i = \omega^i{}_j x^j = -\tfrac{1}{2}\omega^{kl} J_{kl} x^i \tag{1.33}$$

$$J_{kl} = x^k \partial_l - x^l \partial_k \qquad : \qquad J_{kl} x^i = x^k \delta_{il} - x^l \delta_{ik} \tag{1.34}$$

(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 so(n) algebra as differential operators. It is then a simple matter to compute the commutation relations[4]

$$[J_{ij}, J_{kl}] = \delta_{il} J_{jk} - \delta_{ik} J_{jl} - \delta_{jl} J_{ik} + \delta_{jk} J_{il} \ . \tag{1.35}$$

(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.)

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 $1 \le i < j \le n$, 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 (1.35).

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 $g = \mathrm{diag}\,((+1)^p, (-1)^q)$.

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

[4]Note that wrt to the calculation carried out in the O(3,1) group in Chap. 0, § 0.6.2, we have changed our conventions and use here anti-Hermitian generators.

of the circle".) This is obviously a group and an infinite dimensional manifold, which manifests itself in the algebra of infinitesimal diffeomorphisms $z \mapsto z' = z + \epsilon(z)$, generated by differential operators ℓ_n

$$\ell_n = -z^{n+1}\frac{\partial}{\partial z} \ , \qquad n \in \mathbb{Z} \tag{1.36}$$

which satisfy

$$[\ell_n, \ell_m] = (n-m)\ell_{n+m} \ . \tag{1.37}$$

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

$$[L_n, L_m] = (n-m)L_{n+m} + \frac{c}{12}n(n^2-1)\delta_{n,-m} \qquad [c, L_n] = 0 \ . \tag{1.38}$$

(The L_n may be thought of as quantum realizations of the operators ℓ_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 (1.37) 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 $\mathfrak{g}$ and G

Let us examine how properties of G translate in $\mathfrak{g}$.

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 $\mathfrak{g}$ is a subspace $\mathfrak{I}$ of $\mathfrak{g}$ which is stable under multiplication (defined by the Lie bracket) by any element of $\mathfrak{g}$, *i.e.* such that $[\mathfrak{I}, \mathfrak{g}] \subset \mathfrak{I}$. The ideal is called *abelian* si $[\mathfrak{I}, \mathfrak{I}] = \{0\}$.

A Lie algebra $\mathfrak{g}$ is *simple* if $\mathfrak{g}$ has no other ideal than $\{0\}$. It is *semi-simple* if $\mathfrak{g}$ has no other abelian ideal than $\{0\}$.

Example. Consider the Lie algebra of SO(4), denoted so(4), see the formulae given in (1.35) for so(n). It is easy to check that the combinations

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

commute with

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

and that

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

where one sees two commuting copies of so(3). One writes so(4)=so(3)$\oplus$ 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. 0, § 0.6.2. 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

$$G \text{ simple} \Longrightarrow \mathfrak{g} \text{ simple}$$
$$G \text{ semi-simple} \Longrightarrow \mathfrak{g} \text{ semi-simple}$$

but the converse is not true ! Several different Lie groups may have the same Lie algebra, *e.g.* SO(3) which is simple, and SU(2) which is not semi-simple, as seen above in §1.1.7. [5]

1.4.2 Compacity. Complexification

A semi-simple Lie algebra is said to be compact if it is the Lie algebra of a compact Lie group. 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.

At this stage one should 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

$$\begin{array}{ll}
\mathrm{o}(3) & \left\{\begin{array}{ll} X_1 = z\partial_y - y\partial_z & \\ X_2 = x\partial_z - z\partial_x & [X_1,X_2] = y\partial_x - x\partial_y = X_3 \text{ etc} \\ X_3 = y\partial_x - x\partial_y & \end{array}\right. \\
\mathrm{o}(2,1) & \left\{\begin{array}{ll} \widetilde{X}_1 = z\partial_y + y\partial_z & [\widetilde{X}_1,\widetilde{X}_2] = y\partial_x - x\partial_y = \widetilde{X}_3 \\ \widetilde{X}_2 = x\partial_z + z\partial_x & [\widetilde{X}_2,\widetilde{X}_3] = -z\partial_y - y\partial_z = -\widetilde{X}_1 \\ \widetilde{X}_3 = y\partial_x - x\partial_y & [\widetilde{X}_3,\widetilde{X}_1] = -x\partial_z - z\partial_x = -\widetilde{X}_2 \end{array}\right.
\end{array} \tag{1.39}$$

that non isomorphic on the real numbers, but $i\widetilde{X}_1$, $i\widetilde{X}_2$ et $-\widetilde{X}_3$ verify the o(3) algebra. The algebras o(3) and o(2,1) are said to have the same complexified form $\mathfrak{g}_c$, or else, to be two real forms of $\mathfrak{g}_c$, 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 and TD).

The algebras so(4) and so(3,1) studied above and in Chap. 0 provide another example of two algebras, which are two non-isomorphic real forms of the same complexified form. Another example is provided by sp$(2n,\mathbb{R})$ et usp(n). (See Appendix 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.

[5] Beware! Some authors call "simple" any Lie group, whose Lie algebra is simple. This amounts to making a distinction between the concepts 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 of Lie type...

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: $\mathfrak{g} = \mathfrak{g}'$.
– If G is non simply connected, let $\widetilde{G}$ be its universal covering group. G et $\widetilde{G}$ being locally isomorphic, they have the same Lie algebra. Examples: U(1) and $\mathbb{R}$; SO(3) and SU(2); SO(3,1) and $\mathrm{SL}(2,\mathbb{C})$.

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, to 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

$$[t_\alpha, t_\beta] = C_{\alpha\beta}{}^\gamma\, t_\gamma \ , \tag{1.40}$$

are clearly antisymmetric in their two lower indices $C_{\alpha\beta}{}^\gamma = -C_{\beta\alpha}{}^\gamma$. Return to the linear operator $\mathrm{ad}\, X$ defined above in (1.24)

$$\mathrm{ad}\, X\, Z = [X, Z] = \sum x^\alpha z^\beta C_{\alpha\beta}{}^\gamma t_\gamma \ ,$$

and for $X, Y \in \mathfrak{g}$ consider the linear operator $\mathrm{ad}\, X\, \mathrm{ad}\, Y$ which acts in the Lie algebra according to

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

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

$$\sum_\delta \left(C_{\alpha\delta}{}^\epsilon C_{\beta\gamma}{}^\delta + C_{\beta\delta}{}^\epsilon C_{\gamma\alpha}{}^\delta + C_{\gamma\delta}{}^\epsilon C_{\alpha\beta}{}^\delta \right) = 0 \tag{1.41}$$

(note the structure : a cyclic permutation on the three indices α, β, γ with ϵ fixed and summation over the repeated δ) ; and show that this identity may also be expressed as

$$[\mathrm{ad}\, X, \mathrm{ad}\, Y] Z = \mathrm{ad}\, [X, Y] Z \ . \tag{1.42}$$

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

$$(X, Y) := \mathrm{tr}\,(\mathrm{ad}\, X \mathrm{ad}\, Y) = \sum_{\gamma,\delta} C_{\alpha\delta}{}^\gamma C_{\beta\gamma}{}^\delta x^\alpha y^\beta =: g_{\alpha\beta} x^\alpha y^\beta \ , \tag{1.43}$$

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

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

(Symmetry in α, β 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 $\mathrm{ad}\, Z$:

$$\forall X, Y, Z \in \mathfrak{g} \quad ([Z, X], Y) + (X, [Z, Y]) = 0 \tag{1.44}$$

(think of $\mathrm{ad}\, Z$ as an infinitesimal generator acting like a derivative, either on the first term, or on the second). Indeed the first term equals $\mathrm{tr}\,(\mathrm{ad}\, Z \mathrm{ad}\, X \mathrm{ad}\, Y - \mathrm{ad}\, X \mathrm{ad}\, Z \mathrm{ad}\, Y)$ while the second is $\mathrm{tr}\,(\mathrm{ad}\, X \mathrm{ad}\, Z \mathrm{ad}\, Y - \mathrm{ad}\, X \mathrm{ad}\, Y \mathrm{ad}\, Z)$, 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 $C_{\alpha\beta}{}^{\gamma}$, thus defining

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

Let us then show that this $C_{\alpha\beta\gamma}$ is completely antisymmetric in α, β, γ. Given the already known antisymmetry in α, β, it suffices to show that $C_{\alpha\beta\gamma}$ is invariant by cyclic permutations. This follows from (1.44) which may be written in a more symmetric form as

$$(X, [Y, Z]) = (Y, [Z, X]) = (Z, [X, Y]) = C_{\alpha\beta\gamma} x^\alpha y^\beta z^\gamma = C_{\beta\gamma\alpha} y^\beta z^\gamma x^\alpha = C_{\gamma\alpha\beta} z^\gamma x^\alpha y^\beta\,, \tag{1.45}$$

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 Lie 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 $\mathfrak{g}$ is not semi-simple and let us show that $\det g = 0$. Let $\mathfrak{I}$ be an ideal of $\mathfrak{g}$, choose a basis of $\mathfrak{g}$ made of a basis of $\mathfrak{I}$, $\{t_i\}$, $i = 1, \cdots r$, complemented by t_a, $a = r+1, \cdots d$. For $1 \leq i, j \leq r$, compute $g_{ij} = \sum_{\alpha\beta} C_{i\alpha}{}^{\beta} C_{j\beta}{}^{\alpha}$. By definition of an ideal, α and β are themselves between 1 and r, $g_{ij} = \sum_{1\leq k,l \leq r} C_{ik}{}^{l} C_{jl}{}^{k}$. Hence the restriction of the Killing form of $\mathfrak{g}$ to $\mathfrak{I}$ is the Killing form of $\mathfrak{I}$. If moreover the ideal is assumed to be abelian, $g_{ij} = 0$ and $g_{ia} = 0$ (Exercise: check that point!). The form is obviously degenerate ($\det g = 0$). The reciprocal, $\det g = 0 \Rightarrow \mathfrak{g}$ 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 $\langle X, Y \rangle = \sum x^\alpha y^\beta$. For a compact group G, one can make this form invariant by averaging over G : $\varphi(X, Y) := \int \mathrm{d}\mu(g) \langle gXg^{-1}, gYg^{-1} \rangle$. It is invariant $\varphi(gXg^{-1}, gYg^{-1}) = \varphi(X, Y)$, or in infinitesimal form, $\varphi([Z, X], Y] + \varphi(X, [Z, Y]) = 0$, (cf (1.44)). It is also positive definite. Let e_α be a basis which diagonalises it, $\varphi(e_\alpha, e_\beta) = \delta_{\alpha\beta}$. Let us calculate in that basis the matrix of the $\mathrm{ad}\, X$ operator and show that it is antisymmetric, $(\mathrm{ad}\, X)_{\alpha\beta} = -(\mathrm{ad}\, X)_{\beta\alpha}$:

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

Hence the Killing form

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

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

Example. The case of SO(3) or 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* $\mathfrak{g}$ *is a direct sum of simple Lie algebras* $\mathfrak{g}_i$

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

This is a simple consequence of (1.45). Consider a semi-simple algebra $\mathfrak{g}$ with an ideal $\mathfrak{I}$ and call $\mathfrak{C}$ the complement of $\mathfrak{I}$ wrt the Killing form, *i.e.* $(\mathfrak{I},\mathfrak{C}) = 0$. By (1.45), $([\mathfrak{C},\mathfrak{I}],\mathfrak{I}) = (\mathfrak{C},[\mathfrak{I},\mathfrak{I}]) = (\mathfrak{C},\mathfrak{I}) = 0$ (since $\mathfrak{I}$ is a subalgebra), and $([\mathfrak{C},\mathfrak{I}],\mathfrak{C}) = (\mathfrak{I},\mathfrak{C}) = 0$ (since $\mathfrak{I}$ is an ideal), hence $[\mathfrak{C},\mathfrak{I}]$, orthogonal to any element of $\mathfrak{g}$ for the non-degenerate Killing form, vanishes, $[\mathfrak{C},\mathfrak{I}] = 0$, which means that $\mathfrak{g} = \mathfrak{I} \oplus \mathfrak{C}$. Iterating the argument on $\mathfrak{C}$, 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 $\mathfrak{g}$, hence with an invertible Killing form, and a basis $\{t_\alpha\}$ of $\mathfrak{g}$, we define

$$C_2 = \sum_{\alpha,\beta} g^{\alpha\beta} t_\alpha t_\beta \tag{1.46}$$

where $g^{\alpha\beta}$ is the inverse of $g_{\alpha\beta}$, *i.e.* $g_{\alpha\gamma}g^{\gamma\beta} = \delta_\alpha^\beta$.

Formally, this combinaison of the t's, which does not make use of the bracket, does not live in the 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_γ hence with any element of $\mathfrak{g}$. This is the quadratic *Casimir operator*.

$$\begin{aligned}
[C_2, t_\gamma] &= \sum_{\alpha,\beta} g^{\alpha\beta}[t_\alpha t_\beta, t_\gamma] \\
&= \sum_{\alpha,\beta} g^{\alpha\beta}\left(t_\alpha[t_\beta,t_\gamma] + [t_\alpha,t_\gamma]t_\beta\right) \\
&= \sum_{\alpha,\beta,\delta} g^{\alpha\beta} C_{\beta\gamma}{}^{\delta}(t_\alpha t_\delta + t_\delta t_\alpha) \\
&= \sum_{\alpha,\beta,\delta,\kappa} g^{\alpha\beta} g^{\delta\kappa} C_{\beta\gamma\kappa}(t_\alpha t_\delta + t_\delta t_\alpha) \ .
\end{aligned} \tag{1.47}$$

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

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) $\mathbf{J}^2$, 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 $\mathbf{J}^2$ and $\mathbf{K}^2$ in the (complexified) so(3,1)$\simeq$ su(2)$\oplus$su(2) algebra of the Lorentz group (see Chap. 0 § 0.6.2); or P^2 and W^2 in the Poincaré algebra, see Chap. 0, § 0.6.5.

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

$$C_r = g^{\alpha_1\alpha'_1} g^{\alpha_2\alpha'_2} \cdots g^{\alpha_r\alpha'_r} C_{\alpha_1\beta_1}{}^{\beta_2} C_{\alpha_2\beta_2}{}^{\beta_3} \cdots C_{\alpha_r\beta_r}{}^{\beta_1} t_{\alpha'_1} t_{\alpha'_2} \cdots t_{\alpha'_r} \tag{1.48}$$

has a vanishing bracket with any t_γ. What is that C_3 in su(2) ? See Bourbaki ([Bo], chap. 3.52) for a discussion of these general Casimir operators. See also exercice 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. 0, § 0.2.3 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

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

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

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

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

[DNF] Dubrovin, B. A., Fomenko, A. T., Novikov, S. P. *Modern geometry—methods and applications.* Part I. The geometry of surfaces, transformation groups, and fields. Graduate Texts in Mathematics, 93. Springer-Verlag, New York, 1992. Part II. The geometry and topology of manifolds. Graduate Texts in Mathematics, 104. Springer-Verlag, New York, 1985. (a third volume deals with homology…)

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

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

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

[K-S] Y. Kosmann-Schwarzbach, *Groups and symmetries, From Finite Groups to Lie Groups*, Springer 2010.

Group theory for physicists

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

[Co] J.F. Cornwell, *Group theory in physics. An introduction*, Academic Pr. contains much information but sometimes uses a terminology different from the rest of the literature…

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

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

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

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

See also lecture notes of group theory by and for physicists (in French), avalaible on the CCSD server `http://cel.ccsd.cnrs.fr/`, for example
J.-B. Z., *Introduction à la théorie des groupes et de leurs representations*, (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, $i = 1, 2, 3$,

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

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 in terms of Pauli matrices : $e_i \mapsto -i\sigma_i$.

The conjugate of q is the quaternion

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

not to be confused with its complex conjugate

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

Note that $q\bar{q} := |q|^2 = (q^{(0)})^2 + (q^{(1)})^2 + (q^{(2)})^2 + (q^{(3)})^2$, the square norm of the quaternion, and hence $q^{-1} = \bar{q}/|q|^2$ if this norm is non-vanishing.

One may also define the Hermitian conjugate of q as

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

(in accordance with the fact that Pauli matrices are Hermitian).

Note that conjugation and Hermitian conjugation reverse the order of factors

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

A *real quaternion* is a quaternion of the form (A.1) with $q^{(\mu)} \in \mathbb{R}$, 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 $\mathbb{H}$ (from Hamilton).

A.2 Quaternionic matrices

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

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

The *dual* Q^R of a quaternionic matrix Q is the matrix

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

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

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

it is *real quaternionic* if

$$Q^R = Q^\dagger \quad \text{hence} \quad q_{ij} = q_{ij}^* \ , \tag{A.9}$$

i.e. if its elements are real quaternions.

A.3 Symplectic groups $\mathrm{Sp}(2n,\mathbb{R})$ and $\mathrm{USp}(n)$, and the Lie algebras $\mathrm{sp}(2n)$ et $\mathrm{usp}(n)$

Consider the $2n \times 2n$ matrix

$$S = \begin{pmatrix} 0 & \mathbf{1}_N \\ -\mathbf{1}_N & 0 \end{pmatrix} \tag{A.10}$$

and the associated "skew-symmetric" bilinear form

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

The symplectic group $\mathrm{Sp}(2n,\mathbb{R})$ is the group of real $2n \times 2n$ matrices that preserve that form

$$B^T S B = S \ . \tag{A.12}$$

In the basis where $X^T = (x_1, x_{n+1}, x_2, x_{n+2}, \cdots)$, the matrix $S = \mathrm{diag}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \mathrm{diag}\,(-e_2)$ 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 $q_{ij}^{(\alpha)}$ are real for $\alpha = 0,2$ and purely imaginary for $\alpha = 1,3$. This group is non compact. Its Lie algebra $\mathrm{sp}(2n,\mathbb{R})$ 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 $n = 1$, Sp(2,$\mathbb{R}$)=SL(2,$\mathbb{R}$).

A related group is $\mathrm{USp}(n)$, generated by unitary real quaternionic $n \times n$ matrices $Q^R = Q^\dagger = Q^{-1}$. 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 $\mathrm{U}(2n)$. Its Lie algebra $\mathrm{usp}(n)$ is generated by antiselfdual real quaternionic matrices $A = -A^R = -A^\dagger$ (check !). Its dimension is again $n(2n+1)$. For $n = 1$, USp(1)=SU(2).

Expressing the condition on matrices A of $\mathrm{sp}(n,\mathbb{R})$ in terms of quaternions, one sees that the two algebras $\mathrm{sp}(2n,\mathbb{R})$ and $\mathrm{usp}(n)$ have the same complexified algebra, namely $\mathrm{sp}(2n,\mathbb{C})$. Only $\mathrm{usp}(n)$ 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 $\mathcal{V}(x)$ 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 $\mathcal{B}(x)$ *of a point* x : subset of $\mathcal{V}(x)$ such that any $V \in \mathcal{V}(x)$ contains a $W \in \mathcal{B}(x)$. (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.

Consequences :

– any infinite sequence of points in E has an accumulation point in E;

– if E is compact and $f : E \mapsto F$ is continuous, $f(E)$ is compact;

– any continuous real function on a compact space E is bounded.

If E is a subspace of $\mathbb{R}^n$, E compact $\Leftrightarrow$ 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 ——○. More precisely, there exists a collection 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 $f_i(m) = (x^1, x^2, \dots x^n)$ its image in $\mathbb{R}^n$: $(x^1, x^2, \dots x^n)$ 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 fo r 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 $f_j(U_i \cap U_j) \subset \mathbb{R}^n$ 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.

A *Riemann manifold* is a differentiable real manifold on the tangent vectors of which a positive definite inner product has been defined. If the inner product is only assumed to be a non degenerate form of signature $(+1)^p, (-1)^{n-p})$, the manifold is said to be *pseudo-Riemannian.* In local coordinates x^i, we have $X = X^i \frac{\partial}{\partial x^i}$, and the inner product and the squared length element are given by the metric tensor g

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

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,

$$X : \qquad f(x) \mapsto \sum_i X^i \left. \frac{\partial}{\partial x^i} \right|_{x_0} f(x)$$

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

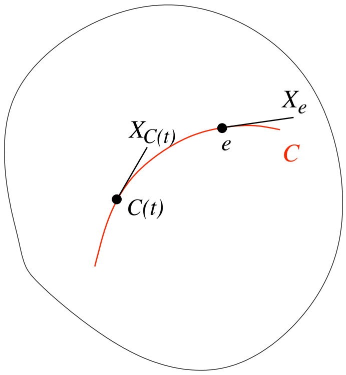


Figure 1.4: The field of tangent vectors to the curve $C(t)$ is a left-invariant vector field

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

$$f \mapsto \left.\frac{\mathrm{d}f(C(t))}{\mathrm{d}t}\right|_{t=0} ,$$

which defines the tangent vector to the curve C at point x_0, also called velocity vector and denoted $C'(t)|_{t=0} = C'(0)$.

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 $\frac{\partial}{\partial x^i}\big|_{x_0}$: 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.

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 invariant vector fields are thus isomorphic, and are both denoted $\mathfrak{g}$.

Conversely, given a tangent vector X_e at e, let

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

be the unique solution to the differential equation

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

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.2) satisfies property (1.17). Note that $C(t)$ satisfies (B.3), 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 (1.21).

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 (0.0.10) (*i.e.* $U = u_0 I - \mathbf{u}.\boldsymbol{\sigma}$, $u \in S^3$ etc) ; if the condition $u_0^2 + \mathbf{u}^2 = 1$ is momentarily relaxed (but $v_0^2 + \mathbf{v}^2 = 1$ maintained), this defines a linear transformation $u \to u'$ which conserves the norm $\det U = u_0^2 + \mathbf{u}^2 = u_0'^2 + \mathbf{u}'^2 = \det U'$. This is thus a rotation of the space $\mathbb{R}^4$ which preserves the natural measure $\mathrm{d}^4 u\, \delta(u^2 - 1)$ 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: $d\mu(U) = d\mu(V.U)$. It is also invariant under $U \to U^{-1}$, since inversion in SU(2) amounts to the restriction to S^3 of the orthogonal transformation $u_0 \to u_0$, $\mathbf{u} \to -\mathbf{u}$ 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 $\mathbf{n}$ (or its two polar angles θ et ϕ) and the rotation angle ψ, one finds

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

normalized for SU(2) to

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

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

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

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

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

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 $ds^2 = \frac{1}{2} \mathrm{tr}\, dU dU^\dagger$, it is invariant by $U \to UV$, $U \to VU$ ou $U \to U^{-1}$, and an invariant integration measure then follows (cf Chap. 0, App. 0). With the parametrization $(\mathbf{n} = (\theta, \phi), \psi)$, one finds

$$ds^2 = \frac{1}{2} \mathrm{tr}\ dU dU^\dagger = \left(d\frac{\psi}{2} \right)^2 + \sin^2 \frac{\psi}{2} \left(d\theta^2 + \sin^2\theta\, d\phi^2 \right) \ , \tag{C.4}$$

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

$$U = e^{-i\alpha \frac{\sigma_3}{2}} e^{-i\beta \frac{\sigma_2}{2}} e^{-i\gamma \frac{\sigma_3}{2}} \tag{C.5}$$

whence

$$ds^2 = \frac{1}{2}\mathrm{tr}\, dUdU^\dagger = \frac{1}{4}\left(d\alpha^2 + 2d\alpha d\gamma\cos\beta + d\gamma^2 + d\beta^2\right) \tag{C.6}$$

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 $U \in \mathrm{U}(n)$ may be diagonalized in the form

$$U = V\Lambda V^\dagger\ , \tag{C.7}$$

with $\Lambda = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_n)$ and the λ_i are in fact of modulus one, $\lambda_j = e^{i\alpha_j}$. These λ_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 Λ. If the latter is generic, with distinct eigenvalues λ_i, V lives in $\mathrm{U}(n)/\mathrm{U}(1)^n$. The natural metric, invariant under $U \mapsto U'U$ or $\mapsto UU'$, reads $\mathrm{tr}\,(\mathrm{d}U\mathrm{d}U^\dagger)$. But $\mathrm{d}U = V(\mathrm{d}\Lambda + [\mathrm{d}X, \Lambda])V^\dagger$, where $\mathrm{d}X := V^\dagger \mathrm{d}V$ is anti-Hermitian (and with no diagonal elements, why?). Thus $\mathrm{tr}\,(\mathrm{d}U\mathrm{d}U^\dagger) = \sum_i |\mathrm{d}\alpha_i|^2 + 2\sum_{i<j} |\mathrm{d}X_{ij}|^2|\lambda_i - \lambda_j|^2$ 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

$$d\mu(U) = \sqrt{\det g}\prod \mathrm{d}\xi^\alpha = \mathrm{const.}\ |\Delta(e^{i\alpha})|^2 \prod \mathrm{d}\alpha_i d\mu(V)\ . \tag{C.8}$$

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

$$\Delta(\lambda) := \prod_{i<j}(\lambda_i - \lambda_j) = \begin{vmatrix} \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \\ \vdots & & & \vdots \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ 1 & 1 & \cdots & 1 \end{vmatrix}\ . \tag{C.9}$$

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

$$d\mu(U) = \mathrm{const.} \prod_{i<j} \sin^2\left(\frac{\alpha_i - \alpha_j}{2}\right) \prod \mathrm{d}\alpha_i \times \text{angular part}\ . \tag{C.10}$$

Note that this radial part of the measure suffices if one has to integrate over the group a function of U which is invariant by $U \to VUV^\dagger$, $V \in$U(n). For example $\int d\mu(U)\,\mathrm{tr}\,P(U)$, with P a polynomial.

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 β 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 $\mathrm{O}(n)$ 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 $G = \mathrm{O}(3)$ 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:

$$S(x) = \{g \in G|\beta(g)x = x\} \ . \tag{1.49}$$

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 $\mathrm{SO}(n)$? 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 $\mathrm{GL}(E)$ 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 $C^{(r)} = \sum_{1\leq i_1, i_2, \cdots i_r \leq n} t_{(i_1 i_2)} t_{(i_2 i_3)} \cdots t_{(i_r i_1)}$ 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
$$\delta S = \int d^d x \, (\partial_\mu a_\nu) \, \Theta^{\mu\nu}(x) \ ; \tag{1.50}$$
- $\Theta_{\mu\nu}$ is conserved: $\partial_\mu \Theta^{\mu\nu}(x) = 0$;
- we assume that $\Theta_{\mu\nu}$ is symmetric in μ , ν.

Prove that if Θ 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 manifold 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

$$ds^2 = g_{\mu\nu}(x) dx^\mu dx^\nu \to ds'^2 = g_{\mu\nu}(x') dx'^\mu dx'^\nu = \alpha(x) ds^2 \tag{1.51}$$

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^μ 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

$$\partial_\mu a_\nu + \partial_\nu a_\mu = \frac{2}{d} g_{\mu\nu} \partial_\rho a^\rho \ . \tag{1.52}$$

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^ν , prove that
$$\partial^2 a_\mu = (2 - d)\partial_\mu D. \tag{1.53}$$

- b) Differentiating (1.53) w.r.t. x^μ, 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^ν, symmetrizing it in μ and ν 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^σ and antisymmetrizing it in ν and σ, prove that

$$\partial_\mu(\partial_\sigma a_\nu - \partial_\nu a_\sigma) = 2(g_{\mu\nu}k_\sigma - g_{\mu\sigma}k_\nu) = \partial_\mu(2k_\sigma x_\nu - 2k_\nu x_\sigma). \tag{1.54}$$

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

$$\partial_\sigma a_\nu - \partial_\nu a_\sigma = (2k_\sigma x_\nu - 2k_\nu x_\sigma) + 2l_{\sigma\nu}\,, \tag{1.55}$$

  which, together with (1.52), gives

$$\partial_\sigma a_\nu = x_\nu k_\sigma - x_\sigma k_\nu + l_{\sigma\nu} + g_{\nu\sigma}k_\rho x^\rho + h g_{\nu\sigma}\ .$$

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

$$a_\nu = k_\sigma x^\sigma x_\nu - \frac{1}{2}x_\sigma x^\sigma k_\nu + l_{\sigma\nu}x^\sigma + h x_\nu + c_\nu \tag{1.56}$$

  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 $x^\mu \to x^\mu/x^2$, 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_μ of translations, $J_{\mu\nu}$ of rotations, D of dilatations and K_μ 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_μ 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 $\mathbf{x}^2 = x_1^2 + \cdots + x_d^2$, on the

[6]This pretty argument is due to Michel Bauer.

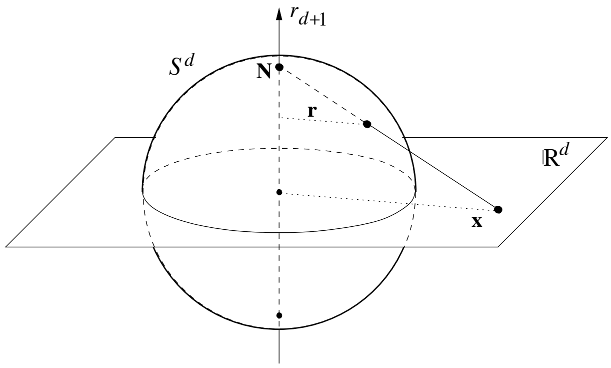


Figure 1.5: Stereographic projection from the North pole

sphere S^d. This sphere is defined by the equation $\mathbf{r}^2 + r_{d+1}^2 = 1$ in the space $\mathbb{R}^{d+1}$, and the mapping is the stereographic projection from the "North pole" $\mathbf{r} = 0,\ r_{d+1} = 1$ (see fig. 1.5). Prove that

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

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

2. The previous sphere is in turn regarded as the section of the light-cone $\mathcal{C}$ in Minkowski space $\mathcal{M}_{d+1,1}$ of equation $z_0^2 - \mathbf{z}^2 - z_{d+1}^2 = 0$ by the hyperplane $z_0 = 1$. Prove that this establishes a one-to-one correspondance 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 $\mathbf{x} \in \mathbb{R}^d$ as a function of $z = (z_0, \mathbf{z}, z_{d+1}) \in \mathcal{C}$ is

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

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 $\mathbf{x}$ in $\mathbb{R}^d$? Show that dilatations of x correspond to "boosts" of rapidity β in the plane (z_0, 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 $(z_0, \mathbf{z}, z_{d+1})$, and such that $\delta\mathbf{z} = \mathbf{a}(z_0 - z_{d+1})$ (to first order in $\mathbf{a}$). To which transformation of $\mathbf{x} \in \mathbb{R}^d$ does it correspond? Compute by exponentiation of T_a the matrix of a finite transformation (Hint: compute the first powers T_a^2, $T_a^3 \ldots$).

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 mentionned 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 frequently encountered 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 $\mathrm{SO}(n)$ acts on the sphre S^{n-1} in a non-linear way, and this is relevant for example in models of ferromagnetism and field theories called non linear σ 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 $\mathrm{GL}(E)$ of linear transformations of E:

$$\begin{array}{rrcl} \forall g \in G & g & \mapsto & D(g) \in \mathrm{GL}(E) \\ \forall g, g' \in G & D(g.g') & = & D(g).D(g') \\ & D(e) & = & I \\ \forall g \in G & D(g^{-1}) & = & (D(g))^{-1} \end{array} \tag{2.1}$$

where I denotes the identity operator in $\mathrm{GL}(E)$. 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 $\in \mathrm{GL}(\mathbb{R})$) 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 $D(g) = D(g') \Leftrightarrow g = g'$. 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, $i = 1, \dots, p$, and associate with any $g \in G$ the representative matrix of $D(g)$, denoted with a curly letter :

$$D(g)e_j = e_i\,\mathcal{D}_{ij}(g) \tag{2.2}$$

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

$$\begin{aligned} D(g.g')e_k &= e_i\,\mathcal{D}_{ik}(g.g') \\ &= D(g)\left(D(g')e_k\right) = D(g)e_j\mathcal{D}_{jk}(g') \\ &= e_i\mathcal{D}_{ij}(g)\mathcal{D}_{jk}(g') \\ \text{hence} \qquad \mathcal{D}_{ik}(g.g') &= \mathcal{D}_{ij}(g)\mathcal{D}_{jk}(g')\ . \end{aligned} \tag{2.3}$$

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

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{2.4}$$

which describe indeed rotations of angle θ around the origin.

The group SU(3) is defined as the set of unitary, unimodular 3×3 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 do 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

$$\forall g \in G \qquad V D(g) = D'(g)V\ . \tag{2.5}$$

Such a V is called an *intertwining operator*, or "intertwiner" in short. If V if 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 operateur $D(g)$:

$$\chi(g) = \operatorname{tr} D(g)\ . \tag{2.6}$$

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

$$\chi(e) = \dim D \; . \tag{2.7}$$

2.1.3 Reducible and irreducible, conjugate, unitary representations...

a. 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 $\dim E_2$

$$\forall g \in G \qquad \mathcal{D}(g) = \begin{pmatrix} \mathcal{D}_1(g) & 0 \\ & \\ 0 & \mathcal{D}_2(g) \end{pmatrix} \; . \tag{2.8}$$

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

$$\mathcal{D}(g) = \begin{pmatrix} \mathcal{D}_1(g) & \mathcal{D}'(g) \\ & \\ 0 & \mathcal{D}_2(g) \end{pmatrix} \; . \tag{2.9}$$

This is the case of representations of the translation group in one dimension. The representation

$$\mathcal{D}(a) = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \tag{2.10}$$

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. (2.9), 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 (2.4) which is irreducible on a space over $\mathbb{R}$ is not over $\mathbb{C}$: it may be rewritten by a (complex) change of basis in the form

$$\begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix} . \tag{2.11}$$

b. 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 (2.3)

$$\mathcal{D}^*_{ik}(g.g') = \mathcal{D}^*_{ij}(g)\mathcal{D}^*_{jk}(g') . \tag{2.12}$$

The representation D is said to be *real* if there exists a basis where $\mathcal{D} = \mathcal{D}^*$. This implies that its character χ is real. Conversely if χ 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

$$\bar{D}(g) = D^{-1\,T}(g) \tag{2.13}$$

or alternatively, $\bar{\mathcal{D}}_{ij}(g) = \mathcal{D}_{ji}(g^{-1})$, which does satisfy (2.3). For a unitary representation, see next paragraph, $\bar{\mathcal{D}}_{ij}(g) = \mathcal{D}^*_{ij}(g)$, and the contragredient representation equals the conjugate. The representations D, D^* and $\bar{D}$ are simultaneously reducible or irreducible.

c. 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 infinite 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 $x, y \in E$

$$\begin{aligned} \langle\, x|y\,\rangle &= \langle\, D(g)x|D(g)y\,\rangle & (2.14)\\ \text{hence} \qquad & D(g)^\dagger D(g) = I & (2.15) \end{aligned}$$

and

$$D(g^{-1}) = D^{-1}(g) = D^\dagger(g) \quad . \tag{2.16}$$

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 $y \in E_2$

$$\langle\, x|D(g)y\,\rangle = \langle\, D(g^{-1})x|y\,\rangle = 0 \tag{2.17}$$

which 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

$$Q = \sum_{g' \in G} D^\dagger(g')D(g') \tag{2.18}$$

which satisfies

$$D^\dagger(g)QD(g) = \sum_{g' \in G} D^\dagger(g'.g)D(g'.g) = Q \tag{2.19}$$

where the "rearrangement" of $\sum_{g'}$ by $\sum_{g'.g}$ has been used (see § 1.2.4). The self-adjoint operator Q is positive definite (why?) and may thus be written

$$Q = V^\dagger V \tag{2.20}$$

with V invertible. (For example, by diagonalisation of the operator Q by a unitary operator, $Q = U\Lambda^2 U^\dagger$, with Λ diagonal real, one may construct the "square root" $V = U\Lambda U^\dagger$.) The intertwiner V defines a representation D' equivalent to D and unitary:

$$\begin{aligned} D'(g) &= VD(g)V^{-1} \\ D'^\dagger(g)D'(g) &= V^{\dagger -1}D^\dagger(g)V^\dagger VD(g)V^{-1} \\ &= V^{\dagger -1}D^\dagger(g)QD(g)V^{-1} \overset{(2.19)}{=} V^{\dagger -1}QV^{-1} = I\,. \end{aligned} \tag{2.21}$$

[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 us 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 space 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 (2.5). 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 result is generally false. A counter-example is given by the representation (2.10) which commutes with matrices $V = \begin{pmatrix}0 & b\\ 0 & 0\end{pmatrix}$.

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 λ; $\lambda \neq 0$ 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 $\mathcal{D}(\theta) = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$ provides counterexamples to both propositions: any matrix $\mathcal{D}(\alpha)$ commutes with $\mathcal{D}(\theta)$ 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 one to apply the argument of Corollary 1: in an irreducible representation, they are multiples of the identity. Thus for SU(2), $\mathbf{J}^2 = j(j+1)I$ 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 carrying representations D_1 and D_2 of a group G. The tensor product [3] $E = E_1 \otimes E_2$ 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. (See Chap. 0 for the example of the group SU(2)). On the vector z above

$$D(g)z = \sum_i D_1(g)x^{(i)} \otimes D_2(g)y^{(i)} \,. \tag{2.22}$$

One readily checks that the character of representation D is the product of characters χ_1 et χ_2 de D_1 et D_2

$$\chi(g) = \chi_1(g)\chi_2(g) \tag{2.23}$$

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

$$\dim D = \dim(E_1 \otimes E_2) = \dim E_1 . \dim E_2 = \dim D_1 . \dim D_2 \tag{2.24}$$

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

$$D \otimes D' = \oplus_j D_j \tag{2.25}$$

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

[3] The reader will find in Appendix D 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 (2.25) in a way showing which of these inequivalent representations appear, *and with which multiplicity*

$$D \otimes D' = \oplus_\rho m_\rho D^{(\rho)} \ . \tag{2.26}$$

A more correct expression would be $E \otimes E' = \oplus_\rho F_\rho \otimes E^{(\rho)}$ where F_ρ is a vector space of dimension m_ρ, the "multiplicity space".

The integers m_ρ are all non negative. The equations (2.25) and (2.26) imply simple rules on characters and dimensions

$$\chi_D.\chi_{D'} = \sum_j \chi_j = \sum_\rho m_\rho \chi^{(\rho)} \tag{2.27}$$

$$\dim D.\dim D' = \sum_j \dim D_j = \sum_\rho m_\rho \dim D^{(\rho)} \ . \tag{2.28}$$

Example: the tensor product of two copies of the euclidean space $\mathbb{R}^3$ does not form an irreducible representation of the rotation group 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),[4] 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

$$x_i y_j = \frac{1}{3}\delta_{ij}\vec{x}.\vec{y} + \frac{1}{2}A_{ij} + \frac{1}{2}S_{ij} \ ; \tag{2.29}$$

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

$$D^{(3)} \otimes D^{(3)} = D^{(1)} \oplus D^{(3)} \oplus D^{(5)} \ . \tag{2.30}$$

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" (see Chap. 0)

$$(j) \otimes (j') = \oplus_{j''=|j-j'|}^{j+j'} (j'') \ . \tag{2.31}$$

[4](such a tensor is "dual" to a vector: $A_{ij} = \epsilon_{ijk} z_k$, $z = x \times y$)

By iteration, one finds

$$D^{(3)} \otimes D^{(3)} \otimes D^{(3)} = D^{(1)} \oplus 3D^{(3)} \oplus 2D^{(5)} \oplus D^{(7)} , \tag{2.32}$$

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 irreducible representations like (2.26, 2.30, 2.32), 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 $m_1 = 1, 1, 3$ 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 SO(3). See also Problem II at the end of this chapter.

Clebsch-Gordan coefficients

Formula (2.25) 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 ρ. 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 τ may appear m_τ times, one must introduce an extra index, $i = 1, \cdots, m_\tau$. One writes

$$e^{(\rho)}_\alpha \otimes e^{(\sigma)}_\beta = \sum_{\tau,\gamma,i} C_{\rho,\alpha;\,\sigma,\beta|\tau_i,\gamma}\, e^{(\tau_i)}_\gamma \ . \tag{2.33}$$

or with notations borrowed from Quantum Mechanics

$$|\rho,\alpha;\sigma,\beta\,\rangle \equiv |\rho\alpha\,\rangle|\sigma\beta\,\rangle = \sum_{\tau,\gamma,i} \langle\,\tau_i\gamma|\rho,\alpha;\,\sigma,\beta\,\rangle\ |\tau_i\gamma\,\rangle \ . \tag{2.34}$$

The coefficients $C_{\rho,\alpha;\,\sigma,\beta|\tau_i,\gamma} = \langle\,\tau_i\gamma|\rho,\alpha;\,\sigma,\beta\,\rangle$ are the *Clebsch-Gordan coefficients*. In contrast with the multiplicities m_ρ in (2.26), they have no reason of being integers, as we saw in Chap. 0 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

$$\sum_{\tau,\gamma,i} \langle\,\tau_i\gamma|\rho,\alpha;\,\sigma,\beta\,\rangle\langle\,\tau_i\gamma|\rho,\alpha';\,\sigma,\beta'\,\rangle^* = \delta_{\alpha,\alpha'}\delta_{\beta,\beta'} \tag{2.35}$$

$$\sum_{\alpha,\beta} \langle\,\tau_i\gamma|\rho,\alpha;\,\sigma,\beta\,\rangle\langle\,\tau'_j\gamma'|\rho,\alpha;\,\sigma,\beta\,\rangle^* = \delta_{\tau,\tau'}\delta_{\gamma,\gamma'}\delta_{i,j} \ . \tag{2.36}$$

This enables us to invert relation (2.34) into

$$|\tau_i\gamma\,\rangle = \sum_{\alpha,\beta} \langle\,\tau_i\gamma|\rho,\alpha;\,\sigma,\beta\,\rangle^* |\rho,\alpha;\,\sigma,\beta\,\rangle \tag{2.37}$$

and justifies the notation

$$\langle\,\rho,\alpha;\,\sigma,\beta|\tau_i\gamma\,\rangle=\langle\,\tau_i\gamma|\rho,\alpha;\,\sigma,\beta\,\rangle^* \tag{2.38}$$

$$|\tau_i\gamma\,\rangle=\sum_{\alpha,\beta}\langle\,\rho,\alpha;\,\sigma,\beta|\tau_i\gamma\,\rangle|\rho,\alpha;\,\sigma,\beta\,\rangle\ . \tag{2.39}$$

Finally, applying a group transformation on both sides of (2.34) and using these relations, one decomposes the product of matrices $\mathcal{D}^{(\rho)}$ and $\mathcal{D}^{(\sigma)}$ in a quite explicit way

$$\mathcal{D}^{(\rho)}_{\alpha\alpha'}\mathcal{D}^{(\sigma)}_{\beta\beta'}=\sum_{\tau,\gamma,\gamma',i}\langle\,\tau_i\gamma|\rho,\alpha;\sigma,\beta\,\rangle^*\,\langle\,\tau_i\gamma'|\rho,\alpha';\sigma,\beta'\,\rangle\,\mathcal{D}^{(\tau_i)}_{\gamma\gamma'}\ . \tag{2.40}$$

We shall see below (§ 2.4.4) 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

$$\forall h\in H\qquad D'(h)=D(h)\ . \tag{2.41}$$

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 down to 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 of 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 $X\in\mathfrak{g}\mapsto d(X)\in\mathrm{End}\,E$ which respects linearity and Lie bracket $X,Y\in\mathfrak{g},\ [X,Y]\mapsto d([X,Y])=[d(X),d(Y)]\in\mathrm{End}\,V$. 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 appropriate bases, *the structure constants are the same in all representations.* More precisely, if t_i is a basis of $\mathfrak{g}$, 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. 0, § 0.2.2, an illustration of this *universality property*. In contrast, Casimir operators take different values in different irreducible representations.

In parallel with the definitions of sect. 2.1.1, one defines the notions of faithful representation of a Lie algebra (its kernel ker $d=\{X|d(X)=0\}$ reduces to the element 0 of $\mathfrak{g}$), 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 into a space E gives a map d of the Lie algebra $\mathfrak{g}$ into the algebra of operators on E. It is obtained by taking the infinitesimal form of $D(g)$, with $g(t)=I+tX$ (or $g=e^{tX}$)

$$d(X):=\frac{d}{dt}\Big|_{t=0}D(g(t))\ , \tag{2.42}$$

or, for t infinitesimal,

$$D(e^{tX})=e^{td(X)}\ . \tag{2.43}$$

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 $e^{tX}e^{uY}e^{-tX}e^{-uY}=e^{Z}$ with $Z=ut[X,Y]+\cdots$, whence

$$\begin{aligned} e^{d(Z)}=D(e^Z) &= D(e^{tX}e^{uY}e^{-tX}e^{-uY})=D(e^{tX})D(e^{uY})D(e^{-tX})D(e^{-uY}) \\ &= e^{td(X)}e^{ud(Y)}e^{-td(X)}e^{-ud(Y)} \\ &= e^{ut[d(X),d(Y)]+\cdots}\ , \end{aligned} \tag{2.44}$$

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 $\mathfrak{g}$ applies in particular to a representation of G which plays a special role, the *adjoint representation* of G into its Lie algebra $\mathfrak{g}$. It is defined by the following action

$$X\in\mathfrak{g}\qquad D^{\mathrm{adj}}(g)(X)=gXg^{-1}\ , \tag{2.45}$$

which we denote $\mathrm{Ad}\,g\,X$. (The right hand side of (2.45) 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 $\mathfrak{g}$ in the space $\mathfrak{g}$, also called adjoint representation. It is obtained by taking the infinitesimal form of (2.45), formally

$g = I + tY$, or by considering the one-parameter subgroup generated by $Y \in \mathfrak{g}$, $g(t) = \exp tY$ and by calculating $\operatorname{Ad} g(t)X = g(t)Xg^{-1}(t) = X + t[Y, X] + O(t^2)$ (cf. Chap. 1 (1.29)), whence

$$\frac{d}{dt}\operatorname{Ad} g(t)X\bigg|_{t=0} = [Y, X] = \operatorname{ad} Y\, X \ . \tag{2.46}$$

where we recover (and justify) our notation ad of Chap. 1.

Exercise: show that matrices T_i defined by $(T_i)_j{}^k = -C_{ij}{}^k$ 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 $\mathfrak{g}$ by antihermitian operators (or matrices). Physicists, who love 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 $\mathfrak{g}$ generates a representation of the unique connected and simply connected group whose Lie algebra is $\mathfrak{g}$. In other words if $X \overset{d}{\mapsto} d(X)$ is a representation of the algebra, $e^X \mapsto e^{d(X)}$ 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 $\mathfrak{g}$, 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 $\mathrm{GL}(E)$) 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. 0 on the concrete cases of SU(2) and $\mathrm{SL}(2, \mathbb{C})$.

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 2π), 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 that its inequivalent irreducible representations are labelled by a discrete index, written in upper position: $D^{(\rho)}$. These representations are a priori of finite or infinite dimension –in fact we shall see below that the dimension n_ρ of $D^{(\rho)}$ is in fact finite. In a finite or countable basis, the matrices $\mathcal{D}^{(\rho)}_{\alpha\beta}$ may be assumed to be unitary, according to the result of § 2.1.3. (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 $x \mapsto e^{ikx}$, but nothing restricts $k \in \mathbb{C}$, except unitarity which imposes $k \in \mathbb{R}$.

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

$$\boxed{\int \frac{d\mu(g)}{v(G)} \mathcal{D}^{(\rho)}_{\alpha\beta}(g) \mathcal{D}^{(\rho')*}_{\alpha'\beta'}(g) = \frac{1}{n_\rho} \delta_{\rho\rho'} \delta_{\alpha\alpha'} \delta_{\beta\beta'}} \tag{2.47}$$

and their characters satisfy thus

$$\boxed{\int \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g) \chi^{(\rho')*}(g) = \delta_{\rho\rho'} \ .} \tag{2.48}$$

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

$$V = \int d\mu(g') \mathcal{D}^{(\rho)}(g') M \mathcal{D}^{(\rho')\dagger}(g') \ . \tag{2.49}$$

The left hand side of (2.47) is (up to a facteur $v(G)$) the derivative with respect to $M_{\beta\beta'}$ of $V_{\alpha\alpha'}$. The representations being unitary, $\mathcal{D}^\dagger(g) = \mathcal{D}(g^{-1})$, it is easy, using the left invariance of the measure $d\mu(g') = d\mu(gg')$, to check that V satisfies

$$V \mathcal{D}^{(\rho')}(g) = \mathcal{D}^{(\rho)}(g) V \tag{2.50}$$

for all $g \in G$. By Schur lemma, the matrix V is thus vanishing if representations ρ and ρ' 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 (2.47) with $\delta_{\rho\rho'} = 0$.
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: $c_1 n_\rho = v(G)\mathcal{D}_{11}(I) = v(G)$, which proves that the dimension n_ρ 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: $c_M n_\rho = v(G) \mathrm{tr}\, M$, which, upon differentiation wrt $M_{\beta\beta'}$, leads to the orthonormality (2.47), qed.

The proposition (2.48) 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 (2.48) 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: $\chi^{(\rho)} = \chi^{(\sigma)} \Longleftrightarrow \rho = \sigma$.

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 measure (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: $\mathcal{D}_{\alpha\beta} \in L^2(G)$. Choosing M as in b), one finds again $\int d\mu(g) = c_1 \,\mathrm{tr}\, I$. In the lhs, the integral over the group ("volume of the group" G) diverges. In the rhs, $\mathrm{tr}\, I$, 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 (2.47) (or (2.48), 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 $\mathcal{D}^{(\rho)}_{\alpha\beta}(g)$ also satisfy a completeness property

$$\boxed{\sum_{\rho,\alpha,\beta} n_\rho \mathcal{D}^{(\rho)}_{\alpha\beta}(g) \mathcal{D}^{(\rho)*}_{\alpha\beta}(g') = v(G) \delta(g, g') \ ,} \tag{2.51}$$

or stated differently

$$\sum_{\rho,\alpha,\beta} n_\rho \mathcal{D}^{(\rho)}_{\alpha\beta}(g) \mathcal{D}^{(\rho)\dagger}_{\beta\alpha}(g') = \sum_\rho n_\rho \chi^{(\rho)}(g.g'^{-1}) = v(G) \delta(g, g') \ , \qquad (2.51)'$$

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 $\mathcal{D}^{(\rho)}_{\alpha\beta}(g)$

$$f(g) = \int d\mu(g')\delta(g,g')f(g') = \sum_{\rho,\alpha,\beta} n_\rho \mathcal{D}^{(\rho)}_{\alpha\beta}(g)\int \frac{d\mu(g')}{v(G)}\mathcal{D}^{(\rho)\dagger}_{\beta\alpha}(g')f(g') =: \sum_{\rho,\alpha,\beta} n_\rho \mathcal{D}^{(\rho)}_{\alpha\beta}(g) f^{(\rho)}_{\alpha\beta} \,. \tag{2.52}$$

This is the *Peter–Weyl theorem*, a non trivial statement that we admit[5]. 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 (2.52):

$$\begin{aligned}
f^{(\rho)}_{\alpha\beta} = \int \frac{d\mu(g')}{v(G)} f(g')\mathcal{D}^{(\rho)\dagger}_{\beta\alpha}(g') &= \int \frac{d\mu(g')}{v(G)} f(hg'h^{-1})\mathcal{D}^{(\rho)\dagger}_{\beta\alpha}(hg'h^{-1}) \qquad \forall h \\
&= \int \frac{d\mu(h)}{v(G)}\frac{d\mu(g')}{v(G)} f(g')\mathcal{D}^{(\rho)\dagger}_{\beta\gamma}(h)\mathcal{D}^{(\rho)\dagger}_{\gamma\delta}(g')\mathcal{D}^{(\rho)\dagger}_{\delta\alpha}(h^{-1}) \\
&= \int \frac{d\mu(g')}{v(G)} f(g')\mathcal{D}^{(\rho)\dagger}_{\gamma\delta}(g')\frac{1}{n_\rho}\delta_{\alpha\beta}\delta_{\gamma\delta} \qquad \text{by (2.47)} \\
&= \frac{1}{n_\rho}\int \frac{d\mu(g')}{v(G)} f(g')\chi^{(\rho)*}(g')\delta_{\alpha\beta}
\end{aligned} \tag{2.53}$$

from which it follows that (2.52) reduces to an expansion on characters, qed.

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

$$\sum_{k=-\infty}^{\infty} e^{ikx}e^{-ikx'} = 2\pi\delta_P(x-x') \tag{2.54}$$

where $\delta_P(x-x') = \sum_{\ell=-\infty}^{\infty}\delta(x-x'-2\pi\ell)$ is the periodic Dirac distribution (alias "Dirac's comb"). Then (2.52) means that any 2π-periodic function (with adequate regularity conditions) may be represented by its Fourier *series*

$$f(x) = \sum_{k=-\infty}^{\infty} e^{ikx} f_k \qquad\qquad f_k = \int_{-\pi}^{\pi}\frac{dx}{2\pi} f(x)e^{-ikx} \,. \tag{2.55}$$

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

$$f(x) = \int_{-\infty}^{\infty} dk \tilde{f}(k)e^{ikx} \qquad\qquad \tilde{f}(k) = \int_{-\infty}^{\infty}\frac{dx}{2\pi} f(x)e^{-ikx} \,. \tag{2.56}$$

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)

$$\mathcal{D}^{(k)}(\alpha) = \begin{pmatrix} \cos k\alpha & -\sin k\alpha \\ \sin k\alpha & \cos k\alpha \end{pmatrix} \quad , \quad k \in \mathbb{N}^* \quad , \quad \chi^{(k)}(\alpha) = 2\cos k\alpha \tag{2.57}$$

What are now the orthogonality and completeness relations?

[5]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

$$\chi = \sum_\rho m_\rho \chi^{(\rho)} \tag{2.58}$$

and multiplicities may be computed by

$$m_\rho = \int \frac{d\mu(g)}{v(G)} \chi(g)\chi^{(\rho)*}(g) \ . \tag{2.59}$$

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

$$\begin{aligned} D^{(\rho)} \otimes D^{(\sigma)} &= \oplus_\tau m_\tau D^{(\tau)} && (2.60)\\ \chi^{(\rho)}\chi^{(\sigma)} &= \sum_\tau m_\tau \chi^{(\tau)} && (2.61)\\ m_\tau &= \int \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g)\chi^{(\sigma)}(g)\chi^{(\tau)*}(g) \ , && (2.62) \end{aligned}$$

and hence the representation τ appears in the product $\rho \otimes \sigma$ with the same multiplicity as σ^* 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 (2.47, 2.48, 2.51) and their consequences (2.58, 2.59, 2.60), 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 $\sum_{g\in G}$:

$$\frac{1}{|G|}\sum_{g\in G}\mathcal{D}^{(\rho)}_{\alpha\beta}(g)\mathcal{D}^{(\rho')*}_{\alpha'\beta'}(g) = \frac{1}{n_\rho}\delta_{\rho\rho'}\delta_{\alpha\alpha'}\delta_{\beta\beta'} \tag{2.63}$$

$$\sum_{\rho,\alpha,\beta}\frac{n_\rho}{|G|}\mathcal{D}^{(\rho)}_{\alpha\beta}(g)\mathcal{D}^{(\rho)*}_{\alpha\beta}(g') = \delta_{g,g'}\ . \tag{2.64}$$

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

$$\sum_\rho n_\rho^2 = |G|\ . \tag{2.65}$$

This follows from the fact that the system of equations (2.63) expresses that the matrix $\mathcal{U}_{\rho,\alpha\beta\,;\,g} := \left(\frac{n_\rho}{|G|}\right)^{\frac12}\mathcal{D}^{(\rho)}_{\alpha\beta}(g)$ 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 $\mathcal{C}_i$ in the group.*

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

$$\frac{1}{|G|}\sum_{i=1}^m|\mathcal{C}_i|\chi^{(\rho)}_i\chi^{(\rho')*}_i = \delta_{\rho\rho'} \tag{2.66a}$$

$$\frac{|\mathcal{C}_i|}{|G|}\sum_{\rho=1}^r\chi^{(\rho)}_i\chi^{(\rho)*}_j = \delta_{ij}\ . \tag{2.66b}$$

(Exercise : derive the second relation from (2.52) and (2.53), applied to a finite group.) But once again, these relations mean that the matrix $\mathcal{K}_{\rho\,i} := \left(\frac{|C_i|}{|G|}\right)^{\frac12}\chi^{(\rho)}_i$ of dimensions $r\times m$ satisfies $\mathcal{K}\mathcal{K}^\dagger = I$, $\mathcal{K}^\dagger\mathcal{K} = I$, thus is a square (and unitary) matrix, $m = r$, 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 (2.66).

We illustrate it on the example of the group T, subgroup of the rotation group 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 π 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 (2.65) that their dimensions can only be $n_\rho = 1,1,1$ and 3. The character table is thus a 4×4 table, of which one row is already known, that of the identity representation D_1, and one column, that of dimensions n_ρ. The spin 1 representation of SO(3) yields a dimension 3 representation of T whose character χ 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, $\|\chi\|^2 = \sum_i\frac{|\mathcal{C}_i|}{|G|}|\chi_i|^2 = 1$ and this character is irreducible. This gives a second row (called D_4). The spin 2 representation of 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 $j + j^2 = -1$.

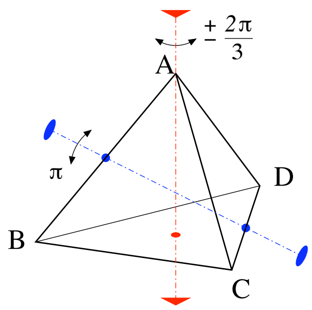


Figure 2.1: A tetrahedron, with two axes of rotation

$\downarrow$ irreps. ρ $\backslash$ Classes C_i $\rightarrow$	$\mathcal{C}(0)$	$\mathcal{C}(\pi)$	$\mathcal{C}(\frac{2\pi}{3})$	$\mathcal{C}(-\frac{2\pi}{3})$
D_1	1	1	1	1
D_2	1	1	j	j^2
D_3	1	1	j^2	j
D_4	3	-1	0	0
$\lvert C_i \rvert$	1	3	4	4

Check that relations (2.66) 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 of 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

$$U(g_1)U(g_2) = e^{i\zeta(g_1,g_2)}U(g_1g_2) \ . \tag{2.67}$$

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

$$\zeta(g_1,g_2)\to\zeta'(g_1,g_2)=\zeta(g_1,g_2)+\alpha(g_1)+\alpha(g_2)-\alpha(g_1g_2)\ . \tag{2.68}$$

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:

$$\forall g_1,g_2,g_3 \qquad (\partial\zeta)(g_1,g_2,g_3):=\zeta(g_1,g_2)+\zeta(g_1g_2,g_3)-\zeta(g_2,g_3)-\zeta(g_1,g_2g_3)=0 \tag{2.69}$$

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

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

For a 1-cochain $\alpha(g)$, $\partial\alpha(g_1,g_2)=\alpha(g_1g_2)-\alpha(g_1)-\alpha(g_2)$, and hence (2.68) 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 ζ is *trivial*, *i.e.* if there exists $\alpha(g)$ such that in (2.68), $\zeta'=0$. In other words, is the 2-cocycle ζ, which is closed ($\partial\zeta=0$) by (2.69), 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 § 2.4.5) by saying that for a semi-simple group G, such as 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 $\widetilde{G}$, its universal covering, give representations up to a phase of G. For example, one recovers that the projective representations of SO(3) (up to a sign) are representations of SU(2). This is also the case of the Lorentz group O(1,3), the universal covering of which is 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[6] of a Hilbert space $\mathcal{H}$, and in which the observables are self-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.*

$$|\psi\,\rangle\to|{}^g\psi\,\rangle\ ,\quad A\to{}^gA \quad \text{such that} \qquad |\langle\,\phi|A|\psi\,\rangle|=|\langle\,{}^g\phi|{}^gA|{}^g\psi\,\rangle|\ . \tag{2.70}$$

One then proves the following theorem

[6]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*

$$|\langle\, \phi|A|\psi\,\rangle| = |\langle\, {}^g\phi|{}^gA|{}^g\psi\,\rangle|\ , \tag{2.71}$$

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.

$$|{}^g\psi\,\rangle = U(g)|\phi\,\rangle\ , \quad {}^gA = U(g)AU^\dagger(g)\ ; \qquad\qquad U(g)U^\dagger(g) = U(g)^\dagger U(g) = I\ . \tag{2.72}$$

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

$$U(\lambda|\phi\,\rangle + \mu|\psi\,\rangle) = \lambda^* U|\phi\,\rangle + \mu^* U|\psi\,\rangle \tag{2.73}$$

and its adjoint is defined by

$$\langle\, \phi|U^\dagger|\psi\,\rangle = \langle\, U\phi|\psi\,\rangle^* = \langle\, \psi|U\phi\,\rangle\ , \tag{2.74}$$

so as to be consistent with linearity

$$\langle\, \lambda\phi|U^\dagger|\psi\,\rangle = \lambda^*\langle\, \phi|U^\dagger|\psi\,\rangle\ . \tag{2.75}$$

If it is also unitary,

$$\langle\, \psi|\phi\,\rangle^* = \langle\, \phi|\psi\,\rangle = \langle\, \phi|U^\dagger U|\psi\,\rangle = \langle\, U\phi|U\psi\,\rangle^*\ , \tag{2.76}$$

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 $|\psi\,\rangle \to |{}^g\psi\,\rangle$ 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

$$\begin{aligned} \langle\, {}^g\phi|{}^gA|{}^g\psi\,\rangle &= \langle\, U\phi|UAU^\dagger|U\psi\,\rangle & (2.77)\\ &= \langle\, \phi|U^\dagger UAU^\dagger U|\psi\,\rangle^{\#} & (2.78)\\ &= \langle\, \phi|A|\psi\,\rangle^{\#} & (2.79)\end{aligned}$$

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 $\mathbf{x}$, but changes the sign of velocities, hence of the momentum vector $\mathbf{p}$

$$\begin{aligned} \mathbf{x}' &= U(T)\mathbf{x}U^\dagger(T) = \mathbf{x} & (2.80)\\ \mathbf{p}' &= U(T)\mathbf{p}U^\dagger(T) = -\mathbf{p}\ . & (2.81)\end{aligned}$$

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

$$\begin{aligned} [x'_j, p'_k] &= -[x_j, p_k] = -i\hbar\delta_{jk} && (2.82)\\ &= U(T)[x_j, p_k]U^\dagger(T) = U(T)i\hbar\delta_{jk}U^\dagger(T) && (2.83) \end{aligned}$$

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

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 $g_1 g_2$ 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 any 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 (two observers) 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 systeme 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

$$[H, U(g)] = 0\ . \tag{2.84}$$

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 ou 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

$$i\hbar\frac{\partial\mathcal{F}(U(g))}{\partial t} = [\mathcal{F}(U(g)), H] = 0 \tag{2.85}$$

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, but in general not simultaneously measurable.

Examples.

Translation group $\longrightarrow P_\mu$ energy–momentum; rotation group $\longrightarrow M_{\mu\nu}$ 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

$$[T_i, T_j] = iC_{ij}^{\ k} T_k \tag{2.86}$$

(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 $\dim E^{(\rho)}$ [7]. If some representation spaces $E^{(\rho)}$ appear with a multiplicity m_ρ 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 us 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 (2.72), the transformation of an operator on $\mathcal{H}$ obeys: $A \to U(g)AU(g)^\dagger$. Suppose we are given a set of such operators, A_α, $\alpha = 1, 2, \cdots$, transforming linearly among themselves, *i.e.* forming a representation:

$$A_\alpha \to U(g)A_\alpha U(g)^\dagger = \sum_{\alpha'} A_{\alpha'}\mathcal{D}_{\alpha'\alpha}(g) \ . \tag{2.87}$$

If the representation D is irreducible, the operators A_α 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_α 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

$$U(g)A_\alpha|\sigma\beta\,\rangle = U(g)A_\alpha U(g)^\dagger U(g)|\sigma\beta\,\rangle = \mathcal{D}^{(\rho)}_{\alpha'\alpha}(g)\mathcal{D}^{(\sigma)}_{\beta'\beta}(g)A_{\alpha'}|\sigma\beta'\,\rangle \tag{2.88}$$

[7] It may happen that the multiplicity of some eigenvalue $\mathcal{E}_\rho$ of H is higher than m_ρ, either because of the existence of a symmetry group larger than G, 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 (2.89), one may decompose on irreducible representations

$$\mathcal{D}^{(\rho)}_{\alpha'\alpha}(g)\mathcal{D}^{(\sigma)}_{\beta'\beta}(g) = \sum_{\tau,\gamma,\gamma',i} \langle\, \tau_i\gamma|\rho,\alpha;\sigma,\beta\,\rangle\langle\, \tau_i\gamma'|\rho,\alpha';\sigma,\beta'\,\rangle^*\,\mathcal{D}^{(\tau_i)}_{\gamma'\gamma}(g)\ . \tag{2.89}$$

Suppose now that the group G is compact (or finite). The representation matrices satisfy the orthogonality property (2.47). One may thus write

$$\begin{aligned}\langle\, \tau\gamma|A_\alpha|\sigma\beta\,\rangle &= \langle\, \tau\gamma|U(g)^\dagger U(g)A_\alpha|\sigma\beta\,\rangle \qquad\qquad \forall g\in G\\ &= \int\frac{d\mu(g)}{v(G)}\langle\, \tau\gamma|U(g)^\dagger U(g)A_\alpha|\sigma\beta\,\rangle\\ &= \int\frac{d\mu(g)}{v(G)}\sum_{\alpha',\beta',\gamma'}\mathcal{D}^{(\tau)*}_{\gamma'\gamma}(g)\langle\, \tau\gamma'|A_{\alpha'}|\sigma\beta'\,\rangle\mathcal{D}^{(\rho)}_{\alpha'\alpha}(g)\mathcal{D}^{(\sigma)}_{\beta'\beta}(g)\\ &= \frac{1}{n_\tau}\sum_{\alpha',\beta',\gamma',i}\langle\, \tau_i\gamma|\rho,\alpha;\sigma,\beta\,\rangle\langle\, \tau_i\gamma'|\rho,\alpha';\sigma,\beta'\,\rangle^*\langle\, \tau_i\gamma'|A_{\alpha'}|\sigma\beta'\,\rangle\ .\end{aligned} \tag{2.90}$$

Introduce the notation

$$\langle\, \tau \parallel A \parallel \sigma\,\rangle_i := \frac{1}{n_\tau}\sum_{\alpha',\beta',\gamma'}\langle\, \tau_i\gamma'|\rho,\alpha';\sigma,\beta'\,\rangle^*\langle\, \tau_i\gamma'|A_{\alpha'}|\sigma\beta'\,\rangle\ . \tag{2.91}$$

It follows that (Wigner–Eckart theorem):

$$\langle\, \tau\gamma|A_\alpha|\sigma\beta\,\rangle = \sum_{i=1}^{m_\tau}\langle\, \tau \parallel A \parallel \sigma\,\rangle_i\langle\, \tau_i\gamma|\rho,\alpha;\sigma,\beta\,\rangle \tag{2.92}$$

in which the "reduced matrix elements" $\langle\, . \parallel A \parallel .\,\rangle_i$ are independent of α,β,γ. The matrix element of the lhs in (2.92) vanishes if the Clebsch-Gordan coefficient is zero, (in particular if the representation τ does not appear in the product of ρ and σ). 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 mentionned at the end of § 2.4.3, when a representation space appears with a multiplicity m_ρ. Labelling by an index $i = 1,\cdots m_\rho$ the various copies of representation ρ, one has thanks to (2.92)

$$\langle\, \rho\alpha i|H|\rho\alpha' i'\,\rangle = \delta_{\alpha\alpha'}\langle\, \rho i \parallel H \parallel \rho i'\,\rangle \tag{2.93}$$

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

$$\langle\, j,m_1|K_1^m|j,m_2\,\rangle = \langle\, j,m_1|J^m|j,m_2\,\rangle\frac{\langle\, \vec{J}.\vec{K}\,\rangle}{j(j+1)} \tag{2.94}$$

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

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 (2.67), 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 $\alpha\beta$

$$i\zeta(I + \alpha t_a, I + \beta t_b) = \alpha\beta z_{ab} \ . \tag{2.95}$$

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

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

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)

$$[T_a, T_b] = C_{ab}{}^c T_c + C_{ab} \qquad [C_{ab}, T_c] = 0 \qquad [C_{ab}, C_{cd}] = 0 \ . \tag{2.96}$$

(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 ζ translates in infinitesimal form into the possibility (or impossibility) of getting rid of the central term by a redefinition of the T

$$T_a \to \widetilde{T}_a = T_a + X_a \qquad [\widetilde{T}_a, \widetilde{T}_b] = C_{ab}{}^c \widetilde{T}_c \ , \tag{2.97}$$

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 (2.97) 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 SU(n), SO(n), 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 $\widetilde{G}$, which are representations up to a phase of G (the group $\pi_1(G) = \widetilde{G}/G$ is represented on $U(1)$). This is the case of the groups SO(n) and their universal covering Spin(n), (for example SO(3)), or of the Lorentz group 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

[Ki] A.A. Kirillov, *Elements of the theory of representations*, Springer.

[Kn] A. Knapp, *Representation Theory of semi-simple groups*, Princeton U. Pr.

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

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

[BrD] T. Bröcker and T. tom Dieck, *Representations of compact Lie groups*, Springer.

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

On projective representations, see

[Ba] V. Bargmann, *Ann. Math.* **59** (1954) 1-46, or

S. Weinberg [Wf] Chap 2.7.

⋆

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

$$z=\sum_\alpha x^{(\alpha)}\otimes y^{(\alpha)} \tag{D.1}$$

with a finite sum over vectors $x^{(\alpha)}\in E,\ y^{(\alpha)}\in F$ (a possible scalar coefficient λ_α 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

$$\begin{aligned} A\otimes B(x\otimes y) &= Ax\otimes By && \text{(D.2)}\\ A\otimes B\sum_\alpha(x^{(\alpha)}\otimes y^{(\alpha)}) &= \sum_\alpha Ax^{(\alpha)}\otimes By^{(\alpha)} && \text{(D.3)}\end{aligned}$$

In particular if E et F have two bases e_i and f_j, $z=x\otimes y=\sum_{i,j}x^iy^je_if_j$, 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

$$(A\otimes B)z=\sum_{i,i',j,j'}A_{ii'}B_{jj'}x^{i'}y^{j'}e_if_j=:(A\otimes B)_{ii';jj'}z^{i'j'}e_i\otimes f_j \tag{D.4}$$

thus

$$(A \otimes B)_{ij;i'j'} = A_{ii'}B_{jj'}\,, \tag{D.5}$$

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, $x \in E \mapsto D(g)x = e_i\mathcal{D}_{ij}x^j$, 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 $\mathcal{D}_{ii'}\mathcal{D}'_{jj'}$.

Another way of saying it is: if x "transforms by the representation D" and y by D', 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

$$(x' \otimes y')^{ij} = x^iy^j = \mathcal{D}_{ii'}\mathcal{D}'_{jj'}x^{i'}y^{j'}\,, \tag{D.7}$$

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 $E_1 \otimes E_2 \otimes \cdots E_p$ and rank p tensors $z^{i_1\cdots i_p}$. This is what we did in Chap. 0, § 0.3.3, in the construction of the representations of SU(2) by symmetrized tensor products of the spin $\frac{1}{2}$ representation, or in § 0.6.4 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 $\mathcal{D}^j$ of SU(2) defined and explicitly constructed in §0.3.2 and 0.3.3.

E.1. Orthogonality, completeness, characters

All unitary representations of SU(2) have been constructed in Chap. 0. Following the discussion of §3, the matrix elements of $\mathcal{D}^j$ 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)

$$\begin{aligned}(2j+1)\int \frac{d\mu(U)}{2\pi^2}\mathcal{D}^j_{mn}(U)\mathcal{D}^{j'*}_{m'n'}(U) &= \delta_{jj'}\delta_{mm'}\delta_{nn'} \\ \sum_{jmn}(2j+1)\mathcal{D}^j_{mn}(U)\mathcal{D}^{j*}_{mn}(U') &= 2\pi^2\delta(U,U')\,.\end{aligned} \tag{E.1}$$

The "delta 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 α, β, γ for example,

$$\delta(U, U') = 8\delta(\alpha - \alpha')\delta(\cos\beta - \cos\beta')\delta(\gamma - \gamma')\,, \tag{E.2}$$

(see Appendix C of Chap. 1). The meaning of equation (E.1) is that functions $\mathcal{D}^j_{mn}(U)$ 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

$$\begin{aligned}\chi_j(U) = \chi_j(\psi) &= \operatorname{tr}\mathcal{D}^j(\mathbf{n},\psi) = \sum_{m=-j}^{j} e^{im\psi} \\ &= \frac{\sin\left(\frac{2j+1}{2}\psi\right)}{\sin\frac{\psi}{2}} \; .\end{aligned} \tag{E.3}$$

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

$$\chi_0(\psi) = 1 \qquad \chi_{\frac{1}{2}}(\psi) = 2\cos\frac{\psi}{2} \qquad \chi_1(\psi) = 1 + 2\cos\psi \quad \text{etc} \; . \tag{E.4}$$

One may then verify all the expected properties

$$\begin{array}{rc} \text{unitarity and reality} & \chi_j(U^{-1}) = \chi_j^*(U) = \chi_j(U) \\ \text{parity and periodicity} & \chi_j(-U) = \chi_j(2\pi+\psi) = (-1)^{2j}\chi_j(U) \\ \text{orthogonality} & \int_0^{2\pi} \mathrm{d}\psi \sin^2\frac{\psi}{2}\chi_j(\psi)\chi_{j'}(\psi) = \pi\delta_{jj'} \\ \text{completeness} & \sum_{j=0,\frac{1}{2},\cdots}\chi_j(\psi)\chi_j(\psi') = \frac{\pi}{\sin^2\frac{\psi}{2}}\delta(\psi-\psi') = \frac{\pi}{2\sin\frac{\psi}{2}}\delta(\cos\frac{\psi}{2} - \cos\frac{\psi'}{2}) \end{array} \tag{E.5}$$

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

Does the multiplicity formula (2.60) lead to the well known formulae (2.31)?

E.2. 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 $\mathcal{D}^j_{m'm}$ and exposes their relation with "special functions" of mathematical physics.

We also noticed that the construction of the *Wigner $\mathcal{D}$ matrices* in § 0.3.3 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 (0.70) of Chap. 0 still holds true

$$P_{jm}(\xi',\eta') = \sum_{m'} P_{jm'}(\xi,\eta)\mathcal{D}^j_{m'm}(A) \; . \tag{0-0.70}$$

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

$$\left(\frac{\partial^2}{\partial a\partial d} - \frac{\partial^2}{\partial b\partial c}\right)(a\xi+c\eta)^{j+m}(b\xi+d\eta)^{j-m} = 0 \tag{E.6}$$

and because of the independance of the $P_{jm}(\xi,\eta)$, the $\mathcal{D}^j_{m'm}(A)$ 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 $A=\rho\,U$, $U\in \mathrm{SU}(2)$, and (E.6) leads to

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

where Δ_4 is the Laplacian in the space $\mathbb{R}^4$ with variables $u_0,\mathbf{u}$, and $a=u_0+iu_3$, $b=u_1+iu_2$. In polar coordinates

$$\Delta_4=\frac{\partial^2}{\partial\rho^2}+\frac{3}{\rho}\frac{\partial}{\partial\rho}+\frac{1}{\rho^2}\Delta_{S^3} \tag{E.8}$$

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

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

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

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

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

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

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

$$\left[\frac{2l+1}{4\pi}\right]^{\frac12}\mathcal{D}^l_{m0}(\phi,\theta,0)=Y_l^{m*}(\theta,\phi)\ . \tag{E.12}$$

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

$$P_l(u) = \frac{1}{2^l l!}\frac{\mathrm{d}^l}{\mathrm{d}u^l}(u^2-1)^l \tag{E.13}$$

$$P_l^m(u) = (1-u^2)^{\frac12 m}\frac{\mathrm{d}^m}{\mathrm{d}u^m}P_l(u) \qquad \text{for } 0\le m\le l\ . \tag{E.14}$$

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

$$P_0=1 \qquad P_1=u \qquad P_2=\frac12(3u^2-1) \qquad P_3=\frac12(5u^3-3u)\ ,\cdots \tag{E.15}$$

while $P_l^0 = P_l$, $P_l^1 = (1-u^2)^{\frac{1}{2}} P_l$, etc. The spherical harmonics are related to Legendre functions $P_l^m(\cos\theta)$ (for $m \ge 0$) by

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

and thus

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

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

$$P_l^{(\alpha,\beta)}(u) = \frac{(-1)^l}{2^l l!}(1-u)^{-\alpha}(1+u)^{-\beta}\frac{\mathrm{d}^l}{\mathrm{d}u^l}\left[(1-u)^{\alpha+l}(1+u)^{\beta+l}\right] \tag{E.18}$$

by

$$d^j_{m'm}(\theta) = \left[\frac{(j+m')!(j-m')!}{(j+m)!(j-m)!}\right]^{\frac{1}{2}} \left(\cos\frac{\theta}{2}\right)^{m+m'} \left(\sin\frac{\theta}{2}\right)^{m-m'} P_{j-m'}^{(m'-m,m'+m)}(\cos\theta)\ . \tag{E.19}$$

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

$$\int_{-1}^{1} \mathrm{d}u(1-u)^\alpha(1+u)^\beta P_j^{(\alpha,\beta)}(u) P_{j'}^{(\alpha,\beta)}(u) = \delta_{jj'}\frac{2^{\alpha+\beta+1}\Gamma(l+\alpha+1)\Gamma(l+\beta+1)}{(2l+\alpha+\beta+1)l!\Gamma(l+\alpha+\beta+1)} \tag{E.20}$$

and satisfy the recursion relation

$$\begin{aligned} &2(l+1)(l+\alpha+\beta+1)(2l+\alpha+\beta)P_{l+1}^{(\alpha,\beta)}(u) \qquad \text{(E.21)}\\ &= (2l+\alpha+\beta+1)[(2l+\alpha+\beta)(2l+\alpha+\beta+2)u+\alpha^2-\beta^2]P_l^{(\alpha,\beta)}(u) \quad - \quad 2(l+\alpha)(l+\beta)(2l+\alpha+\beta+2)P_{l-1}^{(\alpha,\beta)}\ . \end{aligned}$$

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

$$\{(1-u^2)\frac{\mathrm{d}^2}{\mathrm{d}u^2} + [\beta-\alpha-(2+\alpha+\beta)u]\frac{\mathrm{d}}{\mathrm{d}u} + l(l+\alpha+\beta+1)\}P_l^{(\alpha,\beta)}(u) = 0\ . \tag{E.22}$$

The Legendre polynomials correspond to the case $\alpha = \beta = 0$. These relations appear here as related to those of the $\mathcal{D}^j$. 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

$$(\Delta_{S^2} + l(l+1))\,Y_l^m \qquad = 0 \tag{E.23}$$

$$J_z Y_l^m = -i\frac{\partial}{\partial\phi}Y_l^m \quad = mY_l^m \tag{E.24}$$

and may be written as

$$Y_l^m(\theta,\phi) = \frac{(-1)^l}{2^l l!}\sqrt{\frac{(2l+1)(l+m)!}{4\pi(l-m)!}} e^{im\phi}\sin^{-m}\theta\left(\frac{\mathrm{d}}{\mathrm{d}\cos\theta}\right)^{l-m}\sin^{2l}\theta\ . \tag{E.25}$$

(ii) They are normalized to 1 on the unit sphere and more generally satisfy orthogonality and completeness properties

$$\int d\Omega\, Y_l^{m*} Y_{l'}^{m'} = \int_0^{2\pi} d\phi \int_0^{\pi} d\theta\, \sin\theta\, Y_l^{m*} Y_{l'}^{m'} = \delta_{ll'}\delta_{mm'} \tag{E.26}$$

$$\sum_{l=0}^{\infty}\sum_{m=-l}^{l} Y_l^{m*}(\theta,\phi) Y_l^m(\theta',\phi') = \delta(\Omega-\Omega') = \frac{\delta(\theta-\theta')\delta(\phi-\phi')}{\sin\theta} \tag{E.27}$$

$$= \delta(\cos\theta-\cos\theta')\delta(\phi-\phi') \tag{E.28}$$

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

$$Y_l^m(\mathbf{n}') = Y_l^{m'}(\mathbf{n})\mathcal{D}^l(R)_{m'm} \tag{E.29}$$

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

$$Y_l^{m*}(\theta,\phi) = (-1)^m Y_l^{-m}(\theta,\phi) \tag{E.30}$$

and parity

$$Y_l^m(\pi-\theta,\phi+\pi) = (-1)^l Y_l^m(\theta,\phi)\ . \tag{E.31}$$

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 (0.119)

$$e^{\pm i\phi}\left[\pm\frac{\partial}{\partial\theta} + i\mathrm{cotg}\,\theta\frac{\partial}{\partial\phi}\right] Y_l^m = \sqrt{l(l+1)-m(m\pm1)}Y_l^{m\pm1} \tag{E.32}$$

and those coming from the tensor product with the vector representation

$$\sqrt{2l+1}\cos\theta\, Y_l^m = \left(\frac{(l+m)(l-m)}{2l-1}\right)^{\frac12} Y_{l-1}^m + \left(\frac{(l+m+1)(l-m+1)}{2l+3}\right)^{\frac12} Y_{l+1}^m\ . \tag{E.33}$$

More generally, one has a product formula

$$Y_l^m(\theta,\phi)Y_{l'}^{m'}(\theta,\phi) = \sum_L \langle\, lm; l'm' | L, m+m'\,\rangle \left[\frac{(2l+1)(2l'+1)}{4\pi(2L+1)}\right]^{\frac12} Y_L^{m+m'}(\theta,\phi)\ . \tag{E.34}$$

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

$$\frac{2l+1}{4\pi}P_l(\cos\theta) = \sum_{m=-l}^{l} Y_l^m(\mathbf{n})Y_l^{m*}(\mathbf{n}') \tag{E.35}$$

where θ denotes the angle between directions $\mathbf{n}$ and $\mathbf{n}'$. 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

$$(\Delta_{S^2} + l(l+1))\,P_l(\mathbf{n}.\mathbf{n}') = 0$$

as a function of $\mathbf{n}$ or of $\mathbf{n}'$, as well as $(\mathbf{J}+\mathbf{J}')\,P_l=0$ where $\mathbf{J}$ and $\mathbf{J}'$ are generators of rotations of $\mathbf{n}$ and $\mathbf{n}'$ 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 $r'<r$),

$$\frac{1}{|\vec{r}-\vec{r}'|}=\sum_{l=0}^{\infty}\frac{r'^l}{r^{l+1}}P_l(\cos\theta)=\sum_{l,m}\frac{4\pi}{2l+1}\frac{r'^l}{r^{l+1}}Y_l^{m*}(\mathbf{n})Y_l^m(\mathbf{n}')\,. \tag{E.37}$$

The expression of the first Y_l^m may be useful

$$\begin{aligned} Y_0^0 &= \frac{1}{\sqrt{4\pi}} \\ Y_1^0 = \sqrt{\frac{3}{4\pi}}\cos\theta \qquad & Y_1^{\pm 1} = \mp\sqrt{\frac{3}{8\pi}}\sin\theta\, e^{\pm i\phi} \\ Y_2^0 = \sqrt{\frac{5}{16\pi}}\left(3\cos^3\theta-1\right) \qquad Y_2^{\pm 1} &= \mp\sqrt{\frac{15}{8\pi}}\cos\theta\sin\theta\, e^{\pm i\phi} \qquad Y_2^{\pm 2}=\sqrt{\frac{15}{32\pi}}\sin^2\theta\, e^{\pm 2i\phi}\,. \end{aligned} \tag{E.38}$$

E.3. Physical applications

E.3.1. 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

$$\phi(\vec{r})=\frac{1}{\epsilon_0}\sum_{l,m}\frac{1}{2l+1}\frac{Y_l^{m*}(\mathbf{n})}{r^{l+1}}Q_{lm} \tag{E.39}$$

where the Q_{lm}, defined by

$$Q_{lm}=\int d^3r'\rho(\vec{r}')r'^l Y_l^m(\mathbf{n}') \tag{E.40}$$

are the *multipole moments* of the charge distribution ρ. 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 $1/\sqrt{4\pi}$)

$$Q_{00}=\frac{Q}{\sqrt{4\pi}}=\sqrt{4\pi}\int r^2 dr\rho(r) \qquad \phi(r)=\frac{Q}{4\pi\epsilon_0 r}\,.$$

For an arbitrary $\rho(\vec{r})$, the three components of Q_{1m} 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, les Q_{lm} become operators in the Hilbert space of the theory. One may apply the Wigner-Eckart theorem and conclude that

$$\langle\, j_1, m_1|Q_{lm}|j_2, m_2\,\rangle = \langle\, j_1||Q_l||j_2\,\rangle\langle\, j_1, m_1|l, m; j_2, m_2\,\rangle$$

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 $l \le 2j$.

E.3.2. Eigenstates of the angular momentum in Quantum Mechanics

Spherical harmonics may be interpreted as wave functions in coordinates θ, ϕ of the eigenstates of the angular momentum $\vec{L} = \hbar\vec{J} = \hbar\vec{r} \times \vec{\nabla}$

$$Y_l^m(\theta,\phi) = \langle\, \theta, \phi|l, m\,\rangle$$

in analogy with

$$\frac{1}{(2\pi)^{3/2}} e^{i\vec{x}.\vec{p}} = \langle\, \vec{x}|\vec{p}\rangle\ .$$

(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 $\mathbf{n} = (\theta, \phi)$. One writes the scattering amplitude

$$\begin{aligned} \langle\, p, \theta, \phi|\mathcal{T}|p, 0, 0\,\rangle &= \sum_{ll'mm'} Y_l^m(\theta,\phi)\langle\, p, l, m|\mathcal{T}|p, l', m'\,\rangle Y_{l'}^{m'*}(0,0) \\ &= \sum_{lm} Y_l^m(\theta,\phi)\langle\, p, l, m|\mathcal{T}|p, l, m\,\rangle Y_l^{m*}(0,0) \qquad \text{(E.41)} \\ &= \sum_l \frac{2l+1}{4\pi}\mathcal{T}_l(p)P_l(\cos\theta) \end{aligned}$$

using once again the addition formula and $\langle\, plm|\mathcal{T}|pl'm'\,\rangle = \delta_{ll'}\delta_{mm'}\mathcal{T}_l(p)$ 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" $g_1 g_2 g_1^{-1} g_2^{-1}$ 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: $\ker \mathrm{ad} = 0$.

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 $D^{(\rho)} \otimes D^{(\bar\sigma)}$ into irreducible representations?

D. *Chebyshev polynomials*

Consider the expression

$$U_l = \frac{\sin(l+1)\theta}{\sin\theta} \ , \tag{2-113}$$

where l is an integer ≥ 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

$$\frac{2}{\pi}\int_{-1}^{1} \mathrm{d}z\,(1-z^2)^{\frac12}\, U_l(z) U_{l'}(z)$$

and

$$\frac{2}{\pi}\int_{-1}^{1} \mathrm{d}z\,(1-z^2)^{\frac12}\, U_l(z) U_{l'}(z) U_{l''}(z) \ ?$$

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 $(-1)^{m_3}\langle\, l_1, m_1; l_2, m_2 | l_3, -m_3\,\rangle$, with an m independent factor to be determined.

Problem I. Decomposition of an amplitude

Consider two real unitary representations (ρ) and (σ) 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 tr $T^aT^b = -\delta_{ab}$. Consider now the quantity

$$X_{\alpha\beta;\alpha'\beta'} := \sum_{a=1}^{d} T^{(\rho)a}_{\alpha\alpha'} T^{(\sigma)a}_{\beta\beta'}\ . \tag{2-114}$$

To simplify things, we assume that all irreducible representations appearing in the tensor product of representations (ρ) et (σ) 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

$$\left(\mathcal{M}^{(\tau\gamma)}\right)_{\alpha\beta} = \langle\,\tau\gamma|\rho\alpha;\sigma\beta\,\rangle\ . \tag{2-115}$$

1. Recall why these coefficients satisfy orthogonality and completeness relations and write them.

2. Show that it follows that

$$X_{\alpha\beta;\alpha'\beta'} = -\sum_{\tau\gamma}\left(\mathcal{M}^{(\tau\gamma)}\right)_{\alpha\beta}\left(T^{(\rho)a}\mathcal{M}^{(\tau\gamma)}T^{(\sigma)a}\right)_{\alpha'\beta'}\ . \tag{2-116}$$

3. Acting with the infinitesimal generator T^a on the two sides of the relation

$$|\rho\alpha;\sigma\beta\,\rangle = \sum_{\tau,\gamma}\left(\mathcal{M}^{(\tau\gamma)}\right)_{\alpha\beta}|\tau\gamma\,\rangle \tag{2-117}$$

show that one gets

$$\sum_{\gamma'} T^{(\tau)a}_{\gamma\gamma'}\left(\mathcal{M}^{(\tau\gamma')}\right)_{\alpha\beta} = \sum_{\alpha'}\left(\mathcal{M}^{(\tau\gamma)}\right)_{\alpha'\beta}(T^{(\rho)a})_{\alpha'\alpha} + \sum_{m_2'}\left(\mathcal{M}^{(\tau\gamma)}\right)_{\alpha\beta'}(T^{(\sigma)a})_{\beta'\beta} \tag{2-118}$$

or, in terms of matrices of dimensions $\dim(\rho)\times\dim(\sigma)$

$$\sum_{\gamma'} T^{(\tau)a}_{\gamma\gamma'}\mathcal{M}^{(\tau\gamma')} = -T^{(\rho)a}\mathcal{M}^{(\tau\gamma)} + \mathcal{M}^{(\tau\gamma)}T^{(\sigma)a}\ . \tag{2-119}$$

4. Using repeatedly this relation (2-119) in (2-116), show that one finds

$$X_{\alpha\beta;\alpha'\beta'} = \frac{1}{2}\sum_{\tau\gamma}(C_\rho + C_\sigma - C_\tau)\left(\mathcal{M}^{(\tau\gamma)}\right)_{\alpha\beta}\left(\mathcal{M}^{(\tau\gamma)}\right)_{\alpha'\beta'} \qquad (2\text{-}120)$$

where the C are Casimir operators, for example

$$C_\rho = -\sum_a (T^{(\rho)a})^2 \ . \qquad (2\text{-}121)$$

5. Why can one say that “large representations” τ tend to make the coefficient $(C_\rho + C_\sigma - C_\tau)$ increasingly negative? (One may take the example of SU(2) with ρ and σ 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 ξ_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 Fig. 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.

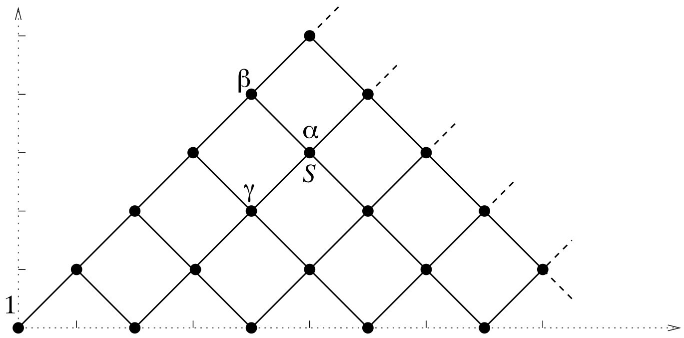


Figure 2.2: A graphical construction of the n_r

(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 $D^{-1\,T}(g)$ 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

$$D = S D^{-1\,T} S^{-1} \tag{2-122}$$

4. Show that (2-122) implies that the bilinear form S is invariant.
   Is this form degenerate?

5. Using (2-122) show that
$$DSS^{-1T} = SS^{-1T}D \ . \tag{2-123}$$

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 *real*, in the latter (S antisymmetric), it is called *pseudoreal* (or *quaternionic*). One may prove that in the former case, there exists a basis on $\mathbb{R}$ in which the representation matrices are real, and that no such basis exists 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 ρ and one denotes $\chi^{(\rho)}(g)$ their character. Let $\chi(g)$ be the character of some arbitrary representation, reductible or not.

   (a) For any function F on the finite G, one denotes $\langle F \rangle$ its group average
   $$\langle F \rangle = \frac{1}{|G|}\sum_{g\in G} F(g) \ . \tag{2-124}$$
   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 ρ, 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 ρ is equivalent to $\bar\rho$ iff
   $$\Big\langle \big(\chi^{(\rho)}(g)\big)^2 \Big\rangle = 1 \ .$$
   Evaluate this expression if ρ 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 $t = \{t^{ij}\}$,
   $$t^{ij} \mapsto t'^{ij} = \cdots$$

(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

$$\chi^{(\rho\otimes\rho)_{\substack{S\\A}}}(g) = \frac{1}{2}\left((\chi^{(\rho)}(g))^2 \pm \chi^{(\rho)}(g^2)\right) \ . \tag{2-125}$$

(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 ρ by

$$\mathrm{ind}(\rho) = \left\langle \chi^{(\rho)}(g^2) \right\rangle \ . \tag{2-126}$$

(a) Using the results of 2., show that one may write

$$\mathrm{ind}(\rho) = \langle\, \chi^{(\rho\otimes\rho)_S}\,\rangle - \langle\, \chi^{(\rho\otimes\rho)_A}\,\rangle \ .$$

(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 ± 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 $Q(h) := \sum_\rho \mathrm{ind}(\rho)\chi^{(\rho)}(h)$. 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 of finite dimension. We want to construct a canonical form of commutation relations modeled on the case of SU(2)

$$[J_z, J_\pm] = \pm J_\pm \qquad [J_+, J_-] = 2J_z \ . \tag{3.1}$$

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.* $\mathrm{ad}\, X = 0 \Rightarrow X = 0$, 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 $\mathfrak{g}$ is the Lie algebra of a Lie group G and if $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$, any conjugate $g\mathfrak{h}g^{-1}$ of $\mathfrak{h}$ by an arbitrary element of G is another Cartan subalgebra.

Let $\mathfrak{h}$ be a Cartan subalgebra, call ℓ its dimension, it is independent of the choice of $\mathfrak{h}$ and it is called the *rank* of $\mathfrak{g}$. 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

$$H_1 = \mathrm{diag}\,(1,-1,0,\cdots,0),\ H_2 = \mathrm{diag}\,(0,1,-1,0,\cdots,0),\cdots,\ H_{n-1} = \mathrm{diag}\,(0,\cdots,0,1,-1)\ . \tag{3.2}$$

An arbitrary matrix of the Lie algebra, (in that representation), (anti-)Hermitian and traceless, 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 $\mathfrak{g}$ is the Lie algebra of a group G, any element of $\mathfrak{g}$ is conjugate by G of an element of $\mathfrak{h}$.

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 $D = \mathrm{diag}\,(\begin{pmatrix} 0 & \mu_j \\ -\mu_j & 0\end{pmatrix}_{j=1,\cdots,n})$ if $N = 2n$ and $D = \mathrm{diag}\,(0, \begin{pmatrix} 0 & \mu_j \\ -\mu_j & 0\end{pmatrix}_{j=1,\cdots,n})$ if $N = 2n+1$, with real μ_j.*

If one allows the complexification of orthogonal matrices, one may fully diagonalise the matrix A in the form $D = \mathrm{diag}\,(\begin{pmatrix} i\mu_j & 0 \\ 0 & -i\mu_j\end{pmatrix}_{j=1,\cdots,n})$ or $D = \mathrm{diag}\,(0, \begin{pmatrix} i\mu_j & 0 \\ 0 & -i\mu_j\end{pmatrix}_{j=1,\cdots,n})$. 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 $\mathfrak{h}$. 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,

$$[\mathrm{ad}\,H_i, \mathrm{ad}\,H_j] = 0\ . \tag{3.3}$$

We may thus diagonalise simultaneously these $\mathrm{ad}\,H_i$. We already know (some?) eigenvectors of vanishing eigenvalue since $\forall i, j,\ \mathrm{ad}\,H_i\,H_j = 0$, and we may complete them to make a basis by finding a set of eigenvectors E_α linearly independent of the H_j

$$\mathrm{ad}\,H_i\,E_\alpha = \alpha_{(i)} E_\alpha \tag{3.4}$$

i.e. a set of elements of $\mathfrak{g}$ such that

$$[H_i, E_\alpha] = \alpha_{(i)} E_\alpha\ , \tag{3.5}$$

[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 $\mathrm{ad}\, H_i$. Since we chose Hermitian $\mathrm{ad}\, H_i$, their eigenvalues $\alpha_{(i)}$ are real. By linearity, for an arbitrary element of $\mathfrak{h}$ written as $H = \sum_i h^i H_i$,

$$\mathrm{ad}\,(H) E_\alpha = \alpha(H) E_\alpha \ , \tag{3.6}$$

and the eigenvalue of $\mathrm{ad}\,(H)$ on E_α 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* α, 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 α is a root, $-\alpha$ in another root;

2. the eigenspace of the eigenvalue α is of dimension 1 (no multiplicity);

3. if α is a root, the only roots of the form $\lambda\alpha$ are $\pm\alpha$;

4. roots α 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_j are diagonalisable, the total number of their eigenvectors E_α 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 α is even and equal to $d - \ell$ (with $\ell = \mathrm{rank}(\mathfrak{g})$). We denote Δ the set of roots.

In the basis $\{H_i, E_\alpha\}$ of $\mathfrak{g}$, the Killing form takes a simple form

$$(H_i, E_\alpha) = 0 \qquad (E_\alpha, E_\beta) = 0 \qquad \text{unless} \quad \alpha + \beta = 0 \ . \tag{3.7}$$

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

$$(H, [H', E_\alpha]) = \mathrm{tr}\,(\mathrm{ad}\, H[\mathrm{ad}\, H', \mathrm{ad}\, E_\alpha]) = \mathrm{tr}\,([\mathrm{ad}\, H, \mathrm{ad}\, H']\, \mathrm{ad}\, E_\alpha) = 0 \tag{3.8}$$

since $[\mathrm{ad}\, H, \mathrm{ad}\, H'] = 0$. It follows that $\forall H, H' \in \mathfrak{h}$, $\alpha(H')(H, E_\alpha) = 0$, hence that $(H, E_\alpha) = 0$. Likewise

$$([H, E_\alpha], E_\beta) = \alpha(H)(E_\alpha, E_\beta) = -(E_\alpha, [H, E_\beta]) = -\beta(H)(E_\alpha, E_\beta) \tag{3.9}$$

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_α 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

$$\forall H \in \mathfrak{h} \quad (H_\alpha, H) := \alpha(H) \ . \tag{3.10}$$

(Or said differently, one solves the linear system $g_{ij}h^j_\alpha = \alpha_{(i)}$ 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

$$\langle\, \alpha, \beta \,\rangle := (H_\alpha, H_\beta) \ , \tag{3.11}$$

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_α among themselves. Using the Jacobi identity, one finds that

$$\mathrm{ad}\, H_i[E_\alpha, E_\beta] = [H_i, [E_\alpha, E_\beta]] = [E_\alpha, [H_i, E_\beta] - [E_\beta, [H_i, E_\alpha]] = (\alpha+\beta)_{(i)}[E_\alpha, E_\beta] \ . \tag{3.12}$$

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 § 3.2.1 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 $\mathrm{ad}\, H_i$ with a vanishing eigenvalue, thus $[E_\alpha, E_{-\alpha}] = H \in \mathfrak{h}$. To determine that H, let us proceed like in (3.9)

$$\begin{aligned}(H_i, [E_\alpha, E_{-\alpha}]) &= \mathrm{tr}\,(\mathrm{ad}\, H_i\, [\mathrm{ad}\, E_\alpha, \mathrm{ad}\, E_{-\alpha}]) = \mathrm{tr}\,([\mathrm{ad}\, H_i, \mathrm{ad}\, E_\alpha]\, \mathrm{ad}\, E_{-\alpha}) \\ &= \alpha_{(i)}(E_\alpha, E_{-\alpha}) = (H_i, H_\alpha)(E_\alpha, E_{-\alpha})\end{aligned} \tag{3.13}$$

hence

$$[E_\alpha, E_{-\alpha}] = (E_\alpha, E_{-\alpha}) H_\alpha \ . \tag{3.14}$$

To recapitulate, we have constructed a canonical basis of the algebra $\mathfrak{g}$

$$\boxed{\begin{aligned}[H_i, H_j] =&\ 0 \\ [H_i, E_\alpha] =&\ \alpha_{(i)} E_\alpha \\ [E_\alpha, E_\beta] =&\ \begin{cases} N_{\alpha\beta} E_{\alpha+\beta} & \text{if } \alpha+\beta \text{ is a root} \\ (E_\alpha, E_{-\alpha})\, H_\alpha & \text{if } \ \alpha+\beta = 0 \\ 0 & \text{otherwise}\end{cases}\end{aligned}} \tag{3.15}$$

Up to that point, the normalisation of the vectors H_i and E_α has not been fixed. It is common to choose, in accord with (3.7)

$$\boxed{(H_i, H_j) = \delta_{ij} \qquad (E_\alpha, E_\beta) = \delta_{\alpha+\beta,0} \ .} \tag{3.15}$$

(Indeed, the restriction of the Killing form to $\mathfrak{h}$, after multiplication by i to make the $\mathrm{ad}\, H_i$ Hermitian, is *positive* definite.) With that normalisation, H_α defined above by (3.10) satisfies also

$$H_\alpha = \alpha . H := \alpha_{(i)} H_i \ . \tag{3.16}$$

Note that E_α, $E_{-\alpha}$ and H_α form an su(2) subalgebra

$$[H_\alpha, E_{\pm\alpha}] = \pm\langle\,\alpha,\alpha\,\rangle\,E_{\pm\alpha} \qquad [E_\alpha, E_{-\alpha}] = H_\alpha\ . \tag{3.17}$$

(This is in fact $H_\alpha/\langle\,\alpha,\alpha\,\rangle$ that we identify with J_z, 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 $\{H_i, E_\alpha, E_{-\alpha}\}$

$$g_{ab} = \begin{pmatrix} \mathbb{I}_\ell & & & & 0 & & \\ & 0 & 1 & & & & \\ & 1 & 0 & & & & \\ 0 & & & \ddots & & & \\ & & & & & 0 & 1 \\ & & & & & 1 & 0 \end{pmatrix} \tag{3.18}$$

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 ℓ, see point 4. in $(*)$ above) generated by the $d-\ell$ roots α inherits the Euclidean metric of $\mathfrak{h}$

$$\langle\,\alpha,\beta\,\rangle := (H_\alpha, H_\beta) = \alpha(H_\beta) = \beta(H_\alpha) = (\alpha.H, \beta.H) = \sum_i \alpha_{(i)}\beta_{(i)}\ , \tag{3.19}$$

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 $\frac{H_\alpha}{\langle\,\alpha,\alpha\,\rangle}$ plays the role of a J_z and has E_β as eigenvectors, $\mathrm{ad}\,H_\alpha E_\beta = \langle\,\alpha,\beta\,\rangle E_\beta$, *i.e.*

$$[H_\alpha, E_\beta] = \langle\,\alpha,\beta\,\rangle E_\beta \tag{3.20}$$

we may conclude that

$$2\frac{\langle\,\alpha,\beta\,\rangle}{\langle\,\alpha,\alpha\,\rangle} = m\ \in\mathbb{Z}\ . \tag{3.21}$$

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_α and "lowering" $E_{-\alpha}$ operators (aka *ladder operators*) on a given eigenvector E_β. We saw that if α and β are two distinct roots, with $\alpha+\beta\neq 0$, it may happen that $\beta\pm\alpha$ are also roots. Let $p\le 0$ be the smallest integer such that $(\mathrm{ad}\,E_{-\alpha})^{|p|}E_\beta$ is non zero, *i.e.* that $\beta + p\alpha$ is a root, and let $q \geq 0$ be the largest integer such that $(\mathrm{ad}\, E_\alpha)^q E_\beta$ 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 *α-chain through β*. Note that the $E_{\beta'}$, when β' runs along that chain, form a basis of a finite dimensional representation of the su(2) algebra generated by H_α and $E_{\pm\alpha}$. According to what we know about these representations of su(2), the lowest and highest eigenvalues of H_α 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)

$$m = -p - q \; . \tag{3.22}$$

This construction also shows that $\beta - m\alpha = \beta + (p+q)\alpha$ is in the α-chain through β, (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

$$|N_{\alpha\beta}| = \sqrt{\frac{1}{2}(1-p)q\langle\, \alpha, \alpha \,\rangle} \; . \tag{3.23}$$

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

$$w_\alpha(x) = x - 2\frac{\langle\, \alpha, x \,\rangle}{\langle\, \alpha, \alpha \,\rangle}\alpha \; . \tag{3.24}$$

This is a reflection in the hyperplane orthogonal to α: $(w_\alpha)^2 = I$, $w_\alpha(\alpha) = -\alpha$, and $w_\alpha(x) = x$ if x is orthogonal to α. 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_α is called a Weyl reflection. By definition the *Weyl group* W is the group generated by the w_α, *i.e.* the set of all possible products of w_α over roots α. Thanks to the remark following (3.22), if α and β 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 Δ, hence a finite group[2].

Example : for the algebra su(n), one finds that $W = \mathcal{S}_n$, 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 α-chain through β, and $\beta_- = \beta + p\alpha$ the lowest one, $w_\alpha(\beta_\pm) = \beta_\mp$ and more generally, the roots of the chain are swapped pairwise under the action of w_α. The chain is thus invariant by w_α. (This is a generalisation of the $m \leftrightarrow -m$ symmetry of the su(2) "multiplets" $(-j, -j+1, \cdots, j-1, j)$.)

---

[2]This 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.

Positive roots, simple roots. Cartan matrix

Roots are not linearly independent in $\mathfrak{h}^*$. One may show that one can partition their set Δ into "positive" and "negative" roots, the opposite of a positive root being negative, and find a basis α_i, $i = 1, \cdots, \ell$ of ℓ *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 α and β are simple roots, $\alpha - \beta$ cannot be a root (why?). The integer p in the previous discussion thus vanishes and $m = -q \le 0$. It follows that $\langle \alpha, \beta \rangle \le 0$.

$$\textit{The scalar product of two simple roots is non positive.} \tag{3.25}$$

We now define the Cartan matrix

$$C_{ij} = 2\frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle} . \tag{3.26}$$

Beware, that matrix is *a priori* non symmetric.[3] Its diagonal elements are 2, its off-diagonal elements are ≤ 0 integers.

One must remember that the scalar product appearing in the numerator of (3.26) is positive definite. According to the Schwarz inequality, $\langle \alpha, \beta \rangle^2 \le \langle \alpha, \alpha \rangle \langle \beta, \beta \rangle$ with equality only if α and β 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 $\langle \alpha_i, \alpha_j \rangle = \|\alpha_i\| \, \|\alpha_j\| \, \cos\widehat{\alpha_i, \alpha_j}$. Then by multiplying or dividing the two equations (3.21) for the pair $\{\alpha_i, \alpha_j\}$, $i \neq j$, namely $C_{ij} = m_i \le 0$ and $C_{ji} = m_j \le 0$, where the property (3.25) above has been taken into account, one finds that if $i \neq j$,

$$\left.\begin{array}{rcl} \cos\widehat{\alpha_i, \alpha_j} & = & -\frac{1}{2}\sqrt{m_i m_j} \\ \dfrac{\| \alpha_i \|}{\| \alpha_j \|} & = & \sqrt{\dfrac{m_i}{m_j}} \end{array}\right\} \qquad \text{with } m_i, m_j \in \mathbb{N} \; , \tag{3.27}$$

and the value -1 of the cosinus is impossible, since $\alpha_i \neq -\alpha_j$ by assumption, so that the only possible values of that cosinus are $0, -\frac{1}{2}, -\frac{\sqrt{2}}{2}, -\frac{\sqrt{3}}{2}$, *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 $?(\text{undetermined}), 1, \sqrt{2}, \sqrt{3}$.

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

$$A_2 : \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \quad B_2 : \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix} \quad G_2 : \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} \quad D_2 : \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} . \tag{3.28}$$

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 (3.26)!

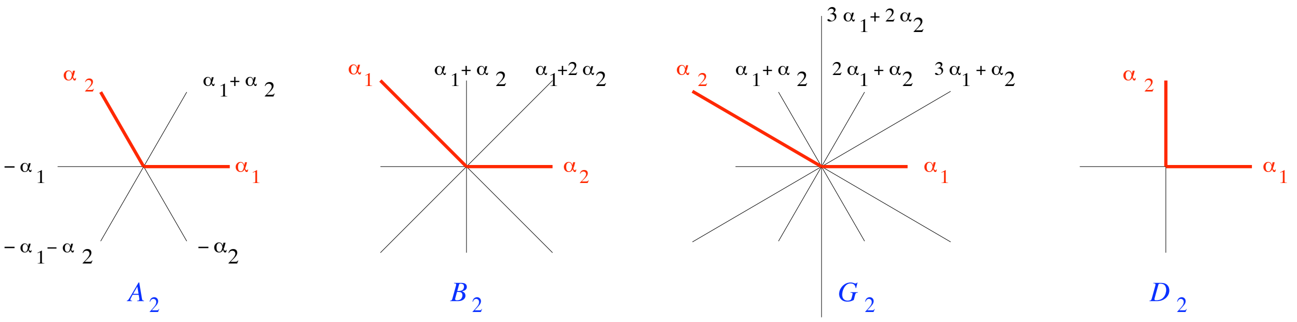


Figure 3.1: Root systems of rank 2. The two simple roots are drawn in thick lines. For the algebras B_2, G_2 et D_2, only positive roots have been labelled.

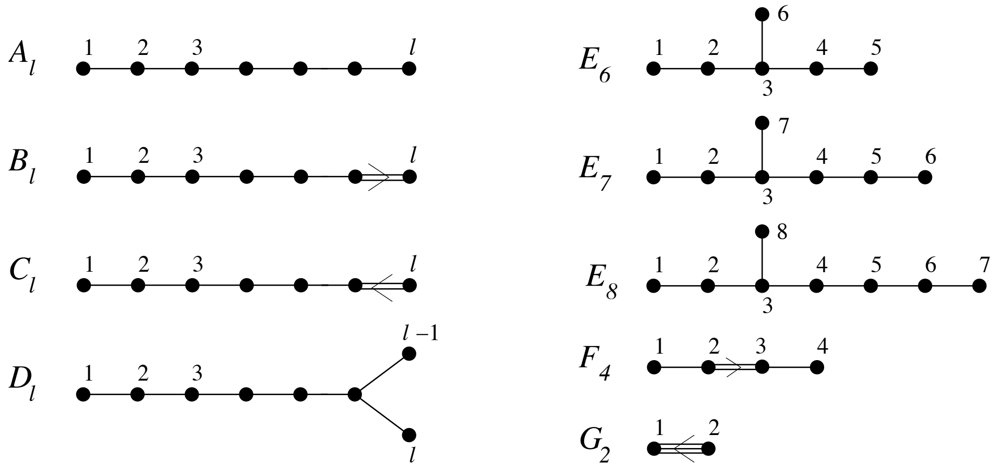


Figure 3.2: Dynkin diagrams

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 diagram

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 $\frac{3\pi}{4}$ resp. $\frac{5\pi}{6}$, 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 !).

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

$$A_\ell, \quad B_\ell, \quad C_\ell, \quad D_\ell, \quad E_6, \quad E_7, \quad E_8, \quad F_4, \quad G_2\,. \tag{3.29}$$

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

$$A_\ell = \mathrm{sl}(\ell+1,\mathbb{C}), \quad B_\ell = \mathrm{so}(2\ell+1,\mathbb{C}), \quad C_\ell = \mathrm{sp}(2\ell,\mathbb{C}), \quad D_\ell = \mathrm{so}(2\ell,\mathbb{C})\ . \tag{3.30}$$

or with their unique compact real form, respectively $A_\ell = \mathrm{su}(\ell+1)$, $B_\ell = \mathrm{so}(2\ell+1)$, $C_\ell = \mathrm{usp}(\ell)$, $D_\ell = \mathrm{so}(2\ell)$.

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 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 et f_i, $i = 1, \cdots, \ell$, be generators attached to simple roots α_i according to

$$e_i = \left(\frac{2}{\langle\,\alpha_i,\alpha_i\,\rangle}\right)^{\frac12} E_{\alpha_i}\ , \qquad f_i = \left(\frac{2}{\langle\,\alpha_i,\alpha_i\,\rangle}\right)^{\frac12} E_{-\alpha_i}\ , \qquad h_i = \frac{2\alpha_i.H}{\langle\,\alpha_i,\alpha_i\,\rangle}\ . \tag{3.31}$$

Their commutation relations read

$$\begin{aligned} [h_i,h_j] &= 0 \\ [h_i,e_j] &= C_{ji}\, e_j \\ [h_i,f_j] &= -C_{ji}\, f_j \\ [e_i,f_j] &= \delta_{ij} h_j \end{aligned} \tag{3.32}$$

(check!). The algebra is generated by the e_i, f_i, h_i and all their commutators, constrained by (3.32) and by the "Serre relations"

$$\begin{aligned} \mathrm{ad}\,(e_i)^{1-C_{ji}} e_j &= 0 \\ \mathrm{ad}\,(f_i)^{1-C_{ji}} f_j &= 0\ . \end{aligned} \tag{3.33}$$

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

$$\alpha_i^\vee := \frac{2}{\langle \alpha_i, \alpha_i \rangle} \alpha_i \ , \tag{3.34}$$

for α_i a simple root, appears frequently, it is given the name of *coroot*. The Cartan matrix may be rewritten as

$$C_{ij} = \langle \alpha_i, \alpha_j^\vee \rangle \ . \tag{3.35}$$

The *highest root* θ 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

$$\theta = \sum_i a_i \alpha_i \quad , \qquad \frac{2}{\langle \theta, \theta \rangle}\theta = \sum_i a_i^\vee \alpha_i^\vee \ , \tag{3.36}$$

called *Kac labels, resp dual Kac labels*, play also a role, in particular through their sums,

$$h = 1 + \sum_i a_i \quad , \qquad h^\vee = 1 + \sum_i a_i^\vee \ . \tag{3.37}$$

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

$$\widehat{C}_{ij} := 2 \frac{\langle \alpha_i, \alpha_j \rangle}{\sqrt{\langle \alpha_i, \alpha_i \rangle \langle \alpha_j, \alpha_j \rangle}} \tag{3.38}$$

yields a spectrum of eigenvalues

$$\text{eigenvalues of } \widehat{C} = \left\{ 4 \sin^2 \left(\frac{\pi}{2h} m_i \right) \right\}, \quad i = 1, \cdots, \ell \ , \tag{3.39}$$

in which a new set of integers m_i appears, the *Coxeter exponents*, satisfying $1 \le m_i \le h-1$ with possible multiplicities. These numbers are relevant for various reasons. They contain useful information on the Weyl group. After addition of 1, (making them ≥ 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 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 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

$$H|\lambda_a\rangle = \lambda(H)|\lambda_a\rangle \tag{3.40}$$

or equivalently

$$H_i|\lambda_a\rangle = \lambda_{(i)}|\lambda_a\rangle\ , \tag{3.41}$$

with an eigenvalue λ 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 λ is real-valued: the weights are real vectors of $\mathfrak{h}^*$. As the eigenvalue λ 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, incuding 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 § 3.2, we shall now show that any weight λ satisfies

$$\forall\alpha, \qquad 2\frac{\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle} = m' \in \mathbb{Z}\ , \tag{3.42}$$

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 λ

$$\lambda + p'\alpha, \cdots, \lambda, \cdots, \lambda + q'\alpha \qquad p' \le 0,\ q' \ge 0\ ,$$

which form a representation of the su(2) subalgebra, and then show that $m' = -p' - q'$.

Let p' be the smallest ≤ 0 integer such that $(E_{-\alpha})^{|p'|}|\lambda_a\rangle \ne 0$, and q' the largest ≥ 0 integer such that $(E_\alpha)^{q'}|\lambda_a\rangle \ne 0$, H_α 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

$$2q' + 2\frac{\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle} = 2j \qquad 2p' + 2\frac{\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle} = -2j\ . \tag{3.43}$$

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

$$2\frac{\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle} = -(q'+p') =: m', \qquad \text{as announced in (3.42).}$$

This chain is invariant under the action of the Weyl reflection w_α. (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 λ 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

$$\mathcal{C}_w = \{\lambda | \langle w\lambda, \alpha_i \rangle \geq 0 \ , \ \forall i = 1, \cdots, \ell\} \ , \tag{3.44}$$

where the α_i are the simple roots. (This is not quite a partition, as some weights belong to the "walls" between chambers.) The *fundamental chamber* is $\mathcal{C}_1$, corresponding to the identity in W. The weights belonging to that fundamental chamber are called *dominant weights.* Any weight may be brought into $\mathcal{C}_1$ 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 $\mathcal{C}_1$.

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_α, 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 Λ, 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 equation (3.43) and $\langle \Lambda, \alpha \rangle = \frac{1}{2}\langle \alpha, \alpha \rangle j > 0$, Λ is thus a dominant weight.

$\triangleright$ *The highest weight of a representation is a dominant weight,* $\Lambda \in \mathcal{C}_1$.

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 λ by

$$\lambda_i = 2 \frac{\langle \lambda, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} \in \mathbb{Z} \tag{3.45}$$

with α_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* Λ_i satisfy by definition

$$2 \frac{\langle \Lambda_j, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = \delta_{ij} \ . \tag{3.46}$$

Their number equals the rank ℓ of the algebra, and they make a basis of $\mathfrak{h}^*$. Each one is the highest weight of an irreducible representation called *fundamental*; hence there are ℓ fundamental representations. We have thus obtained

▷ *Any irreducible representation irreducible is characterised by its highest weight,*

and with a little abuse of notation, we denote (Λ) the irreducible representation of highest weight Λ.

▷ *Any highest weight decomposes on fundamental weights, and its components are its Dynkin labels (3.45),*

$$\Lambda = \sum_{j=1}^{\ell} \lambda_j \Lambda_j \quad , \qquad \lambda_i \in \mathbb{N} \ . \tag{3.47}$$

and any Λ of the form (3.47) is the highest weight of an irreducible representation.

Stated differently, the knowledge of the fundamental weights 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 θ, defined in eq. (3.36).

Weight and root lattices

Generally speaking, given a basis of vectors $e_1, \cdots e_p$ in a p dimensional space, the *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 *weight lattice* P is the lattice generated by the ℓ fundamental weights Λ_i. The root lattice Q is the one generated by the ℓ simple roots α_i. This is a sublattice of P. Any weight of an irreducible representation belongs to P.

One may consider the *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. (Exercise : prove it. Hint : compute the determinants of the Λ_i and of the α_i in the basis of coroots.) In the case of su(n), there are n classes, we shall return to that point later.

One may also introduce the lattice $Q^\vee$ generated by the ℓ coroots $\alpha_i^\vee$ (cf §2.4). It is the "dual" of P, in the sense that $\langle \alpha_i^\vee, \Lambda_j \rangle \in \mathbb{Z}$.

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 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, SU(n)/$\mathbb{Z}_d$ has the su(n) Lie algebra. The case $n = 2$, with SU(2) and 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 *Weyl vector* ρ, defined by any of the two (non trivially!) equivalent formulas

$$\rho \ = \tfrac{1}{2} \textstyle\sum_{\alpha>0} \alpha$$

$$= \sum_j \Lambda_j \ . \tag{3.48}$$

A remarkable formula, due to Weyl, gives the dimension of the representation of highest weight Λ as a product over positive roots

$$\dim(\Lambda) = \prod_{\alpha>0} \frac{\langle \Lambda+\rho, \alpha \rangle}{\langle \rho, \alpha \rangle} \tag{3.49}$$

while the eigenvalue of the quadratic Casimir reads

$$C_2(\Lambda) = \frac{1}{2}\langle \Lambda, \Lambda + 2\rho \rangle \ . \tag{3.50}$$

A related question is that of the trace of generators of $\mathfrak{g}$ in the representation (Λ). Let t_a be the basis of $\mathfrak{g}$ such that $\mathrm{tr}\ t_a t_b = T_A \delta_{ab}$, 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 Λ, one has

$$\mathrm{tr}\ d_\Lambda(t_a) d_\Lambda(t_b) = T_\Lambda \delta_{ab} \ . \tag{3.51}$$

But in that basis, the quadratic Casimir reads $C_2 = \sum_a \left(d_\Lambda(t_a)\right)^2$ hence, taking the trace,

$$\begin{aligned} \mathrm{tr}\, C_2 \ &= \textstyle\sum_a \mathrm{tr}\ \left(d_\Lambda(t_a)\right)^2 = T_\Lambda \sum_a 1 = T_\Lambda \ \dim \mathfrak{g} \\ &= C_2(\Lambda)\, \mathrm{tr}\, I_\Lambda = C_2(\Lambda)\ \dim(\Lambda) \end{aligned} \tag{3.52}$$

whence

$$T_\Lambda = C_2(\Lambda) \frac{\dim(\Lambda)}{\dim \mathfrak{g}} \ , \tag{3.53}$$

a useful formula in calculations (gauge theories ...). In the adjoint representation, $\dim(A) = \dim \mathfrak{g}$, 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 ρ and θ to the dimension of the algebra and the Coxeter number: $\langle \rho, \rho \rangle = \frac{h}{24} \langle \theta, \theta \rangle \dim \mathfrak{g}$.

There is also a formula (Freudenthal) giving the multiplicity of a weight λ within a representation of given highest weight Λ. 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 $\mathfrak{h}$.

Conjugate representation

Given a representation of highest weight Λ, its complex conjugate representation is generally non equivalent. One may characterize its highest weight $\overline{\Lambda}$ thanks to the Weyl group. The non-equivalence of representations (Λ) 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

$$\begin{array}{rcccl} A_\ell = \mathrm{su}(\ell+1) & \lambda_i & \leftrightarrow & \lambda_{\ell+1-i} & \ell > 1 \\ D_{2r+1} = \mathrm{so}(4r+2) & \lambda_\ell & \leftrightarrow & \lambda_{\ell-1} \ , & \ell = 2r+1 \\ E_6 & \lambda_i & \leftrightarrow & \lambda_{6-i}, & i = 1, 2 \ . \end{array} \tag{3.54}$$

3.3.2 Roots and weights of su(n)

Let us construct explicitly the weights and thus the irreducible representations of su(n).

We first pick a convenient parametrization of the space $\mathfrak{h}^*$, which is of dimension $n-1$. Let e_i, $i=1,\cdots n$, be n vectors of $\mathfrak{h}^*=\mathbb{R}^{n-1}$ (hence necessarily dependent), satisfying $\sum_1^n e_i=0$. They are obtained starting from an orthonormal basis $\hat{e}_i$ of $\mathbb{R}^n$ by projecting the $\hat{e}_i$ on an hyperplane orthogonal to $\hat{\rho}:=\sum_{i=1}^n \hat{e}_i$, thus $e_i=\hat{e}_i-\frac{1}{n}\hat{\rho}$. It is convenient to choose the hyperplane $\sum_i^n x^i=1$ in the space $\mathbb{R}^n$. These vectors have scalar products given by

$$\langle\, e_i, e_j\,\rangle=\delta_{ij}-\frac{1}{n}\ . \tag{3.55}$$

In terms of these vectors, the positive roots of su(n)$=A_{n-1}$, whose number equals $|\Delta_+|=n(n-1)/2$, are

$$\alpha_{ij}=e_i-e_j\,,\qquad 1\le i<j\le n\,, \tag{3.56}$$

and the $\ell=n-1$ simple roots are

$$\alpha_i=\alpha_{i\,i+1}=e_i-e_{i+1}\,,\qquad 1\le i\le n-1\ . \tag{3.57}$$

These roots have been normalized by $\langle\,\alpha,\alpha\,\rangle=2$. The sum of positive roots is easily computed to be

$$\begin{aligned} 2\rho\ &=(n-1)e_1+(n-3)e_2+\cdots+(n-2i+1)e_i+\cdots-(n-1)e_n\\ &=(n-1)\alpha_1+2(n-2)\alpha_2+\cdots+i(n-i)\alpha_i+\cdots+(n-1)\alpha_{n-1}. \end{aligned} \tag{3.58}$$

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 (3.57). The fundamental weights Λ_i, $i=1,\cdots,n-1$ are then readily written

$$\Lambda_i\ =\ \sum_{j=1}^{i} e_j\ , \tag{3.59}$$

$$e_1\ =\ \Lambda_1,\ e_i=\Lambda_i-\Lambda_{i-1}\text{ for } i=2,\cdots,n-1,\ e_n=-\Lambda_{n-1} \tag{3.60}$$

with scalar products

$$\langle\,\Lambda_i,\Lambda_j\,\rangle=\frac{i(n-j)}{n}\ ,\quad i\le j\ . \tag{3.61}$$

The Weyl group $W\cong\mathcal{S}_N$ acts on roots and on weights by permuting the e_i:

$$w\in W\leftrightarrow\bar{w}\in\mathcal{S}_N\ :\ w(e_i)=e_{\bar{w}(i)}\,.$$

Dimension of the representation of weight Λ

Combining formulas (3.49) and (3.56), prove the following expression

$$\dim(\Lambda)=\prod_{1\le i<j\le n}\frac{f_i-f_j+j-i}{j-i}\qquad \text{où}\qquad f_i:=\sum_{k=i}^{n-1}\lambda_k,\quad f_n=0. \tag{3.62}$$

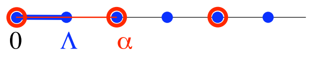


Figure 3.3: Weights of su(2). The positive parts of the weight (small dots) and root (big dots) lattices.

Conjugate representations

If $\Lambda = (\lambda_1, \cdots, \lambda_{n-1})$ is the highest weight of an irreducible representation of 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 quadratic Casimir operator distinguish the representations Λ and $\overline{\Lambda}$.

"n-ality".

There are n congruence classes of P with respect to Q. They are distinguished by the value of

$$\nu(\lambda) := \lambda_1 + 2\lambda_2 + \cdots + (n-1)\lambda_{n-1} \mod n\ , \tag{3.63}$$

to which one may give the unimaginative name of "n-ality", by extension of the "triality" of su(3), see below. The elements of the root lattice thus have $\nu(\lambda) = 0$.

Examples of su(2) and su(3).

In the case of su(2), there is one fundamental weight $\Lambda = \Lambda_1$ and one positive root α, normalised by $\langle \alpha, \alpha \rangle = 2$, hence $\langle \Lambda, \alpha \rangle = 1$, $\langle \Lambda, \Lambda \rangle = \frac{1}{2}$. Thus $\alpha = 2\Lambda$, Λ corresponds to the spin $\frac{1}{2}$ representation, α to spin 1. The weight lattice and the root lattice are easy to draw, see Fig. 3.3. The Dynkin label λ_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 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[5]

$$\dim(\Lambda) = \frac{1}{2}(\lambda_1 + 1)(\lambda_2 + 1)(\lambda_1 + \lambda_2 + 2)\ , \tag{3.64}$$

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 λ_1, λ_2 using the formulas (3.50) and (3.58).

[5] which may be ambiguous; for example, identify on Fig. (3.4) the weight of another representation of dimension 15.

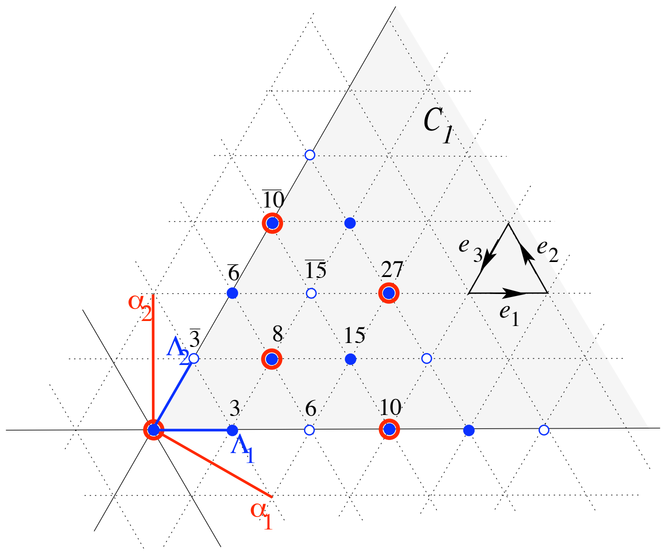


Figure 3.4: Weights of su(3). Only the first Weyl chamber $\mathcal{C}_1$ has been detailed, with some highest weights. The weights of triality 0 (forming the root lattice) are represented by a wide disk, those of triality 1, resp. 2, by a full, resp. open disk.

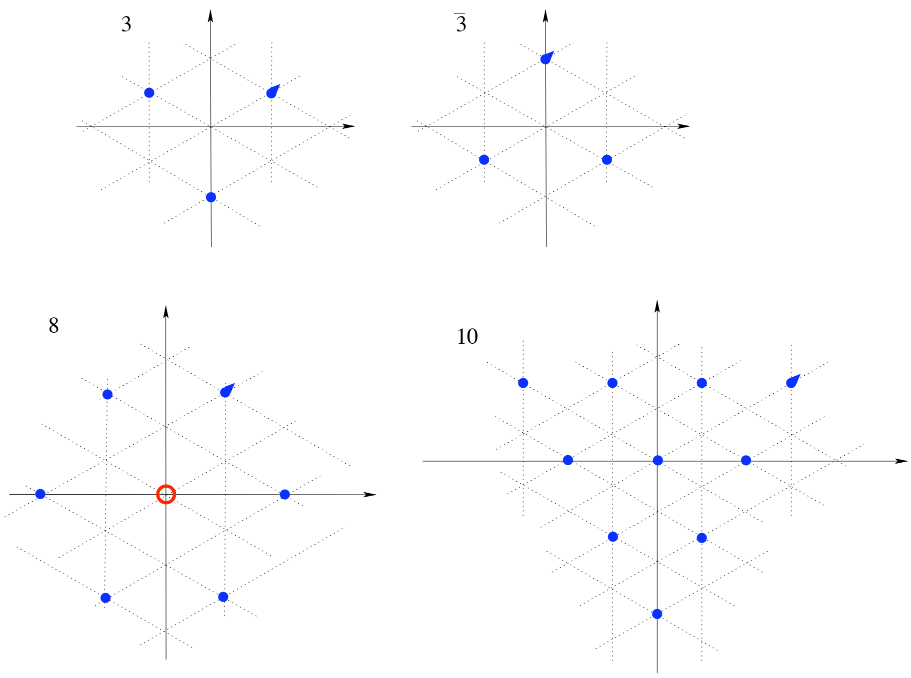


Figure 3.5: The weight diagrams of low lying representations of su(3), denoted by their dimension. Note that a rotation of 30° of the weight lattice has been performed with respect to the previous figure. In each representation, the highest weight is marked by a small indentation. The small dots are weights of multiplicity 1, the wider and open dots have multiplicity 2.

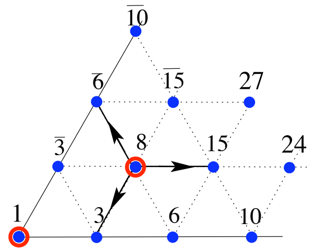


Figure 3.6: Tensor product of the 8 representation by the 3 representation, depicted on the weight diagram of su(3).

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 α_1, and the vertical axis, colinear to Λ_2, will indeed acquire a physical meaning: that of axes of isospin and "hypercharge" coordinates, see next chapter.

Remark. The case of 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 su(n)

3.4.1 Littlewood-Richardson rules

Given two irreducible representations of 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 exist also fairly complex rules giving that decomposition into irreducible representations of a product of two irreducible representations (Λ) and (Λ') of 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 (Λ') is found in an adequate product of fundamental representations, and examining the successive products of representation Λ by these fundamental representations. By the associativity of the tensor product, one brings the original problem back to that of the tensor product of (Λ) by the various fundamental representations.

The latter operation is easy to describe on the weight lattice. Given the highest weight Λ in the first Weyl chamber $\mathcal{C}_1$, the tensor product of (Λ) by the fundamental representation of highest weight Λ_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 Λ and one keeps as highest weights in the decomposition only the weights resulting from this addition that belong to $\mathcal{C}_1$.

Let us illustrate that on the case of 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 $\bar{3}$ (see below, end of § 3.4.2). 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

$$\begin{aligned} 3 \otimes 3 &= \bar{3} \oplus 6 \qquad 3 \otimes \bar{3} = 1 \oplus 8 \\ 3 \otimes 6 &= 8 \oplus 10 \qquad 3 \otimes \bar{6} = \bar{3} \oplus \overline{15} \\ 3 \otimes 8 &= 3 \oplus \bar{6} \oplus 15 \\ 3 \otimes 15 &= 6 \oplus \overline{15} \oplus 24 \end{aligned} \tag{3.65}$$

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 Λ_1, Λ_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, $\tau(3) = 1$, $\tau(15) = 1$, $\tau(6) = 2$, $\tau(\overline{15}) = 2$, etc.

Iterating this procedure, one then computes

$$8 \otimes (1 \oplus 8) = 8 \otimes 3 \otimes \bar{3} = (3 \oplus \bar{6} \oplus 15) \otimes \bar{3} = 1 \oplus 8 \oplus 8 \oplus 8 \oplus 10 \oplus \overline{10} \oplus 27$$

from which one derives the formula

$$8 \otimes 8 = 1 \oplus 8_s \oplus 8_a \oplus 10 \oplus \overline{10} \oplus 27\ . \tag{3.66}$$

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 exists a more elaborate graphical rule for the tensor product of two general highest weight representations (Λ) and (Λ'). 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 $U \in \mathrm{SU}(n)$, $V \mapsto V' = UV$, 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 δ^i_j is invariant

$$\delta'^i_j = U^i{}_{i'} U^{\dagger j'}{}_j \delta^{i'}_{j'} = (UU^\dagger)^i{}_j = \delta^i_j\ .$$

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

$$t'^{\,i_1\cdots i_p}_{\,k_1\cdots k_m}=U^{i_1}_{\ j_1}\cdots U^{i_p}_{\ j_p}U^{\dagger l_1}_{\ \ k_1}\cdots U^{\dagger l_m}_{\ \ k_m}t^{j_1\cdots j_p}_{l_1\cdots l_m}\ . \tag{3.67}$$

• 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 $CUC^{-1}=U^*$, thus $C^{-1}V^*$ transforms like V. Or, since $C_{ij}=\epsilon_{ij}$ and $\epsilon_{i'j'}U^{i'}_{\ i}U^{j'}_{\ j}=\epsilon_{ij}\det U=\epsilon_{ij}$, the antisymmetric tensor ϵ, invariant and invertible ($\epsilon^{ij}=-\epsilon_{ij}$, $\epsilon_{ij}\epsilon^{jk}=\delta_i^k$), 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

$$t^{i_1i_2\cdots i_p}=t^{[i_1,i_2]\cdots i_p}+t^{\{i_1,i_2\}\cdots i_p}$$

with $t^{[i_1,i_2]\cdots i_p}:=\frac{1}{2}(t^{i_1i_2\cdots i_p}-t^{i_2i_1\cdots i_p})$ and $t^{\{i_1,i_2\}\cdots i_p}:=\frac{1}{2}(t^{i_1i_2\cdots i_p}+t^{i_2i_1\cdots i_p})$. The antisymmetric component may be recast as $t^{[i_1,i_2]\cdots i_p}=\epsilon^{i_1i_2}\tilde t^{i_3\cdots i_p}$, with $\tilde t^{i_3\cdots i_p}=-\frac{1}{2}\epsilon_{ab}t^{abi_3\cdots i_p}$, and its rank has thus been reduced[6]. Consequently only tensors that are completely symmetric in all their p upper indices give irreducible representations, and one recovers once again the construction of all irreducible representations of SU(2) by symmetrized tensor products of the representation of dimension 2, see Chap. 0, 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 $\underbrace{\square\square\square}_{p}$. 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

$$\square\square\square\otimes\square=\square\square\square\square\oplus\begin{array}{l}\square\square\square\\ \square\end{array}$$

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.

$$(t^{ijk}u^l-t^{jkl}u^i)=\epsilon^{il}\tilde t^{jk}\qquad,\qquad\tilde t^{jk}=-\epsilon_{ab}t^{ajk}u^b$$

which we represent by erasing the columns with two boxes. Therefore

$$\square\square\square\otimes\square=\square\square\square\square\oplus\square\square$$

[6]It may be useful to recall the identities $\epsilon_{ab}\epsilon_{cd}=\delta_{ac}\delta_{bd}-\delta_{ad}\delta_{bc}$ and hence $\epsilon_{ab}\epsilon_{bc}=-\delta_{ac}$.

where we recognize the familiar rule $j \otimes \frac{1}{2} = (j+\frac{1}{2}) \oplus (j-\frac{1}{2})$.

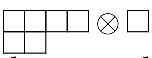
Exercise : reproduce with this method the decomposition rule of $j \otimes j'$.

• 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 $t^{i_1\cdots i_p}_{j_1\cdots j_m}$, $(i., j. = 1, 2, 3)$, transforming under the representation $3^{\otimes p} \otimes \bar{3}^{\otimes m}$. We still have an invariant tensor ϵ, but now of rank 3,

$$\epsilon_{i'j'k'} U^{i'}{}_i U^{j'}{}_j U^{k'}{}_k = \epsilon_{ijk} \det U = \epsilon_{ijk} \ ,$$

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 (3.64), see Exercise E. 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 § 3.4.1), 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  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 3×3 traceless matrices. A tensor of that representation transforms by $t^i_j \mapsto t'^i_j = U^i{}_{i'} U^{*j'}{}_j t^{i'}_{j'}$, or in a matrix form

$$t' = U t U^\dagger \ , \tag{3.68}$$

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 anti-symmetrization operations related to the symmetric group of permutations S_m. We just give a few indications.

Let $E = \mathbb{C}^n$ be the vector space of dimension n. The group $\mathrm{GL}(n,\mathbb{C})$, or $\mathrm{GL}(n)$ in short, is naturally represented in E

$$g \in \mathrm{GL}(n), \quad x \in E \mapsto x' = g.x \ . \tag{3.69}$$

Form the m-th tensor power of E: $F = E^{\otimes m} = E \otimes \cdots \otimes E$. In F, the group $\mathrm{GL}(n)$ acts by a representation, the m-th tensor power of (3.69)

$$g \in \mathrm{GL}(n), \quad D(g)x^{(1)} \otimes \cdots x^{(m)} = g.x^{(1)} \otimes \cdots \otimes g.x^{(m)} \tag{3.70}$$

which is in general reductible. But in F, there is also the action of the symmetric group S_m according to

$$\sigma \in S_m, \quad \widehat{D}(\sigma)x^{(1)} \otimes \cdots x^{(m)} = x^{(\sigma^{-1}1)} \otimes \cdots \otimes x^{(\sigma^{-1}m)} \ . \tag{3.71}$$

Choose a basis e_i in E, and denote g_{ij} the matrix elements of $g \in \mathrm{GL}(n)$ in that basis. The representation of $\mathrm{GL}(n)$ in F has a matrix

$$\mathcal{D}(g)_{\{i_1\cdots i_m\}\{j_1\cdots j_m\}} = \prod_{k=1}^{m} g_{i_k j_k} \tag{3.72}$$

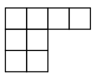
and that of S_m

$$\widehat{\mathcal{D}}(\sigma)_{\{i_1\cdots i_m\}\{j_1\cdots j_m\}} = \prod_{k=1}^{m} \delta_{i_{\sigma k}\, j_k} \ . \tag{3.73}$$

A tensor t, element of F, has components $t^{i_\cdot}$ in that basis and transforms under the action of $g \in \mathrm{GL}(n)$, resp. of $\sigma \in S_m$, into a tensor t' of components $t'^{i_\cdot} = \mathcal{D}_{i_\cdot, j_\cdot} t^{j_\cdot}$, resp. $\widehat{\mathcal{D}}_{i_\cdot, j_\cdot} t^{j_\cdot}$. These two sets of matrices commute

$$\begin{aligned}\sum_{\{j_\cdot\}} \mathcal{D}(g)_{\{i_\cdot\},\{j_\cdot\}}\widehat{\mathcal{D}}(\sigma)_{\{j_\cdot\},\{k_\cdot\}} &= \textstyle\prod_l g_{i_l j_l}\delta_{j_l,k_{\sigma^{-1}l}} = \prod_l g_{i_{\sigma l}k_l} \\ &= \textstyle\sum_{\{j_\cdot\}} \widehat{\mathcal{D}}(\sigma)_{\{i_\cdot\},\{j_\cdot\}}\mathcal{D}(g)_{\{j_\cdot\},\{k_\cdot\}} \ . \end{aligned} \tag{3.74}$$

Define then a *Young diagram.* A Young diagram is made of m boxes set in k rows of non increasing length: $f_1 \geq f_2 \geq \cdots f_k$, $\sum f_i = m$. Here is an example for $m = 8$, with $f_1 = 4, f_2 = 2, f_3 = 2$



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

$$n_Y = \frac{n!}{\prod_i \ell_i!} \prod_{i<j} (\ell_i - \ell_j) \tag{3.75}$$

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 (3.75).

A tensor is said to be of (symmetry) type Y if it transforms by S_m under that representation. The commutation of matrices $\mathcal{D}(g)$ and $\widehat{\mathcal{D}}(\sigma)$, eq. (3.74), then ensures that tensors of type Y form an invariant sub-espace under the action of $\mathrm{GL}(n)$.

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 $\mathrm{GL}(n)$

$$t^{i_1 i_2} = \frac{1}{2}\left(t^{i_1 i_2} + t^{i_2 i_1}\right) + \frac{1}{2}\left(t^{i_1 i_2} - t^{i_2 i_1}\right) \ .$$

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

$$\ytableaushort{123} \quad A = t^{i_1i_2i_3} + t^{i_2i_3i_1} + t^{i_3i_1i_2} + t^{i_2i_1i_3} + t^{i_3i_2i_1} + t^{i_1i_3i_2} \tag{3.76}$$

$$\ytableaushort{1,2,3} \quad B = t^{i_1i_2i_3} + t^{i_2i_3i_1} + t^{i_3i_1i_2} - t^{i_2i_1i_3} - t^{i_3i_2i_1} - t^{i_1i_3i_2} \tag{3.77}$$

$$\ytableaushort{12,3} \quad C_1 = t^{i_1i_2i_3} - t^{i_2i_3i_1} + t^{i_2i_1i_3} - t^{i_3i_2i_1} \tag{3.78}$$

$$\ytableaushort{13,2} \quad D_1 = t^{i_1i_2i_3} - t^{i_2i_1i_3} + t^{i_3i_2i_1} - t^{i_3i_1i_2} \tag{3.79}$$

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:

$$6t^{i_1i_2i_3} = A + B + 2(C_1 + D_1)\ .$$

The labels 1 on C and D recall that under the action of the group S_3, these objects mix with another combination $C_2 = t^{i_1i_3i_2} - t^{i_3i_1i_2} + t^{i_2i_3i_1} - t^{i_1i_2i_3}$, (resp. $D_2 = t^{i_2i_1i_3} + t^{i_2i_3i_1} - t^{i_1i_2i_3} - t^{i_1i_3i_2}$) of t^{ijk} 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 $k \le n$ rows gives rise to a representation. One proves, and we admit, that this representation of GL(n) is irreducible and that its dimension is

$$\dim_Y^{(n)} = \frac{\Delta(f_1 + n - 1, f_2 + n - 2, \cdots, f_n)}{\Delta(n-1, n-2, \cdots, 0)} \tag{3.80}$$

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 as above the lengths of rows of the tableau Y. This is a polynomial of degree $m = \sum f_i$ in n. Compare with (3.62).

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 $1 \le i_1 \le i_2 \le \cdots \le i_n \le n$. One has to arrange in all possible ways $n-1$ $<$ signs between the m indices $i_1, \cdots, i_m$ to mark the successive blocks de 1, 2, ..., n. The seeked dimension is thus the binomial coefficient $\binom{n+m-1}{n} \equiv C^n_{n+m-1}$, in accord in this particular case with (3.80).

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

$$\ydiagram{3} + \ydiagram{1,1,1} + 2\,\ydiagram{2,1}$$

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), ... 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 $k \le n - 1$

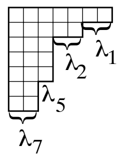
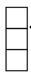


Figure 3.7: Correspondence between a Young diagram and a highest weight (or Dynkin labels). Here $Y \leftrightarrow \Lambda = (2, 2, 0, 0, 1, 0, 2)$

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 $\Lambda_3 =$ . And the correspondence between the highest weight Λ with Dynkin labels λ_i and the tableau Y with rows of length f_i is as follows

$$\Lambda = (\lambda_1, \cdots, \lambda_{n-1}) \leftrightarrow Y = (f_i = \sum_{j=i}^{n-1} \lambda_j) \ . \tag{3.81}$$

In other words, λ_k is the number of columns of Y of height k, see Fig. 3.7.

$\star$

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.

On octonions and exceptional groups, look at the exhaustive article by John C. Baez, *Bull. Amer. Math. Soc.* 39 (2002), 145-205 (also available on line); or P. Ramond, *Group Theory, A Physicist's Survey*, Cambridge 2010.

For the classification of real forms, see S. Helgason, *Differential Geometry, Lie groups and Symmetric spaces*, Academic Press, 1978, or Kirillov, *op. cit.* in chap. 2 .

Appendix F. The classical algebras of type A, B, C, D

F.1 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 $\sum_1^N e_i = 0$, $\langle e_i, e_j \rangle = \delta_{ij} - \frac{1}{N}$.
Roots $\alpha_{ij} = e_i - e_j$, $i \neq j = 1, \cdots N$; positive roots $\alpha_{ij} = e_i - e_j$, $i < j$; their number $|\Delta_+| = N(N-1)/2$; simple roots $\alpha_i := \alpha_{i\,i+1} = e_i - e_{i+1}$ $i = 1, \cdots, N-1$.
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

$$\begin{aligned} 2\rho &= (N-1)e_1 + (N-3)e_2 + \cdots + (N-2i+1)e_i + \cdots - (N-1)e_N \\ &= (N-1)\alpha_1 + 2(N-2)\alpha_2 + \cdots + i(N-i)\alpha_i + \cdots + (N-1)\alpha_{N-1}. \end{aligned} \tag{F.1}$$

$$\text{Cartan matrix } \langle \alpha_i, \alpha_j \rangle = \begin{cases} 2 & \text{if } i = j \\ -1 & \text{if } i = j \pm 1 \\ 0 & \text{otherwise} \end{cases}$$

Fundamental weights Λ_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: $W \equiv \mathcal{S}_N$ acts on the weights by permuting the e_i: $w \in W \leftrightarrow \bar{w} \in \mathcal{S}_N$: $w(e_i) = e_{\bar{w}(i)}$

Coxeter exponents $\{1, 2, \cdots, N-1\}$.

F.2 so($2l+1$) $= B_l$, $l \geq 2$

Rank $= l$, dimension $l(2l+1)$, Coxeter number $h = 2l$, dual Coxeter number $h^\vee = 2l-1$
e_i, $i = 1, \cdots, l$, $\langle e_i, e_j \rangle = \delta_{ij}$ a basis of $\mathbb{R}^l$.
Roots $\pm e_i$, $1 \leq i \leq l$ and $\pm e_i \pm e_j$, $1 \leq i < j \leq l$. Basis of simple roots $\alpha_i = e_i - e_{i+1}$, $i = 1, \cdots, l-1$, and $\alpha_l = e_l$.
Positive roots

$$\begin{aligned} e_i &= \sum_{i \leq k \leq l} \alpha_k, \quad 1 \leq i \leq l\,, \\ e_i - e_j &= \sum_{i \leq k < j} \alpha_k, \quad 1 \leq i < j \leq l\,, \\ e_i + e_j &= \sum_{i \leq k < j} \alpha_k + 2 \sum_{j \leq k \leq l} \alpha_k, \quad 1 \leq i < j \leq l\,, \end{aligned} \tag{F.2}$$

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

$$\begin{aligned} 2\rho &= (2l-1)e_1 + (2l-3)e_2 + \cdots + (2l-2i+1)e_i + \cdots + 3e_{l-1} + e_l \\ &= (2l-1)\alpha_1 + 2(2l-2)\alpha_2 + \cdots + i(2l-i)\alpha_i + \cdots + l^2\alpha_l. \end{aligned} \tag{F.3}$$

$$\text{Cartan matrix } \langle \alpha_i, \alpha_j^\vee \rangle = \begin{cases} 2 & \text{if } 1 \leq i = j \leq l \\ -1 & \text{if } 1 \leq i = (j \pm 1) \leq l-1 \\ -2 & \text{if } i = l-1\,, j = l \\ -1 & \text{if } i = l\,, j = l-1 \\ 0 & \text{otherwise} \end{cases}$$

Fundamental weights $\Lambda_i = \sum_{j=1}^{i} e_j$, $i = 1, \cdots, l-1$, $\Lambda_l = \frac{1}{2}\sum_{j=1}^{l} e_j$; hence $e_1 = \Lambda_1 = (1, 0, \cdots, 0)$, $e_i = \Lambda_i - \Lambda_{i-1} = (0, \cdots, -1, 1, 0 \cdots)$, $i = 2, \cdots, l-1$, $e_l = 2\Lambda_l - \Lambda_{l-1} = (0, \cdots, 0, -1, 2)$.
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: $W \equiv \mathcal{S}_l \ltimes (\mathbb{Z}_2)^l$, of order $2^l.l!$, acts on the weights by permuting the e_i and $e_i \mapsto (\pm 1)_i e_i$.

Coxeter exponents $\{1, 3, 5, \cdots, 2l-1\}$.

F.3. sp$(2l) = C_l$, $l \geq 2$

Rank $= l$, dimension $l(2l+1)$, Coxeter number $h = 2l$, dual Coxeter number $h^\vee = l+1$
e_i, $i = 1, \cdots, l$, $\langle e_i, e_j \rangle = \frac{1}{2}\delta_{ij}$ a basis of $\mathbb{R}^l$ (**Beware !** factor 2 to enforce the normalisation $\theta^2 = 2$). Basis of simple roots $\alpha_i = e_i - e_{i+1}$, $i = 1, \cdots, l-1$, 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

$$\begin{aligned} e_i - e_j &= \sum_{i \leq k < j} \alpha_k, \quad 1 \leq i < j \leq l\,, \\ e_i + e_j &= \sum_{i \leq k < j} \alpha_k + 2\sum_{j \leq k < l} \alpha_k + \alpha_l, \quad 1 \leq i < j \leq l\,, \\ 2e_i &= 2\sum_{i \leq k < l} \alpha_k + \alpha_l, \quad 1 \leq i \leq l\,, \end{aligned} \tag{F.4}$$

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

$$\begin{aligned} 2\rho &= 2le_1 + (2l-2)e_2 + \cdots + (2l-2i+2)e_i + \cdots + 4e_{l-1} + 2e_l \\ &= 2l\alpha_1 + 2(2l-1)\alpha_2 + \cdots + i(2l-i+1)\alpha_i + \cdots + (l-1)(l-2)\alpha_{l-1} + \frac{1}{2}l(l+1)\alpha_l. \end{aligned} \tag{F.5}$$

$$\text{Cartan matrix } \langle \alpha_i, \alpha_j^\vee \rangle = \begin{cases} 2 & \text{if } 1 \leq i = j \leq l \\ -1 & \text{if } 1 \leq i = (j \pm 1) \leq l-1 \\ -1 & \text{if } i = l-1\,, j = l \\ -2 & \text{if } i = l\,, j = l-1 \\ 0 & \text{otherwise} \end{cases}$$

Fundamental weights $\Lambda_i = \sum_{j=1}^{i} e_j$, $i = 1, \cdots, l$, hence $e_1 = \Lambda_1 = (1, 0, \cdots, 0)$, $e_i = \Lambda_i - \Lambda_{i-1} = (0, \cdots, -1, 1, 0 \cdots)$, $i = 2, \cdots, l$.
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: $W \equiv \mathcal{S}_l \ltimes (\mathbb{Z}_2)^l$, of order $2^l.l!$, acts on the weights by permuting the e_i and $e_i \mapsto (\pm 1)_i e_i$.

F.4. so$(2l) = D_l$, $l \geq 3$

Rank $= l$, dimension $l(2l-1)$, Coxeter number $=$ dual Coxeter number $h = 2l-2 = h^\vee$
e_i, $i = 1, \cdots, l$, $\langle e_i, e_j \rangle = \delta_{ij}$ a basis of $\mathbb{R}^l$.
Basis of simple roots $\alpha_i = e_i - e_{i+1}$, $i = 1, \cdots, l-1$, and $\alpha_l = e_{l-1} + e_l$.

Positive roots

$$\begin{aligned} e_i - e_j &= \sum_{i\le k<j} \alpha_k, \quad 1 \le i < j \le l\,, \\ e_i + e_j &= \sum_{i\le k<j} \alpha_k + 2 \sum_{j\le k<l-1} \alpha_k + \alpha_{l-1} + \alpha_l, \quad 1 \le i < j \le l-1\,, \\ e_i + e_l &= \sum_{i\le k\le l-2} \alpha_k + \alpha_l, \quad 1 \le i \le l-1\,, \end{aligned} \tag{F.6}$$

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

$$\begin{aligned} 2\rho &= 2(l-1)e_1 + 2(l-2)e_2 + \cdots + 2e_{l-1} \\ &= 2(l-1)\alpha_1 + 2(2l-3)\alpha_2 + \cdots + i(2l-i-1)\alpha_i + \cdots + \frac{l(l-1)}{2}(\alpha_{l-1} + \alpha_l). \end{aligned} \tag{F.7}$$

Weyl group: $W \equiv \mathcal{S}_l \ltimes (\mathbb{Z}_2)^{l-1}$, of order $2^{l-1}.l!$, acts on the weights by permuting the e_i and $e_i \mapsto (\pm 1)_i e_i$, with $\prod_i (\pm 1)_i = 1$.

Coxeter exponents $\{1, 3, 5, \cdots, 2l-3, l-1\}$, with $l-1$ appearing twice if l is even.

Cartan matrix $\langle\, \alpha_i, \alpha_j \,\rangle = \begin{cases} 2 & \text{if } 1 \le i = j \le l \\ -1 & \text{if } 1 \le i = (j \pm 1) \le l-2 \\ -1 & \text{if } (i,j) = (l-2, l) \text{ or } (l, l-2) \\ 0 & \text{otherwise} \end{cases}$

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 $i = 1, \cdots, l-2$; $\Lambda_{l-1} = \frac{1}{2}(e_1 + \cdots + e_{l-1} - e_l) = \frac{1}{2}(\alpha_1 + 2\alpha_2 + \cdots + (l-2)\alpha_{l-2}) + \frac{l}{2}\alpha_{l-1} + \frac{l-2}{2}\alpha_l$; $\Lambda_l = \frac{1}{2}(e_1 + \cdots + e_{l-1} + e_l) = \frac{1}{2}(\alpha_1 + 2\alpha_2 + \cdots + (l-2)\alpha_{l-2}) + \frac{l-2}{2}\alpha_{l-1} + \frac{l}{2}\alpha_l$.

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 $\mathrm{ad}\,(H)$ on such a vector X; conclude that $\mathrm{ad}\,(H)\mathrm{ad}\,(H')X = \sum_{\alpha\in\Delta} x^\alpha \alpha(H)\alpha(H')E_\alpha$ and taking into account that the eigenspace of each root α has dimension 1, cf point (∗) 2. of § 3.1.2, that the Killing form reads

$$(H, H') = \mathrm{tr}\,(\mathrm{ad}\,(H)\mathrm{ad}\,(H')) = \sum_{\alpha\in\Delta} \alpha(H)\alpha(H') \ . \tag{3.82}$$

2. One wants to show that roots α defined by (3.5) or (3.6) generate all the dual space $\mathfrak{h}^*$ of the Cartan subalgebra $\mathfrak{h}$. Prove that if it were not so, there would exist an element H of $\mathfrak{h}$ such that

$$\forall\alpha \in \Delta \quad \alpha(H) = 0 \ . \tag{3.83}$$

Using (3.82) 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 (3.83), show that H would commute with all H_i and all the E_α, 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

$$N_{\alpha\beta} = -N_{-\alpha,-\beta} \ . \tag{3.84}$$

2. Consider three roots satisfying $\alpha + \beta + \gamma = 0$. Writing the Jacobi identity for the triplet E_α, E_β, E_γ, show that $\alpha_{(i)}N_{\beta\gamma} + \text{cycl.} = 0$. Derive from it the relation

$$N_{\alpha\beta} = N_{\beta,-\alpha-\beta} = N_{-\alpha-\beta,\alpha} \ . \tag{3.85}$$

3. Considering the α-chain through β and the two integers p et q defined in § 3.2.1, write the Jacobi identity for E_α, $E_{-\alpha}$ and $E_{\beta+k\alpha}$, with $p \le k \le q$, 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 (3.85), show that the previous equation may be recast as

$$\langle\, \alpha, \beta + k\alpha \,\rangle = f(k) - f(k-1) \ . \tag{3.86}$$

4. What are $f(q)$ et $f(q-1)$? Show that the recursion relation (3.86) is solved by

$$f(k) = -(N_{\alpha,\beta+k\alpha})^2 = (k-q)\langle\, \alpha, \beta + \frac{1}{2}(k+q+1)\alpha \,\rangle \ . \tag{3.87}$$

What is $f(p-1)$? Show that the expression (3.21) is recovered. Show that (3.87) is in accord with (3.23). The sign of the square root is still to be determined . . . , see [Gi].

C. *Study of the $B_l = so(2l+1)$ and G_2 algebras*

1. $\underline{so(2l+1) = B_l,\ l \ge 2}$

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 2×2 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

$$\Delta = \{\pm e_i\ ,\ 1 \le i \le l\} \cup \{\pm e_i \pm e_j\ ,\quad 1 \le i < j \le l\}$$

What is the cardinal of Δ? Δ 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

$$\begin{aligned} e_i &= \sum_{i\le k\le l} \alpha_k, \quad 1 \le i \le l\,, \\ e_i - e_j &= \sum_{i\le k<j} \alpha_k, \quad 1 \le i < j \le l\,, \\ e_i + e_j &= \sum_{i\le k<j} \alpha_k + 2\sum_{j\le k\le l} \alpha_k, \quad 1 \le i < j \le l\,, \end{aligned} \tag{3.88}$$

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 ρ of positive roots.

h. The Weyl is the ("semi-direct") product $W \equiv \mathcal{S}_l \ltimes (\mathbb{Z}_2)^l$, which acts on the e_i (hence on weights and roots) by permutation and by independant sign changes $e_i \mapsto (\pm 1)_i e_i$. What is its order? In the case of B_2, check that assertion and draw the first Weyl chamber.

i. Show that the vectors $\Lambda_i = \sum_{j=1}^{i} e_j$, $i = 1, \cdots, l-1$, $\Lambda_l = \frac{1}{2}\sum_{j=1}^{l} e_j$ 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. $\underline{G_2}$

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 $\alpha_1 = e_1 - e_2$ and $\alpha_2 = -2e_1 + e_2 + e_3$ 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 ρ, 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

$$\Lambda_1 = 2\alpha_1 + \alpha_2 \qquad \Lambda_2 = 3\alpha_1 + 2\alpha_2$$

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. *Root systems. Folding of Dynkin diagrams*

One considers the simple roots α_i of the algebra su$(2n)$, numbered as in the lectures. (Beware! we do say $2n$!)

1. What is the rank of that algebra? What are the $\langle \alpha_i, \alpha_j \rangle$? Draw the corresponding Dynkin diagram. What is the symmetry of that diagram?

2. One then defines $\beta_i = (\alpha_i + \alpha_{2n-i})/\sqrt{2}$, for $i = 1, \cdots, n-1$ and $\beta_n = \alpha_n/\sqrt{2}$. Calculate the $\langle \beta_i, \beta_j \rangle$.

3. Show that the β form a root system and identify it.

4. More generally, any system of simple roots with the same lengths may be "folded" according to a possible symmetry of its Dynkin diagram and then gives rise to another Dynkin diagram. *With no calculation*, which diagram should be obtained in that manner, starting from the E_6 diagram?

E. *Dimensions of SU(3) representations*

We admit that the construction of § 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 δ^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 (3.64).

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 ϕ by a linear representation, and to each element g of the group corresponds a matrix or operator $D(g)$, independent of the point x

$$\phi(x) \mapsto D(g)\phi(x) \ . \tag{4.1}$$

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_i^μ of vanishing divergence (see *Chapitre* 0[1]) and to conserved

[1] Here and below in this chapter, *Chapitre* 0 refers to Chap. 0 of the French version, especially in its discussion of Noether's currents.

charges $Q_i = \int d\mathbf{x}\, j^0(\mathbf{x}, t)$, $i = 1, \cdots, \dim G$. These charges act on fields as infinitesimal generators, classically in the sense of the Poisson bracket, $\{Q_i, \phi(x)\}\delta\alpha^i = \delta\phi(x)$, 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

$$\phi(x) \mapsto e^{iq\alpha}\phi(x)\ ; \quad \phi^\dagger(x) \mapsto e^{-iq\alpha}\phi^\dagger(x)\,,$$

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* explicitely. 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

$$\partial_\mu j_i^\mu(x) = \frac{\partial \mathcal{L}(x)}{\partial \alpha^i}\ , \tag{4.2}$$

(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 $\mathbb{Z}_2$ symmetry under $x \to -x$, 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.

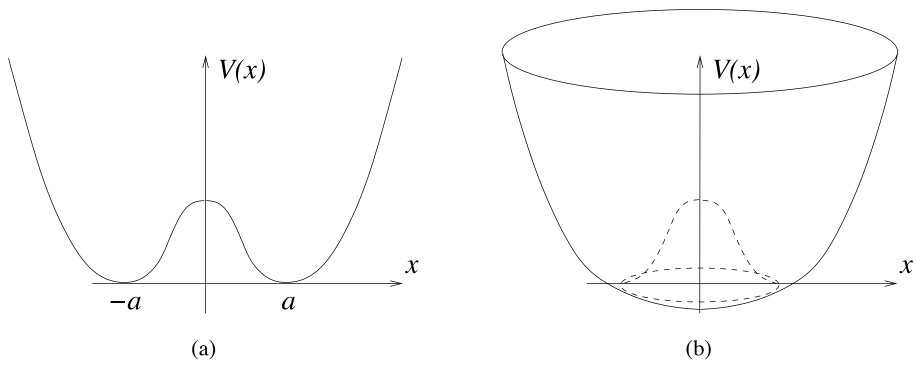


Figure 4.1: Potentials (a) with a "double well" ; (b) "mexican hat"

▷ *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 $\boldsymbol{\phi} = \{\phi^i\}$

$$\mathcal{L} = \frac{1}{2}(\partial\boldsymbol{\phi})^2 - \frac{1}{2}m^2\boldsymbol{\phi}^2 - \frac{\lambda}{4}(\boldsymbol{\phi}^2)^2 \tag{4.3}$$

is invariant under the O(n) rotation group. The Noether current $j^a_\mu = \partial_\mu\phi^i(T^a)_{ij}\phi^j$ (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 $V = \frac{1}{2}m^2\boldsymbol{\phi}^2 + \frac{\lambda}{4}(\boldsymbol{\phi}^2)^2$ is no longer at $\boldsymbol{\phi}^2 = 0$ but at some value v^2 of $\boldsymbol{\phi}^2$ such that $-m^2 = \lambda v^2$, see Fig. 4.1(b). The field ϕ "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

$$\langle\, 0|\boldsymbol{\phi}|0\,\rangle = v\hat{n}\ . \tag{4.4}$$

This "vev" breaks the initial invariance group $G = \mathrm{O}(n)$ down to its subgroup H that leaves invariant the vector $\langle\, 0|\boldsymbol{\phi}|0\,\rangle = v\hat{n}$, hence a group isomorphic to O$(n-1)$. The fact that a vacuum expectation value of a non invariant field be non zero, $\langle\, 0|\boldsymbol{\phi}|0\,\rangle \neq 0$, 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 $\boldsymbol{\phi} = (v+\sigma)\hat{n} + \boldsymbol{\pi}$, where $\boldsymbol{\pi}$ denote the $n-1$ components of the field $\boldsymbol{\phi}$ orthogonal to $\langle\, \boldsymbol{\phi}\,\rangle = v\hat{n}$ and determine the terms of $V(\sigma, \boldsymbol{\pi})$ that are linear and quadratic in the fields σ and $\boldsymbol{\pi}$; check that the linear term in σ vanishes (minimum of the potential), that σ has a non-zero mass term, but that the $\boldsymbol{\pi}$ 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, $G = \mathrm{O}(n)$, $H = \mathrm{O}(n-1)$, $d(G) - d(H) =$

$n-1$.

Let us give a simple proof of that theorem in the case of a Lagrangian field theory. We write $\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - V(\phi)$ with quite generic notations, ϕ 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 $\delta^a\phi_i$, $a = 1, \cdots, \dim G$. 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)

$$\frac{\partial V}{\partial\phi_i}\frac{\partial\delta^a\phi_i}{\partial\phi_j} + \frac{\partial^2 V}{\partial\phi_i\partial\phi_j}\delta^a\phi_i = 0$$

and evaluate it at $\phi(x) = v$, a (constant, x-independent) minimum of the potential : the first term vanishes, the second tells us

$$\frac{\partial^2 V}{\partial\phi_i\partial\phi_j}\Big|_{\phi=v}\delta^a v_i = 0\ , \tag{4.5}$$

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 φ 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" (eigenvectors 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

$$\mathcal{L} = \bar\psi i\partial\!\!\!/\psi + g(\bar\psi\gamma^\mu\psi)(\bar\psi\gamma_\mu\psi)\ , \tag{4.6}$$

where $\psi = \{\psi_\alpha\}_{\alpha=1,\cdots,N}$ is a N-component vector of 4-spinor fields. Note the absence of a mass term $\bar\psi\psi$ in (4.6). That Lagrangian is invariant under the action of two types of infinitesimal transformations

$$\begin{aligned}\delta_A\psi(x) &= \delta A\psi(x) \\ \delta_B\psi(x) &= \delta B\gamma_5\psi(x)\ ,\end{aligned} \tag{4.7}$$

where the matrices A and B are infinitesimal $N\times N$ antihermitian, that act on the "flavor" indices α but on spinor indices and hence commute with γ matrices. Recall that γ_5 is Hermitian and anticommutes with the γ_μ and check that $\delta_A\bar\psi = -\bar\psi\delta A$, $\delta_B\bar\psi = \bar\psi\delta B\gamma_5$. The conserved Noether currents are respectively

$$J^a_\mu = \bar\psi T^a\gamma_\mu\psi \qquad J^{a(5)}_\mu = \bar\psi T^a\gamma_5\gamma_\mu\psi\ , \tag{4.8}$$

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 γ_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 $(\gamma_5)^2 = I$ and that $\frac{1}{2}(I \pm \gamma_5)$ are thus projectors; one has thus $\bar\psi_L = \psi_L^\dagger \gamma_0 = \frac{1}{2}\bar\psi(I+\gamma_5)$, etc, and

$$\mathcal{L} = \bar\psi_L i\partial\!\!\!/ \psi_L + \bar\psi_R i\partial\!\!\!/ \psi_R + g(\bar\psi_L\gamma_\mu\psi_L + \bar\psi_R\gamma_\mu\psi_R)(\bar\psi_L\gamma^\mu\psi_L + \bar\psi_R\gamma^\mu\psi_R)$$

which is clearly invariant under the finite unitary transformations $\psi_L \to U_1\psi_L$, $\psi_R \to U_2\psi_R$, with U_1, $U_2 \in \mathrm{U}(N)$. The group of *chiral symmetry* is thus $\mathrm{U}(N)\times\mathrm{U}(N)$.

If we now introduce a mass term $\delta\mathcal{L} = -m\bar\psi\psi$ (which "couples" the left and right components ψ_L and ψ_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

$$\partial^\mu J_\mu^{a(5)}(x) \propto m\bar\psi T^a\gamma_5\psi \ . \tag{4.9}$$

The residual symmetry group is $\mathrm{U}(N)$, "diagonal" subgroup of $\mathrm{U}(N)\times\mathrm{U}(N)$ (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 σ and $\boldsymbol{\pi}$

$$\mathcal{L} = \bar\psi\big(i\partial\!\!\!/ + g(\sigma + i\boldsymbol{\pi}.\boldsymbol{\tau}\gamma_5)\big)\psi + \frac{1}{2}\left((\partial\boldsymbol{\pi})^2 + (\partial\sigma)^2\right) - \frac{1}{2}m^2(\sigma^2+\boldsymbol{\pi}^2) - \frac{\lambda}{4}(\sigma^2+\boldsymbol{\pi}^2)^2 \ , \tag{4.10}$$

in which the Pauli matrices have been exceptionally denoted by $\boldsymbol{\tau}$ not to confuse them with the field σ. The symmetry group is $\mathrm{U}(2)\times\mathrm{U}(2)$, with fields ψ_L, ψ_R and $\sigma + i\boldsymbol{\pi}.\boldsymbol{\tau}$ transforming respectively by the representations $(\frac{1}{2},0)$, $(0,\frac{1}{2})$ and $(\frac{1}{2},\frac{1}{2})$ of $\mathrm{SU}(2)\times\mathrm{SU}(2)$ (see exercise A). If $m^2 < 0$, the field $\phi = (\sigma,\boldsymbol{\pi})$ develops a non-zero vev, that may be oriented in the direction σ if one has initially introduced a small explicit breaking term $\delta\mathcal{L} = c\sigma$, 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 σ' has now a vanishing vev, one sees that the fermions have acquired a mass $m_\psi = -gv$, whereas the $\boldsymbol{\pi}$ are massless. This Lagrangian, the σ-model of Gell-Mann–Lévy, has been proposed as a model explaining the chiral symmetry breaking and the low mass of the π 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 $\partial_\mu J_5^\mu \neq 0$. 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. 0) that if the weak and electromagnetic interactions are neglected, *hadrons*, *i.e.* particules subject to strong interactions such as proton and neutron, π 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

$$Q = \frac{1}{2}\mathcal{B} + I_z \qquad \text{[for SU(2)]} \tag{4.11}$$

where a new quantum number $\mathcal{B}$ appears, the *baryonic charge*, supposed to be (additively) conserved in *all* interactions (until further notice). $\mathcal{B}$ is 0 for π mesons, 1 for "baryons" as proton or neutron, -1 for their antiparticles, 4 for an α 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 Λ^0 and $+1$ for the K^+ and the K^0, the reaction $p+\pi^- \to \Lambda^0+K^0$ conserves strangeness, whereas the observed decay $\Lambda^0 \to p+\pi^-$ violates that conservation law, as it proceeds through weak interactions. Relation (4.11) must be modified into the *Gell-Mann–Nishima relation*

$$Q = \frac{1}{2}\mathcal{B} + \frac{1}{2}S + I_z = \frac{1}{2}Y + I_z \ , \tag{4.12}$$

where we introduced the *hypercharge* Y, which, *at this stage*, equals $Y = \mathcal{B} + S$.

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 mesurable 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) is the natural candidate, as it possesses an irreducible 8-dimensional representation (see also 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 λ_a that play the role of Pauli matrices σ_i for su(2). These matrices are normalised by

$$\mathrm{tr}\,\lambda_a\lambda_b = 2\delta_{ab} \ . \tag{4.13}$$

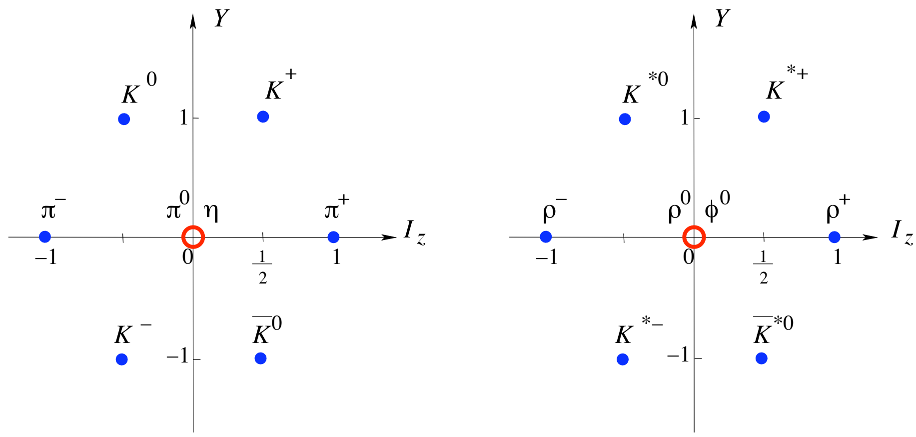


Figure 4.2: Octets of pseudoscalar ($J^P = 0^-$) and of vector mesons ($J^P = 1^-$)

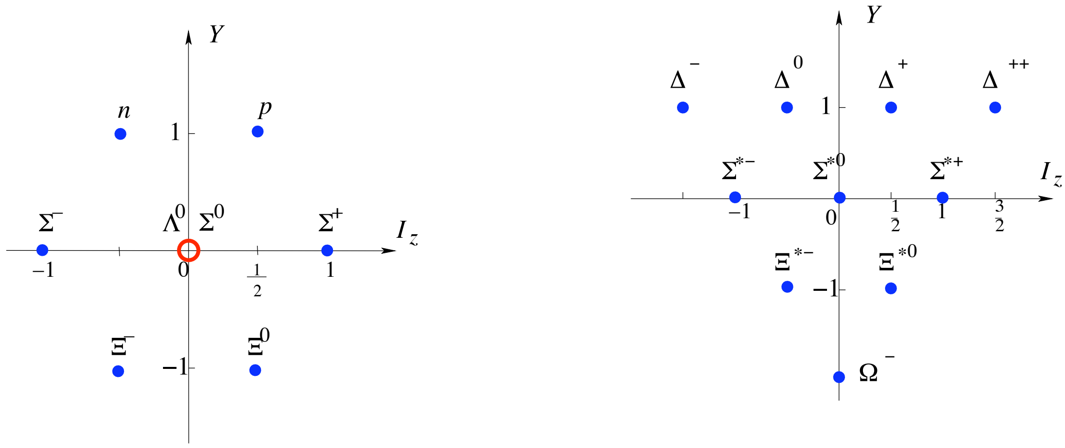


Figure 4.3: Baryon octet ($J^P = \frac{1}{2}^+$) and decuplet ($J^P = \frac{3}{2}^+$)

λ_1 and λ_2, λ_4 and λ_5, λ_6 and λ_7 have the same matrix elements as σ_1 and σ_2 at the $*$ locations $\begin{pmatrix} . & * & . \\ * & . & . \\ . & . & . \end{pmatrix}$, $\begin{pmatrix} . & . & * \\ . & . & . \\ * & . & . \end{pmatrix}$ and $\begin{pmatrix} . & . & . \\ . & . & * \\ . & * & . \end{pmatrix}$ respectively, where dots stand for zeros. The two generators of the Cartan algebra are

$$\lambda_3 = \begin{pmatrix} 1 & . & . \\ . & -1 & . \\ . & . & . \end{pmatrix} \qquad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & . & . \\ . & 1 & . \\ . & . & -2 \end{pmatrix} . \tag{4.14}$$

The charges I_z and Y are then representatives in the representation under study of $\frac{1}{2}\lambda_3$ and $\frac{1}{\sqrt{3}}\lambda_8$. See exercise B for the change of coordinates from (λ_1, λ_2) (Dynkin labels of some representation, not to be confused with the above matrices !) to (I_z, Y).

The matrices λ_a satisfy commutation relations

$$[\lambda_a, \lambda_b] = 2if_{abc}\lambda_c \tag{4.15}$$

with structure constants (real and completely antisymmetric) f_{abc} of the su(3) Lie algebra. It is useful to also consider the anticommutators

$$\{\lambda_a, \lambda_b\} = \frac{4}{3}\delta_{ab} + 2d_{abc}\lambda_c . \tag{4.16}$$

Thanks to (4.13), (4.15) and (4.16) may be rewritten as $\mathrm{tr}\,([\lambda_a,\lambda_b]\lambda_c)=4if_{abc}$, $\mathrm{tr}\,(\{\lambda_a,\lambda_b\}\lambda_c)=4d_{abc}$. 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 $\mathcal{B}=1$ are easily identified. Contrary to isospin symmetry, the SU(3) symmetry [2] 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 $\bar{3}$. 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

$$\Phi=\begin{pmatrix}\frac{1}{\sqrt2}\pi^0-\frac{1}{\sqrt6}\eta & \pi^+ & K^+\\ \pi^- & -\frac{1}{\sqrt2}\pi^0-\frac{1}{\sqrt6}\eta & K^0\\ K^- & \bar{K}^0 & \sqrt{\frac23}\eta\end{pmatrix}\,, \tag{4.17}$$

and

$$\Psi=\begin{pmatrix}\frac{1}{\sqrt2}\Sigma^0-\frac{1}{\sqrt6}\Lambda & \Sigma^+ & p\\ \Sigma^- & -\frac{1}{\sqrt2}\Sigma^0-\frac{1}{\sqrt6}\Lambda & n\\ \Xi^- & \Xi^0 & \sqrt{\frac23}\Lambda\end{pmatrix}\,. \tag{4.18}$$

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

$$Q=I_z+\frac12 Y=\frac13\begin{pmatrix}2&0&0\\0&-1&0\\0&0&-1\end{pmatrix}\qquad Y=\frac13\begin{pmatrix}1&0&0\\0&1&0\\0&0&-2\end{pmatrix} \tag{4.19}$$

act in the adjoint representation by commutation and one has indeed

$$[Q,\Phi]=\begin{pmatrix}0&\pi_+&K_+\\-\pi_-&0&0\\-K_-&0&0\end{pmatrix}\qquad [Y,\Phi]=\begin{pmatrix}0&0&K_+\\0&0&K_0\\-K_-&-\bar{K}_0&0\end{pmatrix}\,.$$

[2] said to be "of flavour", according to the modern terminology, but called "unitary symmetry" or "eightfold way" at the time of Gell-Mann and Ne'eman...

Exercises : (i) with no further calculation, what is $[I_z, \Phi]$? Check.
(ii) Compute $\mathrm{tr}\,\Phi^2$, 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,

$$8 \otimes 8 = 1 \oplus 8 \oplus 8 \oplus 10 \oplus \overline{10} \oplus 27 \,. \tag{4.20}$$

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 Φ and Ψ. 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 Ψ and Φ are written as traceless 3×3 matrices as in (4.17) and (4.18), $\Psi = \{\psi_j{}^i\}$ and $\Phi = \{\phi_k{}^i\}$, these two couplings read

$$\mathrm{tr}\,\bar\Psi\Psi\Phi = \bar\psi_j{}^i \psi_i{}^k \phi_k{}^j \quad \text{and} \quad \mathrm{tr}\,\bar\Psi\Phi\Psi = \bar\psi_j{}^i \phi_i{}^k \psi_k{}^j \tag{4.21}$$

(this compact writing omits indices of Dirac spinors, a possible γ_5 matrix, etc). An often preferred expression uses the sum and difference of these two terms, hence $\mathrm{tr}\,\bar\Psi[\Phi,\Psi]$ and $\mathrm{tr}\,\bar\Psi\{\Phi,\Psi\}$, 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 $\mathcal{N}$ and $\mathcal{P}$: $\mathcal{N}_i + \mathcal{P}_i \to \mathcal{N}_f + \mathcal{P}_f$? (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

$$\langle \mathcal{N}_f \mathcal{P}_f | \mathcal{T} | \mathcal{N}_i \mathcal{P}_i \rangle = \sum_{r=1}^{8} A_r(s,t) \, \langle (I, I_z, Y)_{(\mathcal{N}_f)}, (I, I_z, Y)_{(\mathcal{P}_f)} | r, (I, I_z, Y)_{(r)} \rangle \langle r, (I, I_z, Y)_{(r)} | (I, I_z, Y)_{(\mathcal{N}_i)}, (I, I_z, Y)_{(\mathcal{P}_i)} \rangle \tag{4.22}$$

(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 Φ_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 $\mathrm{tr}\,(\Phi_{P1}\Phi_{P2}\Phi_{P3}\Phi_{P4})$ and $\mathrm{tr}\,(\Phi_{P1}\Phi_{P2})\,\mathrm{tr}\,(\Phi_{P3}\Phi_{P4})$ 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

$$\mathcal{L}_{em} = -qj^\mu A_\mu \tag{4.23}$$

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 $Q = \int d^3x j_0(\mathbf{x}, t)$, 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^μ 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

$$\langle\, B|j_\mu(x)|B'\,\rangle = e^{ikx}\bar{u}(F_e^{BB'}(k^2)\gamma_\mu + F_m^{BB'}(k^2)\sigma_{\mu\nu}k^\nu)u' \tag{4.24}$$

where $\bar{u}$ 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 $F_m^{BB}(0)$ 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 8×8 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

$$F_{e,m}^{BB'}(k^2) = F_{e,m}^{(1)}(k^2)\,\mathrm{tr}\,\bar{B}QB' + F_{e,m}^{(2)}(k^2)\,\mathrm{tr}\,\bar{B}B'Q$$

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 $\mathrm{tr}\,\bar{B}QB'$ means the coefficient of $\bar{B}B'$ in the matrix trace $\mathrm{tr}\,\bar{\Psi}Q\Psi$, and likewise for $\mathrm{tr}\,\bar{B}B'Q$. For example, the magnetic moment of the neutron $\mu(n)$ is proportional to the magnetic term in

$\bar{n}n$, 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

$$\begin{aligned} \mu(n) &= \mu(\Xi^0) = 2\mu(\Lambda) = -2\mu(\Sigma^0) \qquad & \mu(\Sigma^+) &= \mu(p) \\ \mu(\Xi^-) &= \mu(\Sigma^-) = -(\mu(p) + \mu(n)) \qquad & \mu(\Sigma^0 \to \Lambda) &= \frac{\sqrt{3}}{2}\mu(n) \ , \end{aligned} \tag{4.25}$$

where the last quantity is the *transition magnetic moment* $\Sigma^0 \to \Lambda$. 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 $\Sigma^0 \to \Lambda$ is determined from the cross-section $\Lambda \to \Sigma^0$ in the Coulomb field of a heavy nucleus. One reads in tables

$$\begin{aligned} \mu(p) &= 2.792847351 \pm 0.000000028\,\mu_N \qquad & \mu(n) &= -1.9130427 \pm 0.0000005\,\mu_N \\ \mu(\Lambda) &= -0.613 \pm 0.004\,\mu_N & |\mu(\Sigma^0 \to \Lambda)| &= 1.61 \pm 0.08\,\mu_N \\ \mu(\Sigma^+) &= 2.458 \pm 0.010\,\mu_N & \mu(\Sigma^-) &= -1.160 \pm 0.025\,\mu_N \\ \mu(\Xi^0) &= -1.250 \pm 0.014\,\mu_N & \mu(\Xi^-) &= -0.6507 \pm 0.0025\,\mu_N \end{aligned} \tag{4.26}$$

where μ_N is the nuclear magneton, $\mu_N = \frac{e\hbar}{2m_p} = 3.152\ 10^{-14}\,\mathrm{MeV\ T}^{-1}$.

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 ΔM transforms under the representation 8, more precisely, since it must have vanishing isospin and hypercharge, that it transforms like the η or Λ 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 $8 = 3 \otimes \bar{3} \ominus 1$) ; 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 3×3 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 $G_{11} = I_z + \frac{1}{2}Y = Q$, 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, $\mathrm{cof}G_{33} = \frac{1}{4}Y^2 - I_z^2 - 2I_+I_- = \frac{1}{4}Y^2 - \vec{I}^2$.

One gets in that way a mass formula for any representation (any multiplet)

$$M = m_1 + m_2 Y + m_3(I(I+1) - \frac{1}{4}Y^2) \tag{4.27}$$

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[3]. For the decuplet, show that the same formula gives equal mass differences between the four particles Δ, Σ^*, Ξ^* and Ω^-. The latter result led to an accurate prediction of the existence and mass of the Ω^- particle, which was regarded as one of the major achievments of SU(3). For the octet of pseudoscalar mesons, the mass formula is empirically better verified in terms of the square masses

$$m_K^2 = \frac{3m_\eta^2 + m_\pi^2}{4} \ .$$

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

Quarks	u	d	s	$\bar{u}$	$\bar{d}$	$\bar{s}$
Isospin I_z	$\frac{1}{2}$	$-\frac{1}{2}$	0	$-\frac{1}{2}$	$\frac{1}{2}$	0
Baryonic charge $\mathcal{B}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$
Strangeness S	0	0	-1	0	0	1
Hypercharge Y	$\frac{1}{3}$	$\frac{1}{3}$	$-\frac{2}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$
Electric charge Q	$\frac{2}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

Table 1. 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, $3 \otimes \bar{3} = 1 \oplus 8$

[3]The observed masses of these hadrons are $M_N \approx 939$ MeV/c^2, $M_\Lambda = 1116$ MeV/c^2, $M_\Sigma \approx 1195$ MeV/c^2, $M_\Xi \approx 1318$ MeV/c^2 ; 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.

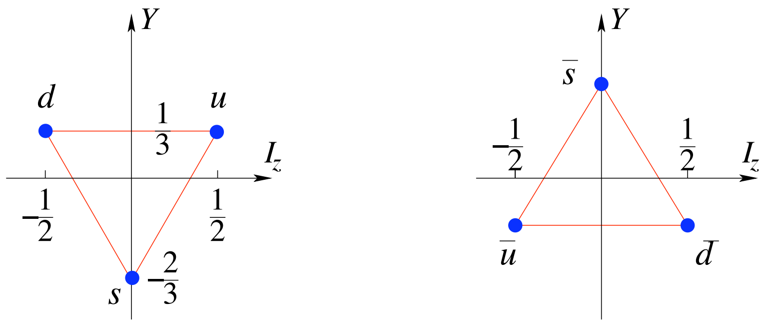


Figure 4.4: The triplets of quarks and antiquarks.

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

$$
\begin{aligned}
&p = uud,\ n = udd, \quad \Omega^- = sss, \qquad \Delta^{++} = uuu,\ \cdots,\ \Delta^- = ddd, \\
&\pi^+ = u\bar{d},\ \pi^0 = \frac{(u\bar{u} - d\bar{d})}{\sqrt{2}},\ \pi^- = d\bar{u}\,, \qquad \eta_8 = \frac{(u\bar{u} + d\bar{d} - 2s\bar{s})}{\sqrt{6}}, \quad K^+ = u\bar{s},\ K^0 = d\bar{s} \quad \text{etc.}
\end{aligned}
\tag{4.29}
$$

The quark model interprets the singlet that appears in the product $3 \times \bar{3}$ as a bound state $\eta_1 = \frac{(u\bar{u}+d\bar{d}+s\bar{s})}{\sqrt{3}}$. The physically observed particles η (masse 548 MeV) and η' (958 MeV) result from a "mixing" (*i.e.* a linear combination) due to SU(3) breaking interactions of these η_1 and η_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)

$$\mathcal{L}_{\text{Fermi}} = -\frac{G}{\sqrt{2}} J^\rho(x) J^\dagger_\rho(x) \tag{4.30}$$

where G is the Fermi constant, whose value (in units where $\hbar = c = 1$) is

$$G = (1,026 \pm 0,001) \times 10^{-5} m_p^{-2}\ . \tag{4.31}$$

(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, $\mathcal{L}_{\text{Fermi}}$ offers a good description of physics, whence the name "effective".) The current J_ρ is the sum of a leptonic and a hadronic contributions

$$J_\rho(x) = l_\rho(x) + h_\rho(x) \tag{4.32}$$

The leptonic current

$$l_\rho(x) = \bar{\psi}_e(x)\gamma_\rho(1-\gamma_5)\psi_{\nu_e} + \bar{\psi}_\mu(x)\gamma_\rho(1-\gamma_5)\psi_{\nu_\mu} \quad [+\bar{\psi}_\tau(x)\gamma_\rho(1-\gamma_5)\psi_{\nu_\tau}]$$

is the sum of contributions of the lepton families (or *generations*), e, μ (and τ that we omit in this first approach). The hadronic current, if one restricts to the first two generations, reads

$$h_\rho = \cos\theta_C\, h_\rho^{(\Delta S=0)} + \sin\theta_C\, h_\rho^{(\Delta S=1)} \tag{4.33}$$

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,

$$\begin{aligned} h_\rho^{(\Delta S=0)} &= (V_\rho^1 - iV_\rho^2) - (A_\rho^1 - iA_\rho^2) &(4.34)\\ h_\rho^{(\Delta S=1)} &= (V_\rho^4 - iV_\rho^5) - (A_\rho^4 - iA_\rho^5)\ . &(4.35)\end{aligned}$$

The vector currents $V_\rho^{1,2,3}$ are the Noether currents of isospin, the other components of V_ρ 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 $G\langle\, p|h_\rho^{(\Delta S=0)}|n\,\rangle = \bar u_p\gamma_\rho(G_V(q^2) - G_A(q^2)\gamma_5)u_n$ 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 than the combination $j_\rho = V_\rho^3 + \frac{1}{\sqrt{3}}V_\rho^8$. In the quark model, these hadronic currents have the form

$$V_\rho^a(x) = \bar q(x)\frac{\lambda^a}{2}\gamma_\rho\, q(x) \qquad A_\rho^a(x) = \bar q(x)\frac{\lambda^a}{2}\gamma_\rho\gamma_5\, q(x)\ . \tag{4.36}$$

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 two 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 is 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 of 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

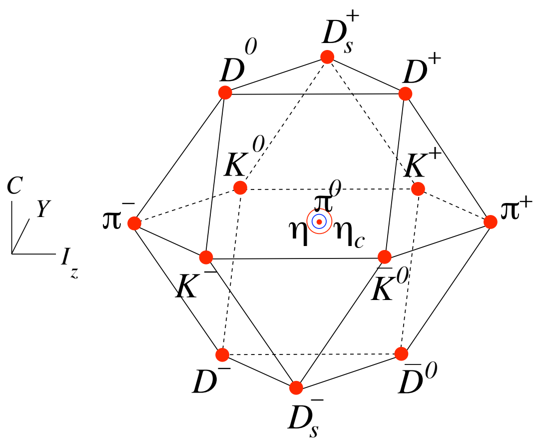


Figure 4.5: Mesons of spin $J^P = 0^-$ of the representation 15 of SU(4)

group SU(6) is very strongly broken, as attested by masses of the 6 quarks[4]

$$\begin{aligned} m_u &\approx 1.5 - 4\,\mathrm{MeV}\,, \qquad m_d \approx 4 - 8\,\mathrm{MeV}\,, \qquad m_s \approx 80 - 130\,\mathrm{MeV} \\ m_c &\approx 1.15 - 1.35\,\mathrm{GeV}\,, \qquad m_b \approx 4 - 5\,\mathrm{GeV}\,, \qquad m_t \approx 175\,\mathrm{GeV} \end{aligned} \tag{4.37}$$

and this limits its usefulness. One may however rewrite (4.12) in the form

$$Q = \frac{1}{2}Y + I_z \qquad Y = \mathcal{B} + S + C + B + T$$

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

Quarks	u	d	s	c	b	t
Isospin I_z	$\frac{1}{2}$	$-\frac{1}{2}$	0	0	0	0
Baryonic charge $\mathcal{B}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
Strangeness S	0	0	-1	0	0	0
Charm C	0	0	0	1	0	0
Beauty B	0	0	0	0	-1	0
Truth T	0	0	0	0	0	1
Hypercharge Y	$\frac{1}{3}$	$\frac{1}{3}$	$-\frac{2}{3}$	$\frac{4}{3}$	$-\frac{2}{3}$	$\frac{4}{3}$
Electric charge Q	$\frac{2}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{2}{3}$

Table 2. Quantum numbers of quarks $u,\ d,\ s,\ c,\ b,\ t$

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 $\mathrm{SU}(3)_c$.

[4]One 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.

Considerations leading to that triplicating hypothesis are first the study of the Δ^{++} 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 $\pi^0 \to 2\gamma$ is proportional to the sum $\sum Q^2 I_z$ over the set of fermionic constituents of the π^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 $\mathrm{SU}(3)_c$ 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 $\mathrm{SU}(3)_c$, 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 $\mathrm{SU}(3)_c$ 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^-, ν_e), (μ^-, ν_μ), (τ^-, ν_τ). 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.

In the discussion of SU(3) breakings, I followed
S. Coleman, *Aspects of Symmetry*, Cambridge Univ. Press 1985.

For a more recent presentation of flavor physics, see
K. Huang, *Quarks, Leptons and Gauge Fields*, World Scientific 1992.

All the properties of particles mentionned in this chapter may be found in the tables of the Particle Data Group, on line on the site `http://pdg.lbl.gov/2012/reviews/contents_sports.html`

⋆

Exercises for chapter 4

A. *Sigma model and chiral symmetry breaking*
Consider the Lagrangian (4.10) and define $W = \sigma + i\boldsymbol{\pi}\boldsymbol{\tau}$.

1. Compute $\det W$. Show that one may write $\mathcal{L}$ in terms of $\psi_{L,R}$ and W as

$$\mathcal{L} = \bar{\psi}_R i\partial\!\!\!/ \psi_R + \bar{\psi}_L i\partial\!\!\!/ \psi_L + g(\bar{\psi}_L W \psi_R + \bar{\psi}_R W^\dagger \psi_L) + \mathcal{L}_K - \frac{1}{2}m^2 \det W - \frac{\lambda}{4}(\det W)^2$$

where $\mathcal{L}_K$ is the kinetic term of the fields $(\sigma, \boldsymbol{\pi})$. One may also give that term the form $\mathcal{L}_K = \frac{1}{2}(\det \partial_0 W - \sum_{i=1}^3 \det \partial_i W)$ (which looks a bit odd, but which is indeed Lorentz invariant!).

2. Show that $\mathcal{L}$ is invariant under transformations of SU(2) × SU(2) with $\psi_L \to U\psi_L$, $\psi_R \to V\psi_R$, provided W transforms in a way to be specified. Justify the assertion made in §4.1.2 : ψ_L, ψ_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 σ, $\langle \sigma \rangle = v$, show that the field ψ acquires a mass $M = -gv$.

B. *Changes of basis in SU(3)*
In SU(3), write the change of basis which transforms the weights Λ_1, Λ_2 of Chap. 3 into the axes used in figures 2, 3 and 4. Derive the transformation of the coordinates (λ_1, λ_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 $\mathcal{BD} \to \mathcal{BD}$, where $\mathcal{B}$ and $\mathcal{D}$ 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

$$A^3 - (\operatorname{tr} A)A^2 + \frac{1}{2}\left((\operatorname{tr} A)^2 - \operatorname{tr} A^2\right)A - \det A = 0$$

implies a relation between $\operatorname{tr} A^4$ and $(\operatorname{tr} A^2)^2$.

b. If the group SU(3) acts on A by $A \to UAU^\dagger$, 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 $x_1x_2x_3x_4$. Show that this gives an identity of the form (Burgoyne's identity)

$$\sum_P \mathrm{tr}\,(A_{P1}A_{P2}A_{P3}A_{P4}) = a \sum_P \mathrm{tr}\,(A_{P1}A_{P2})\,\mathrm{tr}\,(A_{P3}A_{P4}) \tag{4.38}$$

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 $A_1, \cdots, A_4$, invariant under the action of SU(3) $A_i \to UA_iU^\dagger$ 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 Φ of the *pseudoscalar* meson octet, see (4.17).

a. Why is it natural to impose that this Lagrangian be even in the field Φ ?

b. Using the results of Problem 1., write the most general form of an SU(3) invariant Lagrangian, of degree less or equal to 4 (for renormalizability) and even in Φ.

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 tr Φ^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 Φ 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 $\mathcal{L}_{em}(x) = -qj^\mu(x)A_\mu(x)$. If $|B\,\rangle$ is a baryon state, one should thus compute

$$\delta M_B = \langle\, B|(\int d^4x\, \mathcal{L}_{em})^2|B\,\rangle\ . \tag{4.39}$$

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

$$(8 \otimes 8)_S = 1 \oplus 8 \oplus 27 \quad . \tag{4.40}$$

d. i) What is then the number of invariant amplitudes contributing to δM_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 Ψ, $\bar{\Psi}$ 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 $\Delta_{em} M = \alpha \mathrm{tr}\, \bar{\Psi} Q^2 \Psi + \beta \mathrm{tr}\, \bar{\Psi}\Psi Q^2 + \gamma \mathrm{tr}\, \bar{\Psi} Q \Psi Q$, (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 $\bar{n}n$, etc, and check the relation

$$M_{\Xi^-} - M_{\Xi^0} = M_{\Sigma^-} - M_{\Sigma^+} + M_p - M_n \ . \tag{4.41}$$

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 ?

⋆

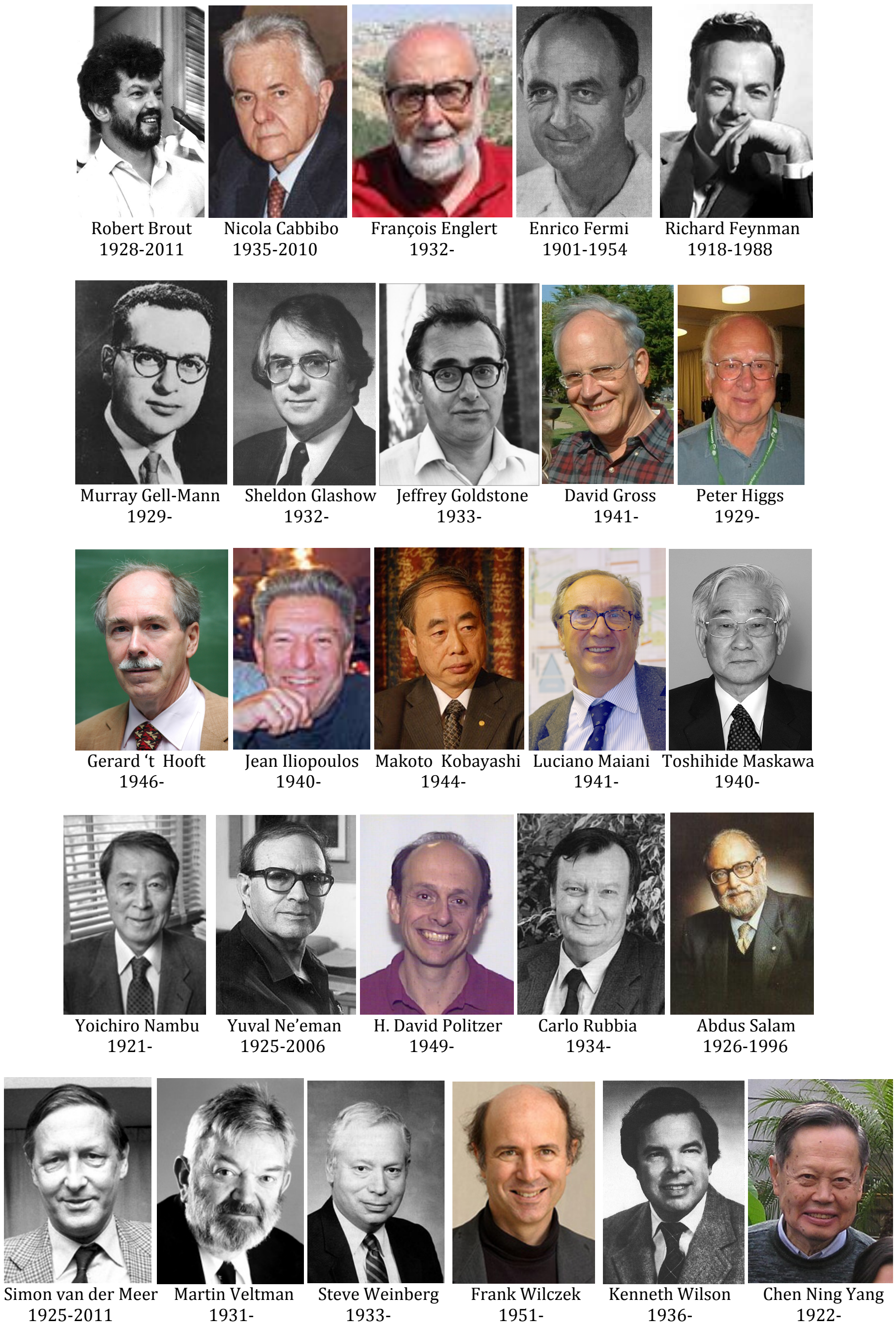


Figure 4.6: Some of the physicists mentionned in the second part of these notes.

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

$$\mathcal{L} = \bar\psi(i\partial\!\!\!/ - eA\!\!\!/ - m)\psi - \frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) \tag{5.1}$$

is invariant under infinitesimal gauge transformations

$$\begin{aligned} \delta A_\mu(x) &= -\partial_\mu \delta\alpha(x) \\ \delta\psi(x) &= ie\delta\alpha(x)\psi(x)\ , \end{aligned} \tag{5.2}$$

since the electromagnetic field tensor

$$F_{\mu\nu} = (\partial_\mu A_\nu - \partial_\nu A_\mu)$$

is invariant, and the combination

$$iD\!\!\!\!/\,\psi(x) := (i\partial\!\!\!/ - eA\!\!\!/)\psi(x)$$

also transforms as ψ. The finite form of these transformations is readily written

$$\begin{aligned} A_\mu(x) &\mapsto A_\mu(x) - \partial_\mu \alpha(x) \\ \psi(x) &\mapsto e^{ie\alpha(x)}\psi(x) \ , \end{aligned} \tag{5.3}$$

which shows that the transformations give a local (*i.e.* x dependent) version of those of the group U(1) or $\mathbb{R}$ (depending whether one identifies or not α 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 ψ 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 ϕ, whose contribution to the Lagrangian reads

$$\delta\mathcal{L} = [(\partial_\mu - iqA_\mu)\phi^*]\,[(\partial^\mu + iqA_\mu)\phi] - V(\phi^*\phi) \tag{5.4}$$

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 ψ be a field (which we denote as a fermion field, but this is irrelevant) transforming under G by some representation $\mathcal{D}$. 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

$$\delta\psi(x) = T_a \delta\alpha^a \psi(x) \ . \tag{5.5}$$

(In this section, we denote t_a the corresponding matrices in the adjoint representation.) To extend the notion of local transformation, we need a gauge field A_μ, which allows to construct a covariant derivative $D_\mu\psi$. It is natural to consider that A_μ lives in the Lie algebra of G, as it is associated with infinitesimal transformations of the group, and hence it carries indices of the adjoint representation

$$A_\mu(x) = \{A^a_\mu(x)\} \tag{5.6}$$

or equivalently, is represented in any representation by the antihermitian matrix[2]

$$A_\mu(x) = T_a A^a_\mu(x) \ . \tag{5.7}$$

The covariant derivative reads

$$D_\mu\psi(x) := (\partial_\mu - A_\mu(x))\psi(x) \ , \tag{5.8}$$

[1] An additional term in the Lagrangan 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.

or, componentwise

$$D_\mu \psi_A(x) := \left(\partial_\mu \delta_{AB} - A^a_\mu(x)\, (T_a)_A{}^B\right) \psi_B(x)\ . \tag{5.9}$$

That covariant derivative does transform as ψ, just like in the abelian case, provided one imposes that A_μ transforms according to

$$\begin{aligned} \delta A^a_\mu(x) &= \partial_\mu \delta\alpha^a(x) + C_{bc}{}^a \delta\alpha^b(x) \\ &= \left(\partial_\mu \delta_{ab} - A^c_\mu(x)(t_c)_a{}^b\right) \delta\alpha^b(x) = (D_\mu \delta\alpha)^a(x)\ . \end{aligned} \tag{5.10}$$

The term $\partial_\mu \delta\alpha^a(x)$ notwithstanding, one sees that $\{A^a_\mu\}$ transforms as the adjoint representation (whose matrices are $(t_c)_a{}^b = -C_{bc}{}^a$). Lastly a field tensor transforming in a covariant way (*i.e.* without any inhomogeneous term in $\partial\delta\alpha^a(x)$) may be constructed

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - [A_\mu, A_\nu] \tag{5.11}$$

or in components

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - C_{bc}{}^a A^b_\mu A^c_\nu\ . \tag{5.12}$$

One proves, after some algebra and using the Jacobi identity, that

$$\delta F^a_{\mu\nu}(x) = C_{bc}{}^a \delta\alpha^b(x) F^c_{\mu\nu}(x)\ , \tag{5.13}$$

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,

$$\begin{aligned} \psi(x) &\mapsto \mathcal{D}(g(x))\psi(x) \\ A_\mu = A^a_\mu T_a &\mapsto \mathcal{D}(g(x))(-\partial_\mu + A_\mu(x))\mathcal{D}(g^{-1}(x))\ , \end{aligned} \tag{5.14}$$

(with $\mathcal{D}$ the representation carried by ψ), and for the covariant derivative acting on ψ,

$$D_\mu \psi(x) \mapsto \mathcal{D}(g(x)) D_\mu \psi(x) \tag{5.15}$$

or equivalently[3]

$$D_\mu \mapsto \mathcal{D}(g(x)) D_\mu \mathcal{D}(g^{-1}(x))\ . \tag{5.16}$$

Now one verifies easily that in a given representation

$$[D_\mu, D_\nu] = -F_{\mu\nu} := -F^a_{\mu\nu} T_a \tag{5.17}$$

from which follows that $F_{\mu\nu}(x) \mapsto \mathcal{D}(g(x)) F_{\mu\nu} \mathcal{D}(g^{-1}(x))$, and in particular, in the adjoint representation, the finite transformation of $F_{\mu\nu} = F^a_{\mu\nu} t_a$ is

$$F_{\mu\nu}(x) \mapsto g(x) F_{\mu\nu}(x) g^{-1}(x)\ , \tag{5.18}$$

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 ∂_μ contained in D_μ acts on everything sitting on its right, whereas in the second equation (5.14), it acts only on $\mathcal{D}(g^{-1}(x))$.

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.*

$$F_{\mu\nu} = 0 \iff A_\mu(x) = (\partial_\mu g(x))\, g^{-1}(x) \ . \tag{5.19}$$

(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 $\Leftarrow$ 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

$$\gamma(C) := P \exp\left(\int_C dx^\mu A_\mu(x)\right) \tag{5.20}$$

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)

$$\gamma(C) \mapsto g(x)\gamma(C)g^{-1}(x_0) \ . \tag{5.21}$$

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 $\gamma_{\mathcal{D}}(C) \mapsto \mathcal{D}(g(x))\gamma_{\mathcal{D}}(C)\mathcal{D}(g^{-1}(x_0))$.
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 $A_\mu \to g(x)(-\partial_\mu + A_\mu(x))g^{-1}(x)$, show that $\gamma(C) \to g(x + dx)\gamma(C)g^{-1}(x)$. The result for a finite curve follows by recombining these infinitesimal elements.

Given an objet, like the field ψ, 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

$$\gamma(C) \approx \exp \frac{1}{2} \int_{\mathcal{S}} dx^\mu \wedge dx^\nu F_{\mu\nu} \ , \tag{5.22}$$

where the integration is carried out on an infinitesimal surface $\mathcal{S}$ of boundary C.
Exercise: Prove that statement by considering an elementary square circuit extending from x along the coordinate axes μ and ν: $(x \to x + dx^\mu \to x + dx^\mu + dx^\nu \to x + dx^\nu \to x)$, and expand to second order in dx to find $\gamma(C) \approx 1 + dx^\mu dx^\nu F_{\mu\nu}$ (with no summation over μ, ν). Hint: the use of the commutator formula of Chap. 1, (1.22), 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

$$W(C) = \mathrm{tr}\ \gamma(C) = \mathrm{tr}\, P \exp \oint dx^\mu A_\mu(x) \tag{5.23}$$

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 $\mathrm{tr}\, F_{\mu\nu}F^{\mu\nu}$, $\bar\psi(i\partial\!\!\!/ - A\!\!\!/)\psi$ 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 $A = A^a T_a$ 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 § 5.3.1 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 content. 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 ψ, 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 ψ via the minimal coupling reads

$$\mathcal{L} = \frac{1}{2\mathrm{g}^2}\mathrm{tr}\,(F_{\mu\nu}F^{\mu\nu}) + \bar\psi\big(i(\partial\!\!\!/ - A\!\!\!/) - m\big)\psi\ , \tag{5.24}$$

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 $F_{\mu\nu} = F^a_{\mu\nu}T_a$. 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, $\mathrm{tr}\, T_a T_b = -T_R \delta_{ab}$, 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 $\operatorname{tr} T_a T_b = -\frac{1}{2}\delta_{ab}$. 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 (5.24) ressembles very much the Lagrangian of the abelian case (5.1), after a change $iA \to \mathrm{g}A$ 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)×SU(2)×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 either to introduce "soft" mechanisms of (spontaneous) breaking of gauge invariance to remedy it, or to invoke confinement to hide the unseen massless gluons;
- 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_μ 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 β 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

$$(\partial_\mu A_\nu - \partial_\nu A_\mu)^2 \quad \text{or in Fourier space} \quad A_\mu(-k)(k^\mu k^\nu - k^2 g^{\mu\nu})A_\nu(k)$$

is degenerate, thus non invertible, which reflects gauge invariance. Consequently the propagator of the field A_μ 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).

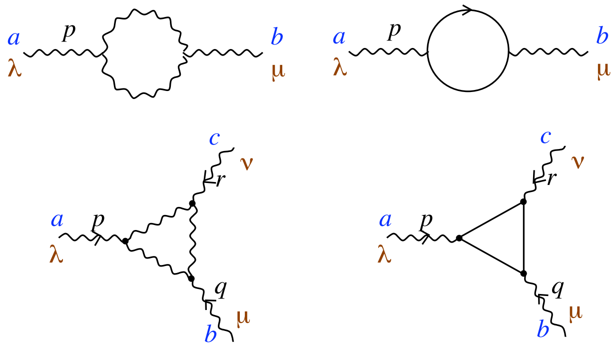


Figure 5.1: Some one-loop diagrams in a gauge theory

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 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

$$\mathrm{g} \mapsto \mathrm{g}_0 = \left(1 - \frac{\mathrm{g}^2}{(4\pi)^2}\left(\frac{11}{3}C_2 - \frac{4}{3}T_f\right)\log\frac{\Lambda}{\mu}\right)\mathrm{g}\ , \tag{5.25}$$

where Λ is a scale of ultraviolet "cutoff" and μ 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

$$\mathcal{L}_{\rm Fermi} = -\frac{G}{\sqrt{2}}J^\rho(x)J^\dagger_\rho(x) \tag{5.26}$$

gives a good description of the low energy physics of weak interactions: leptonic processes like $\bar\nu_e e^- \to \bar\nu_e e^-$ or $\bar\nu_\mu\mu^-$, semi-leptonic ones like $\pi^+ \to \mu^+\nu_\mu$ or the β decay $n \to pe^-\bar\nu_e$, or non-leptonic ones : $\Lambda \to p\pi^-$, $K^0 \to \pi\pi$, 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 σ 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 σ must decrease in each partial wave like $1/s$. A violation of unitarity by the Born term is

[4] G. 't Hooft and M. Veltman, Nobel prize 1999

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 $\mathcal{L}_{\rm Fermi}$ as an approximation of a theory where the charged current J^ρ is coupled to a charged vector field W of mass M, in the large mass limit[5]. Consider the new Lagrangian

$$\mathcal{L}_{\rm int.boson} = gJ^\rho(x)W^\dagger_\rho(x) + \text{h.c.} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + M^2 W^\dagger_\rho W^\rho \ . \tag{5.27}$$

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 $\mathcal{L}_{\rm Fermi}$ provided

$$\frac{G}{\sqrt{2}} = \frac{g^2}{M^2} \tag{5.28}$$

which relates the new coupling g to the Fermi constant G. Is the theory (5.27) 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

$$-i\,\frac{g_{\mu\nu} - \frac{k_\mu k_\nu}{M^2}}{k^2 - M^2} \tag{5.29}$$

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 ϕ acquires a vev $\langle\, \phi \,\rangle = v/\sqrt{2} \neq 0$. Reparametrizing the field ϕ according to

$$\phi(x) = e^{iq\theta(x)/v}\frac{v + \varphi(x)}{\sqrt{2}} \tag{5.30}$$

with v real and φ hermitian, and accompanying it by a U(1) gauge transformation

$$\begin{aligned} \phi(x) &\mapsto \phi'(x) = e^{-iq\theta(x)/v}\phi(x) = \frac{v + \varphi(x)}{\sqrt{2}} \\ A_\mu(x) &\mapsto A'_\mu(x) = A_\mu(x) + \frac{1}{v}\partial_\mu\theta(x) \end{aligned} \tag{5.31}$$

together with the corresponding transformation of other charged fields (ψ ...), one sees that the Lagrangian $\delta\mathcal{L}$ of (5.4) reads

$$\delta\mathcal{L} \ = \ (\partial_\mu - iqA'_\mu)\phi'(\partial^\mu + iqA'_\mu)\phi' - V(\phi'^2)$$

[5]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).

$$= \frac{1}{2}|(\partial_\mu - iqA'_\mu)\varphi|^2 + \frac{1}{2}q^2v^2A'_\mu A'^\mu - V\left(\frac{1}{2}(v+\varphi)^2\right) . \qquad (5.32)$$

Finally, one sees that the spontaneous breaking of the U(1) symmetry by the boson field ϕ leads to the appearance of a mass term for the gauge field A'_μ ! One also notes that the field θ 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_μ ; the total number of degrees of freedom of these fields is thus not modified: we started with 2 transverse modes of the massless electomagnetic 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 ϕ is coupled to a fermion field ψ by a term of the type $\bar\psi\phi\psi$, the appearance of its "vev" gives rise to a mass term $\frac{qv}{\sqrt{2}}\bar\psi\psi$ for the ψ.

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 : $G = \mathrm{SU}(2)\times\mathrm{U}(1)$, $H = \mathrm{U}(1)$ (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*: $\mathrm{U}(1)\times\mathrm{SU}(2)\times\mathrm{SU}(3)$, 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 $\mathrm{SU}(3) \equiv \mathrm{SU}(3)_c$ is the gauge group of color (see Chap. 4, § 4.3.2). The gauge fields A_μ 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, ψ_{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 $\mathrm{SU}(3)_c$ 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 (5.24), with fermionic mass terms depending on flavor, generated by the electroweak sector.

Asymptotic freedom

Knowing the coupling constant renormalization , (5.25), one may compute the corresponding beta function. One finds[6]

$$\beta(\mathrm{g}) = -\Lambda \frac{\partial}{\partial \Lambda} \mathrm{g}(\Lambda)|_{\mathrm{g}_0} = -\frac{\mathrm{g}^3}{(4\pi)^2}\left(\frac{11}{3}C_2 - \frac{4}{3}T_f\right) + \mathrm{O}(\mathrm{g}^5) \tag{5.33}$$

It thus appears that this beta function is negative in the vicinity of $\mathrm{g} = 0$, 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 $\mathrm{g} = 0$ is an attractive ultraviolet fixed point of the renormalization group: $\frac{d\mathrm{g}^2(\lambda)}{d\log\lambda} < 0 \Rightarrow \mathrm{g}^2(\lambda) \sim (2b\log\lambda)^{-1} \to 0$ as $\lambda \to \infty$, with $b =$ coefficient of the term $-\mathrm{g}^3$ in (5.33). 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 $\mathrm{SU}(3)_c$ 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 to be 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 from 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 σ 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 σ, 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

[6]David J. Gross, H. David Politzer, Frank Wilczek, Nobel prize 2004

scale of the lattice spacing a, thus, according to asymptotic freedom, $g_0^2 = g^2(\Lambda = 1/a) \to 0$). These Monte Carlo computations allow to determine numerically the coefficient κ in V, or *string tension*.

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 is accumulating data, precise calculations of QCD contributions are of fundamental importance: "new physics" may be identified only if the background of the Standard Model is perfectly known. Moreover the study of hadronization of quarks and gluons, of high energy "deep inelastic" scattering and of other hadronic phenomena remains a very hot subject and a crucial point where experiment confronts theory.

5.3.2 The electro-weak sector, a sketch

The group $U(1) \times SU(2)$ describes the electro-weak interactions (Glashow–Salam–Weinberg model[7]). Generators of these groups SU(2) and U(1) are referred to as weak isopin 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_μ^a, W_μ^i and B_μ the gauge fields of SU(3), SU(2) and 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_μ and B_μ in a different way. One writes the covariant derivative of one of these fields as

$$D_\mu\psi = (\partial_\mu - g_3 A_\mu^a T_a - g_2 W_\mu^j t_j - i\frac{g_1}{2} y B_\mu)\psi \tag{5.34}$$

where T_a, resp. t_j denote the infinitesimal generators of SU(3) and SU(2) in the representation of ψ; the representations assigned to each field, either lepton or quark, left or right, are the triplet representation of $SU(3)_c$ for quarks and the trivial one for leptons, of course, and for the electroweak part, are given in the following Table

Quarks & Leptons	(ν_L^e, e_L)	ν_R^e	e_R	(u_L, d_L)	u_R	d_R
Weak isospin t_z	$(\frac{1}{2}, -\frac{1}{2})$	0	0	$(\frac{1}{2}, -\frac{1}{2})$	0	0
Weak hypercharge y	$(-1, -1)$	0	-2	$(\frac{1}{3}, \frac{1}{3})$	$\frac{4}{3}$	$-\frac{2}{3}$
Electric charge $Q = \frac{1}{2}y + t_z$	$(0, -1)$	0	-1	$(\frac{2}{3}, -\frac{1}{3})$	$\frac{2}{3}$	$-\frac{1}{3}$

Table 1. Weak quantum numbers of leptons ν^e and e and of quarks u, d.
Things repeat themselves identically for the next generations.

A remarkable consequence of the use of SU(2) as a symmetry group of weak interactions is that, beside the two charged currents $J_\mu^{1,2}$ (or $J_\mu^\pm$) of Fermi theory, a third component J_μ^3 appears. This neutral current, which is *not* the electromagnetic current and which is coupled to the gauge field W_μ^3, is necessarily present and contributes for example to the $e^-\nu_\mu \to e^-\nu_\mu$ 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.

[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

The group $\mathrm{U}(1)_{em}$ of electromagnetism is identified thanks to the charges of the fields. There is a "mixing" of the initial U(1) factor and of a U(1) subgroup of SU(2). This mixing is characterized by an angle θ_W, called *Weinberg angle*: if the U(1) and SU(2) gauge fields are denoted B_μ and W_μ respectively, the electromagnetic field is $A_\mu^{em} = \cos\theta_W B_\mu + \sin\theta_W W_\mu^3$, while the orthogonal combination corresponds to another neutral vector field called Z^0.

Let us examine the "neutral current" terms that couple for example the electron and its neutrino to neutral boson vectors W^3 and B. They are read off the covariant derivatives (5.34) with quantum numbers of Table 1

$$\frac{1}{2} i\big[\bar{e}_L(-\mathrm{g}_2 W_\mu^3 - \mathrm{g}_1 B_\mu)\gamma^\mu e_L + \bar{e}_R(-2\mathrm{g}_1 B_\mu)\gamma^\mu e_R + \bar{\nu}_e(\mathrm{g}_2 W_\mu^3 - \mathrm{g}_1 B_\mu)\gamma^\mu \nu_e\big]$$

The rotation $W^3 = \cos\theta_W Z^0 + \sin\theta_W A$, $B = -\sin\theta_W Z^0 + \cos\theta_W A$ must be such that the electric charge e (coupling to A) is the same for e_L and e_R and zero for ν_e. One finds

$$2e = \mathrm{g}_2\sin\theta_W + \mathrm{g}_1\cos\theta_W = 2\mathrm{g}_1\cos\theta_W \quad \text{et} \quad \mathrm{g}_2\sin\theta_W - \mathrm{g}_1\cos\theta_W = 0$$

which are indeed compatible and give

$$\tan\theta_W = \frac{\mathrm{g}_1}{\mathrm{g}_2} \qquad e = \mathrm{g}_1\cos\theta_W = \mathrm{g}_2\sin\theta_W\ . \tag{5.35}$$

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, $(\mathrm{g}_1,\mathrm{g}_2)\mapsto(e,\theta_W)$ 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): $\Phi = \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix}$, of weak isospin $\frac{1}{2}$ and weak hypercharge $y=+1$, and $D_\mu\Phi = (\partial_\mu - \mathrm{g}_2 W_\mu^i\frac{\tau_i}{2} - \frac{1}{2}\mathrm{g}_1 B_\mu)\Phi$. The field Φ is endowed with a potential $V(\Phi)$ with a "mexican hat" shape, which is responsible of the spontaneous breaking of $\mathrm{U}(1)\times\mathrm{SU}(2)$ into $\mathrm{U}(1)_{em}$, 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 φ (and the "Higgs particle" it creates) that seems to have been discovered this year in the 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 $\mathrm{U}(1)\times\mathrm{SU}(2)$ by the field Φ occurs in a direction that preserves $\mathrm{U}(1)_{em}$. (Or more exactly the direction of that breaking determines what is called $\mathrm{U}(1)_{em}$.) One writes, by generalizing (5.30) to the SU(2) group of generators $i\frac{\tau^j}{2}$ (τ^j = Pauli matrices)

$$\Phi(x) = e^{i\xi_j(x)\frac{\tau^j}{2v}}\begin{pmatrix}0\\ \frac{v+\varphi(x)}{\sqrt{2}}\end{pmatrix}\ ,$$

which is accompanied by a gauge transformation, which causes the fields ξ_j to disappear and gives the fields W and B the quadratic mass terms

$$\mathcal{L}_{(2)} = \frac{1}{8}v^2[(\mathrm{g}_1 B - \mathrm{g}_2 W^3)^2 + \mathrm{g}_2^2((W^1)^2 + (W^2)^2)]$$

As expected the component $Z^0 = (\mathrm{g}_1 B - \mathrm{g}_2 W^3)/\sqrt{\mathrm{g}_1^2+\mathrm{g}_2^2}$ becomes massive, and so do $W^{1,2}$, whereas the orthogonal combination $A = (\mathrm{g}_2 B + \mathrm{g}_1 W^3)/\sqrt{\mathrm{g}_1^2+\mathrm{g}_2^2}$ remains massless. One finds

$$M_{W^\pm} = \frac{1}{2}v\mathrm{g}_2 \qquad M_{Z^0} = \frac{1}{2}v\sqrt{\mathrm{g}_1^2+\mathrm{g}_2^2} \tag{5.36}$$

and using (5.35), the relation $\frac{G}{\sqrt{2}} = \frac{g_2^2}{8M_W^2}$ read off the Lagrangian and the experimental values of e and of $G = 10^{-5}m_p^2$, one computes

$$M_{W^\pm} \approx \frac{38}{\sin\theta_W}\mathrm{GeV} \qquad M_{Z^0} = \frac{M_W}{\cos\theta_W} \approx \frac{38}{\sin\theta_W\cos\theta_W}\mathrm{GeV}\ .$$

These expressions then undergo small perturbative corrections. Lastly the mass of the famous Higgs boson φ is *not* predicted by the theory. Successive experiments have excluded wider and wider regions, leaving for the possible values of the mass more and more narrow "windows", from the range 100–200 GeV down to a window between 120 and 130 GeV. The latest results from LHC seem to point to a Higgs boson of mass between 125 and 126 GeV. Confirmation that the observed events qualify as a Higgs particle –as for spin, decay modes, etc– could be available in the coming months.

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 $M_{W^\pm} = 80.4\,\mathrm{GeV}$ and $M_{Z^0} = 91.2\,\mathrm{GeV}$ are compatible with the following value of the Weinberg angle

$$\sin^2\theta_W \approx 0.23\ , \tag{5.37}$$

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

$$\mathcal{L} = -\frac{1}{4}\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu} + \sum_{\substack{left\ and\ right\\ quarks\ \&\ leptons}} \bar\psi\gamma^\mu D_\mu\psi + |D\Phi|^2 - V(\Phi) + \mathrm{Higgs-fermions\ couplings}\ , \tag{5.38}$$

where $\mathbf{F}_{\mu\nu}$ denote the gauge field tensors of $\mathbf{A}$, $\mathbf{W}$ and B. Note that the $\mathrm{SU}(2)\times\mathrm{U}(1)$ 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 Φ 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),

$$\mathcal{L}_Y = -Y^d_{ij}\bar\psi_{Li}.\Phi\, d_{Rj} - Y^u_{ij}\bar\psi_{Li}.\widetilde\Phi^\dagger u_{Rj} + \mathrm{h.c.}\ , \tag{5.39}$$

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 Φ and $\widetilde\Phi^\dagger = \begin{pmatrix}\phi^{0\dagger}\\ -\phi^{+\dagger}\end{pmatrix}$ with quark doublets

$$\psi_{Li} = \begin{pmatrix}u_i\\ d_i\end{pmatrix}_L = \left(\begin{pmatrix}u\\ d\end{pmatrix}_L, \begin{pmatrix}c\\ s\end{pmatrix}_L, \begin{pmatrix}t\\ b\end{pmatrix}_L\right)\ .$$

Couplings of the same type appear between leptons and scalar fields.

The vev $v/\sqrt{2}$ of ϕ^0 then gives rise to a "mass matrix". A complication of the theory described by (5.38) 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

[9]Carlo Rubbia and Simon van der Meer, Nobel prize 1984

(5.38) : 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

$$J_\mu = (\bar{u}\bar{c}\bar{t})_L \gamma_\mu M \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L \tag{5.40}$$

with M the *unitary Cabibbo-Kobayashi-Maskawa matrix*[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 δ and θ_{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 achievments 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 require little amendments to the Lagrangian (5.38), 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 which seems to have been discovered 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 susbsection 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 at high energy (see next subsection) and also the existence

[10]M. Kobayashi, T. Maskawa, Nobel prize 2008, with Y. Nambu

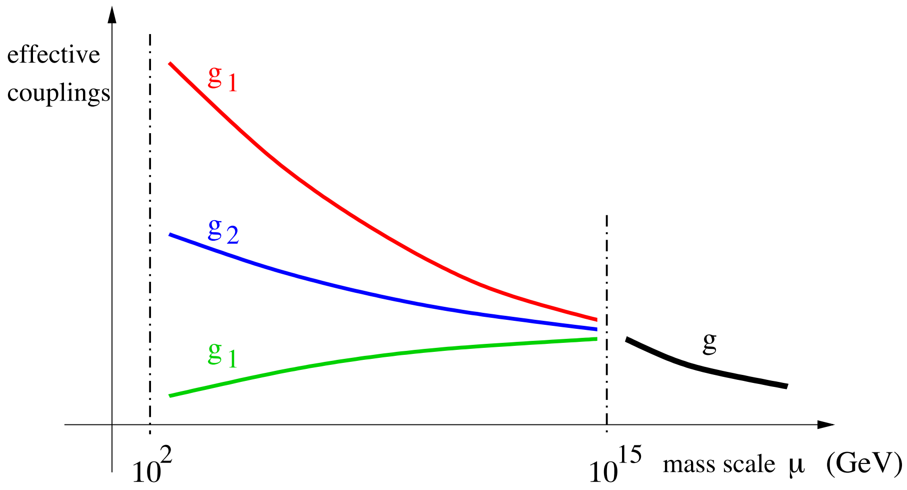


Figure 5.2: Schematic evolutions of the three effective couplings of the standard model and of that of a grand-unified theory

of supersymmetric partners for all known particles. On that issue too, results from LHC might confirm or infirm different scenarii.

5.4.2 Grand-unified theories or GUTs

An empirical observation is that the three coupling constants $\mathrm{g}_1, \mathrm{g}_2, \mathrm{g}_3$, 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 $10^{15 \text{ or } 16}$GeV. This was a very strong incentive towards a grand-unification, see 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)× SU(2)× 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 dim SU(5)= 24 gauge fields.

The main reason of that choice of SU(5) comes from the number of chiral fermions per 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 accomodate 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 $\bar{5}$) and 10 (or $\overline{10}$).

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 $\bar{5}$ decomposes into representations $(\bar{3}, 1) \oplus (1, 2)$ and contains antiquarks $\bar{d}_L$ and left leptons e_L^- and ν_e ; the 10 decomposes into $(1, 1) \oplus (3, 2) \oplus (\bar{3}, 1)$ 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) $\to$ SU(3) $\times$ SU(2) $\times$ U(1).

The breakdown SU(5) $\to$ 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 $\mathrm{g}_3, \mathrm{g}_2, \mathrm{g}_1$ 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 relate the couplings to the U(1) and SU(2) gauge fields and predict the Weinberg angle: one finds $\sin^2\theta = \frac{3}{8}$, ... 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 $X^\rho(\bar{d}\gamma_\rho e^+ + \bar{u}^c\gamma_\rho u)$, with one of the new gauge fields (the matrices of the generators have been omitted), allows proton decay $p = d\,uu \to d\bar{d}e^+ = \pi^0 e^+$, 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) $\to$ SU(3) $\times$ SU(2) $\times$ U(1) $\to$ SU(3) $\times$ U(1) at two very different scales...

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);

but

- 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

$$\mathcal{L} = i\bar{\psi}(\partial\!\!\!/ - A\!\!\!/)\frac{(1-\gamma_5)}{2}\psi$$

which contains a term $A^\mu_a J_{\mu a}$ with $J_{\mu a} = \bar{\psi}T_a\frac{(1-\gamma_5)}{2}\psi$. Classically that current $J_{\mu a}$ 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

$$D_\mu J^\mu = \frac{i}{24\pi^2}\partial_\mu \epsilon^{\mu\nu\rho\sigma}\mathrm{tr}\, T_a(A_\nu\partial_\rho A_\sigma + \frac{1}{2}A_\nu A_\rho A_\sigma)\ .$$

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

$$d_{abc} = \mathrm{tr}\,(T_a\{T_b, T_c\})$$

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, $T_a^* = CT_aC^{-1}$. 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 $\mathrm{tr}\,(t_3^2 Q)$ 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 $\bar{5}$ and 10 cancel one another.

⋆

Further references for Chapter 5

On geometric aspects of gauge theory and an introduction to the theory of fiber bundles, see for example M. Daniel and C. Viallet, *The geometric setting of gauge theories of the Yang-Mills type*, *Rev. Mod. Phys.* **52** (1980) 175-197.

On gauge theories, Yang-Mills, the standard model, etc, one may consult any book of quantum field theory posterior to 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 (available on the ICFP Master website).

For a historical review of these developments, see *Broken Symmetries*, note of the Nobel Committee, 2008,
http://nobelprize.org/nobel_prizes/physics/laureates/2008/phyadv08.pdf

For a detailed review of the Standard Model and a compilation of all known properties of elementary particles, see *The Review of Particle Physics*, on
http://pdg.lbl.gov/ already cited in Chap. 4.

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

$$D_{\mu a}{}^{b} F_{\nu\rho b} t^{a} = [D_\mu, F_{\nu\rho}] = \partial_\mu F_{\nu\rho} - [A_\mu, F_{\nu\rho}] \;.$$

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 $\not{D} = \not{\partial} - \not{A}$ acting on Dirac fermions in a representation R. One wants to compute $\not{D}^2$. 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 $\not{D}^2$ 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 $\mathfrak{g}$ 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

$$(X, Y)^{(R)} = \mathrm{tr}\,(T_a T_b) x^a y^b$$

if $X = x^a t_a$ and $Y = y^b t_b \in \mathfrak{g}$ (with summation over repeated indices).

a) Prove that this form is invariant in the sense that

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

One recalls that any invariant bilinear form on a simple Lie algebra is a multiple of the Killing form.

b) Prove that one may choose a basis of t_a and hence of T_a such that

$$\mathrm{tr}\,(T_a T_b) = -T_R \delta_{ab}$$

with T_R a coefficient that depends on the representation.

c) What is the sign of T_R ?

d) Consider then the quadratic Casimir operator

$$C_2^{(R)} = -\sum_\alpha (T_a)^2 \;.$$

On how many values of a does one sum in that expression?

e) Recall why $C_2^{(R)}$ is a multiple of the identity in the representation space of R

$$C_2^{(R)} = c_2(R)\,\mathbb{I}\ .$$

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

$$\mathrm{tr}\,\sum_a (T_a)^2\ .$$

i) To what does this relation boil down for the adjoint representation adjoint of G ?

j) Normalize the (antihermitian) generators of SU(N) in such a way that in the defining representation, $\mathrm{tr}\, T_a T_b = -\frac{1}{2}\delta_{ab}$, thus $T_f = \frac{1}{2}$. Is this verified by infinitesimal generators $i\frac{\sigma_a}{2}$ 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 (3.50) of Chap. 3 $c_2(\Lambda) = \frac{1}{2}\langle\,\Lambda, \Lambda + 2\rho\,\rangle$ may be rewritten as $c_2(\Lambda) = \frac{1}{2}\left(\langle\,\Lambda+\rho, \Lambda+\rho\,\rangle - \langle\,\rho,\rho\,\rangle\right)$, thus for SU($N$), using expressions (3.48) and (3.61) 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 θ in 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 $c_2(\mathrm{adj})$.

f) What is the value of T_{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

$$d_{\alpha\beta\gamma} = \mathrm{tr}\ (T_\alpha(T_\beta T_\gamma + T_\gamma T_\beta))\ .$$

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_α are antihermitian, one may find a unitary matrix U such that the complex conjugate of each T_α 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) group 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 Φ 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_μ ? 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 ρ.

1. Show that the orthogonality relations of $\mathcal{D}^{(\rho)}$ imply the following formulas:

$$\int_G \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g.g_1.g^{-1}.g_2) = \frac{1}{n_\rho}\chi^{(\rho)}(g_1)\chi^{(\rho)}(g_2)\ , \tag{5.41}$$

and

$$\int_G \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g.g_1)\chi^{(\sigma)}(g^{-1}.g_2) = \frac{\delta_{\rho,\sigma}}{n_\rho}\chi^{(\rho)}(g_1.g_2)\ . \tag{5.42}$$

Recall why a representation of G may always be regarded as unitary and show that then

$$\chi^{(\rho)}(g^{-1}) = \chi^{(\bar\rho)}(g) = (\chi^{(\rho)}(g))^* \,, \tag{5.43}$$

where $\bar\rho$ is the complex conjugate representation of ρ.
We make a frequent use of these three relations in the following.

2. Let χ be the character of a real representation r (not necessarily irreducible) of G, β a real parameter.
a) Show that one may expand $\exp\beta\chi(g)$ on characters of irreducible representations of G according to

$$e^{\beta\chi(g)} = \sum_\rho n_\rho b_\rho \chi^{(\rho)}(g) \,,$$

with functions $b_\rho(\beta)$.
Express the function $b_\rho(\beta)$ in terms of a group integral.
Using (5.43), show that the functions $b_\rho(\beta)$ are real, $b_\rho(\beta) = (b_\rho(\beta))^* = b_{\bar\rho}(\beta)$.
b) Show that b_ρ is non vanishing provided representation ρ appears in some tensor power $r^{\otimes n}$.
c) For $G = \mathrm{SU}(2)$ and $r = (j = \frac12)$, the representation of spin $\frac12$, is condition **b)** satisfied for any ρ ? Why ?
If $r = (j = 1)$, what are the representations for which b_ρ is a priori zero ?
d) For G =SU(3) and $\chi = \chi^{(3)} + \chi^{(\bar 3)}$, show that b_ρ is non zero for all ρ.
For $\beta \to 0$, what is the leading behaviour of $b_a(\beta)$ when $\beta \to 0$ if a denotes the adjoint representation of SU(3) ? More generally what is the leading behaviour of $b_\rho(\beta)$ where ρ 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 :

$$g_p = g_{ij}\cdot g_{jk}\cdot g_{kl}\cdot g_{li}$$

and the "energy" of a configuration of these variables is given by

$$E = -\sum_{\text{plaquettes } p} \chi(g_p) \tag{5.44}$$

where χ is, like in question 2, the character of some real representation r of the group. The Boltzmann weight is thus

$$e^{-\beta E} = \prod_p e^{\beta\chi(g_p)} \qquad , \qquad \beta = \frac{1}{kT}$$

and the partition function reads

$$Z = \prod_{\text{links}\,\ell} \int_G \frac{\mathrm{d}\mu(g_\ell)}{v(G)} \prod_{\text{plaquettes}} e^{\beta\chi(g_p)} \,. \tag{5.45}$$

a) Show that the energy E is invariant by redefinition of g_{ij} as $g_{ij} \mapsto g_i.g_{ij}.g_j^{-1}$, where $g_i \in G$, (this is a *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.
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

$$g_{ij} \equiv P \exp \int_{l=\vec{ij}} A_\mu dx^\mu$$

- For a small lattice spacing a, show, by using for example the BCH formula and by expanding to the first non vanishing order that

$$g_p = \exp\left(a^2 F_{\mu\nu} + \mathrm{o}(a^2)\right)$$

where μ and ν denote the directions of the edges of plaquette p. (One is here interested in an Euclidean version of gauge theory, and position of indices μ, ν is irrelevant.) Show then that the energy E_p (5.44) 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 χ.

- Explain why the parameter β 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 $\mathcal{N}$ 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_ℓ 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_ℓ along a closed oriented curve C for some irreducible representation σ of G

$$W^{(\sigma)}(C) := \langle\, \chi^{(\sigma)}(g_C)\, \rangle = \frac{1}{Z} \prod_{\text{liens}\,\ell} \int_G \frac{\mathrm{d}\mu(g_l)}{v(G)} \chi^{(\sigma)}\Big(\prod_{\ell \in C} g_\ell\Big) \prod_p e^{\beta\chi(g_p)} \,. \tag{5.46}$$

c) 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

$$e^{\beta\chi(g_p)} = \sum_\rho n_\rho b_\rho \, \chi^{(\rho)}(g_p) \,. \tag{5.47}$$

d) One inserts in (5.45) or (5.46) the expansion (5.47) for each plaquette. Show that if two plaquettes share one link ℓ, formulas of part 1 permit an integration over the variable g_ℓ of that link and that the two representations carried by the two adjacent plaquettes are then identical.

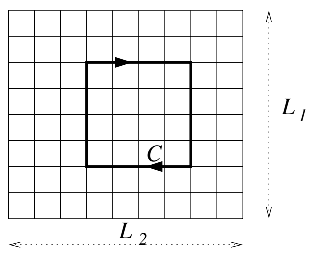


Figure 5.3: Square lattice in 2 d

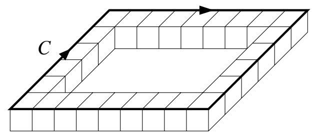


Figure 5.4: A tubular configuration contributing to the Wilson loop

Using repeatedly these formulas of part 1, show that one may integrate over all variables g_ℓ and that

$$Z = b_1^{\mathcal{N}} \qquad\qquad W^{(\sigma)}(C) = n_\sigma \left(\frac{b_\sigma}{b_1}\right)^A \tag{5.48}$$

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.
e) One now consider the case of dimension $d = 3$. Variables g_ℓ are attached to links of a cubic lattice. Energy is again given (5.44), 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 (5.41) and (5.42) on all variables g_ℓ leads to the following expression

$$W^{(\sigma)}(C)\Big|_{\text{tube}} = \sum_\rho \left(\frac{b_\rho}{b_1}\right)^P \int_G \frac{d\mu(g)}{v(G)} \chi^{(\rho)}(g)\chi^{(\rho)}(g^{-1})\chi^{(\sigma)}(g) \tag{5.49}$$

where P is the number of plaquettes making the tube.
– Under which condition $\mathcal{C}$ on representation σ of the loop C is the contribution of representation ρ to the right hand side of (5.49) non vanishing?
– Give an example for G=SU(2) of representations σ for which this condition $\mathcal{C}$ is *never* satisfied for any ρ, and hence these tubular configurations do not contribute.
– Inversely give an example (again for SU(2)) of a possible choice of σ which satisfies it.

We admit that at high temperature, (small β), the dominant contribution to $W^{(\sigma)}(C)$ is of type (5.49) if condition $\mathcal{C}$ may be satisfied, and of type (5.48) 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 σ 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 (5.48), or from the contribution to (5.49) due to representation ρ. 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 β ("high temperature") expansion. The fact that $W^{(\sigma)}(C)$ decays like x^A ($x = b_\sigma/b_1 < 1$ for β 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