

Symmetries in Quantum Mechanics and Particle Physics

Dirk H. Rischke

ECT* Doctoral Training Program 2018

Contents

1	Summary	1
2	Symmetries in Quantum Mechanics	2
2.1	Space translations	2
2.2	Time translations	4
2.3	Rotations	6
3	Introduction to Group Theory	7
3.1	General remarks	7
3.2	Representation of groups	9
3.3	Lie groups	12
3.4	Simple and semi-simple Lie groups	22
3.5	Casimir operators, multiplets, Schur's lemma	24
4	The group $(SU(3), \cdot)$	32
4.1	Generators	32
4.2	Structure constants	34
4.3	Subalgebras and subgroups	35
4.4	Multiplets	36
4.5	Construction of multiplets from the fundamental representation	46
5	Unitary symmetries of the strong interaction	51
5.1	$(U(1), \cdot)$ symmetry of quantum electrodynamics	51
5.2	$(SU(3), \cdot)$ color symmetry of quantum chromodynamics	52
5.3	$(SU(N_f), \cdot)$ flavor symmetry	57
5.4	Isospin symmetry	58
5.5	Strangeness and $(SU(3), \cdot)$ flavor symmetry	61
5.6	Charm and $(SU(4), \cdot)$ flavor symmetry	69
6	The Poincaré group	75
6.1	Characterization of the Poincaré group	75
6.2	Generators	76
6.3	The Poincaré algebra	78
6.4	Casimir operators	80

1 Summary

The notion of **symmetry**, i.e., the **invariance** of a theory under certain **transformations** of its degrees of freedom, plays an important, if not the most important, role in the mathematical formulation of the laws of Nature.

The notion of symmetry is already known from lecture courses on “Mechanics”. According to **Noether’s theorem**, the **invariance of the action** under **coordinate transformations**

$$\begin{aligned} t &\longrightarrow t'(t) , \\ \vec{q}(t) &\longrightarrow \vec{q}'(\vec{q}(t), t) , \end{aligned}$$

leads to **integrals of motion**, i.e., to **conservation laws**. A transformation which leaves the action invariant is called **symmetry transformation** or **symmetry**.

As we shall see in this lecture course, the notion of symmetry can be applied to **quantum mechanical systems**, too. Also in this case, we will be lead to **integrals of motion**. However, we will not derive them via the invariance of the action, as in Noether’s theorem, but we will directly investigate the behavior of **Schrödinger’s equation** under symmetry transformations (Chapter 2). To understand the mathematical structure of symmetry transformations in quantum mechanics, we will give a short introduction to **group theory** (Chapter 3).

We will see that symmetries of a theory are not restricted to the invariance under coordinate transformations. At least of equal importance is the class of so-called **internal symmetries**, i.e., transformations which are not related to the space-time coordinates of the system. Of fundamental importance for physics is the **group of unitary transformations in three dimensions**, $SU(3)$ (Chapter 4), and its **applications in the theory of strong interactions** in particle physics (Chapter 5).

We will conclude this lecture course (again if time allows) by a discussion of the **Poincaré group**, which is of eminent importance in relativistic quantum field theory (Chapter 6), because all theories of Nature (which are well-established up to now) are formulated to respect this symmetry.

2 Symmetries in Quantum Mechanics

2.1 Space translations

Consider the Hilbert-space state $|\psi\rangle$. We can assign a coordinate-space **wave function** to this state according to

$$|\psi\rangle \longrightarrow \psi(\vec{r}) \equiv \langle \vec{r} | \psi \rangle .$$

Let us now consider a **translation in space**

$$\vec{r} \longrightarrow \vec{r}' = \vec{r} + \vec{a} , \quad (2.1)$$

where $\vec{a} = \overrightarrow{const.}$. If we demand that physics is **invariant** under such space translations, i.e., that the space translation (2.1) is a **symmetry transformation**, then we must have for the transformed wave function

$$\psi'(\vec{r}') \equiv \psi(\vec{r}) , \quad (2.2)$$

i.e., the transformed wave function ψ' at the transformed position \vec{r}' has to be equal to the original wave function ψ at the original position \vec{r} . Using Eq. (2.1) we may write Eq. (2.2) as follows

$$\psi'(\vec{r} + \vec{a}) = \psi(\vec{r}) \quad \text{oder} \quad \psi'(\vec{r}) = \psi(\vec{r} - \vec{a}) . \quad (2.3)$$

Using the so-called **spatial translation operator** $\hat{U}_r(\vec{a})$ the last equation can be expressed as follows

$$\psi'(\vec{r}) = \hat{U}_r(\vec{a}) \psi(\vec{r}) . \quad (2.4)$$

If a space translation is a symmetry transformation, it cannot influence the probability with which a quantum mechanical state occurs. According to the **conservation of probability** we therefore have

$$\begin{aligned} \langle \psi' | \psi' \rangle &= \int d^3\vec{r}' \langle \psi' | \vec{r}' \rangle \langle \vec{r}' | \psi' \rangle = \int d^3\vec{r} \psi'^*(\vec{r}) \psi'(\vec{r}) = \int d^3\vec{r} \left[\hat{U}_r(\vec{a}) \psi(\vec{r}) \right]^* \hat{U}_r(\vec{a}) \psi(\vec{r}) \\ &= \int d^3\vec{r} \psi^*(\vec{r}) \hat{U}_r^\dagger(\vec{a}) \hat{U}_r(\vec{a}) \psi(\vec{r}) \equiv \int d^3\vec{r} \psi^*(\vec{r}) \psi(\vec{r}) = \langle \psi | \psi \rangle . \end{aligned} \quad (2.5)$$

This means that $\hat{U}_r(\vec{a})$ must be a **unitary** operator,

$$\hat{U}_r^\dagger(\vec{a}) \hat{U}_r(\vec{a}) \equiv \mathbb{1} \quad \iff \quad \hat{U}_r^\dagger(\vec{a}) = \hat{U}_r^{-1}(\vec{a}) . \quad (2.6)$$

Let us now explicitly determine this operator. Without restriction of generality we consider a translation in x -direction, $\vec{a} = (a, 0, 0)^T$, and expand the right-hand side of Eq.

(2.3) in terms of a Taylor series around the point \vec{r} ,

$$\begin{aligned}\psi(\vec{r} - \vec{a}) &= \psi(\vec{r}) - a \frac{\partial}{\partial x} \psi(\vec{r}) + \frac{1}{2!} a^2 \frac{\partial^2}{\partial x^2} \psi(\vec{r}) - \frac{1}{3!} a^3 \frac{\partial^3}{\partial x^3} \psi(\vec{r}) + \dots \\ &= \sum_{n=0}^{\infty} \frac{(-a)^n}{n!} \frac{\partial^n}{\partial x^n} \psi(\vec{r}) \equiv \exp\left(-a \frac{\partial}{\partial x}\right) \psi(\vec{r}) \\ &= \exp\left(-\frac{i}{\hbar} a \hat{p}^x\right) \psi(\vec{r}) \equiv \exp\left(-\frac{i}{\hbar} \vec{a} \cdot \hat{\vec{p}}\right) \psi(\vec{r}),\end{aligned}$$

where we have used the momentum operator $\hat{\vec{p}} = -i\hbar\vec{\nabla}$. Comparison with Eq. (2.4) yields

$$\hat{U}_r(\vec{a}) \equiv \exp\left(-\frac{i}{\hbar} \vec{a} \cdot \hat{\vec{p}}\right). \quad (2.7)$$

This form of the space-translation operator holds also for arbitrary vectors \vec{a} . The momentum operator $\hat{\vec{p}}$ is **hermitean** (since the momentum is a physical observable), $\hat{\vec{p}}^\dagger \equiv \hat{\vec{p}}$, and thus the unitarity (2.6) of the space-translation operator $\hat{U}_r(\vec{a})$ is immediately obvious. On the one hand we have

$$\hat{U}_r^\dagger(\vec{a}) = \exp\left(+\frac{i}{\hbar} \vec{a} \cdot \hat{\vec{p}}^\dagger\right) = \exp\left[-\frac{i}{\hbar} (-\vec{a}) \cdot \hat{\vec{p}}\right] = \hat{U}_r(-\vec{a}),$$

where we used that $\vec{a} \in \mathbb{R}^3$. On the other hand, since a translation by $-\vec{a}$ reverses the translation by $+\vec{a}$,

$$\hat{U}_r(-\vec{a}) \hat{U}_r(\vec{a}) \equiv \mathbb{1},$$

one has

$$\hat{U}_r^\dagger(\vec{a}) = \hat{U}_r(-\vec{a}) = \hat{U}_r^{-1}(\vec{a}), \quad \text{q.e.d.} \quad (2.8)$$

If the space translation (2.1) is a symmetry transformation, the Hamilton operator has to be **invariant**, too,

$$\hat{H} \longrightarrow \hat{H}' \equiv \hat{H} \quad (\vec{r} \longrightarrow \vec{r}' = \vec{r} + \vec{a}). \quad (2.9)$$

The transformed wave function (2.2) obeys the (in general time-dependent) Schrödinger equation,

$$i\hbar \frac{\partial}{\partial t} \psi'(t, \vec{r}') = \hat{H} \psi'(t, \vec{r}'). \quad (2.10)$$

If we insert Gl. (2.4) (a possible time dependence of the wave function does not influence this equation), we obtain

$$i\hbar \frac{\partial}{\partial t} \hat{U}_r(\vec{a}) \psi(t, \vec{r}) \equiv \hat{U}_r(\vec{a}) i\hbar \frac{\partial}{\partial t} \psi(t, \vec{r}) = \hat{H} \hat{U}_r(\vec{a}) \psi(t, \vec{r}), \quad (2.11)$$

since temporal and spatial derivatives commute with each other. Since the original wave function also fulfills the time-dependent Schrödinger equation,

$$i\hbar \frac{\partial}{\partial t} \psi(t, \vec{r}) = \hat{H} \psi(t, \vec{r}), \quad (2.12)$$

we can write Eq. (2.11) as

$$\hat{U}_r(\vec{a}) \hat{H} \psi(t, \vec{r}) = \hat{H} \hat{U}_r(\vec{a}) \psi(t, \vec{r}) , \quad (2.13)$$

or, since the wave function $\psi(t, \vec{r})$ was arbitrary,

$$\hat{U}_r(\vec{a}) \hat{H} = \hat{H} \hat{U}_r(\vec{a}) \iff [\hat{U}_r(\vec{a}), \hat{H}] = 0 , \quad (2.14)$$

where we introduced the **commutator** of two operators \hat{A}, \hat{B} ,

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} . \quad (2.15)$$

Thus, a space translation is a **symmetry transformation**, if the space-translation operator $\hat{U}_r(\vec{a})$ commutes with the Hamilton operator of the system.

This must also be valid for **infinitesimal** space translations, $|\vec{a}| \ll 1$. For such one can expand the space-translation operator (2.7) in terms of a Taylor series,

$$\hat{U}_r(\vec{a}) \simeq \mathbb{1} - \frac{i}{\hbar} \vec{a} \cdot \hat{\vec{p}} , \quad (2.16)$$

where we can neglect terms of order $O(a^2)$. Inserting this into Eq. (2.14) yields

$$\left[\mathbb{1} - \frac{i}{\hbar} \vec{a} \cdot \hat{\vec{p}}, \hat{H} \right] \equiv \frac{i}{\hbar} \vec{a} \cdot [\hat{\vec{p}}, \hat{H}] = 0 \iff [\hat{\vec{p}}, \hat{H}] = 0 , \quad (2.17)$$

since the infinitesimal translation vector \vec{a} was arbitrary. We conclude that, for **systems invariant under space translations**, the **momentum operator** commutes with the Hamilton operator. Then, however, the **momentum** is an **integral of motion**, i.e., we have **momentum conservation**, and one can find a system of eigenstates of \hat{H} which are simultaneously eigenstates of $\hat{\vec{p}}$.

2.2 Time translations

We now consider **translations in time**,

$$t \longrightarrow t' = t + a , \quad (2.18)$$

where $a = \text{const.}$. We demand that physics remains **invariant** under such translations in time, i.e., the time translation (2.18) is a **symmetry transformation**. Then the transformed wave function must fulfill (we suppress a potential spatial dependence, since it does not play any role in the following)

$$\psi'(t') = \psi(t) , \quad (2.19)$$

i.e., the transformed wave function ψ' at the shifted time t' must be identical to the original wave function ψ at the original time t . If we insert Eq. (2.18), we can write this in the following way,

$$\psi'(t+a) = \psi(t) \iff \psi'(t) = \psi(t-a) \equiv \hat{U}_t(a) \psi(t) , \quad (2.20)$$

where, in analogy to the preceding section, we have again expressed the effect of the time translation by an operator, the so-called **time-translation operator** $\hat{U}_t(a)$. In complete analogy to the derivation of the space-translation operator (2.7) we now obtain

$$\hat{U}_t(a) = \exp\left(-a \frac{\partial}{\partial t}\right) \equiv \exp\left(\frac{i}{\hbar} a \hat{E}\right), \quad (2.21)$$

with the **energy operator** $\hat{E} \equiv i\hbar\partial/\partial t$. The unitarity of the time-translation operator follows from the hermiticity of the energy operator, $\hat{E}^\dagger \equiv \hat{E}$ (energy eigenvalues are always real-valued), and from $a \in \mathbb{R}$.

Acting with the operator (2.21) on states which fulfill the time-dependent Schrödinger equation (2.12) yields the identity $\hat{E} \equiv \hat{H}$, such that we can also write the time-translation operator (2.21) as

$$\hat{U}_t(a) \equiv \exp\left(\frac{i}{\hbar} \hat{H} a\right). \quad (2.22)$$

Thus, it is identical to the **time-evolution operator** well-known from lectures on quantum mechanics, for Hamilton operators which do not explicitly depend on time. (Note that the time translation (2.18) affects a time evolution of the state $\psi(t)$ into the state $\psi(t-a)$, cf. Eq. (2.20), i.e., backwards in time by an amount $-a$.)

Symmetry under time translations means that also the Hamilton operator remains **invariant**,

$$\hat{H} \longrightarrow \hat{H}' \equiv \hat{H} \quad (t \longrightarrow t' = t + a). \quad (2.23)$$

The time-translated state $\psi'(t)$ of course fulfills the time-dependent Schrödinger equation

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \psi'(t) &= \hat{H} \psi'(t) \\ \iff i\hbar \frac{\partial}{\partial t} \hat{U}_t(a) \psi(t) &\equiv \hat{U}_t(a) i\hbar \frac{\partial}{\partial t} \psi(t) = \hat{U}_t(a) \hat{H} \psi(t) = \hat{H} \hat{U}_t(a) \psi(t), \end{aligned} \quad (2.24)$$

where we have used Eq. (2.20) and the original Schrödinger equation (2.12) for the state $\psi(t)$. Since this holds for an arbitrary wave function, we get

$$\hat{U}_t(a) \hat{H} = \hat{H} \hat{U}_t(a) \iff \left[\hat{U}_t(a), \hat{H}\right] = 0, \quad (2.25)$$

in analogy to Eq. (2.14) for space translations. If the **system is invariant under time translations**, the time-translation operator commutes with the Hamilton operator.

Since this holds also for infinitesimal time translations, $|a| \ll 1$, after Taylor-expanding the time-translation operator (2.21) we obtain

$$\left[\hat{E}, \hat{H}\right] = \left[i\hbar \frac{\partial}{\partial t}, \hat{H}\right] = 0 \iff \frac{\partial \hat{H}}{\partial t} = 0, \quad (2.26)$$

i.e., the Hamilton operator **must not depend explicitly on time**. For systems which are invariant under time translations, the time-translation operator is thus identical to the time-evolution operator, cf. Eq. (2.22). The **conserved quantity** associated with this symmetry is the **total energy** of the system.

2.3 Rotations

Let us consider a **rotation in space** by an **infinitesimal** angle $\delta\vec{\phi}$ (around an axis defined by the direction of $\delta\vec{\phi}$),

$$\vec{r} \longrightarrow \vec{r}' = \vec{r} + \delta\vec{\phi} \times \vec{r}. \quad (2.27)$$

We demand that this is a **symmetry transformation**, i.e.,

$$\psi'(\vec{r}') = \psi(\vec{r}) \iff \psi'(\vec{r}) = \psi(\vec{r} - \delta\vec{\phi} \times \vec{r}). \quad (2.28)$$

For infinitesimal rotation angles, $|\delta\vec{\phi}| \ll 1$ we may again expand the right-hand side in terms of a Taylor series, which we can truncate at order $O(\delta\phi^2)$,

$$\begin{aligned} \psi'(\vec{r}) &= \psi(\vec{r}) - \left(\delta\vec{\phi} \times \vec{r}\right) \cdot \vec{\nabla} \psi(\vec{r}) + O(\delta\phi^2) \\ &\simeq \psi(\vec{r}) - \delta\vec{\phi} \cdot \left(\vec{r} \times \vec{\nabla}\right) \psi(\vec{r}) \equiv \left[\mathbb{1} - \frac{i}{\hbar} \delta\vec{\phi} \cdot \left(\vec{r} \times \hat{\vec{p}}\right)\right] \psi(\vec{r}) \\ &= \left(\mathbb{1} - \frac{i}{\hbar} \delta\vec{\phi} \cdot \hat{\vec{L}}\right) \psi(\vec{r}), \end{aligned} \quad (2.29)$$

where we have used the definitions of the triple product, of the momentum operator, $\hat{\vec{p}} = -i\hbar\vec{\nabla}$, and of the **angular-momentum operator**, $\hat{\vec{L}} = \vec{r} \times \hat{\vec{p}}$. We again demand that the transformation (2.28) is affected by an operator,

$$\psi'(\vec{r}) = \psi(\vec{r} - \delta\vec{\phi} \times \vec{r}) = \hat{U}_R(\delta\vec{\phi}) \psi(\vec{r}). \quad (2.30)$$

A comparison with Eq. (2.29) shows that the **infinitesimal rotation operator** has the form

$$\hat{U}_R(\delta\vec{\phi}) = \mathbb{1} - \frac{i}{\hbar} \delta\vec{\phi} \cdot \hat{\vec{L}}. \quad (2.31)$$

These are the first two terms of the Taylor expansion of the exponential function. The rotation operator for **arbitrary** (not necessarily infinitesimal) rotation angles is thus defined as

$$\hat{U}_R(\vec{\phi}) \equiv \exp\left(-\frac{i}{\hbar} \vec{\phi} \cdot \hat{\vec{L}}\right). \quad (2.32)$$

The unitarity of the rotation operator follows from the hermiticity of the angular-momentum operator, $\hat{\vec{L}}^\dagger = \hat{\vec{L}}$, and from $\vec{\phi} \in \mathbb{R}^3$.

For **systems which are invariant under rotations** the Hamilton operator must also remain invariant,

$$\hat{H} \longrightarrow \hat{H}' = \hat{H} \quad (\vec{r} \longrightarrow \vec{r}' = \vec{r} + \vec{\phi} \times \vec{r}). \quad (2.33)$$

By the same arguments as in the two preceding sections one then shows that

$$\left[\hat{U}_R(\vec{\phi}), \hat{H}\right] = 0, \quad (2.34)$$

i.e., the rotation operator commutes with the Hamilton operator. For infinitesimal rotations this is equivalent to

$$\left[\hat{\vec{L}}, \hat{H}\right] = 0. \quad (2.35)$$

This means that the **angular momentum is conserved** and that one can find a common system of eigenstates of \hat{H} , \hat{L}^2 and \hat{L}^z .

3 Introduction to Group Theory

3.1 General remarks

Definition: A set

$$G = \{a, b, c, \dots\}$$

is a **group**, if there exists an **operation** (“**multiplication**”), which combines elements $a, b \in G$,

$$a \circ b,$$

with the following properties:

(i) **Closure:** $\forall a, b \in G$ also $a \circ b \in G$.

One says that G is **closed** under the operation (multiplication) which combines group elements.

(ii) **Identity element:** $\exists e \in G$, such that $\forall a \in G$ it holds that $a \circ e = e \circ a = a$, i.e., a so-called **identity element** exists.

(iii) **Inverse element:** $\forall a \in G \exists a^{-1} \in G$, such that $a^{-1} \circ a = a \circ a^{-1} = e$, i.e., for each group element exists an **inverse element**.

(iv) **Associativity:** $\forall a, b, c \in G$ it holds that $(a \circ b) \circ c = a \circ (b \circ c)$, the so-called **law of associativity of group multiplication**.

The group multiplication is essential to the definition of the group, hence one also denotes the group by (G, \circ) , i.e., the set of elements G which are combined by the group multiplication \circ .

Definition: A group (G, \circ) is called **abelian**, if $\forall a, b \in G$ it holds that $a \circ b = b \circ a$, i.e., the group multiplication fulfills the **law of commutativity**.

Definition: A group is called **continuous**, if it is possible to characterize the elements of the group by continuous **parameters** $\vec{t} = (t_1, t_2, \dots)^T \in \mathbb{R}^n$, i.e.,

$$G = \{a(\vec{t}), b(\vec{t}), c(\vec{t}), \dots\}.$$

Definition: A group is called **continuously connected**, if one can transform every group element into another one by a continuous change of the parameters \vec{t} .

Example 1: The group of space translations

$$(G_r, \cdot) = \left(\left\{ \hat{U}_r(\vec{a}), \vec{a} \in \mathbb{R}^3 \right\}, \cdot \right)$$

(with the standard multiplication “.” of group elements as group multiplication) is a **continuously connected, abelian group**.

It is **continuously connected**, since one can generate any arbitrary element of the group by a continuous change of the parameters $\vec{a} = (a^x, a^y, a^z)^T$.

It is **abelian**, since two arbitrary translations commute with each other. It is sufficient to show this for **infinitesimal** translations, because, since the group is continuously connected, one can represent any arbitrary **finite** translation by an **infinite series** of **infinitesimal** translations. For two infinitesimal translations by the constant vectors $\delta\vec{a}_1$ and $\delta\vec{a}_2$ we have

$$\begin{aligned} \hat{U}_r(\delta\vec{a}_1)\hat{U}_r(\delta\vec{a}_2) &= \left(\mathbb{1} - \frac{i}{\hbar} \delta\vec{a}_1 \cdot \hat{\vec{p}} \right) \left(\mathbb{1} - \frac{i}{\hbar} \delta\vec{a}_2 \cdot \hat{\vec{p}} \right) \\ &= \mathbb{1} - \frac{i}{\hbar} (\delta\vec{a}_1 + \delta\vec{a}_2) \cdot \hat{\vec{p}} - \frac{1}{\hbar^2} \delta a_1^i \delta a_2^j \hat{p}^i \hat{p}^j \\ &= \mathbb{1} - \frac{i}{\hbar} (\delta\vec{a}_1 + \delta\vec{a}_2) \cdot \hat{\vec{p}} - \frac{1}{\hbar^2} \delta a_1^i \delta a_2^j \hat{p}^j \hat{p}^i \\ &= \left(\mathbb{1} - \frac{i}{\hbar} \delta\vec{a}_2 \cdot \hat{\vec{p}} \right) \left(\mathbb{1} - \frac{i}{\hbar} \delta\vec{a}_1 \cdot \hat{\vec{p}} \right) \equiv \hat{U}_r(\delta\vec{a}_2)\hat{U}_r(\delta\vec{a}_1), \end{aligned}$$

where we used the fact that the components of the momentum operator **commute with each other**,

$$[\hat{p}^i, \hat{p}^j] = 0, \quad \text{q.e.d.} \quad (3.1)$$

Example 2: In contrast, the **group of rotations in space**,

$$(G_R, \cdot) = \left(\left\{ \hat{U}_R(\vec{\phi}), \vec{\phi} \in S^2 \right\}, \cdot \right), \quad (3.2)$$

(S^2 is the two-dimensional unit sphere around the origin of three-dimensional space) is **continuously connected**, but **non-abelian**, since

$$[\hat{L}^i, \hat{L}^j] = i\hbar \epsilon^{ijk} \hat{L}_k. \quad (3.3)$$

The commutation residue is responsible for the fact that

$$\hat{U}_R(\vec{\phi}_1)\hat{U}_R(\vec{\phi}_2) \neq \hat{U}_R(\vec{\phi}_2)\hat{U}_R(\vec{\phi}_1).$$

Definition: A continuous group is called **compact**, if the set of all values assumed by the parameters \vec{t} , the so-called **group manifold**, is compact.

Example: Rotations in two space dimensions, i.e., in a plane, are affected by **orthogonal** (2×2) matrices

$$O(\phi) \equiv \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix}$$

with determinant $\det O(\phi) = \cos^2 \phi + \sin^2 \phi = +1$. These matrices form the **group of special orthogonal (2×2) matrices**,

$$(SO(2), \cdot) = (\{O(\phi), \phi \in [0, 2\pi]\}, \cdot), \quad (3.4)$$

with the standard matrix multiplication “ \cdot ” as group multiplication. It is a **one-parameter** group, with the rotation angle ϕ as group parameter. Since each rotation maps onto itself after rotating by an angle 2π , one can restrict the group manifold to the **compact** interval $[0, 2\pi]$. Hence, $(SO(2), \cdot)$ is a compact group.

Definition: A group (G, \circ) is called **homomorphic** to a group (H, \square) , if there exists a function $f : G \rightarrow H$ with the property

$$f(a \circ b) = f(a) \square f(b) , \quad \forall a, b \in G , \quad f(a), f(b) \in H .$$

In other words, it does not matter whether one first multiplies the two elements $a, b \in G$ and then applies the map f , or whether one first applies the map f and then multiplies the two elements $f(a), f(b) \in H$. This ensures that $f(e_G) = e_H$ and $f(a^{-1}) = f^{-1}(a) \forall a \in G$. The function f is called **group homomorphism**. Its properties are such that it preserves the structure of the group (G, \circ) and maps it onto the group (H, \square) .

Definition: Two groups (G, \circ) and (H, \square) are called **locally isomorphic**, if the group homomorphism f is a **bijection** (a “one-to-one correspondence”) between group elements on subsets $U \subset G, V \subset H$,

$$\forall a \in U \subset G , \quad b \in V \subset H : \quad b = f(a) , \quad a = f^{-1}(b) .$$

Example: $(SO(2), \cdot)$ is isomorphic to the **group of unitary transformations in one dimensions**,

$$(U(1), \cdot) = (\{e^{i\phi}, \phi \in [0, 2\pi]\} , \cdot) , \quad (3.5)$$

i.e., multiplication with a “phase factor”, a complex number of modulus 1. This can be interpreted as a rotation in the **complex plane**. The existence of a bijection f is now evident: one simply assigns each element $O(\phi) \in SO(2)$ to the corresponding element $e^{i\phi} \in U(1)$.

Definition: If $U = G, V = H$, one speaks of **global isomorphy**.

Example: Since the group manifold $[0, 2\pi]$ is the **same** for $(SO(2), \cdot)$ and $(U(1), \cdot)$, these groups are not only locally but even **globally isomorphic**.

Definition: A group is called **finite**, if it has a finite number of elements, otherwise it is called **infinite**. The **order** of a finite group is the number of its elements.

3.2 Representation of groups

Lecture 2

Definition: A **representation** of a group (G, \circ) is a **map** D of elements $g \in G$ onto the set L of **linear operators** $D(g)$,

$$\begin{aligned} D : G &\rightarrow L , \\ g &\mapsto D(g) , \end{aligned}$$

with the properties

(i) $D(e) \equiv \mathbb{1}$,

the identity element is mapped onto the unit operator, and

(ii) $D(a \circ b) = D(a) \cdot D(b) \equiv D(a)D(b)$,

the group multiplication corresponds to the standard multiplication of linear operators. One can easily check that D is a **group homomorphism** between the group (G, \circ) and (a particular subset of) the set of linear operators L .

Example 1: $e^{i\phi}$ is strictly speaking not an element of the group $(U(1), \cdot)$, but the **representation** of such an element. In physics, one often uses the terms “representation of a group element” and “group element” synonymously.

Example 2: The linear operator $\hat{U}_r(\vec{a}) = \exp\left(-\frac{i}{\hbar} \vec{a} \cdot \hat{\vec{p}}\right)$ is strictly speaking the **representation** of an element of the group (G_r, \cdot) of space translations, namely that which affects a translation of the position vector \vec{r} by the vector \vec{a} .

Example 3: The **cyclic group of order 3**, (Z_3, \circ) , is defined by the following link table:

◦		e		a		b
e		e		a		b
a		a		b		e
b		b		e		a

Table 3.1: Link table of (Z_3, \circ) .

The link table obviously provides closure of the group. One also realizes that $b = a^{-1}$, such that the existence of an inverse is guaranteed. The validity of the law of associativity can also be readily proven using the link table. The group (Z_3, \circ) is not a continuous, but a so-called **discrete** group. Moreover it is **abelian**, as one can convince oneself using the link table. A **representation** of Z_3 is the set of phase factors

$$D(e) = 1, \quad D(a) = e^{2\pi i/3}, \quad D(b) = e^{-2\pi i/3} \equiv e^{4\pi i/3},$$

One readily checks that this representation, with the standard multiplication as group multiplication, has the same properties as given in the link table.

Definition: The **dimension** of a representation is the dimension of the space, onto which the linear operators of the representation act.

Example: In the above representations of $U(1)$ or Z_3 this is one dimension, hence these representations are one-dimensional.

Definition: The **regular representation** $\bar{D}(G)$ of a group (G, \circ) corresponds to assigning to each $g_j \in G$ a state vector $|g_j\rangle$ of an orthonormal basis of a Hilbert space \mathcal{H} ,

$$\forall g_j \in G : \quad g_j \mapsto |g_j\rangle, \quad \langle g_i | g_j \rangle = \delta_{ij} \quad \forall |g_i\rangle, |g_j\rangle \in \mathcal{H}, \quad (3.6)$$

and by defining the action of an arbitrary element $g_i \in G$ onto the state vector $|g_j\rangle$ via a linear map

$$\bar{D}(g_i) |g_j\rangle = |g_i \circ g_j\rangle \quad \forall g_i, g_j \in G. \quad (3.7)$$

Example: Let us consider (Z_3, \circ) . First we have to assign Hilbert-space states to the group elements,

$$\begin{aligned} g_1 \equiv e &\mapsto |g_1\rangle \equiv |e\rangle, \\ g_2 \equiv a &\mapsto |g_2\rangle \equiv |a\rangle, \\ g_3 \equiv b &\mapsto |g_3\rangle \equiv |b\rangle, \end{aligned}$$

with $\langle g_i | g_j \rangle = \delta_{ij}$. The regular representation follows according to the definition (3.7) and with the help of the link table 3.1:

$$\begin{aligned} \bar{D}(e)|e\rangle &= |e \circ e\rangle = |e\rangle, & \bar{D}(a)|e\rangle &= |a \circ e\rangle = |a\rangle, & \bar{D}(b)|e\rangle &= |b \circ e\rangle = |b\rangle, \\ \bar{D}(e)|a\rangle &= |e \circ a\rangle = |a\rangle, & \bar{D}(a)|a\rangle &= |a \circ a\rangle = |b\rangle, & \bar{D}(b)|a\rangle &= |b \circ a\rangle = |e\rangle, \\ \bar{D}(e)|b\rangle &= |e \circ b\rangle = |b\rangle, & \bar{D}(a)|b\rangle &= |a \circ b\rangle = |e\rangle, & \bar{D}(b)|b\rangle &= |b \circ b\rangle = |a\rangle. \end{aligned}$$

The **dimension of the regular representation** of a group corresponds to the **order** of the group. The order of the cyclic group (Z_3, \circ) is 3 and thus $\dim \bar{D}(Z_3) = 3$.

The regular representation (3.7) can also be written as a **matrix representation**. To this end one defines

$$[\bar{D}(g)]_{ij} \equiv \langle g_i | \bar{D}(g) | g_j \rangle \quad \forall g \in G, \quad |g_i\rangle, |g_j\rangle \in \mathcal{H}. \quad (3.8)$$

Example: For (Z_3, \circ) we have

$$[\bar{D}(e)] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad [\bar{D}(a)] = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad [\bar{D}(b)] = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

In the matrix representation the group multiplication is the ordinary matrix multiplication,

$$\begin{aligned} [\bar{D}(g_i \circ g_j)]_{k\ell} &= \langle g_k | \bar{D}(g_i \circ g_j) | g_\ell \rangle = \langle g_k | \bar{D}(g_i) \bar{D}(g_j) | g_\ell \rangle \\ &= \sum_m \langle g_k | \bar{D}(g_i) | g_m \rangle \langle g_m | \bar{D}(g_j) | g_\ell \rangle = \sum_m [\bar{D}(g_i)]_{km} [\bar{D}(g_j)]_{m\ell}. \end{aligned} \quad (3.9)$$

In general, matrix multiplication is **not** commutative. One readily checks, however, that this is still true for (Z_3, \circ) (since it is an abelian group).

In general, **matrix representations** of a group are again groups themselves, which are (globally) **isomorphic** to the group. Therefore, in physics one uses the term “group” synonymous with “matrix representation of a group” or “representation of a group”.

Example: Let us consider the group of space translations, (G_r, \cdot) . We have already encountered a **representation** of (G_r, \cdot) ,

$$D(G_r) = \left\{ \hat{U}_r(\vec{a}), \vec{a} \in \mathbb{R}^3 \right\},$$

with the **linear operator** $\hat{U}_r(\vec{a}) = \exp\left(-\frac{i}{\hbar} \vec{a} \cdot \hat{\vec{p}}\right)$. (Obviously, we have previously been somewhat sloppy with the use of the term “group” and “representation of a group”; we

had introduced the “group” of space translations as a possible representation of this group in the form of the linear operators $\hat{U}_r(\vec{a})$.) A **matrix representation** of (G_r, \cdot) can be obtained by taking an arbitrary orthonormal basis $\{|\psi_n\rangle, n = 0, 1, 2, \dots\}$ and constructing matrices from the representation $D(g(\vec{a}))$ of the group element $g(\vec{a}) \in G_r$,

$$[D(g(\vec{a}))]_{ij} = \langle \psi_i | D(g(\vec{a})) | \psi_j \rangle \equiv \langle \psi_i | \hat{U}_r(\vec{a}) | \psi_j \rangle = \langle \psi_i | \exp\left(-\frac{i}{\hbar} \vec{a} \cdot \hat{\vec{p}}\right) | \psi_j \rangle .$$

By inserting complete sets of position-space eigenstates this can be written as

$$\begin{aligned} [D(g(\vec{a}))]_{ij} &= \int d^3\vec{r}' d^3\vec{r} \langle \psi_i | \vec{r}' \rangle \langle \vec{r}' | \exp\left(-\frac{i}{\hbar} \vec{a} \cdot \hat{\vec{p}}\right) | \vec{r} \rangle \langle \vec{r} | \psi_j \rangle \\ &= \int d^3\vec{r}' d^3\vec{r} \psi_i^*(\vec{r}') \delta^{(3)}(\vec{r}' - \vec{r}) \exp\left(-\vec{a} \cdot \vec{\nabla}_r\right) \psi_j(\vec{r}) \\ &= \int d^3\vec{r} \psi_i^*(\vec{r}) \exp\left(-\vec{a} \cdot \vec{\nabla}_r\right) \psi_j(\vec{r}) , \end{aligned}$$

where we have used the momentum operator in coordinate representation and the orthonormality of the position-space eigenstates.

3.3 Lie groups

Lie groups are continuous groups with $N \in \mathbb{N}$ real-valued parameters $\vec{\alpha} = (\alpha_1, \dots, \alpha_N)^T \in \mathbb{R}^N$, whose elements (in the representation as linear operators) can be written in the form

$$\hat{U}(\vec{\alpha}) = \exp\left(-\frac{i}{\hbar} \sum_{j=1}^N \alpha_j \hat{X}_j\right) . \quad (3.10)$$

Obviously,

$$\hat{U}(\vec{0}) = \mathbb{1} , \quad (3.11)$$

the origin in parameter space is mapped onto the identity element (or its corresponding representation as linear operator, i.e., the unit operator).

Examples:

- (i) The elements of the group of space translations, (G_r, \cdot) , have a representation in the form of Eq. (3.10), with $N = 3$, $\hat{X}_j = \hat{p}^j$ and $\alpha_j = a^j$, $j = x, y, z$, cf. Eq. (2.7).
- (ii) The elements of the rotation group, (G_R, \cdot) , have a representation in the form of Eq. (3.10), with $N = 3$, $\hat{X}_j = \hat{L}^j$ and $\alpha_j = \phi^j$, $j = x, y, z$, cf. Eq. (2.32).

The so-called **generators** \hat{X}_j of the Lie group are defined as derivatives of a group element with respect to the parameters at the location of the identity element,

$$\hat{X}_j = i\hbar \left. \frac{\partial}{\partial \alpha_j} \hat{U}(\vec{\alpha}) \right|_{\vec{\alpha}=0} . \quad (3.12)$$

Infinitesimal transformations are obtained from the Taylor expansion of the exponential function for infinitesimal parameters, $|\delta\alpha_i| \ll 1$, $i = 1, \dots, N$,

$$\hat{U}(\delta\vec{\alpha}) = \mathbb{1} - \frac{i}{\hbar} \sum_{j=1}^N \delta\alpha_j \hat{X}_j + O(\delta\alpha^2). \quad (3.13)$$

Any **finite** transformation can be written as an **infinite sequence** of **infinitesimal** transformations. To see this, let $\alpha_j = n\delta\alpha_j$, $n \in \mathbb{N}$, $j = 1, \dots, N$, and compute

$$\begin{aligned} \lim_{n \rightarrow \infty} [\hat{U}(\delta\vec{\alpha})]^n &= \lim_{n \rightarrow \infty} \left[\mathbb{1} - \frac{i}{\hbar} \sum_{j=1}^N \delta\alpha_j \hat{X}_j \right]^n = \lim_{n \rightarrow \infty} \left[\mathbb{1} - \frac{i}{\hbar} \sum_{j=1}^N \frac{\alpha_j}{n} \hat{X}_j \right]^n \\ &= \exp \left(-\frac{i}{\hbar} \sum_{j=1}^N \alpha_j \hat{X}_j \right) \equiv \hat{U}(\vec{\alpha}), \end{aligned}$$

where we have used the identity

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n} \right)^n = e^x,$$

with $x = -(i/\hbar) \sum_{j=1}^N \alpha_j \hat{X}_j$.

Consequences:

- (i) The generators are **linearly independent**.

Proof: The identity element is unique, so that from the condition

$$\hat{U}(\vec{\alpha}) = \mathbb{1}$$

follows that

$$\alpha_1 = \dots = \alpha_N \equiv 0.$$

This also holds for infinitesimal parameters $\delta\alpha_j$, i.e., from the condition

$$\hat{U}(\delta\vec{\alpha}) = \mathbb{1} - \frac{i}{\hbar} \sum_{j=1}^N \delta\alpha_j \hat{X}_j + O(\delta\alpha^2) = \mathbb{1} \quad (3.14)$$

follows that

$$\delta\alpha_1 = \dots = \delta\alpha_N \equiv 0.$$

By subtracting the $\mathbb{1}$ on both sides of Eq. (3.14) and after neglecting quadratically small terms, this condition can be written as

$$\sum_{j=1}^N \delta\alpha_j \hat{X}_j = 0. \quad (3.15)$$

However, if from this condition follows that all $\delta\alpha_j = 0$, then all \hat{X}_j must be linearly independent, generalizing the definition of linear independence from vectors (cf. lectures on linear algebra) to operators, q.e.d.

(ii) For **unitary** Lie groups the generators \hat{X}_j are **hermitean**.

Proof: For unitary Lie groups we have

$$\hat{U}^\dagger(\vec{\alpha}) = \hat{U}^{-1}(\vec{\alpha}) . \quad (3.16)$$

For infinitesimal transformations this leads to

$$\begin{aligned} \mathbb{1} + \frac{i}{\hbar} \sum_{j=1}^N \delta\alpha_j^* \hat{X}_j^\dagger + O(\delta\alpha^2) &= \mathbb{1} - \frac{i}{\hbar} \sum_{j=1}^N (-\delta\alpha_j) \hat{X}_j + O(\delta\alpha^2) \\ \iff \sum_{j=1}^N \delta\alpha_j \hat{X}_j^\dagger &= \sum_{j=1}^N \delta\alpha_j \hat{X}_j \\ \implies \hat{X}_j^\dagger &= \hat{X}_j , \quad j = 1, \dots, N , \end{aligned}$$

where in the second step we have used that the parameters are real-valued, $\delta\alpha_j^* = \delta\alpha_j$, and in the last step that they can be chosen arbitrarily, q.e.d.

Example 1: The elements of the group $U(1)$ have the representation $e^{i\phi}$, with the (single) parameter $\phi \in \mathbb{R}$ and the (single) generator $\hat{X} \equiv \mathbb{1} \equiv 1$.

Example 2: The group $SU(2)$ is the group of **special unitary (2×2) matrices**, i.e., the unitary (2×2) matrices with determinant $+1$. A representation of the elements of $SU(2)$ is

$$\hat{U}(\vec{\alpha}) = \exp\left(-\frac{i}{\hbar} \vec{\alpha} \cdot \hat{S}\right) , \quad (3.17)$$

where the parameters are $\alpha_j \in \mathbb{R}$, $j = 1, 2, 3$, and $\hat{S} = \frac{\hbar}{2} \hat{\sigma}$ is the **spin operator**, which is proportional to the vector of **Pauli matrices** $\hat{\sigma} = (\hat{\sigma}_1, \hat{\sigma}_2, \hat{\sigma}_3)^T$,

$$\hat{\sigma}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \hat{\sigma}_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad \hat{\sigma}_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \quad (3.18)$$

For this reason, the group $SU(2)$ is also called **spin group**.

Lecture 3

From Eq. (3.17) it is obvious that $SU(2)$ is a Lie group. But are the elements (3.17) in fact unitary (2×2) matrices with determinant $+1$? The unitarity is immediately obvious from the fact that the parameters α_j are real-valued and \hat{S} is a hermitean operator (since spin is a physical observable, or because the Pauli matrices are hermitean, respectively),

$$\hat{U}^\dagger(\vec{\alpha}) = \exp\left(\frac{i}{\hbar} \vec{\alpha} \cdot \hat{S}^\dagger\right) = \exp\left(-\frac{i}{\hbar} (-\vec{\alpha}) \cdot \hat{S}\right) = \hat{U}^{-1}(\vec{\alpha}) .$$

In order to show that $\hat{U}(\vec{\alpha})$ is a (2×2) matrix, we first compute

$$(\vec{\alpha} \cdot \hat{\sigma})^{2n} = \left[(\vec{\alpha} \cdot \hat{\sigma})^2\right]^n = \left[\vec{\alpha} \cdot \vec{\alpha} + i \hat{\sigma} \cdot (\vec{\alpha} \times \vec{\alpha})\right]^n = \alpha^{2n} \mathbb{1} , \quad (3.19)$$

$$(\vec{\alpha} \cdot \hat{\sigma})^{2n+1} = (\vec{\alpha} \cdot \hat{\sigma})^{2n} \vec{\alpha} \cdot \hat{\sigma} = \alpha^{2n} \vec{\alpha} \cdot \hat{\sigma} = \alpha^{2n+1} \frac{\vec{\alpha}}{\alpha} \cdot \hat{\sigma} , \quad (3.20)$$

where we have used the identity

$$\hat{\sigma} \cdot \vec{a} \hat{\sigma} \cdot \vec{b} = \vec{a} \cdot \vec{b} \mathbb{1} + i \hat{\sigma} \cdot (\vec{a} \times \vec{b}) , \quad (3.21)$$

readily proven by employing the commutation and anticommutation relations of the Pauli matrices,

$$[\hat{\sigma}_i, \hat{\sigma}_j] = 2i \epsilon_{ijk} \hat{\sigma}_k \quad (3.22)$$

$$\{\hat{\sigma}_i, \hat{\sigma}_j\} = 2\delta_{ij} \mathbb{1} . \quad (3.23)$$

With Eqs. (3.19), (3.20) we can write the Taylor expansion of an element (3.17) of the spin group as

$$\begin{aligned} \hat{U}(\vec{\alpha}) &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i}{2}\right)^n (\vec{\alpha} \cdot \hat{\sigma})^n \\ &= \sum_{n=0}^{\infty} \frac{1}{(2n)!} \left(-\frac{i}{2}\right)^{2n} (\vec{\alpha} \cdot \hat{\sigma})^{2n} + \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} \left(-\frac{i}{2}\right)^{2n+1} (\vec{\alpha} \cdot \hat{\sigma})^{2n+1} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \left(\frac{\alpha}{2}\right)^{2n} \mathbb{1} - i \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \left(\frac{\alpha}{2}\right)^{2n+1} \frac{\vec{\alpha}}{\alpha} \cdot \hat{\sigma} \\ &= \cos \frac{\alpha}{2} \mathbb{1} - i \sin \frac{\alpha}{2} \frac{\vec{\alpha}}{\alpha} \cdot \hat{\sigma} . \end{aligned} \quad (3.24)$$

This is a linear combination of (2×2) matrices and thus itself a (2×2) matrix. Finally, we compute the determinant of $\hat{U}(\vec{\alpha})$:

$$\ln \det \hat{U}(\vec{\alpha}) = \text{Tr} \ln \hat{U}(\vec{\alpha}) = \text{Tr} \left(-\frac{i}{2} \vec{\alpha} \cdot \hat{\sigma}\right) = -\frac{i}{2} \vec{\alpha} \cdot \text{Tr} \hat{\sigma} \equiv 0 ,$$

since the Pauli matrices are traceless. Then we have

$$\det \hat{U}(\vec{\alpha}) = +1 , \quad \text{q.e.d. .}$$

From the closure of the group under multiplication we now derive another important relation for the generators of a Lie group. In general,

$$\exp \left(-\frac{i}{\hbar} \sum_{j=1}^N \alpha_j \hat{X}_j\right) \exp \left(-\frac{i}{\hbar} \sum_{k=1}^N \beta_k \hat{X}_k\right) \neq \exp \left[-\frac{i}{\hbar} \sum_{i=1}^N (\alpha_i + \beta_i) \hat{X}_i\right] .$$

However, it is always true that multiplying two group elements yields another element of the group,

$$\exp \left(-\frac{i}{\hbar} \sum_{j=1}^N \alpha_j \hat{X}_j\right) \exp \left(-\frac{i}{\hbar} \sum_{k=1}^N \beta_k \hat{X}_k\right) = \exp \left(-\frac{i}{\hbar} \sum_{i=1}^N \delta_i \hat{X}_i\right) .$$

3 Introduction to Group Theory

We now compute δ_i . To this end we take the logarithm on both sides,

$$\sum_{i=1}^N \delta_i \hat{X}_i = i\hbar \ln \left[\exp \left(-\frac{i}{\hbar} \sum_{j=1}^N \alpha_j \hat{X}_j \right) \exp \left(-\frac{i}{\hbar} \sum_{k=1}^N \beta_k \hat{X}_k \right) \right],$$

and expand the exponential function up to second order in the parameters α_j, β_k ,

$$\begin{aligned} \sum_{i=1}^N \delta_i \hat{X}_i &\simeq i\hbar \ln \left[\left(\mathbb{1} - \frac{i}{\hbar} \sum_{j=1}^N \alpha_j \hat{X}_j - \frac{1}{2\hbar^2} \sum_{j,k=1}^N \alpha_j \alpha_k \hat{X}_j \hat{X}_k \right) \right. \\ &\quad \times \left. \left(\mathbb{1} - \frac{i}{\hbar} \sum_{\ell=1}^N \beta_\ell \hat{X}_\ell - \frac{1}{2\hbar^2} \sum_{\ell,m=1}^N \beta_\ell \beta_m \hat{X}_\ell \hat{X}_m \right) \right] \\ &\simeq i\hbar \ln \left[\mathbb{1} - \frac{i}{\hbar} \sum_{j=1}^N (\alpha_j + \beta_j) \hat{X}_j - \frac{1}{2\hbar^2} \sum_{j,k=1}^N (\alpha_j \alpha_k + 2\alpha_j \beta_k + \beta_j \beta_k) \hat{X}_j \hat{X}_k \right]. \end{aligned}$$

Now we also expand the logarithm up to second order in the parameters α_j, β_k , using the formula

$$\ln(1+x) = x - \frac{x^2}{2} + O(x^3),$$

and obtain

$$\begin{aligned} \sum_{i=1}^N \delta_i \hat{X}_i &\simeq i\hbar \left[-\frac{i}{\hbar} \sum_{j=1}^N (\alpha_j + \beta_j) \hat{X}_j - \frac{1}{2\hbar^2} \sum_{j,k=1}^N (\alpha_j \alpha_k + 2\alpha_j \beta_k + \beta_j \beta_k) \hat{X}_j \hat{X}_k \right. \\ &\quad \left. + \frac{1}{2\hbar^2} \sum_{j,k=1}^N (\alpha_j \alpha_k + \alpha_j \beta_k + \beta_j \alpha_k + \beta_j \beta_k) \hat{X}_j \hat{X}_k \right] \\ &= \sum_{j=1}^N (\alpha_j + \beta_j) \hat{X}_j - \frac{i}{2\hbar} \sum_{j,k=1}^N (\alpha_j \beta_k - \beta_j \alpha_k) \hat{X}_j \hat{X}_k. \end{aligned}$$

If we move the first term on the right-hand side to the left-hand side and exchange the summation indices in the second term, $\beta_j \alpha_k \hat{X}_j \hat{X}_k \rightarrow \beta_k \alpha_j \hat{X}_k \hat{X}_j$, we obtain

$$2i\hbar \sum_{i=1}^N (\delta_i - \alpha_i - \beta_i) \hat{X}_i = \sum_{j,k=1}^N \alpha_j \beta_k [\hat{X}_j, \hat{X}_k]. \quad (3.25)$$

If all generators commute with each other,

$$[\hat{X}_j, \hat{X}_k] = 0 \quad \forall j, k = 1, \dots, N,$$

it is now obvious that

$$\delta_i = \alpha_i + \beta_i, \quad i = 1, \dots, N,$$

due to the linear independence of the generators. Then, however, we have

$$\begin{aligned} \exp\left(-\frac{i}{\hbar} \sum_{j=1}^N \alpha_j \hat{X}_j\right) \exp\left(-\frac{i}{\hbar} \sum_{k=1}^N \beta_k \hat{X}_k\right) &= \exp\left[-\frac{i}{\hbar} \sum_{i=1}^N (\alpha_i + \beta_i) \hat{X}_i\right] \\ &\equiv \exp\left(-\frac{i}{\hbar} \sum_{k=1}^N \beta_k \hat{X}_k\right) \exp\left(-\frac{i}{\hbar} \sum_{j=1}^N \alpha_j \hat{X}_j\right), \end{aligned} \quad (3.26)$$

i.e., the group elements themselves **commute**. In this case we therefore have an **abelian group**.

However, in general not all generators commute with each other. Let us denote

$$2i\hbar(\delta_i - \alpha_i - \beta_i) \equiv i\hbar \gamma_i \quad (3.27)$$

and compute the γ_i . Obviously, they have to be proportional to the α_j as well as to the β_k , since they vanish if **either** all α_j **or** all β_k vanish, cf. Eq. (3.25). We therefore make the following **Ansatz**:

$$\gamma_i = \sum_{j,k=1}^N f_{jki} \alpha_j \beta_k, \quad (3.28)$$

with some constants f_{jki} . Inserting this into Eq. (3.25) we obtain

$$i\hbar \sum_{i,j,k=1}^N f_{jki} \alpha_j \beta_k \hat{X}_i = \sum_{j,k=1}^N \alpha_j \beta_k [\hat{X}_j, \hat{X}_k], \quad (3.29)$$

or, since α_j, β_k can assume arbitrary values,

$$[\hat{X}_j, \hat{X}_k] = i\hbar \sum_{i=1}^N f_{jki} \hat{X}_i \equiv i\hbar f_{jki} \hat{X}_i, \quad (3.30)$$

where we have again used Einstein's sum convention. These commutation relations define an **algebra** for the generators of the Lie group, the so-called **Lie algebra**. The constants f_{jki} are the so-called **structure constants** of the group.

Example 1: For the rotation group (G_R, \cdot) , Eq. (3.2), the three generators are the components of the **angular momentum operator**, which fulfill the **angular momentum algebra** (3.3). Therefore we have

$$f_{jki} \equiv \epsilon^{jki}.$$

Example 2: For the spin group $(SU(2), \cdot)$ with elements (3.17) the generators are the components of the **spin operator**. On account of the commutation relation (3.22) for the Pauli matrices, these obey the Lie algebra

$$[\hat{S}^j, \hat{S}^k] = \frac{\hbar^2}{4} [\hat{\sigma}_j, \hat{\sigma}_k] = \frac{\hbar^2}{4} 2i \epsilon_{jki} \hat{\sigma}_i \equiv i\hbar \epsilon^{jki} \hat{S}^i, \quad (3.31)$$

i.e., the same Lie algebra as the rotation group (G_R, \cdot) .

Properties of the structure constants:

(i) Because of $[\hat{X}_j, \hat{X}_k] = -[\hat{X}_k, \hat{X}_j]$ we have

$$f_{jki} = -f_{kji} . \quad (3.32)$$

For $(SU(2), \cdot)$ and (G_R, \cdot) this is automatically fulfilled due to the antisymmetry of the Levi-Civita tensor ϵ^{ijk} .

(ii) For **unitary** groups the structure constants are **real-valued**. Namely, using the hermiticity of the generators we compute

$$\begin{aligned} [\hat{X}_j, \hat{X}_k]^\dagger &= [\hat{X}_k^\dagger, \hat{X}_j^\dagger] = [\hat{X}_k, \hat{X}_j] = -[\hat{X}_j, \hat{X}_k] = -i\hbar f_{jki} \hat{X}_i \\ \iff -i\hbar f_{jki}^* \hat{X}_i^\dagger &= -i\hbar f_{jki}^* \hat{X}_i = -i\hbar f_{jki} \hat{X}_i . \end{aligned}$$

Due to the linear independence of the \hat{X}_i we immediately conclude that

$$f_{jki}^* = f_{jki} , \quad \text{q.e.d.} . \quad (3.33)$$

(iii) **Jacobi identity:**

$$f_{ijl} f_{lkm} + f_{jkl} f_{lim} + f_{kil} f_{ljm} = 0 . \quad (3.34)$$

Proof: We have the following identity for the generators:

$$\begin{aligned} &[[\hat{X}_i, \hat{X}_j], \hat{X}_k] + [[\hat{X}_j, \hat{X}_k], \hat{X}_i] + [[\hat{X}_k, \hat{X}_i], \hat{X}_j] \\ &= \hat{X}_i \hat{X}_j \hat{X}_k - \hat{X}_j \hat{X}_i \hat{X}_k - \hat{X}_k \hat{X}_i \hat{X}_j + \hat{X}_k \hat{X}_j \hat{X}_i \\ &+ \hat{X}_j \hat{X}_k \hat{X}_i - \hat{X}_k \hat{X}_j \hat{X}_i - \hat{X}_i \hat{X}_j \hat{X}_k + \hat{X}_i \hat{X}_k \hat{X}_j \\ &+ \hat{X}_k \hat{X}_i \hat{X}_j - \hat{X}_i \hat{X}_k \hat{X}_j - \hat{X}_j \hat{X}_k \hat{X}_i + \hat{X}_j \hat{X}_i \hat{X}_k \equiv 0 . \end{aligned}$$

Using the Lie algebra (3.30) this is identical to

$$\begin{aligned} 0 &= i\hbar \left\{ f_{ijl} [\hat{X}_l, \hat{X}_k] + f_{jkl} [\hat{X}_l, \hat{X}_i] + f_{kil} [\hat{X}_l, \hat{X}_j] \right\} \\ &= -\hbar^2 (f_{ijl} f_{lkm} + f_{jkl} f_{lim} + f_{kil} f_{ljm}) \hat{X}_m . \end{aligned}$$

Due to the linear independence of the generators the term in parentheses has to vanish, i.e., the Jacobi identity (3.34) holds, q.e.d.

Example: The structure constants of the groups $(SU(2), \cdot)$ and (G_R, \cdot) , respectively, are given by the components of the Levi-Civita tensor. From the lectures on classical mechanics we already know that these obey the Jacobi identity (3.34).

Lie groups have the special property that, on account of Eq. (3.10), knowledge of the generators suffices to determine **all** group elements. A **representation of the generators** then defines a **representation of the group**. A particularly important representation is the so-called **adjoint representation**, where the generators are determined by the structure constants,

$$(\hat{X}_i)_{jk} = -i\hbar f_{ijk} . \quad (3.35)$$

This is a **matrix representation** of the generator \hat{X}_i ; the (jk) element of the i th generator is (up to a factor $-i\hbar$) given by the structure constant f_{ijk} . Due to Eq. (3.10) it is obvious that, for matrix-valued generators, also the group elements are matrix-valued. The **adjoint representation** thus also defines a **matrix representation of the group**.

We now show that the adjoint representation (3.35) is indeed a possible representation of the generators, because also in this particular representation the latter obey the Lie algebra of the group. We start with the Jacobi identity (3.34), which we write using Eqs. (3.32) and (3.35) in the form

$$\begin{aligned} \hbar^2 f_{ij\ell} f_{\ell km} &= -\hbar^2 (f_{kil} f_{\ell jm} + f_{jkl} f_{\ell im}) \\ &= -\hbar^2 [(-f_{ik\ell})(-f_{j\ell m}) + f_{jkl}(-f_{i\ell m})] \\ \iff i\hbar f_{ij\ell} (\hat{X}_\ell)_{km} &= (\hat{X}_i)_{k\ell} (\hat{X}_j)_{\ell m} - (\hat{X}_j)_{k\ell} (\hat{X}_i)_{\ell m} \\ &\equiv \left([\hat{X}_i, \hat{X}_j] \right)_{km}. \end{aligned}$$

This is indeed the correct Lie algebra (3.30) for the (km) element of the generators in matrix representation.

Example: The adjoint representation of the generators $\hat{L}^x, \hat{L}^y, \hat{L}^z$ of the rotation group (G_R, \cdot) reads

$$(\hat{L}^i)_{jk} = -i\hbar \epsilon_{ijk}. \quad (3.36)$$

Therefore, the generators are the (3×3) matrices

$$\hat{L}^x = -i\hbar \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad \hat{L}^y = -i\hbar \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \hat{L}^z = -i\hbar \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (3.37)$$

Let us for example consider a rotation around the z axis,

$$\exp\left(-\frac{i}{\hbar} \phi \hat{L}^z\right) = \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar} \phi\right)^n (\hat{L}^z)^n. \quad (3.38)$$

In the adjoint representation,

$$(\hat{L}^z)^{2n} = (-i\hbar)^{2n} \begin{pmatrix} (-1)^n & 0 & 0 \\ 0 & (-1)^n & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (\hat{L}^z)^{2n+1} = (-i\hbar)^{2n+1} (-1)^n \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Using this result Eq. (3.38) becomes

$$\begin{aligned} \exp\left(-\frac{i}{\hbar} \phi \hat{L}^z\right) &= \mathbb{1} + \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n)!} \phi^{2n} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \phi^{2n+1} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (\cos \phi - 1) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \sin \phi \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \equiv \hat{D}_{\vec{e}_z}(-\phi). \end{aligned}$$

This is just the **rotation matrix** for rotations of **three-dimensional vectors** by an angle $-\phi$ around the z axis. Analogously, using the generators \hat{L}^x and \hat{L}^y in the adjoint representation, one can compute the rotation matrices for rotations around the x and y axes.

Lecture 4

Rotation matrices in three dimensions form the group of **special, orthogonal (3×3) matrices with determinant +1**, or short **($SO(3), \cdot$)**. We have just shown that **($SO(3), \cdot$)** is **isomorphic** to the **adjoint representation** of the rotation group **(G_R, \cdot)**. Since we use the terms “representation of a group” and “group” synonymously, we can also say that **($SO(3), \cdot$)** is isomorphic to the rotation group **(G_R, \cdot)**, or that **($SO(3), \cdot$)** is the rotation group. It can be shown that **($SO(3), \cdot$)** is (locally) isomorphic to **($SU(2), \cdot$)**. Thus, **($SU(2), \cdot$)** is also (locally) isomorphic to **(G_R, \cdot)**.

Since the structure constants of the group **($SU(2), \cdot$)** are the same as those of **(G_R, \cdot)**, Eq. (3.37) represents also the **adjoint representation** of the generators \hat{S}^i of **($SU(2), \cdot$)**. Note the difference to the so-called **fundamental representation** $\hat{S}^i = \frac{\hbar}{2} \hat{\sigma}^i$ of these generators. For **($SU(2), \cdot$)** the fundamental representation consists of (2×2) matrices, while the adjoint is formed of (3×3) matrices.

In general, the form of the structure constants depends on the choice for the generators. However, the Lie algebra (3.30) must always be fulfilled, thus the structure constants are **always antisymmetric** in the **first two** indices, cf. Eq. (3.32). However, once can choose the generators in such a way that the structure constants are **completely antisymmetric in all indices**,

$$f_{jki} = f_{kij} = f_{ijk} = -f_{kji} = -f_{jik} = -f_{ikj} . \tag{3.39}$$

Remark: For the rotation group, **(G_R, \cdot)**, and thus also for **($SO(3), \cdot$)** and **($SU(2), \cdot$)**, this is automatically fulfilled, since $f_{jki} = \epsilon_{jki}$, and the Levi-Civita tensor is completely antisymmetric in all indices.

Proof: We need a basis in which the generators fulfill

$$\text{Tr} \left(\hat{X}_i \hat{X}_j \right) = \lambda_i \delta_{ij} , \tag{3.40}$$

where we do not sum over i on the right-hand side. This equation can be interpreted in the sense that the generators are **orthogonal** in this choice of basis. By suitably normalizing the generators one can moreover achieve that

$$\text{Tr} \left(\hat{X}_i \hat{X}_j \right) = \frac{\hbar^2}{2} \delta_{ij} . \tag{3.41}$$

Using the Lie algebra (3.30) we now compute

$$-\frac{2i}{\hbar^3} \text{Tr} \left\{ \left[\hat{X}_j, \hat{X}_k \right] \hat{X}_i \right\} = \frac{2}{\hbar^2} f_{jkl} \text{Tr} \left(\hat{X}_l \hat{X}_i \right) = f_{jkl} \delta_{li} = f_{jki} . \tag{3.42}$$

On the other hand, the left-hand side of this equation is identical to

$$\begin{aligned} -\frac{2i}{\hbar^3} \text{Tr} \left(\hat{X}_j \hat{X}_k \hat{X}_i - \hat{X}_k \hat{X}_j \hat{X}_i \right) &= -\frac{2i}{\hbar^3} \text{Tr} \left(\hat{X}_k \hat{X}_i \hat{X}_j - \hat{X}_i \hat{X}_k \hat{X}_j \right) \\ &= -\frac{2i}{\hbar^3} \text{Tr} \left\{ \left[\hat{X}_k, \hat{X}_i \right] \hat{X}_j \right\} \equiv f_{kij} , \end{aligned} \tag{3.43}$$

where we used the cyclic permutability of operators under the trace and the result of the computation which had lead to Eq. (3.42). Comparison with Eq. (3.42) already yields the first equality in Eq. (3.39). Starting from Eq. (3.43) and performing a second cyclic permutation under the trace we also obtain the second equality in Eq. (3.39). The remaining equalities can now be obtained from the previous ones using property (3.32).

It remains to prove whether one can find a basis in which Eq. (3.40) is fulfilled. To this end, consider a linear transformation L of the generators (a “change of basis”), such that

$$\hat{X}_i \longrightarrow \hat{X}'_i = L_{ik} \hat{X}_k . \quad (3.44)$$

In this new basis, the commutation relation (3.30) reads

$$\begin{aligned} [\hat{X}'_i, \hat{X}'_j] &= L_{ik} L_{j\ell} [\hat{X}_k, \hat{X}_\ell] = L_{ik} L_{j\ell} i\hbar f_{k\ell m} \hat{X}_m \\ &= i\hbar L_{ik} L_{j\ell} f_{k\ell m} L_{mn}^{-1} L_{nr} \hat{X}_r = i\hbar L_{ik} L_{j\ell} f_{k\ell m} L_{mn}^{-1} \hat{X}'_n \\ &\equiv i\hbar f'_{ijn} \hat{X}'_n , \end{aligned}$$

where, from the first to the second line, we used the invertibility of the linear transformation,

$$L_{mn}^{-1} L_{nr} \equiv \delta_{mr}$$

and then, to obtain the next equality, the definition (3.44) of the change of basis. Finally, the last identity arises from demanding that also the new generators \hat{X}'_i must obey the Lie algebra (3.30), with the **new structure constants** f'_{ijn} . We read off the following transformation behavior for the structure constants:

$$f_{ijn} \longrightarrow f'_{ijn} \equiv L_{ik} L_{j\ell} f_{k\ell m} L_{mn}^{-1} . \quad (3.45)$$

These new structure constants define a new **adjoint representation** of the generators:

$$\begin{aligned} (\hat{X}'_i)_{jk} &= -i\hbar f'_{ijk} = -i\hbar L_{im} L_{jn} f_{mnr} L_{rk}^{-1} \\ &= L_{im} L_{jn} (\hat{X}_m)_{nr} L_{rk}^{-1} = L_{im} \left(L \hat{X}_m L^{-1} \right)_{jk} , \end{aligned}$$

where we used the adjoint representation of the old generators \hat{X}_m . In the adjoint representation we now compute

$$\begin{aligned} \text{Tr} \left(\hat{X}'_i \hat{X}'_j \right) &= L_{im} L_{jn} \text{Tr} \left(L \hat{X}_m L^{-1} L \hat{X}_n L^{-1} \right) = L_{im} L_{jn} \text{Tr} \left(\hat{X}_m \hat{X}_n \right) \\ &= L_{im} \text{Tr} \left(\hat{X}_m \hat{X}_n \right) L_{nj}^T . \end{aligned} \quad (3.46)$$

For unitary groups the structure constants are real-valued, thus the matrices

$$M'_{ij} \equiv \text{Tr} \left(\hat{X}'_i \hat{X}'_j \right) , \quad M_{mn} \equiv \text{Tr} \left(\hat{X}_m \hat{X}_n \right) ,$$

are also real-valued. Moreover, because of the cyclic permutability under the trace, these matrices are symmetric. Therefore, they can be diagonalized with the help of **orthogonal**

matrices. Choosing for the linear transformation L that orthogonal matrix $O \equiv L$ which diagonalizes M , we obtain from Eq. (3.46) the result

$$\lambda_i \delta_{ij} = M'_{ij} = O_{im} M_{mn} O_{nj}^T .$$

Thus, there exists indeed a basis (which one obtains by a suitable linear transformation $L \equiv O$) in which

$$\text{Tr} \left(\hat{X}'_i \hat{X}'_j \right) \equiv \lambda_i \delta_{ij} , \quad \text{q.e.d. .}$$

3.4 Simple and semi-simple Lie groups

Definition: An invariant subalgebra \mathcal{I} , a so-called **ideal**, is a set of generators

$$\mathcal{I} = \left\{ \hat{X}_i, i = 1, \dots, M \right\} , \quad M < N ,$$

of the set of all generators

$$\left\{ \hat{Y}_j, j = 1, \dots, N \right\} ,$$

with the property

$$[\hat{Y}_j, \hat{X}_i] \in \mathcal{I} \quad \forall i = 1, \dots, M, \quad j = 1, \dots, N, \quad (3.47)$$

i.e., the commutator between an arbitrary generator \hat{Y}_j with any generator from the ideal is again proportional to (some linear combination of) the generators from the ideal.

Definition: An invariant subgroup $U \subset G$ is the set of all group elements $u \in G$ for which holds

$$g^{-1} u g \in U \quad \forall g \in G .$$

Theorem: An ideal defines an invariant subgroup \mathcal{N} , a so-called **normal Lie subgroup**,

$$\mathcal{N} = \left\{ h = \exp \left(-\frac{i}{\hbar} \alpha_i \hat{X}_i \right) \right\} \subset G .$$

In other words, for

$$\begin{aligned} h &= e^{-iX} \in \mathcal{N}, \quad X \equiv \frac{1}{\hbar} \alpha_i \hat{X}_i, \\ g &= e^{-iY} \in G, \quad Y \equiv \frac{1}{\hbar} \beta_j \hat{Y}_j, \end{aligned}$$

we have

$$g^{-1} h g \in \mathcal{N} .$$

Proof: We have

$$\begin{aligned} g^{-1} h g &= e^{iY} e^{-iX} e^{-iY} = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} e^{iY} X^n e^{-iY} \\ &= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} (e^{iY} X e^{-iY})^n \equiv \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} X'^n = e^{-iX'} \equiv h' , \end{aligned}$$

where we defined $X' \equiv e^{iY} X e^{-iY}$. We now have to show that $X' \in \mathcal{I}$, since then $h' \in \mathcal{N}$. To this end consider the following operator,

$$X'(\epsilon) \equiv e^{i\epsilon Y} X e^{-i\epsilon Y} .$$

Obviously, $X'(0) \equiv X$ and $X'(1) \equiv X'$. A Taylor expansion of $X'(\epsilon)$ around $\epsilon = 0$ yields

$$\begin{aligned} X'(\epsilon) &= X + \epsilon \left. \frac{\partial}{\partial \epsilon} X'(\epsilon) \right|_{\epsilon=0} + \frac{\epsilon^2}{2} \left. \frac{\partial^2}{\partial \epsilon^2} X'(\epsilon) \right|_{\epsilon=0} + \dots \\ &= X + \epsilon \left[iY e^{i\epsilon Y} X e^{-i\epsilon Y} + e^{i\epsilon Y} X (-iY) e^{-i\epsilon Y} \right]_{\epsilon=0} \\ &\quad + \frac{\epsilon^2}{2} \left(-Y^2 e^{i\epsilon Y} X e^{-i\epsilon Y} + 2Y e^{i\epsilon Y} X Y e^{-i\epsilon Y} - e^{i\epsilon Y} X Y^2 e^{-i\epsilon Y} \right)_{\epsilon=0} + \dots \\ &= X + i\epsilon [Y, X] - \frac{\epsilon^2}{2} (Y^2 X - 2Y X Y + X Y^2) + \dots \\ &= X + i\epsilon [Y, X] - \frac{\epsilon^2}{2} [Y, [Y, X]] + \dots . \end{aligned}$$

From this follows that

$$X' = X'(1) = X + i[Y, X] - \frac{1}{2} [Y, [Y, X]] + \dots .$$

However, since

$$\begin{aligned} X &= \frac{1}{\hbar} \alpha_i \hat{X}_i \in \mathcal{I} , \\ [Y, X] &= \frac{1}{\hbar^2} \alpha_i \beta_j [\hat{Y}_j, \hat{X}_i] \in \mathcal{I} , \\ [Y, [Y, X]] &= \frac{1}{\hbar} \beta_j [\hat{Y}_j, [Y, X]] \in \mathcal{I} , \quad \text{etc.} , \end{aligned}$$

we obtain that $X' \in \mathcal{I}$, q.e.d.

Definition: A so-called **abelian normal Lie subgroup** is an invariant Lie subgroup where **all** elements commute with each other,

$$\forall h, h' \in \mathcal{N} : [h, h'] = 0 .$$

It is obvious that the ideal \mathcal{I} corresponding to \mathcal{N} has to be a so-called **abelian ideal**,

$$[\hat{X}_i, \hat{X}_j] = 0 , \quad \forall i, j = 1, \dots, M .$$

Definition: A Lie group is called **simple**, if it **does not have a normal Lie subgroup**. Its corresponding Lie algebra is called **simple**, if it **does not have an ideal** (except for the Lie algebra itself, $M = N$, and the zero).

Definition: A Lie group is called **semi-simple**, if it **does not have an abelian normal Lie subgroup**. Its corresponding Lie algebra is called **semi-simple**, if it **does not have an abelian ideal**.

Example 1: Let us consider the rotation group (G_R, \cdot) (or $(SO(3), \cdot)$ or $(SU(2), \cdot)$). The generators obey the commutation relations (3.3),

$$\begin{aligned} [\hat{L}^x, \hat{L}^y] &= i\hbar \hat{L}^z, \\ [\hat{L}^y, \hat{L}^z] &= i\hbar \hat{L}^x, \\ [\hat{L}^z, \hat{L}^x] &= i\hbar \hat{L}^y, \end{aligned}$$

i.e., there is **no** ideal. (G_R, \cdot) is therefore not only a semi-simple but even a **simple** Lie group.

Example 2: Let us consider the **product group** $(SU(2) \times SO(3), \cdot)$. The generators of this group are the components \hat{S}^i of the spin operator (for $SU(2)$) and the components \hat{L}^j of the angular momentum operator (for $SO(3)$). We have the commutation relations

$$\begin{aligned} [\hat{S}^i, \hat{S}^j] &= i\hbar \epsilon^{ijk} \hat{S}^k, \\ [\hat{L}^i, \hat{L}^j] &= i\hbar \epsilon^{ijk} \hat{L}^k, \\ [\hat{S}^i, \hat{L}^j] &= 0. \end{aligned}$$

Thus, there are two non-abelian ideals, the set of the \hat{S}^i and the set of the \hat{L}^j . Therefore, the product group $(SU(2) \times SO(3), \cdot)$ is **semi-simple**.

Definition: The **rank** of a semi-simple Lie group is equal to the number of **mutually commuting generators** of its Lie algebra.

Example 1: The group of space translations (G_r, \cdot) : A representation of the group elements is given by Eq. (2.7). The generators are the components \hat{p}^i of the momentum operator. All three generators mutually commute $[\hat{p}^i, \hat{p}^j] = 0$, $i, j = 1, 2, 3$. Hence, $\text{rank}(G_r, \cdot) = 3$.

Example 2: The rotation group (G_R, \cdot) : A representation of the group elements is given by Eq. (2.32). The generators obey the angular momentum algebra (3.3). Each generator \hat{L}^i commutes only with itself, $\text{rank}(G_R, \cdot) = 1$.

Definition: The so-called **Cartan subalgebra** is that algebra which contains the **maximum** number of mutually commuting generators. Consequently, the rank of a semi-simple Lie group is equal to the number of generators of the Cartan subalgebra.

Lecture 5

3.5 Casimir operators, multiplets, Schur's lemma

Definition: A **Casimir operator** \hat{C} is an operator which is not identical to the unit operator and which commutes with **all** group elements,

$$[\hat{C}, g] = 0 \quad \forall g \in G. \quad (3.48)$$

Example: Consider the rotation group (G_R, \cdot) . The **square** of the angular momentum operator, $\hat{\vec{L}}^2 = \sum_{j=1}^3 \hat{L}_j^2$ commutes with all generators \hat{L}^i ,

$$[\hat{\vec{L}}^2, \hat{L}^i] = 0, \quad i = 1, 2, 3, \quad (3.49)$$

and therefore also with all group elements $\hat{U}_R(\vec{\phi}) = \exp\left(-\frac{i}{\hbar} \vec{\phi} \cdot \hat{\vec{L}}\right)$ (in their representation as linear operators),

$$[\hat{\vec{L}}^2, \hat{U}_R(\vec{\phi})] = 0. \quad (3.50)$$

Thus, $\hat{\vec{L}}^2$ is a Casimir operator of the group (G_R, \cdot) (in fact, it is the only one, see below). Casimir operators are usually not generators, but functions of the generators,

$$\hat{C} = \hat{C}(\hat{X}_1, \hat{X}_2, \dots, \hat{X}_N). \quad (3.51)$$

In general, they are **not uniquely** determined. However, for **unitary** groups they can always be chosen to be **hermitean** operators.

Proof: For any given Casimir operator \hat{C} also \hat{C}^\dagger is a Casimir operator, since

$$[\hat{X}_i, \hat{C}^\dagger] = [\hat{X}_i^\dagger, \hat{C}^\dagger] = [\hat{C}, \hat{X}_i]^\dagger = 0,$$

where we have used the hermiticity of the generators for unitary groups. Let us then define the new Casimir operator

$$\hat{\hat{C}} \equiv \hat{C} + \hat{C}^\dagger, \quad (3.52)$$

which is hermitean by definition, $\hat{\hat{C}} \equiv \hat{\hat{C}}^\dagger$, q.e.d.

Of great importance is the following theorem (which we quote without proof), since it relates the rank of a semi-simple Lie group to the maximum number of Casimir operators.

Racah's theorem: Every semi-simple Lie group (G, \circ) of rank r has r Casimir operators $\hat{C}_i(\hat{X}_1, \dots, \hat{X}_N)$, $i = 1, \dots, r$.

Example: $\text{rank}(G_R, \cdot) = \text{rank}(SU(2), \cdot) = \text{rank}(SO(3), \cdot) = 1$, thus there is **exactly one** Casimir operator,

$$\hat{C}_1 = \hat{C}_1(\hat{L}^x, \hat{L}^y, \hat{L}^z) \equiv \sum_{j=1}^3 \hat{L}_j^2 = \hat{\vec{L}}^2.$$

The physical meaning of the Casimir operators is due to the fact that their **eigenvectors** span the so-called **multiplets** of the group, e.g. for (G_R, \cdot)

$$\hat{\vec{L}}^2 |\ell\rangle = \hbar^2 \ell(\ell + 1) |\ell\rangle. \quad (3.53)$$

In this case, the multiplets are the $(2\ell + 1)$ -fold degenerate states for a given angular-momentum quantum number ℓ . One denotes the multiplet for

$$\begin{aligned} \ell = 0 & \quad \text{as singlet,} \\ \ell = \frac{1}{2} & \quad \text{as doublet,} \\ \ell = 1 & \quad \text{as triplet,} \\ \ell = \frac{3}{2} & \quad \text{as quadruplet, etc.} \end{aligned}$$

We also allowed half-integer values for the angular momentum. These cannot occur for orbital angular momenta (they are always integer-valued), but for spin (which can also assume half-integer values).

3 Introduction to Group Theory

The meaning of a multiplet is clear for the rotation group (G_R, \cdot) . But what is a multiplet in general, i.e., for an arbitrary group? To clarify this, we need the following

Definition: An subspace U of a Hilbert space \mathcal{H} which remains **invariant** under transformations of a symmetry group (G, \circ) is called an **invariant subspace**. It consists of the set of all states which remains **invariant** under the action of linear operators $D(g)$, $g \in G$, i.e.,

$$\forall |\psi\rangle \in U, \quad \forall g \in G : \quad D(g) |\psi\rangle = |D(g) \psi\rangle \equiv |\psi'\rangle \in U .$$

In other words, the linear operators $D(g)$ of a symmetry group transform the states of an invariant subspace only among themselves, but do not lead to states outside of the invariant subspace. In this case one speaks of a **reducible subspace** and of a **reducible representation** $D(g)$ of group elements $g \in G$.

Example: Consider eigenstates of angular momentum ℓ and of its z component m , $|\ell m\rangle$. An invariant subspace U with respect to the rotation group (G_R, \cdot) is spanned e.g. by

$$U = \{|00\rangle, |11\rangle, |10\rangle, |1-1\rangle\} , \quad (3.54)$$

because acting with the operators (2.32) on these states does not lead outside of this subspace. It is sufficient to show this for infinitesimal transformations,

$$\begin{aligned} \hat{U}_R(\delta\vec{\phi}) &= \mathbb{1} - \frac{i}{\hbar} \delta\vec{\phi} \cdot \hat{\vec{L}} = \mathbb{1} - \frac{i}{\hbar} \left(\delta\phi^x \hat{L}^x + \delta\phi^y \hat{L}^y + \delta\phi^z \hat{L}^z \right) \\ &= \mathbb{1} - \frac{i}{\hbar} \left(\delta\phi^z \hat{L}^z + \frac{\delta\phi^x - i\delta\phi^y}{2} \hat{L}_+ + \frac{\delta\phi^x + i\delta\phi^y}{2} \hat{L}_- \right) , \end{aligned}$$

where we used the **ladder operators**

$$\hat{L}_\pm \equiv \hat{L}^x \pm i \hat{L}^y .$$

The operator \hat{L}^z simply measures the value of m , while \hat{L}_\pm raises or lowers the value of m by one unit, respectively. Since

$$\hat{L}_\pm |\ell \pm \ell\rangle = 0 ,$$

the ladder operators do not lead outside of the invariant subspace U .

Definition: An invariant subspace $M \subset \mathcal{H}$ is **irreducible**, if it does not contain any further invariant subspaces (except for M itself). In this case the linear operators $D(g)$, which act on states of M are in an **irreducible representation** of the group elements $g \in G$.

Definition: A **multiplet** is an irreducible invariant subspace.

Example: The invariant subspace (3.54) of the preceding example is **reducible**, because it contains **two** irreducible invariant subspaces, the **singlet**

$$M_0 = \{|00\rangle\} ,$$

and the **triplet**

$$M_1 = \{|11\rangle, |10\rangle, |1-1\rangle\} .$$

It is easy to see that acting with linear operators $\hat{U}_R(\vec{\phi})$ onto M_0, M_1 does not lead to states outside of these multiplets.

Definition: A representation $D(g)$ is **completely reducible**, if it can be written in block-diagonal form,

$$D(g) = \begin{pmatrix} D_0(g) & 0 & \cdots \\ 0 & D_1(g) & \\ \vdots & & \ddots \end{pmatrix} , \quad (3.55)$$

where all $D_i(g)$ are **irreducible**, i.e., they act only onto the i th multiplet M_i .

Example:

$$\hat{L}^2 = \begin{pmatrix} 0 & 0 & \cdots & & \\ 0 & \hbar^2 \frac{3}{4} \mathbb{1}_2 & 0 & \cdots & \\ \vdots & 0 & \hbar^2 2 \mathbb{1}_3 & 0 & \cdots \\ & \vdots & 0 & \hbar^2 \frac{15}{4} \mathbb{1}_4 & \\ & & \vdots & & \ddots \end{pmatrix}$$

is a completely reducible representation of \hat{L}^2 . The individual block matrices are the irreducible representations of \hat{L}^2 for the respective multiplets.

In Chapter 2 we had seen that a system is invariant under space and time translations, as well as under rotations, of the **generators** of the respective groups $(G_r, \cdot), (G_t, \cdot)$, and (G_R, \cdot) **commute** with the **Hamilton operator** of the system. Let us repeat this argument for arbitrary Lie groups and then draw consequences for the Casimir operators of the group.

Let (G, \circ) be a Lie group and $(D(G), \cdot) = (\{D(g), g \in G\}, \cdot)$ a representation of this group, with

$$D(g) = \hat{U}(\vec{\alpha}) = \exp\left(-\frac{i}{\hbar} \alpha_j \hat{X}_j\right) .$$

The Hilbert-space states of the system obey the time-dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle . \quad (3.56)$$

The state which has been transformed under (G, \circ) is

$$|\psi'(t)\rangle = D(g) |\psi(t)\rangle = \hat{U}(\vec{\alpha}) |\psi(t)\rangle . \quad (3.57)$$

We apply the transformation $\hat{U}(\vec{\alpha})$ to the Schrödinger equation (3.56),

$$\begin{aligned} \hat{U} i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle &= i\hbar \frac{\partial}{\partial t} \hat{U} |\psi(t)\rangle \equiv i\hbar \frac{\partial}{\partial t} |\psi'(t)\rangle \\ &= \hat{U} \hat{H} |\psi(t)\rangle = \hat{U} \hat{H} \hat{U}^{-1} \hat{U} |\psi(t)\rangle \equiv \hat{U} \hat{H} \hat{U}^{-1} |\psi'(t)\rangle , \end{aligned}$$

where we have used the definition (3.57) of the transformed Hilbert-space state several times. If the system is **invariant** under the Lie group (G, \circ) , i.e., (G, \circ) is a **symmetry group** of the system, the transformed state also obeys the Schrödinger equation, with **the very same** Hamilton operator, $\hat{H}' = \hat{H}$, i.e.,

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} |\psi'(t)\rangle &= \hat{U} \hat{H} \hat{U}^{-1} |\psi'(t)\rangle \equiv \hat{H} |\psi'(t)\rangle \\ \iff \hat{U} \hat{H} \hat{U}^{-1} &= \hat{H} \iff \hat{U} \hat{H} = \hat{H} \hat{U} \\ \iff [\hat{U}, \hat{H}] &= 0, \end{aligned}$$

or, for **infinitesimal** transformations,

$$[\hat{X}_j, \hat{H}] = 0 \quad \forall j = 1, \dots, N. \quad (3.58)$$

Since the Casimir operators are **functions** of the generators, cf. Eq. (3.51), it is obvious that also these operators commute with the Hamilton operator,

$$[\hat{C}_i, \hat{H}] = 0 \quad \forall i = 1, \dots, r, \quad (3.59)$$

where r denotes the **total number** of Casimir operators (according to Racah's theorem, for semi-simple groups r is identical to the rank of the group). According to their definition, all Casimir operators also mutually commute with each other,

$$[\hat{C}_i, \hat{C}_j] = 0 \quad \forall i, j = 1, \dots, r. \quad (3.60)$$

For now, the maximum set of commuting operators (which have the properties of Casimir operators, i.e., they commute with all generators of the group) is then

$$\{\hat{C}_1, \dots, \hat{C}_r, \hat{H}\}.$$

Since \hat{H} fulfills all requirements for a Casimir operator of the system, but **cannot be a new** Casimir operator, i.e., one which is **linearly independent** from the other \hat{C}_i (since r is already the maximum number of possible Casimir operators), \hat{H} must be a **function** of the Casimir operators of the system,

$$\hat{H} = \hat{H}(\hat{C}_1, \dots, \hat{C}_r).$$

If this function can be expanded in terms of a Taylor series, \hat{H} can be simply expressed (to leading order in the Casimir operators) as a **linear combination** of the Casimir operators of the system, plus the unit matrix $\mathbb{1}$, which always commutes with all $D(g) \in D(G)$,

$$\hat{H} = c_0 \mathbb{1} + \sum_{i=1}^r c_i \hat{C}_i + O(\hat{C}_i^2). \quad (3.61)$$

Example: Hydrogen atom. Because of the central potential the system is invariant under **rotations**, i.e., under the group (G_R, \cdot) . As we know, the only Casimir operator of this group is \hat{L}^2 . Indeed, in spherical coordinates the Hamilton operator reads

$$\hat{H}(\vec{r}) = \left[-\frac{\hbar^2}{2m} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + V(r) \right] \mathbb{1} + \frac{1}{2mr^2} \hat{L}^2,$$

cf. lectures on quantum mechanics. We thus identify

$$c_0 \equiv -\frac{\hbar^2}{2m} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + V(r),$$

$$c_1 \equiv \frac{1}{2mr^2}.$$

The coefficients c_0, c_1 are still (operator-valued) functions of radial distance r , which, however, is invariant under G_R .

For now, the maximum set of **linearly independent**, mutually commuting operators is then the set of the Casimir operators of the system,

$$\{\hat{C}_1, \dots, \hat{C}_r\}.$$

Therefore, we can choose a system of states which are simultaneously **eigenfunctions of all Casimir operators**. These states are characterized by quoting the respective **eigenvalues**,

$$|C_1, \dots, C_r\rangle.$$

This is, however, not yet a **pure** state, i.e., a state where **all possibly measurable** quantum numbers are fixed. Namely, there can be further operators \hat{A}_j , $j = 1, \dots, s$, which are **not** Casimir operators, but which **still** commute with the Hamilton operator and with all Casimir operators of the system.

Example: In the hydrogen atom the generator \hat{L}^z is such an operator. It commutes with \hat{H} as well as with \hat{L}^2 . (We have excluded spin from this consideration. In principle, \hat{S}^2 is another Casimir operator, which, however, does not occur in \hat{H} , and \hat{S}^z is a further operator which commutes with \hat{H} .)

Pure states can therefore be characterized by the eigenvalues C_i of the Casimir operators \hat{C}_i and the eigenvalues α_j of the operators \hat{A}_j ,

$$|\psi\rangle \equiv |C_1, \dots, C_r, \alpha_1, \dots, \alpha_s\rangle. \quad (3.62)$$

Theorem: The states of a multiplet are **degenerate** with respect to **all** Casimir operators.

Proof: Let us consider a pure state (3.62), which belongs to a certain multiplet. In the following, we will use the abbreviation $\alpha \equiv \{\alpha_1, \dots, \alpha_s\}$ for the set of eigenvalues of the operators \hat{A}_j . According to the definition of a multiplet, also

$$|\psi'\rangle = \hat{U} |\psi\rangle = \hat{U} |C_1, \dots, C_r, \alpha\rangle \equiv |C'_1, \dots, C'_r, \alpha'\rangle$$

is a state of the very **same** multiplet. It remains to be shown that

$$C'_i \equiv C_i \quad \forall i = 1, \dots, r.$$

To this end we compute using $[\hat{C}_i, \hat{U}] = 0$ that

$$\begin{aligned} \hat{C}_i |C'_1, \dots, C'_r, \alpha'\rangle &\equiv C'_i |C'_1, \dots, C'_r, \alpha'\rangle \\ &= \hat{C}_i \hat{U} |C_1, \dots, C_r, \alpha\rangle = \hat{U} \hat{C}_i |C_1, \dots, C_r, \alpha\rangle \\ &= \hat{U} C_i |C_1, \dots, C_r, \alpha\rangle = C_i \hat{U} |C_1, \dots, C_r, \alpha\rangle \\ &= C_i |C'_1, \dots, C'_r, \alpha'\rangle, \end{aligned}$$

3 Introduction to Group Theory

i.e., indeed $C'_i = C_i$, q.e.d.

This theorem means that the various multiplets can be **uniquely** characterized by the eigenvalues C_i of the Casimir operators.

Example: The various multiplets of different angular momenta are characterized by specifying ℓ ; $\hbar^2\ell(\ell + 1)$ is the eigenvalue of \hat{L}^2 on the multiplet characterized by ℓ .

Corollary: Since $\hat{H} = \hat{H}(\hat{C}_1, \dots, \hat{C}_r)$, all states of a multiplet are **energetically degenerate**.

At the end of this chapter we prove another important theorem, which consists of two parts:

Schur's lemma (I): Let (G, \circ) be a semi-simple unitary Lie group and \hat{H} the Hamilton operator of a given system. If (G, \circ) is a symmetry group of the system, i.e., if the dynamics is invariant under transformations $D(g) \in D(G)$, i.e.,

$$[D(g), \hat{H}] = 0 \quad \forall g \in G ,$$

then transitions between different multiplets are forbidden.

Proof: Because of Eq. (3.59) we have for states

$$|C_1, \dots, C_r, \alpha\rangle \in M , \quad |C'_1, \dots, C'_r, \alpha'\rangle \in M' ,$$

which belong to **different** multiplets $M, M', M \neq M'$,

$$\begin{aligned} 0 &= \langle C'_1, \dots, C'_r, \alpha' | [\hat{C}_i, \hat{H}] |C_1, \dots, C_r, \alpha\rangle \\ &= \langle C'_1, \dots, C'_r, \alpha' | \hat{C}_i \hat{H} - \hat{H} \hat{C}_i |C_1, \dots, C_r, \alpha\rangle \\ &= \langle C'_1, \dots, C'_r, \alpha' | \hat{C}_i^\dagger \hat{H} - \hat{H} \hat{C}_i |C_1, \dots, C_r, \alpha\rangle \\ &= (C'_i - C_i) \langle C'_1, \dots, C'_r, \alpha' | \hat{H} |C_1, \dots, C_r, \alpha\rangle \\ &= (C'_i - C_i) E(C_1, \dots, C_r) \langle C'_1, \dots, C'_r, \alpha' |C_1, \dots, C_r, \alpha\rangle , \end{aligned}$$

where we used the hermiticity of the Casimir operators in the step from the second to the third line. This equation can only be fulfilled if either (i) $C'_i = C_i$ or (ii)

$$\langle C'_1, \dots, C'_r, \alpha' |C_1, \dots, C_r, \alpha\rangle = 0$$

(without loss of generality we can assume that the energy $E(C_1, \dots, C_r)$ of the states does not vanish). According to the assumption the two states belong to **different** multiplets and these are **uniquely** characterized by the eigenvalues of the Casimir operators. Thus, **at least for one** Casimir operator, e.g. \hat{C}_j , we have must have $C'_j \neq C_j$. Then, the equation can only be satisfied if case (ii) applies, i.e., the **transition matrix element** between the two states must vanish, q.e.d.

Remark: If we consider the next-to-last line of the preceding proof, we can also formulate this fact in another way. Obviously, the Hamilton operator \hat{H} **does not possess** a piece which induces a **transition** between states of different multiplets. If there were such

a piece, then $\langle C'_1, \dots, C'_r, \alpha' | C_1, \dots, C_r, \alpha \rangle \neq 0$. This can only happen, if \hat{H} contains parts which **explicitly break the symmetry**.

Schur's lemma (II): The irreducible representations of the Casimir operators are proportional to the unit matrix.

Proof: Let us consider a multiplet M of d states,

$$M \equiv \{|C_1, \dots, C_r, j\rangle, j = 1, \dots, d\},$$

where we numbered the states of the multiplet (which differ by the values of the quantum numbers $\alpha \equiv \{\alpha_1, \dots, \alpha_s\}$) by the index j . Obviously, $\dim M = d$. An irreducible representation of an operator $D(g)$ has the $(d \times d)$ matrix representation

$$[D(g)]_{jk} = \langle C_1, \dots, C_r, j | D(g) | C_1, \dots, C_r, k \rangle.$$

The irreducible representation of the Casimir operator \hat{C}_i is therefore

$$\begin{aligned} [\hat{C}_i]_{jk} &= \langle C_1, \dots, C_r, j | \hat{C}_i | C_1, \dots, C_r, k \rangle \\ &= C_i \langle C_1, \dots, C_r, j | C_1, \dots, C_r, k \rangle \equiv C_i \delta_{jk}, \end{aligned}$$

where we have used the orthonormality of the states of the multiplet. Thus we have

$$[\hat{C}_i] = C_i \mathbb{1}_{d \times d}, \quad \text{q.e.d. .}$$

4 The group $(SU(3), \cdot)$

The group $(SU(3), \cdot)$, i.e., the group of special unitary transformations in three dimensions, is of paramount importance for the theory of the strong interaction, one of the fundamental forces of Nature. Therefore, we discuss this group separately and in more detail in this chapter.

4.1 Generators

The group $(SU(N), \cdot)$ of special unitary transformations in N dimensions possesses $N^2 - 1$ generators.

Example: The group $(SU(2), \cdot)$ has $N^2 - 1 = 3$ generators, the three components \hat{S}^a of the **spin operator**. These read in the **fundamental representation**

$$\hat{S}^a = \frac{\hbar}{2} \hat{\sigma}_a, \quad a = 1, 2, 3, \quad (4.1)$$

with the **Pauli matrices** (3.18). Obviously, the generators in fundamental representation are **(2×2) matrices**, i.e., they act on objects (Hilbert-space states) which are **two-dimensional**.

We now construct the $N^2 - 1 = 8$ generators of $(SU(3), \cdot)$ in the **fundamental representation**, i.e., in the representation where they are **(3×3) matrices**, and therefore act on **three-dimensional** state vectors. To this end we generalize the Pauli matrices to three dimensions. In analogy to $(SU(2), \cdot)$, cf. Eq. (4.1), we write for the generators of $(SU(3), \cdot)$

$$\hat{T}_a \equiv \frac{\hbar}{2} \hat{\lambda}_a, \quad a = 1, \dots, 8. \quad (4.2)$$

Note that \hat{T}_a as well as $\hat{\lambda}_a$ have to be **hermitean**,

$$\hat{T}_a = \hat{T}_a^\dagger, \quad \hat{\lambda}_a = \hat{\lambda}_a^\dagger, \quad (4.3)$$

since $(SU(3), \cdot)$ is a **unitary** Lie group. The (3×3) matrices $\hat{\lambda}_a$ play the same role for $(SU(3), \cdot)$ as the Pauli matrices $\hat{\sigma}_a$ for $(SU(2), \cdot)$. The first three $\hat{\lambda}_a$ are trivial extensions of the Pauli matrices to three dimensions,

$$\hat{\lambda}_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \hat{\lambda}_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \hat{\lambda}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (4.4)$$

The hermiticity of $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3$ is guaranteed because of the hermiticity of the Pauli matrices,

For the next four $\hat{\lambda}_a$ we take the structure of the first two Pauli matrices as an example and displace the non-trivial elements by one row and one column, respectively,

$$\hat{\lambda}_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \hat{\lambda}_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \hat{\lambda}_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \hat{\lambda}_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}. \quad (4.5)$$

Also these matrices are obviously hermitean.

For the last generator we need a matrix with a non-vanishing (33) element, because otherwise we cannot construct all possible $SU(3)$ matrices with the generators, in particular not those which have non-vanishing (33) elements. We make the Ansatz

$$\hat{\lambda}_8 = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \beta \end{pmatrix}, \quad (4.6)$$

with **real-valued** constants α, β , since $\hat{\lambda}_8$ must be hermitean. Elements of the group $(SU(N), \cdot)$ have the property that their determinant assumes the value +1, therefore

$$0 = \ln \det \hat{U} = \text{Tr} \ln \hat{U} = -\frac{i}{\hbar} \alpha_a \text{Tr} \hat{T}_a, \quad (4.7)$$

i.e., the generators of $(SU(N), \cdot)$ are **trace-free**. Since this must also hold for $\hat{T}_8 = \frac{\hbar}{2} \hat{\lambda}_8$, we must have $\beta = -2\alpha$ in our Ansatz (4.6). The constant α can be determined from the **orthogonality** (3.41) of the generators,

$$\text{Tr} \left(\hat{T}_a \hat{T}_b \right) = \frac{\hbar^2}{2} \delta_{ab}, \quad (4.8)$$

or, for $\hat{\lambda}_a$,

$$\text{Tr} \left(\hat{\lambda}_a \hat{\lambda}_b \right) = 2 \delta_{ab}. \quad (4.9)$$

One easily checks that our Ansatz (4.6) fulfills this relation for $a = 8$ and $b = 1, \dots, 7$. We determine the constant α from this equation for the choice $a = b = 8$,

$$\text{Tr} \left(\hat{\lambda}_8 \hat{\lambda}_8 \right) = \text{Tr} \begin{pmatrix} \alpha^2 & 0 & 0 \\ 0 & \alpha^2 & 0 \\ 0 & 0 & 4\alpha^2 \end{pmatrix} = 6\alpha^2 \equiv 2 \quad \implies \quad \alpha = \frac{1}{\sqrt{3}}.$$

The final result for $\hat{\lambda}_8$ thus reads

$$\hat{\lambda}_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \quad (4.10)$$

The matrices defined through Eqs. (4.4), (4.5), and (4.10) are called **Gell-Mann matrices** (after their inventor Murray Gell-Mann). Due to their orthogonality they are linearly independent and thus span a basis in the space of trace-free hermitean (3×3) matrices.

4.2 Structure constants

The Lie algebra of $(SU(3), \cdot)$ is

$$[\hat{T}_a, \hat{T}_b] = i\hbar f_{abc} \hat{T}_c, \quad (4.11)$$

or in terms of the Gell-Mann matrices, respectively,

$$[\hat{\lambda}_a, \hat{\lambda}_b] = 2i f_{abc} \hat{\lambda}_c. \quad (4.12)$$

One determines the structure constants from Eq. (3.42) using the explicit form of the Gell-Mann matrices,

$$f_{abc} = \frac{2}{i\hbar^3} \text{Tr} \left\{ [\hat{T}_a, \hat{T}_b] \hat{T}_c \right\} \equiv \frac{1}{4i} \text{Tr} \left\{ [\hat{\lambda}_a, \hat{\lambda}_b] \hat{\lambda}_c \right\}. \quad (4.13)$$

This relation insures that the structure constants are then **completely antisymmetric**. They read explicitly:

abc	123	147	156	246	257	345	367	458	678
f_{abc}	1	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{2}$

For all other combinations of indices abc the structure constants vanish.

The generators of $(SU(3), \cdot)$ (or, more generally, of $(SU(N), \cdot)$) also obey **anti-commutation relations**,

$$\{\hat{T}_a, \hat{T}_b\} = \frac{\hbar^2}{3} \delta_{ab} \mathbb{1}_3 + \hbar d_{abc} \hat{T}_c, \quad (4.14)$$

or in terms of the Gell-Mann matrices, respectively,

$$\{\hat{\lambda}_a, \hat{\lambda}_b\} = \frac{4}{3} \delta_{ab} \mathbb{1}_3 + 2 d_{abc} \hat{\lambda}_c. \quad (4.15)$$

The **completely symmetric** structure constants d_{abc} can be determined from the relation

$$d_{abc} = \frac{2}{\hbar^3} \text{Tr} \left(\{\hat{T}_a, \hat{T}_b\} \hat{T}_c \right) \equiv \frac{1}{4} \text{Tr} \left(\{\hat{\lambda}_a, \hat{\lambda}_b\} \hat{\lambda}_c \right). \quad (4.16)$$

(We leave the proof of this equation as homework exercise.) The non-vanishing symmetric structure constants read explicitly:

abc	118	146	157	228	247	256	338	344	355	366	377
d_{abc}	$\frac{1}{\sqrt{3}}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$

abc	448	558	668	778	888
d_{abc}	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$

4.3 Subalgebras and subgroups

The algebra of $(SU(3), \cdot)$ has several **subalgebras**, which generate **subgroups** of $(SU(3), \cdot)$:

(i) The set

$$\{\hat{T}_1, \hat{T}_2, \hat{T}_3\},$$

constitutes an $(SU(2), \cdot)$ **subalgebra**, since $f_{123} = 1 \equiv \epsilon_{123}$,

$$[\hat{T}_a, \hat{T}_b] = i\hbar f_{abc} \hat{T}_c \equiv i\hbar \epsilon_{abc} \hat{T}_c, \quad a, b, c = 1, 2, 3.$$

This subalgebra generates an $(SU(2), \cdot)$ **subgroup** of $(SU(3), \cdot)$,

$$SU(2) \subset SU(3).$$

(ii) A second subalgebra consists of the set

$$\{\hat{T}_2, \hat{T}_5, \hat{T}_7\}.$$

If we consider the explicit form of these generators,

$$\begin{aligned} \hat{T}_2 &= \frac{\hbar}{2} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \frac{1}{2} (-i\hbar) \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \implies (\hat{T}_2)_{ij} = \frac{1}{2} (-i\hbar \epsilon_{3ij}), \\ \hat{T}_5 &= \frac{\hbar}{2} \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} = -\frac{1}{2} (-i\hbar) \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \implies (\hat{T}_5)_{ij} = -\frac{1}{2} (-i\hbar \epsilon_{2ij}), \\ \hat{T}_7 &= \frac{\hbar}{2} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} = \frac{1}{2} (-i\hbar) \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \implies (\hat{T}_7)_{ij} = \frac{1}{2} (-i\hbar \epsilon_{1ij}), \end{aligned}$$

it becomes obvious that, up to a factor of $1/2$ for \hat{T}_2 and \hat{T}_7 and a factor of $-1/2$ for \hat{T}_5 , respectively, we face the **generators of $(SO(3), \cdot)$ in the adjoint representation!** If we define

$$\hat{J}_3 \equiv 2\hat{T}_2, \quad \hat{J}_2 \equiv -2\hat{T}_5, \quad \hat{J}_1 \equiv 2\hat{T}_7, \quad (4.17)$$

we obtain from the commutation relation

$$[\hat{T}_a, \hat{T}_b] = i\hbar f_{abc} \hat{T}_c, \quad a, b, c = 2, 5, 7,$$

with the value of the structure constant $f_{257} = 1/2$ the following commutation relation for the generators \hat{J}_i of $(SO(3), \cdot)$ defined through Eq. (4.17),

$$[\hat{J}_i, \hat{J}_j] = i\hbar \epsilon_{ijk} \hat{J}_k, \quad i, j, k = 1, 2, 3.$$

We thus have identified an $(SO(3), \cdot)$ **subalgebra** of $(SU(3), \cdot)$,

$$SO(3) \subset SU(3).$$

4 The group $(SU(3), \cdot)$

(iii) The **Cartan subalgebra** consists of

$$\{\hat{T}_3, \hat{T}_8\},$$

since $[\hat{T}_3, \hat{T}_8] = 0$. It follows that $\text{rank}(SU(3), \cdot) = 2$. The two **Casimir operators** are

$$\hat{C}_1 = f_{abc} \hat{T}_a \hat{T}_b \hat{T}_c, \quad \hat{C}_2 = d_{abc} \hat{T}_a \hat{T}_b \hat{T}_c. \quad (4.18)$$

We leave the proof as homework exercise.

4.4 Multiplets

Multiplets are determined by the eigenvalues C_1, C_2 of the two Casimir operators (4.18) of $(SU(3), \cdot)$. The question is, how many states belong to a multiplet of given (C_1, C_2) and how to distinguish them from each other. To this end we need to quantify further quantum numbers, which belong to operators which commute with the Casimir operators (but not with all generators) of $(SU(3), \cdot)$.

To answer this for $(SU(3), \cdot)$, let us step back one step and consider the group $(SU(2), \cdot)$. The sole Casimir operator is $\hat{C}_1 = \hat{J}^2$ and the multiplets are characterized by the eigenvalues of \hat{J}^2 , i.e., $\hbar^2 j(j+1)$, or simply the value of j . Furthermore, \hat{J}^2 commutes with \hat{J}_z , which has eigenvalues $\hbar m$. Thus, the individual states of a multiplet can be uniquely characterized by the value of j (labelling the multiplet) and the value of m (labelling the individual state within the multiplet), $|j m\rangle$. For given j there are $2j+1$ states which differ by the value of m , $m = -j, -j+1, \dots, j-1, j$.

Obviously, \hat{J}_z is the generator of the **Cartan subalgebra** of $(SU(2), \cdot)$. In the case of $(SU(3), \cdot)$ it therefore stands to reason that we also use the generators of the Cartan subalgebra, \hat{T}_3 and \hat{T}_8 , to distinguish between the states of a given multiplet. Since \hat{T}_3, \hat{T}_8 commute with each other and per definition also with the two Casimir operators \hat{C}_1, \hat{C}_2 , we have a set of four mutually commuting operators, the eigenvalues of which uniquely determine the states of a given multiplet,

$$|C_1, C_2, T_3, T_8\rangle. \quad (4.19)$$

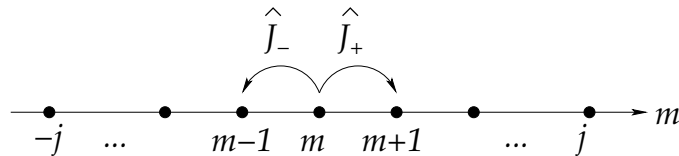
For $(SU(2), \cdot)$ we had defined **ladder operators**

$$\hat{J}_\pm \equiv \hat{J}_x \pm i \hat{J}_y,$$

which allow to jump between adjacent states of a multiplet,

$$\hat{J}_\pm |j m\rangle \sim |j m \pm 1\rangle.$$

This can be graphically depicted as follows:



For $(SU(3), \cdot)$ this is quite analogous, but here the states of a multiplet differ by **two** quantum numbers, T_3 and T_8 , and not only by one, m . One can therefore raise or lower quantum numbers not only in one direction, but in two, as it is schematically (but, as we shall see, not quite correctly) depicted in Fig. 4.1:

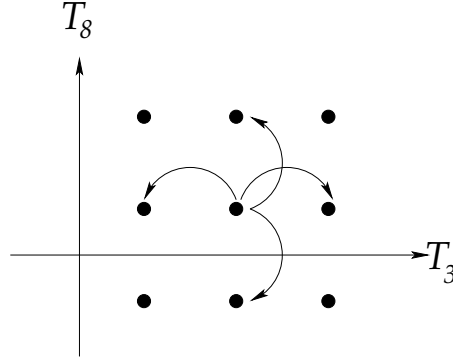


Figure 4.1: Schematic action of ladder operators on $(SU(3), \cdot)$ multiplets.

In order to see how this works precisely, let us first define suitable **ladder operators**:

$$\begin{aligned}\hat{T}_{\pm} &\equiv \hat{T}_1 \pm i\hat{T}_2, \\ \hat{V}_{\pm} &\equiv \hat{T}_4 \pm i\hat{T}_5, \\ \hat{U}_{\pm} &\equiv \hat{T}_6 \pm i\hat{T}_7.\end{aligned}\tag{4.20}$$

In addition, we define the so-called **hypercharge operator**

$$\hat{Y} \equiv \frac{2}{\sqrt{3}}\hat{T}_8 = \frac{\hbar}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.\tag{4.21}$$

The commutation relations for the ladder operators read:

$$[\hat{T}_3, \hat{T}_{\pm}] = \pm\hbar\hat{T}_{\pm},\tag{4.22}$$

$$[\hat{T}_+, \hat{T}_-] = 2\hbar\hat{T}_3.\tag{4.23}$$

These relations are the same as in the angular momentum algebra for \hat{L}^z and the ladder operators \hat{L}_{\pm} . This, in turn, has the consequence that the operators $\{\hat{T}_{\pm}, \hat{T}_3\}$ define an $(SU(2), \cdot)$ subalgebra (which we had already mentioned above). Furthermore we have:

$$[\hat{T}_3, \hat{V}_{\pm}] = \pm\frac{\hbar}{2}\hat{V}_{\pm},\tag{4.24}$$

$$[\hat{T}_3, \hat{U}_{\pm}] = \mp\frac{\hbar}{2}\hat{U}_{\pm},\tag{4.25}$$

$$[\hat{V}_+, \hat{V}_-] = 2\hbar\left(\frac{1}{2}\hat{T}_3 + \frac{3}{4}\hat{Y}\right) \equiv 2\hbar\hat{V}_3,\tag{4.26}$$

$$[\hat{U}_+, \hat{U}_-] = 2\hbar\left(-\frac{1}{2}\hat{T}_3 + \frac{3}{4}\hat{Y}\right) \equiv 2\hbar\hat{U}_3,\tag{4.27}$$

4 The group $(SU(3), \cdot)$

where the right-hand sides of the last two equations represent the definitions of the operators \hat{V}_3 and \hat{U}_3 . We furthermore compute:

$$[\hat{Y}, \hat{T}_\pm] = 0, \quad (4.28)$$

$$[\hat{Y}, \hat{V}_\pm] = \pm \hbar \hat{V}_\pm, \quad (4.29)$$

$$[\hat{Y}, \hat{U}_\pm] = \pm \hbar \hat{U}_\pm, \quad (4.30)$$

$$[\hat{T}_\pm, \hat{V}_\pm] = 0, \quad (4.31)$$

$$[\hat{T}_\pm, \hat{U}_\mp] = 0, \quad (4.32)$$

$$[\hat{U}_\pm, \hat{V}_\pm] = 0, \quad (4.33)$$

$$[\hat{T}_\pm, \hat{V}_\mp] = \mp \hbar \hat{U}_\mp, \quad (4.34)$$

$$[\hat{T}_\pm, \hat{U}_\pm] = \pm \hbar \hat{V}_\pm, \quad (4.35)$$

$$[\hat{U}_\pm, \hat{V}_\mp] = \pm \hbar \hat{T}_\mp, \quad (4.36)$$

$$[\hat{T}_3, \hat{Y}] = 0, \quad (4.37)$$

The proof of Eqs. (4.22) – (4.37) is left as homework exercise.

With the help of the relations (4.24), (4.25), (4.29), and (4.30) we also derive

$$[\hat{V}_3, \hat{V}_\pm] = \frac{1}{2} [\hat{T}_3, \hat{V}_\pm] + \frac{3}{4} [\hat{Y}, \hat{V}_\pm] = \pm \frac{\hbar}{4} \hat{V}_\pm \pm \frac{3}{4} \hbar \hat{V}_\pm \equiv \pm \hbar \hat{V}_\pm, \quad (4.38)$$

$$[\hat{U}_3, \hat{U}_\pm] = -\frac{1}{2} [\hat{T}_3, \hat{U}_\pm] + \frac{3}{4} [\hat{Y}, \hat{U}_\pm] = \pm \frac{\hbar}{4} \hat{U}_\pm \pm \frac{3}{4} \hbar \hat{U}_\pm \equiv \pm \hbar \hat{U}_\pm. \quad (4.39)$$

Together with Eqs. (4.26) and (4.27) these equations can be interpreted in the way that the operators $\{\hat{V}_\pm, \hat{V}_3\}$ and $\{\hat{U}_\pm, \hat{U}_3\}$ form **two additional** $(SU(2), \cdot)$ subalgebras. Since \hat{V}_3 and \hat{U}_3 both depend on \hat{T}_3 , these are, however, **not independent** from the $(SU(2), \cdot)$ subalgebra spanned by $\{\hat{T}_\pm, \hat{T}_3\}$. As we have already mentioned, there is only a **single** $(SU(2), \cdot)$ subalgebra contained in $(SU(3), \cdot)$.

We now need to clarify how the ladder operators \hat{T}_\pm, \hat{V}_\pm , and \hat{U}_\pm affect the states of a multiplet. To this end we consider the eigenstates (4.19), or by replacing T_8 by $Y \equiv \sqrt{3} T_8/2$, the eigenstates

$$|C_1, C_2, T_3, Y\rangle \equiv |T_3 Y\rangle,$$

where we have abbreviated the list of arguments, since C_1, C_2 cannot be changed by the ladder operators for any given multiplet of $(SU(3), \cdot)$. By definition we have

$$\begin{aligned} \hat{T}_3 |T_3 Y\rangle &= T_3 |T_3 Y\rangle, \\ \hat{Y} |T_3 Y\rangle &= Y |T_3 Y\rangle. \end{aligned}$$

We now conclude that

(i) because of Eq. (4.24),

$$\begin{aligned} [\hat{T}_3, \hat{V}_\pm] |T_3 Y\rangle &= \hat{T}_3 \hat{V}_\pm |T_3 Y\rangle - \hat{V}_\pm T_3 |T_3 Y\rangle = \pm \frac{\hbar}{2} \hat{V}_\pm |T_3 Y\rangle \\ \iff \hat{T}_3 \hat{V}_\pm |T_3 Y\rangle &= \left(T_3 \pm \frac{\hbar}{2} \right) \hat{V}_\pm |T_3 Y\rangle. \end{aligned} \quad (4.40)$$

This equation means that \hat{V}_\pm raises/lowers the eigenvalue T_3 of a state $|T_3 Y\rangle$ by $\hbar/2$.

(ii) because of Eq. (4.25)

$$\begin{aligned} [\hat{T}_3, \hat{U}_\pm] |T_3 Y\rangle &= \hat{T}_3 \hat{U}_\pm |T_3 Y\rangle - \hat{U}_\pm T_3 |T_3 Y\rangle = \mp \frac{\hbar}{2} \hat{U}_\pm |T_3 Y\rangle \\ \iff \hat{T}_3 \hat{U}_\pm |T_3 Y\rangle &= \left(T_3 \mp \frac{\hbar}{2}\right) \hat{U}_\pm |T_3 Y\rangle. \end{aligned} \quad (4.41)$$

This equation means that \hat{U}_\pm lowers/raises the eigenvalue T_3 of a state $|T_3 Y\rangle$ by $\hbar/2$.

(iii) because of Eq. (4.29)

$$\begin{aligned} [\hat{Y}, \hat{V}_\pm] |T_3 Y\rangle &= \hat{Y} \hat{V}_\pm |T_3 Y\rangle - \hat{V}_\pm Y |T_3 Y\rangle = \pm \hbar \hat{V}_\pm |T_3 Y\rangle \\ \iff \hat{Y} \hat{V}_\pm |T_3 Y\rangle &= (Y \pm \hbar) \hat{V}_\pm |T_3 Y\rangle. \end{aligned} \quad (4.42)$$

This means that \hat{V}_\pm raises/lowers the eigenvalue Y of a state $|T_3 Y\rangle$ by \hbar .

(iv) because of Eq. (4.30)

$$\begin{aligned} [\hat{Y}, \hat{U}_\pm] |T_3 Y\rangle &= \hat{Y} \hat{U}_\pm |T_3 Y\rangle - \hat{U}_\pm Y |T_3 Y\rangle = \pm \hbar \hat{U}_\pm |T_3 Y\rangle \\ \iff \hat{Y} \hat{U}_\pm |T_3 Y\rangle &= (Y \pm \hbar) \hat{U}_\pm |T_3 Y\rangle. \end{aligned} \quad (4.43)$$

This means that \hat{U}_\pm raises/lowers the eigenvalue Y of a state $|T_3 Y\rangle$ by \hbar .

(v) \hat{T}_\pm raises/lowers the eigenvalue T_3 of a state $|T_3 Y\rangle$ by \hbar .

(vi) because of Eq. (4.28) \hat{T}_\pm does not change the eigenvalue Y of a state $|T_3 Y\rangle$.

To summarize,

(i) \hat{T}_\pm raises/lowers T_3 by \hbar and leaves Y unchanged.

(ii) \hat{V}_\pm raises/lowers T_3 by $\hbar/2$ and Y by \hbar .

(iii) \hat{U}_\pm lowers/raises T_3 by $\hbar/2$ and raises/lowers Y by \hbar .

This can be graphically depicted as shown in Fig. 4.2. The red line is the so-called **T–line**. It defines the direction along which the ladder operators \hat{T}_\pm act. These change the value of T_3 by $\pm\hbar$ and leave the value of Y unchanged. The blue line is the so-called **V–line**. It defines the direction along which the ladder operators \hat{V}_\pm act. They change T_3 by $\pm\hbar/2$ and simultaneously Y by $\pm\hbar$. Finally, the green line is the so-called **U–line**. It defines the direction along which the ladder operators \hat{U}_\pm act. These change the value of T_3 by $\mp\hbar/2$ and simultaneously Y by $\pm\hbar$.

The way the ladder operators act has the consequence that the states of an $(SU(3), \cdot)$ multiplets do not form a regular lattice in the $(T_3 - Y)$ plane, as shown in Fig. 4.1, but one where states on a line of fixed Y are shifted in T_3 direction by $\hbar/2$ with respect to states on a line with $Y \pm \hbar$, as shown in Fig. 4.2.

This information suffices to determine the **shape of $(SU(3), \cdot)$ multiplets** in the $(T_3 - Y)$ plane. To this end let us first remember that

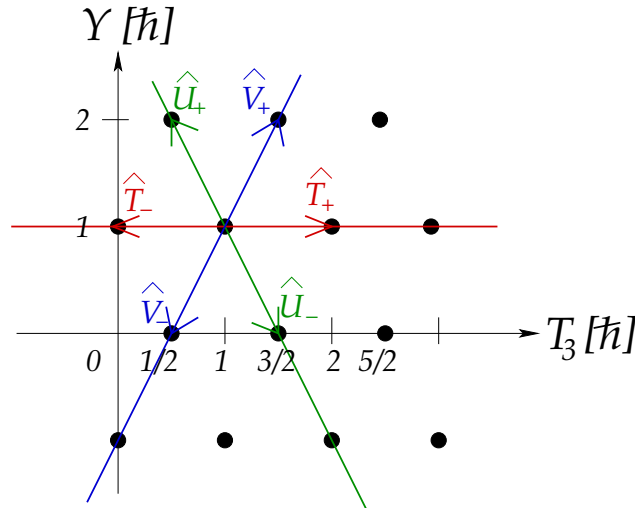


Figure 4.2: Action of the ladder operators \hat{T}_\pm , \hat{V}_\pm , and \hat{U}_\pm on states of an $(SU(3), \cdot)$ multiplet.

- (i) $\{\hat{T}_+, \hat{T}_-, \hat{T}_3\}$ forms an $(SU(2), \cdot)$ subalgebra of $(SU(3), \cdot)$, which transforms states of $(SU(2), \cdot)$ multiplets (as parts of $(SU(3), \cdot)$ multiplets) among themselves. Graphically, these $(SU(2), \cdot)$ multiplets are located along the (red) T -lines in Fig. 4.2 and the ladder operators \hat{T}_\pm lead from one state on a T -line to the next. While the value of Y always remains the same for these $(SU(2), \cdot)$ multiplets, the value of T_3 runs between $-T_3^{\max}$ and $+T_3^{\max}$. This, in turn, implies that these $(SU(2), \cdot)$ multiplets must be located **mirror-symmetrically** with respect to the Y axis (i.e., the line $T_3 = 0$).
- (ii) $\{\hat{V}_+, \hat{V}_-, \hat{V}_3\}$ also constitutes an $(SU(2), \cdot)$ subalgebra of $(SU(3), \cdot)$, which transforms the states of $(SU(2), \cdot)$ multiplets (as parts of $(SU(3), \cdot)$ multiplets) among themselves. However, these $(SU(2), \cdot)$ multiplets now lie along the (blue) V -lines in Fig. 4.2. When jumping between states with the ladder operators \hat{V}_\pm , the value of V_3 varies between $-V_3^{\max}$ and $+V_3^{\max}$. This, in turn, implies that these $(SU(2), \cdot)$ multiplets must have **as many states left as right** of the line $V_3 = 0$ or, if we use the definition (4.26) of \hat{V}_3 , of the straight line $Y = -\frac{2}{3}T_3$.
- (iii) $\{\hat{U}_+, \hat{U}_-, \hat{U}_3\}$ also constitutes an $(SU(2), \cdot)$ subalgebra of $(SU(3), \cdot)$, which transforms the states of $(SU(2), \cdot)$ multiplets (as parts of $(SU(3), \cdot)$ multiplets) among themselves. However, these $(SU(2), \cdot)$ multiplets now lie along the (green) U -lines in Fig. 4.2. When jumping between states with the ladder operators \hat{U}_\pm , the value of U_3 varies between $-U_3^{\max}$ and $+U_3^{\max}$. This, in turn, implies that these $(SU(2), \cdot)$ multiplets must have **as many states left as right** of the line $U_3 = 0$ or, if we use the definition (4.27) of \hat{U}_3 , of the straight line $Y = \frac{2}{3}T_3$.

The symmetry axes thus identified are shown in Fig. 4.3. The three-fold symmetry only allows **three possible geometrical shapes** for $(SU(3), \cdot)$ multiplets in the $(T_3 - Y)$ plane:

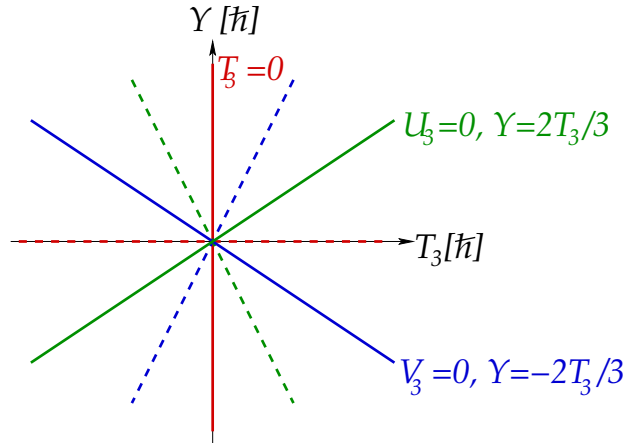


Figure 4.3: Symmetry axes of $(SU(3), \cdot)$ multiplets in the $(T_3 - Y)$ plane (full lines), as well as T -, V -, and U -lines (dashed lines).

- (a) a single state, the so-called **singlet** located at the **origin** of the $(T_3 - Y)$ plane, i.e., for $T_3 = Y = 0$,
- (b) **triangles** centered at the origin of the $(T_3 - Y)$ plane,
- (c) **hexagons** centered at the origin of the $(T_3 - Y)$ plane.

The simplest multiplets which fulfill criteria (a) and (b) are shown in Figs. 4.4 (a–g).

Remarks:

- (i) The number of states of a given multiplet is determined by the following operation: beginning from the upper right (or lower left) corner move with the ladder operators \hat{T}_- (\hat{T}_+) or \hat{V}_- (\hat{V}_+), respectively, to the adjacent state. From there use \hat{T}_- (\hat{T}_+), \hat{V}_- (\hat{V}_+), or \hat{U}_+ (\hat{U}_-) to move to the next possible state inside a given triangle. One cannot leave the triangle, but in this way one can also reach states inside the triangle. Finally, this prescription defines a lattice of states on the sides and inside the triangle. Counting the states, one obtains the total number of states in a given multiplet.
- (ii) Counting along T -lines, the **triplet** $[3]$ contains a T -singlet and a T -doublet or, counting along V -lines, a V -singlet and a V -doublet or, counting along U -lines, a U -singlet and a U -doublet, respectively. An analogous argument holds for the multiplet which is **conjugate** to the triplet, the so-called **anti-triplet** $[\bar{3}]$.
- (iii) The **sextet** $[6]$ contains a T -singlet, a T -doublet, and a T -triplet, or a V -singlet, a V -doublet, and a V -triplet, or a U -singlet, a U -doublet, and a U -triplet, respectively. An analogous argument holds for the **anti-sextet** $[\bar{6}]$.

4 The group $(SU(3), \cdot)$

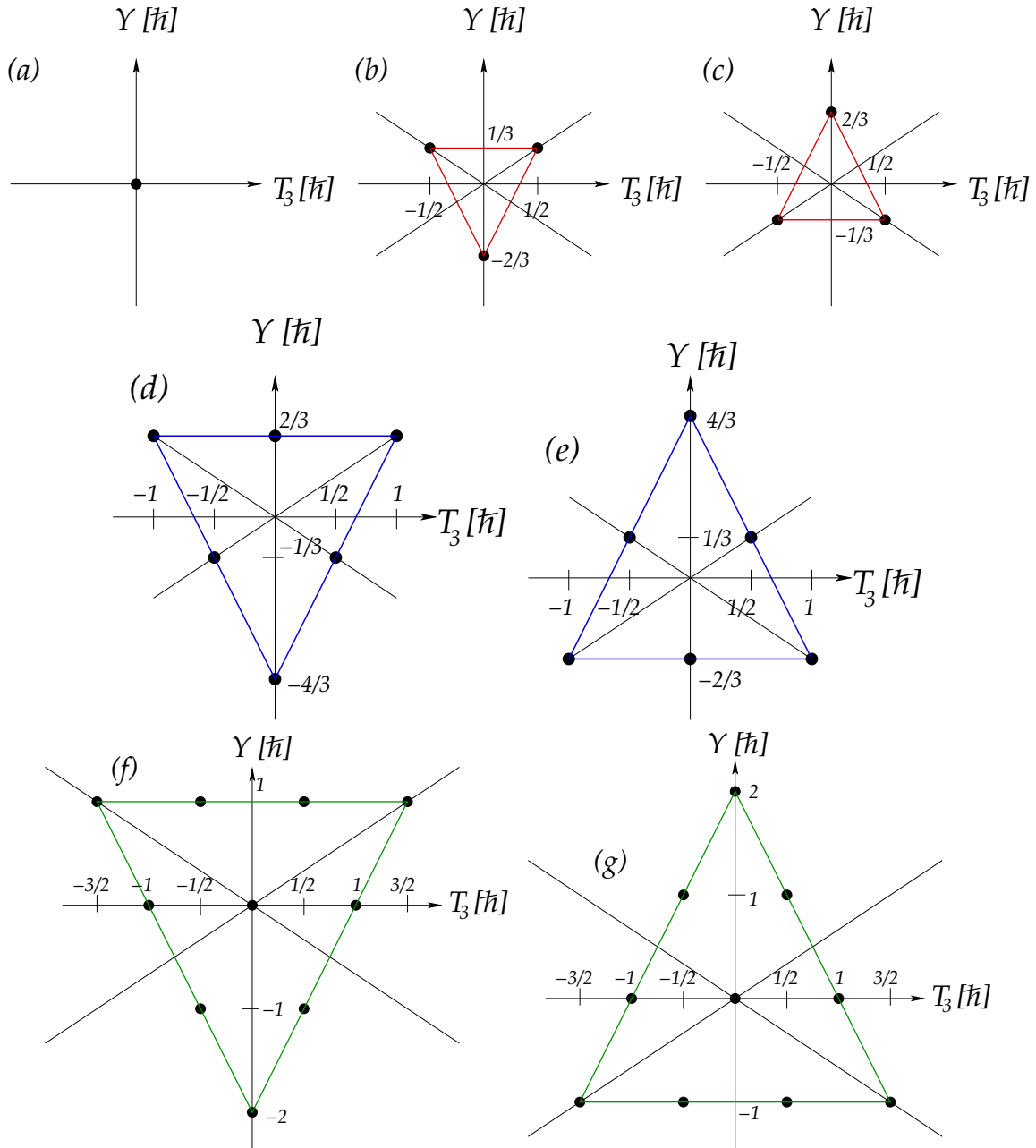


Figure 4.4: (a) Singlet $[1]$, (b) triplet $[3]$, (c) anti-triplet $[\bar{3}]$, (d) sextet $[6]$, (e) anti-sextet $[\bar{6}]$, (f) decuplet $[10]$, and (g) anti-decuplet $[\bar{10}]$.

- (iv) The **decuplet** $[10]$ contains a T -singlet, a T -doublet, a T -triplet, and a T -quadruplet, or a V -singlet, a V -doublet, a V -triplet, and a V -quadruplet, or a U -singlet, a U -doublet, a U -triplet, and a U -quadruplet, respectively. An analogous argument holds for the **anti-decuplet** $[\bar{10}]$.
- (v) The triplet $[3]$ with three possible states corresponds to the **fundamental representation** of $(SU(3), \cdot)$. We will see in the next section that we can generate the

singlet, the anti-triplet, as well as **all higher-dimensional** multiplets by coupling a suitable number of triplets together.

The simplest multiplet obeying criterium (c) is the **octet [8]** shown in Fig. 4.5. This is the **adjoint representation** of $(SU(3), \cdot)$.

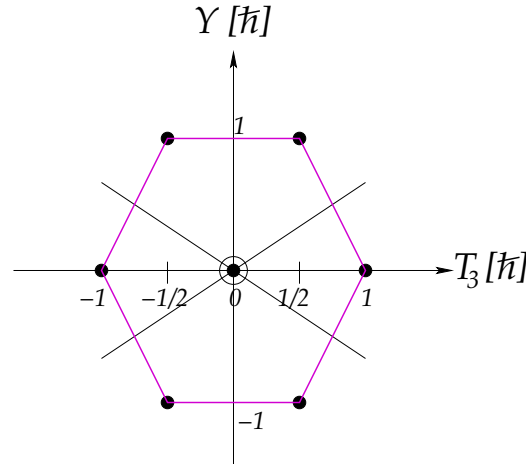


Figure 4.5: Octet [8].

Remarks:

- (i) The **anti-octet** $[\bar{8}]$ has the **same** shape as the octet. One says that $[\bar{8}]$ and $[8]$ **self-conjugate**.
- (ii) From Fig. 4.5 one observes that the octet contains two T -doublets and a T -triplet, or two V -doublets and a V -triplet, or two U -doublets and a U -triplet, respectively.
- (iii) This yields only **seven** states. So why does one speak of an “octet”, i.e., a multiplet with **eight** states? The reason is that the origin is **doubly occupied** with two **different** states, namely with one state of the triplet and an **additional** singlet. This is indicated by an additional circle around the origin in Fig. 4.5.

Lecture 7

What is the reason for this **double occupancy**? Starting from the state $|T_3 = 1, Y = 0\rangle$, one can reach the origin in three different ways:

$$\begin{aligned}
 |00\rangle_{\text{I}} &\sim \hat{T}_- |10\rangle, \\
 |00\rangle_{\text{II}} &\sim \hat{V}_- \hat{U}_+ |10\rangle \sim \hat{V}_- \left| \frac{1}{2}, 1 \right\rangle, \\
 |00\rangle_{\text{III}} &\sim \hat{U}_+ \hat{V}_- |10\rangle \sim \hat{U}_+ \left| \frac{1}{2}, -1 \right\rangle.
 \end{aligned}$$

These different ways are shown graphically in Fig. 4.6.

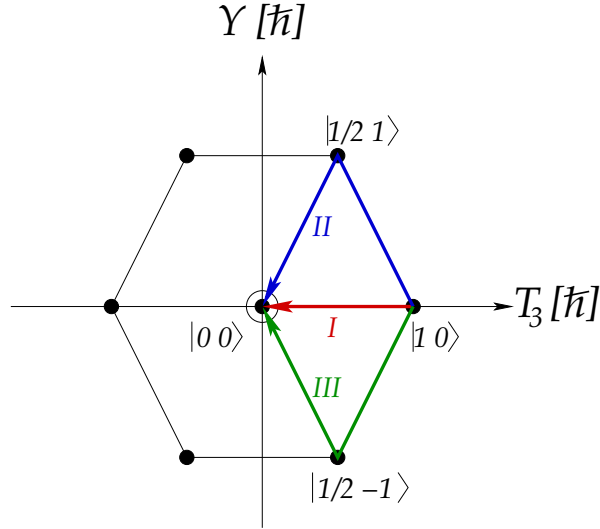


Figure 4.6: The three different ways to reach the origin $|00\rangle$ starting from $|10\rangle$.

However, because of Eq. (4.36) $\hat{U}_+ \hat{V}_- = \hat{V}_- \hat{U}_+ + \hbar \hat{T}_-$, way III is not linearly independent of way I and way II,

$$|00\rangle_{\text{III}} \sim \hat{U}_+ \hat{V}_- |10\rangle = \hat{V}_- \hat{U}_+ |10\rangle + \hbar \hat{T}_- |10\rangle = \alpha |00\rangle_{\text{II}} + \beta |00\rangle_{\text{I}}.$$

Thus, the octet has **two linearly independent states**, $|00\rangle_{\text{I}}$ and $|00\rangle_{\text{II}}$, which are located at the origin.

The rule of multiple occupancy or **degeneracy** of states with the same quantum numbers T_3 and Y can be generalized to arbitrary multiplets. In general, the degeneracy of states on each **inner** shell is by one unit larger than those of the states on the adjacent outer shell. The degeneracy of the outermost shell is always equal to one. An example is shown in Fig. 4.7.

This rule holds as long as a shell is not a triangle. Inside a triangle the degeneracy does not increase any more, but stays constant. The reason is that there is now only **one** linearly independent way to reach a point inside a triangle. To understand this, let us consider the decuplet, Fig. 4.8.

At first glance, one would think that, starting from state $|10\rangle$, there are three different ways to reach the origin $|00\rangle$, just like for the octet. For the same reasons as for the octet, at least two of them should be **linearly independent**, e.g.

$$|00\rangle_{\text{I}} \sim \hat{T}_- |10\rangle \quad \text{and} \quad |00\rangle_{\text{II}} \sim \hat{V}_- \hat{U}_+ |10\rangle,$$

so that the origin is again **doubly** occupied. However, it holds

$$|10\rangle \sim \hat{V}_- \left| \frac{3}{2} 1 \right\rangle$$

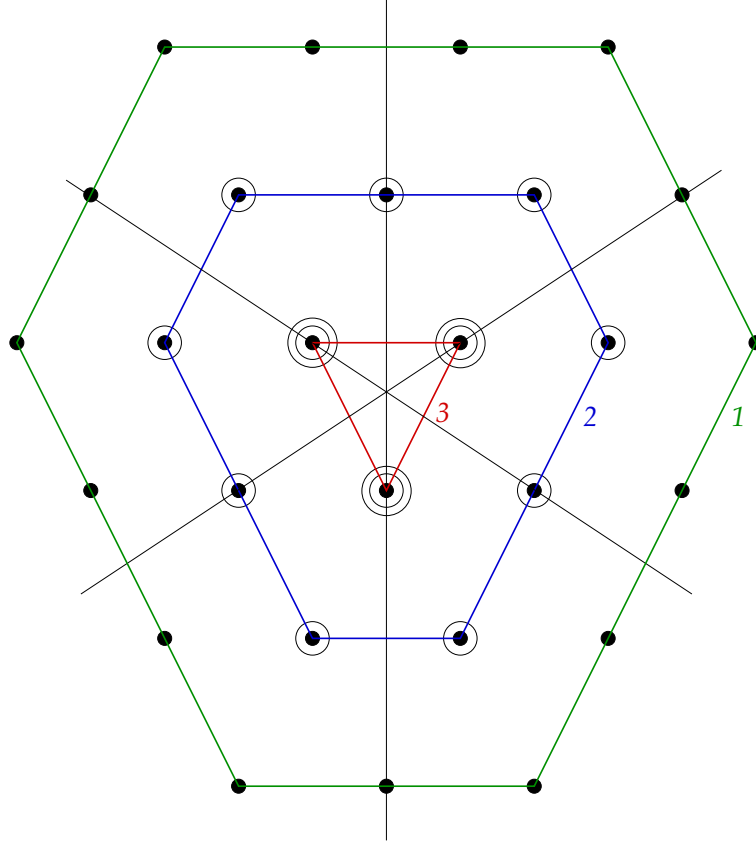


Figure 4.7: Starting from the outermost shell, the degeneracy of states increases by one as one moves from one inner shell to the next.

and therefore

$$\begin{aligned}
 |00\rangle_{\text{II}} &\sim \hat{V}_- \hat{U}_+ \hat{V}_- \left| \frac{3}{2} 1 \right\rangle = \hat{V}_- \left(\hat{V}_- \hat{U}_+ + \hbar \hat{T}_- \right) \left| \frac{3}{2} 1 \right\rangle \\
 &= (\hat{V}_-)^2 \hat{U}_+ \left| \frac{3}{2} 1 \right\rangle + \hbar \hat{V}_- \hat{T}_- \left| \frac{3}{2} 1 \right\rangle \equiv \hbar \hat{V}_- \hat{T}_- \left| \frac{3}{2} 1 \right\rangle \\
 &= \hbar \hat{T}_- \hat{V}_- \left| \frac{3}{2} 1 \right\rangle = \hat{T}_- (\hbar \hat{V}_- \left| \frac{3}{2} 1 \right\rangle) \\
 &\sim \hat{T}_- |10\rangle \sim |00\rangle_{\text{I}}.
 \end{aligned} \tag{4.44}$$

Here we have made use of Eqs. (4.31) and (4.36), and employed the fact that the state $|\frac{3}{2} 1\rangle$ is the state with the maximum hypercharge $Y = 1$ in the multiplet, such that $\hat{U}_+ |\frac{3}{2} 1\rangle \equiv 0$, since this operation would lead outside the multiplet, which is not allowed. Equation (4.44) thus shows that $|00\rangle_{\text{II}} \sim |00\rangle_{\text{I}}$, i.e., these two states are **not linearly independent** from each other, thus **not different**. Another way to see this is, starting from $|\frac{3}{2} 1\rangle$, to apply either $\hat{T}_- \hat{V}_-$ or $\hat{V}_- \hat{T}_-$ in order to reach $|00\rangle$. Because of Eq. (4.31) this is, however, the same, thus not linearly independent.

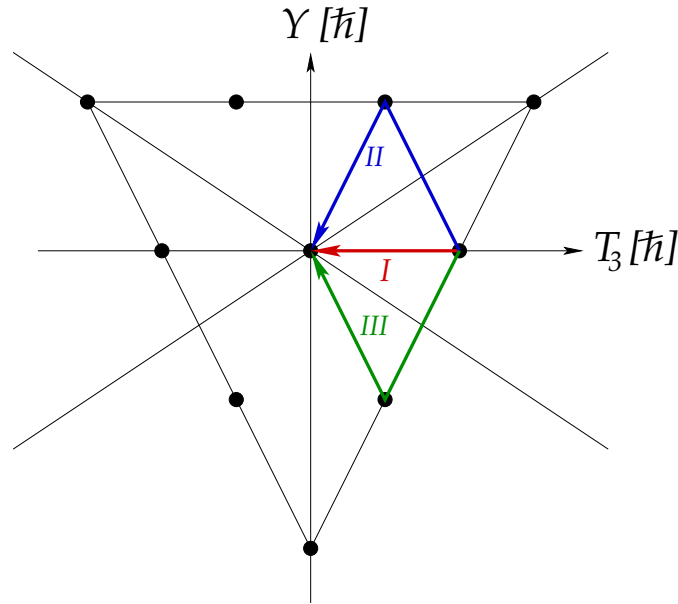


Figure 4.8: The three ways to reach the origin $|00\rangle$, starting from state $|10\rangle$.

4.5 Construction of multiplets from the fundamental representation

In the theory of coupling of angular momenta, two angular momenta (or spins, respectively) are coupled or “added” to form states with a certain total angular momentum. From the mathematical point of view, one takes the product space of the two angular momenta and decomposes it in terms of **multiplets** for given **total angular momentum** (or **spin**, respectively). This process is called **reduction**.

It stands to reason that, by coupling together sufficiently many **smallest non-trivial** angular momenta (or spins, respectively), it should be possible to generate **all higher-dimensional multiplets**. This is indeed the case: by coupling together sufficiently many **fundamental representations** of $(SU(2), \cdot)$, i.e., **doublets**, all higher multiplets for given angular momentum (or spin, respectively) can be generated. We show this at hand of two examples:

$$\begin{aligned}
 \text{(i)} \quad j_1 = \frac{1}{2} \quad \text{with} \quad j_2 = \frac{1}{2} &\implies j = 0 \quad \text{and} \quad j = 1 & (4.45) \\
 \uparrow \quad \otimes \quad \uparrow &\implies \uparrow\downarrow \quad \text{and} \quad \uparrow & .
 \end{aligned}$$

This relation can be simply written with the notation “doublet = [2]”, “singlet = [1]”, and “triplet = [3]” as

$$[2] \otimes [2] = [1] \oplus [3] . \quad (4.46)$$

4.5 Construction of multiplets from the fundamental representation

This process can be continued, by coupling the doublet to a triplet:

$$(ii) \quad j_1 = \frac{1}{2} \quad \text{with} \quad j_2 = 1 \quad \implies \quad j = \frac{1}{2} \quad \text{and} \quad j = \frac{3}{2} \quad (4.47)$$

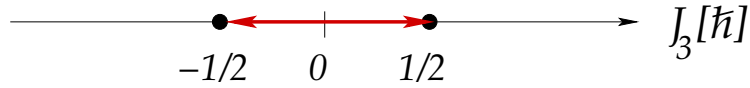
$$\quad \quad \quad \uparrow \quad \otimes \quad \uparrow \quad \implies \quad \uparrow \quad \downarrow \quad \text{and} \quad \uparrow \quad .$$

With the notation “quadruplet = [4]” this can be succinctly written as

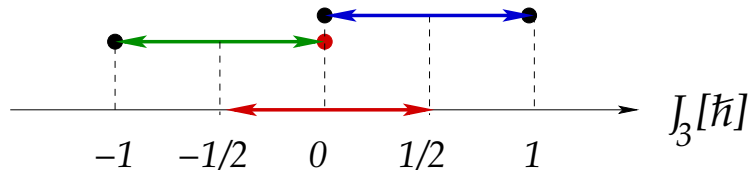
$$[2] \otimes [3] = [2] \oplus [4] . \quad (4.48)$$

By coupling a doublet or triplet to the quadruplet, one can continue this process and in this way generate all higher-dimensional multiplets.

There is a **graphical method** to perform this reduction. To this end, we draw the first doublet onto the J_3 axis:

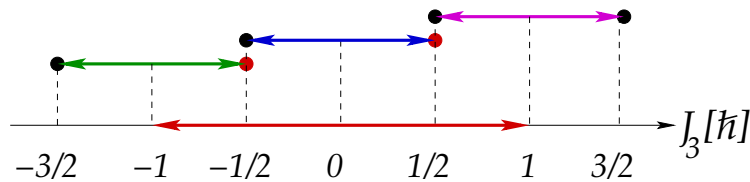


Now we put the second doublet with its **center** ($J_3 = 0$) onto each of the two states $J_3 = 1/2$ and $J_3 = -1/2$ (in units of \hbar) of the first doublet:



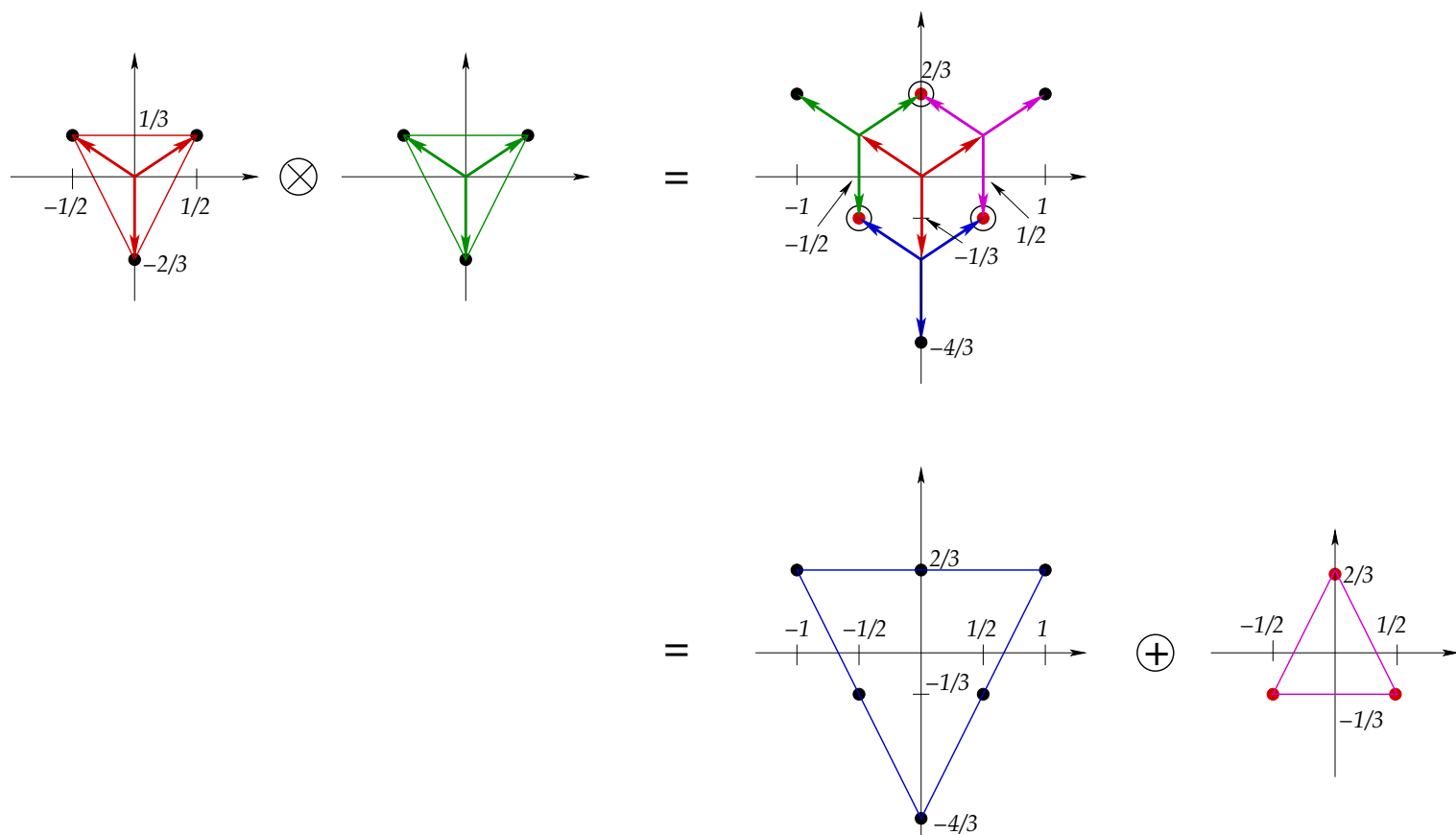
This generates **once** the states $J_3 = 1$ and $J_3 = -1$, as well as **twice** the state $J_3 = 0$. This just corresponds to **one** singlet, $J = 0$ or [1] (red dot), and **one** triplet, $J = 1$ or [3] (black dots).

If one repeats this with another doublet and the triplet, one obtains a doublet, $J = 1/2$ or [2] (red dots), and a quadruplet, $J = 3/2$ or [4] (black dots):



Analogously, one can generate alle ($SU(3), \cdot$) multiplets by coupling, starting from the fundamental representation, the triplet. As a first example, we consider the coupling of two triplets:

4 The group $(SU(3), \cdot)$

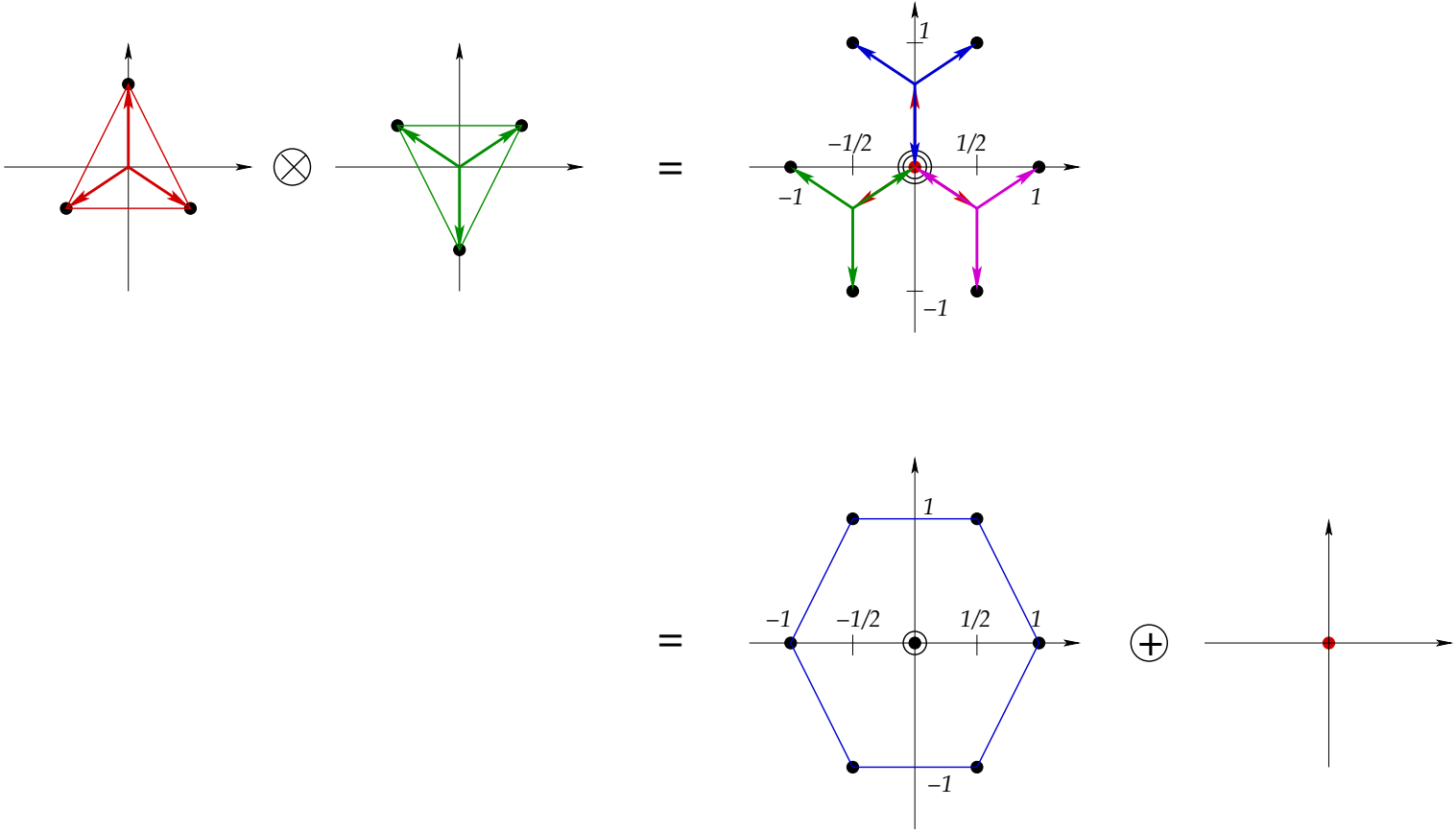


This yields a sextet and an anti-triplet,

$$[3] \otimes [3] = [\bar{3}] \oplus [6]. \quad (4.49)$$

4.5 Construction of multiplets from the fundamental representation

Now we couple the anti-triplet to another triplet:

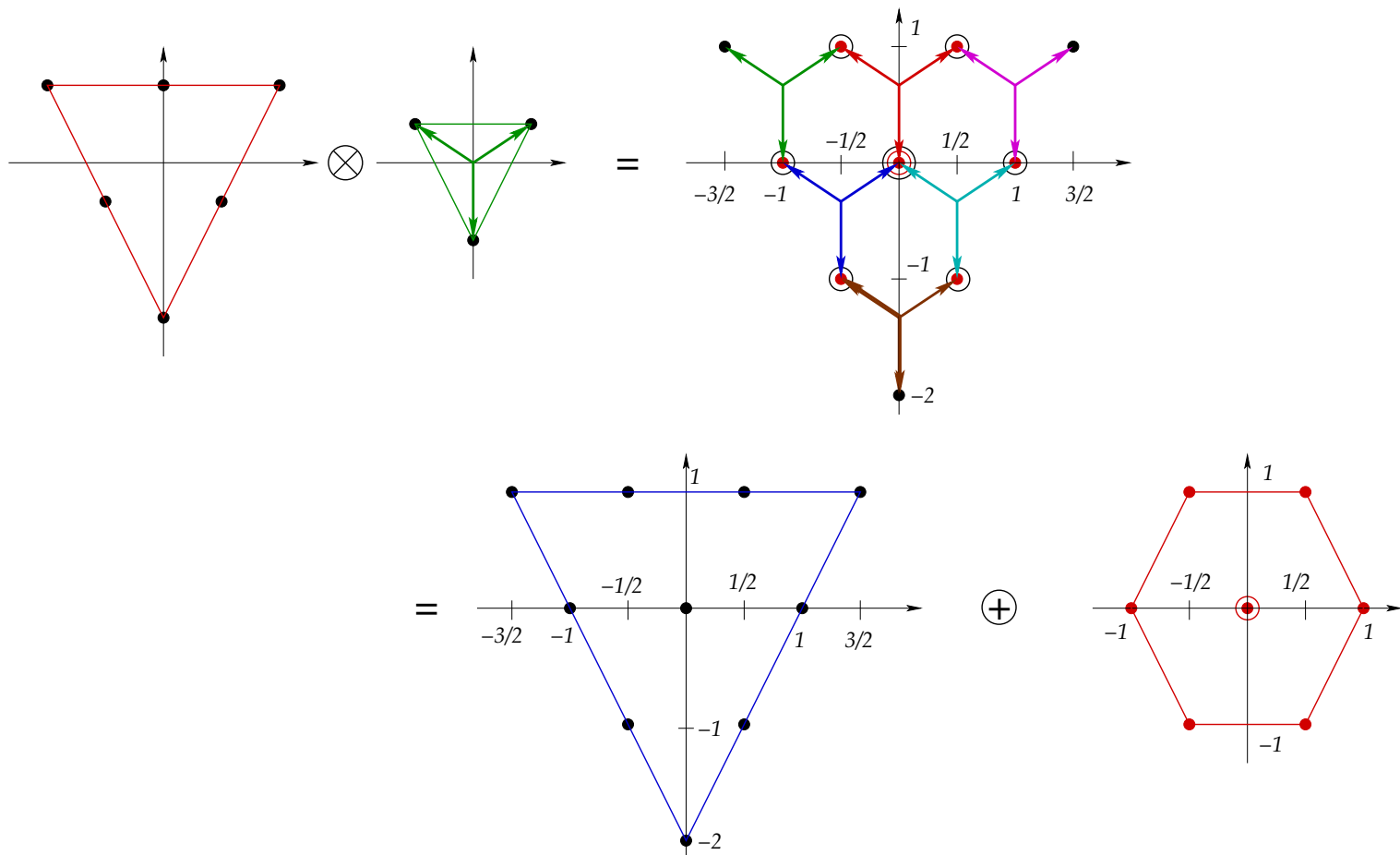


This yields an octet and a singlet,

$$[\bar{3}] \otimes [3] = [1] \oplus [8]. \quad (4.50)$$

Starting from the sextet, we obtain by coupling it to another triplet:

4 The group $(SU(3), \cdot)$



This yields a decuplet and an octet,

$$[3] \otimes [6] = [8] \oplus [10] . \quad (4.51)$$

We can use rules (4.49), (4.50), and (4.51) together with the law of associativity to determine more complex couplings, e.g.

$$[3] \otimes [3] \otimes [3] = [3] \otimes ([\bar{3}] \oplus [6]) = ([3] \otimes [\bar{3}]) \oplus ([3] \otimes [6]) = [1] \oplus [8] \oplus [8] \oplus [10] . \quad (4.52)$$

5 Unitary symmetries of the strong interaction

5.1 ($U(1), \cdot$) symmetry of quantum electrodynamics

Symmetries play a prominent role in modern theories of the Forces of Nature. This is especially true for the theory of the strong interaction. Before turning to this theory, however, let us first consider the (simpler) theory of electromagnetism, **quantum electrodynamics (QED)**, which describes the interactions between **electrons** and **photons**. The respective **Lagrangian** reads (in natural units, $\hbar = c = 1$)

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + \bar{\psi}(i\not{D} - m)\psi, \quad (5.1)$$

where $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ is the **field-strength tensor** of the electromagnetic field, with the 4-vector potential A^μ , the so-called **gauge field**, and the **covariant derivative**

$$D_\mu = \partial_\mu - ieA_\mu. \quad (5.2)$$

The quantity ψ is the 4-spinor of the electron and m its mass. The first term in the Lagrangian (5.1) is the so-called **gauge-field term**, which describes the dynamics of the gauge field A^μ , in this case the photon field. The second term is the so-called **matter term**, which describes the dynamics of the matter fields, in this case the electron.

The Lagrangian (5.1) is **invariant** under so-called ($U(1), \cdot$) **gauge transformations**,

$$\psi \longrightarrow \psi' = \hat{U} \psi, \quad (5.3)$$

$$A^\mu \longrightarrow A'^\mu = \hat{U} A^\mu \hat{U}^{-1} - \frac{i}{e} (\partial^\mu \hat{U}) \hat{U}^{-1}, \quad (5.4)$$

where

$$\hat{U} \equiv e^{ie\Lambda(X)} \in U(1) \quad (5.5)$$

is a **space-time dependent** phase factor, and at the same time a **representation of an element of the group ($U(1), \cdot$)**. For elements of ($U(1), \cdot$) Eq. (5.4) simplifies as follows:

$$A^\mu \longrightarrow A'^\mu = A^\mu + \partial^\mu \Lambda.$$

The invariance of the Lagrangian (5.1) under the transformations (5.3), (5.4) is seen as follows:

- (i) The field-strength tensor is invariant under the transformation (5.4),

$$F^{\mu\nu} \longrightarrow F'^{\mu\nu} = \partial^\mu (A^\nu + \partial^\nu \Lambda) - \partial^\nu (A^\mu + \partial^\mu \Lambda) = F^{\mu\nu},$$

as long as $\Lambda(X)$ is twice continuously differentiable, $\partial^\mu \partial^\nu \Lambda = \partial^\nu \partial^\mu \Lambda$. Consequently, also the gauge-field term in Eq. (5.1) is invariant.

(ii) The covariant derivative of the matter field transforms as the matter field itself,

$$\begin{aligned} D_\mu \psi &\longrightarrow D'_\mu \psi' = (\partial_\mu - ie A'_\mu) \psi' = [\partial_\mu - ie A_\mu - ie (\partial_\mu \Lambda)] e^{ie\Lambda} \psi \\ &= e^{ie\Lambda} [\partial_\mu + ie (\partial_\mu \Lambda) - ie A_\mu - ie (\partial_\mu \Lambda)] \psi \\ &= e^{ie\Lambda} (\partial_\mu - ie A_\mu) \psi \equiv e^{ie\Lambda} D_\mu \psi . \end{aligned}$$

Therefore, also the matter term in Eq. (5.1) is invariant,

$$\bar{\psi}(i\mathcal{D} - m)\psi \longrightarrow \bar{\psi}'(i\mathcal{D}' - m)\psi' = \bar{\psi}e^{-ie\Lambda}e^{ie\Lambda}(i\mathcal{D} - m)\psi = \bar{\psi}(i\mathcal{D} - m)\psi .$$

Gauge transformations are also termed **local symmetry transformations** and one speaks of **local invariance** under these transformations or, short, of a **local symmetry**. The special case $\Lambda(X) = \text{const.}$ corresponds to **global symmetry transformations** or **global invariance** or **global symmetry**, respectively.

5.2 ($SU(3), \cdot$) color symmetry of quantum chromodynamics

The theory of the strong interaction is **quantum chromodynamics (QCD)** (greek: $\chi\rho\tilde{\omega}\mu\alpha = \text{color}$). It describes the interaction between **quarks** and **gluons**. The Lagrangian of QCD looks quite similar to that of QED, Eq. (5.1),

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{2} \text{Tr}(\mathcal{F}^{\mu\nu} \mathcal{F}_{\mu\nu}) + \bar{\Psi}(i\mathcal{D} - \hat{m})\Psi . \quad (5.6)$$

Here,

$$\mathcal{F}^{\mu\nu} \equiv F_a^{\mu\nu} \hat{T}_a \quad (5.7)$$

is the **matrix-valued field-strength tensor**, with the eight **generators** \hat{T}_a of $(SU(3), \cdot)$ in the **fundamental representation**, i.e., as (3×3) matrices, cf. Eq. (4.2). Ψ is the quark field and \hat{m} its mass.

Using Eq. (3.41) the gauge-field term of QCD can be brought into a similar form as that of QED,

$$\frac{1}{2} \text{Tr}(\mathcal{F}^{\mu\nu} \mathcal{F}_{\mu\nu}) = F_a^{\mu\nu} F_{\mu\nu}^a \frac{1}{2} \text{Tr}(\hat{T}_a \hat{T}_a) = \frac{1}{4} F_a^{\mu\nu} F_{\mu\nu}^a \delta_{ab} = \frac{1}{4} F_a^{\mu\nu} F_{\mu\nu}^a .$$

The difference to QED is that there are now **eight** different field-strength tensors $F_a^{\mu\nu}$, corresponding to the **eight colors** of the gluon fields. There is, however, another, less obvious, difference between the field-strength tensors of QCD and that of QED. Up to a factor i/g , where g is the **coupling constant** of the strong interaction, the field-strength tensor can be defined in terms of the **commutator**,

$$\mathcal{F}^{\mu\nu} \equiv \frac{i}{g} [D^\mu, D^\nu] , \quad (5.8)$$

of two covariant derivatives

$$D_\mu = \partial_\mu - ig \mathcal{A}_\mu , \quad (5.9)$$

5.2 ($SU(3), \cdot$) color symmetry of quantum chromodynamics

where the covariant derivative is **matrix-valued**, just as the field-strength tensor, with the **matrix-valued 4-vector potential**, or **gauge field**, respectively,

$$\mathcal{A}^\mu \equiv A_a^\mu \hat{T}_a . \quad (5.10)$$

The eight 4-vector potentials A_a^μ correspond to the eight gluon fields. We compute the commutator (5.8) explicitly,

$$\begin{aligned} \frac{i}{g} [D^\mu, D^\nu] &= \frac{i}{g} [(\partial^\mu - ig \mathcal{A}^\mu)(\partial^\nu - ig \mathcal{A}^\nu) - (\partial^\nu - ig \mathcal{A}^\nu)(\partial^\mu - ig \mathcal{A}^\mu)] \\ &= \frac{i}{g} [\partial^\mu \partial^\nu - ig(\partial^\mu \mathcal{A}^\nu) - ig \mathcal{A}^\nu \partial^\mu - ig \mathcal{A}^\mu \partial^\nu - g^2 \mathcal{A}^\mu \mathcal{A}^\nu \\ &\quad - \partial^\nu \partial^\mu + ig(\partial^\nu \mathcal{A}^\mu) + ig \mathcal{A}^\mu \partial^\nu + ig \mathcal{A}^\nu \partial^\mu + g^2 \mathcal{A}^\nu \mathcal{A}^\mu] \\ &= \partial^\mu \mathcal{A}^\nu - \partial^\nu \mathcal{A}^\mu - ig[\mathcal{A}^\mu, \mathcal{A}^\nu] \\ &= \partial^\mu \mathcal{A}^\nu - \partial^\nu \mathcal{A}^\mu - ig A_b^\mu A_c^\nu [\hat{T}^b, \hat{T}^c] \\ &= (\partial^\mu A_a^\nu - \partial^\nu A_a^\mu + g f_{abc} A_b^\mu A_c^\nu) \hat{T}_a , \end{aligned}$$

where we used the Lie algebra (4.11) of $(SU(3), \cdot)$ in the last step. Comparing this result with Eq. (5.7), we realize that

$$F_a^{\mu\nu} \equiv \partial^\mu A_a^\nu - \partial^\nu A_a^\mu + g f_{abc} A_b^\mu A_c^\nu . \quad (5.11)$$

The **non-abelian** nature of the group $(SU(3), \cdot)$ causes additional terms in the field-strength tensor of the a th gluon, which depend on the gluon fields with colors b and c . These **non-abelian terms** lead to **3- and 4-gluon interaction terms** in the gauge-field term of the QCD Lagrangian,

$$\begin{aligned} -\frac{1}{4} F_a^{\mu\nu} F_{\mu\nu}^a &= -\frac{1}{2} \partial^\mu A_a^\nu (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a) \\ &\quad - g f_{abc} \partial^\mu A_a^\nu A_\mu^b A_\nu^c - \frac{g^2}{4} f_{abc} f_{ade} A_b^\mu A_c^\nu A_\mu^d A_\nu^e , \end{aligned} \quad (5.12)$$

where we used the antisymmetry of the structure constants f_{abc} . These **self-interactions** of the gluon fields lead to physically very different properties of QCD as compared to QED. For instance, QCD is an **asymptotically free** theory, while QED is not. A more detailed discussion is topic of lectures on quantum field theory.

Let us make a remark concerning the matter term in the QCD Lagrangian: the matrix-valued covariant derivative implies that the quark spinors Ψ are not only Dirac 4-spinors, but simultaneously **3-vectors** in so-called **color space**. The three components of these vectors symbolize the **quark colors**, usually termed **red**, **green**, and **blue**. Moreover, there are **six different quark flavors**: **up**, **down**, **strange**, **charm**, **bottom**, and **top**. Therefore, quark spinors have $4 \cdot 3 \cdot 6 \equiv 4N_c N_f = 72$ components, where we denoted the number of (fundamental) quark colors as N_c and the number of flavors as N_f . The quark mass \hat{m} in the QCD Lagrangian is not a single number, but a diagonal $[(4N_c N_f) \times (4N_c N_f)]$

5 Unitary symmetries of the strong interaction

matrix in Dirac, color, and flavor space,

$$\hat{m} = \begin{pmatrix} \hat{m}_u & 0 & \cdots & & & & \\ 0 & \hat{m}_d & 0 & \cdots & & & \\ \vdots & 0 & \hat{m}_s & 0 & \cdots & & \\ & \vdots & 0 & \hat{m}_c & 0 & \cdots & \\ & & \vdots & 0 & \hat{m}_b & 0 & \\ & & & \vdots & 0 & \hat{m}_t & \end{pmatrix},$$

where the individual flavor matrices $\hat{m}_i = m_i \mathbb{1}_{4N_c}$, with the mass m_i of quark flavor i .

Quite analogously to the case of QED, the QCD Lagrangian (5.6) is **invariant** under **local** ($SU(N_c), \cdot$) transformations in color space,

$$\Psi \longrightarrow \Psi' = \hat{U} \Psi, \quad (5.13)$$

$$\mathcal{A}^\mu \longrightarrow \mathcal{A}'^\mu = \hat{U} \mathcal{A}^\mu \hat{U}^{-1} - \frac{i}{g} (\partial^\mu \hat{U}) \hat{U}^{-1}, \quad (5.14)$$

where

$$\hat{U} = \exp \left[ig \Lambda_a(X) \hat{T}_a \right] \in SU(3) \quad (5.15)$$

is a representation of an element of the group ($SU(3), \cdot$). The invariance can be seen as follows:

(i) The covariant derivative transforms under gauge transformations as follows:

$$\begin{aligned} D_\mu = \partial_\mu - ig \mathcal{A}_\mu &\longrightarrow D'_\mu = \partial_\mu - ig \mathcal{A}'_\mu \\ &= \partial_\mu - ig \left[\hat{U} \mathcal{A}_\mu \hat{U}^{-1} - \frac{i}{g} (\partial_\mu \hat{U}) \hat{U}^{-1} \right] \\ &= \partial_\mu + \left[\partial_\mu (\hat{U} \hat{U}^{-1}) \right] - ig \hat{U} \mathcal{A}_\mu \hat{U}^{-1} - (\partial_\mu \hat{U}) \hat{U}^{-1} \\ &= \partial_\mu + \hat{U} (\partial_\mu \hat{U}^{-1}) + (\partial_\mu \hat{U}) \hat{U}^{-1} - ig \hat{U} \mathcal{A}_\mu \hat{U}^{-1} \\ &\quad - (\partial_\mu \hat{U}) \hat{U}^{-1} \\ &= \hat{U} \left[\hat{U}^{-1} \partial_\mu + (\partial_\mu \hat{U}^{-1}) - ig \mathcal{A}_\mu \hat{U}^{-1} \right] \\ &= \hat{U} (\partial_\mu - ig \mathcal{A}_\mu) \hat{U}^{-1} \equiv \hat{U} D_\mu \hat{U}^{-1}. \end{aligned} \quad (5.16)$$

Therefore, also the matrix-valued field-strength tensor transforms as

$$\begin{aligned} \mathcal{F}_{\mu\nu} \equiv \frac{i}{g} [D_\mu, D_\nu] &\longrightarrow \mathcal{F}'^{\mu\nu} = \frac{i}{g} [D'_\mu, D'_\nu] = \frac{i}{g} (D'_\mu D'_\nu - D'_\nu D'_\mu) \\ &= \frac{i}{g} \left(\hat{U} D_\mu \hat{U}^{-1} \hat{U} D_\nu \hat{U}^{-1} - \hat{U} D_\nu \hat{U}^{-1} \hat{U} D_\mu \hat{U}^{-1} \right) \\ &= \frac{i}{g} \hat{U} [D_\mu, D_\nu] \hat{U}^{-1} \equiv \hat{U} \mathcal{F}_{\mu\nu} \hat{U}^{-1}. \end{aligned} \quad (5.17)$$

5.2 ($SU(3), \cdot$) color symmetry of quantum chromodynamics

Thus, the gauge-field term in the QCD Lagrangian is **invariant** under local $(SU(N_c), \cdot)$ transformations,

$$\begin{aligned} \frac{1}{2} \text{Tr} (\mathcal{F}^{\mu\nu} \mathcal{F}_{\mu\nu}) &\longrightarrow \frac{1}{2} \text{Tr} (\mathcal{F}'^{\mu\nu} \mathcal{F}'_{\mu\nu}) \\ &= \frac{1}{2} \text{Tr} (\hat{U} \mathcal{F}^{\mu\nu} \hat{U}^{-1} \hat{U} \mathcal{F}_{\mu\nu} \hat{U}^{-1}) \equiv \frac{1}{2} \text{Tr} (\mathcal{F}^{\mu\nu} \mathcal{F}_{\mu\nu}) , \end{aligned} \quad (5.18)$$

because we may cyclically permute terms under the trace.

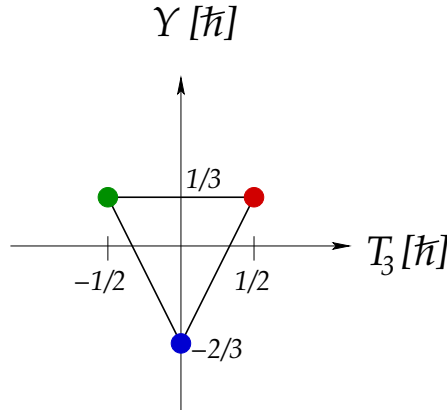
- (ii) The matter term is also invariant under local $(SU(N_c), \cdot)$ transformations, because the covariant derivative of the quark spinor transforms with the help of Eq. (5.16) as

$$D_\mu \Psi \longrightarrow D'_\mu \Psi' = \hat{U} D_\mu \hat{U}^{-1} \hat{U} \Psi = \hat{U} D_\mu \Psi , \quad (5.19)$$

and thus

$$\bar{\Psi}(i\not{D} - \hat{m})\Psi \longrightarrow \bar{\Psi}'(i\not{D}' - \hat{m})\Psi' = \bar{\Psi} \hat{U}^{-1} \hat{U} (i\not{D} - \hat{m})\Psi \equiv \bar{\Psi}(i\not{D} - \hat{m})\Psi . \quad (5.20)$$

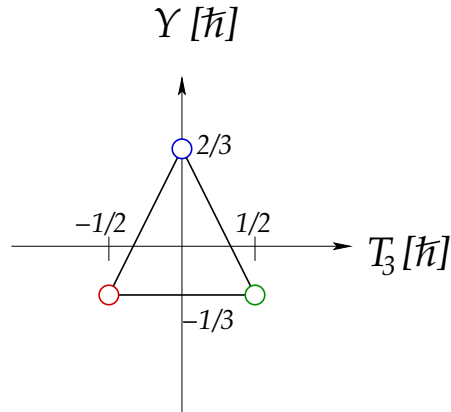
The **gauge invariance** of QCD, i.e., the **symmetry** under **local** $(SU(N_c), \cdot)$ **transformations** implies that quarks as well as gluons can be grouped into **multiplets** of $(SU(N_c), \cdot)$. This must be the case, since quarks can also transform amongst each other, but not into gluons or other objects. The same holds for gluons. The $N_c = 3$ quark colors are states of the fundamental representation, the **triplet**:



The colors are assigned as follows to eigenstates $|T_3 Y\rangle$ of “color spin” T_3 and “color hypercharge” Y :

$$\begin{aligned} |r\rangle \equiv |”red”\rangle &\equiv \left| \frac{1}{2} \frac{1}{3} \right\rangle , \\ |g\rangle \equiv |”green”\rangle &\equiv \left| -\frac{1}{2} \frac{1}{3} \right\rangle , \\ |b\rangle \equiv |”blue”\rangle &\equiv \left| 0 -\frac{2}{3} \right\rangle . \end{aligned}$$

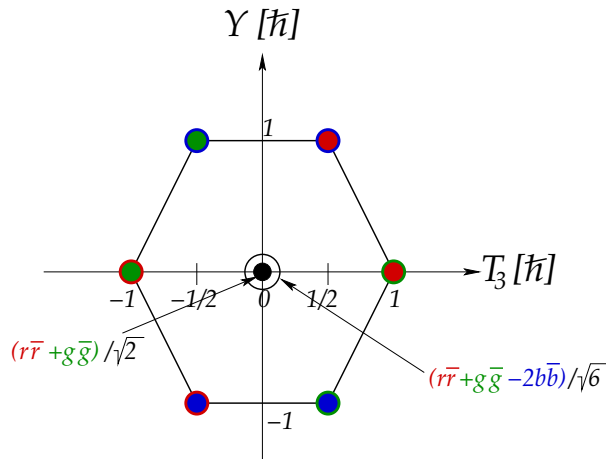
Anti-quarks differ from quarks in **all** charge quantum numbers, and thus also in **color**. Therefore, anti-quarks occupy states of the **anti-triplet**,



The assignment of colors to $|T_3 Y\rangle$ eigenstates is the following:

$$\begin{aligned}
 |\bar{r}\rangle &\equiv |\text{"anti-red"}\rangle &\equiv & \left| -\frac{1}{2} \quad -\frac{1}{3} \right\rangle, \\
 |\bar{g}\rangle &\equiv |\text{"anti-green"}\rangle &\equiv & \left| \frac{1}{2} \quad -\frac{1}{3} \right\rangle, \\
 |\bar{b}\rangle &\equiv |\text{"anti-blue"}\rangle &\equiv & \left| 0 \quad \frac{2}{3} \right\rangle.
 \end{aligned}$$

The eight **gluons** occupy states of the adjoint representation, the **octet**. Since the octet can be generated by coupling a triplet and an anti-triplet, cf. Eq. (4.50), one can imagine gluons to carry combinations of colors and anti-colors.



A peculiarity of QCD is the so-called **color confinement**: particles which are subject to the strong interaction, so-called **hadrons**, appear in Nature always as **singlet** (“white” objects) of the $(SU(N_c), \cdot)$ color symmetry, **color-charged** objects such as quarks and gluons are **never** observed in isolation. The confinement of color charges is an experimentally observed fact, but so far could not be shown rigorously on the basis of the theory of QCD.

The fact that hadrons are always color-singlets means that they must consist of **several** quarks and gluons, and in particular of such couplings between these particles which allow a singlet, e.g. as in Eqs. (4.50) and (4.52). Color-singlets which consist of a quark and an anti-quark are called **mesons**, color-singlets which consist of three (anti-)quarks are called **(anti-)baryons**. We will discuss concrete examples in the next sections, but first we discuss another symmetry of QCD, the so-called **flavor symmetry**.

5.3 ($SU(N_f), \cdot$) flavor symmetry

In the case of vanishing mass, left- and right-handed Dirac fields,

$$\psi_{r,\ell} \equiv \mathcal{P}_{r,\ell} \psi \equiv \frac{1 \pm \gamma_5}{2} \psi, \quad \bar{\psi}_{r,\ell} \equiv \bar{\psi} \mathcal{P}_{\ell,r} \equiv \bar{\psi} \frac{1 \mp \gamma_5}{2}, \quad (5.21)$$

where $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$, decouple in the Dirac Lagrangian,

$$\bar{\psi} i\not{\partial} \psi = \bar{\psi}_r i\not{\partial} \psi_r + \bar{\psi}_\ell i\not{\partial} \psi_\ell. \quad (5.22)$$

(In order to show this, use $\{\gamma_5, \gamma^\mu\} = 0$, as well as the projector properties $\mathcal{P}_{r,\ell}^2 = \mathcal{P}_{r,\ell}$, $\mathcal{P}_r \mathcal{P}_\ell = \mathcal{P}_\ell \mathcal{P}_r = 0$, $\mathcal{P}_r + \mathcal{P}_\ell = \mathbb{1}_4$.) This also holds for the QCD Lagrangian (5.6), which is then **invariant** under **global unitary transformations** of right- and left-handed quark fields,

$$\begin{aligned} \Psi_{r,\ell} &\longrightarrow \Psi'_{r,\ell} = \hat{U}_{r,\ell} \Psi_{r,\ell}, \\ \hat{U}_{r,\ell} &= \exp\left(-i\alpha_{r,\ell}^a \hat{T}_a\right) \in U(N_f)_{r,\ell}, \quad \alpha_{r,\ell}^a = \text{const.}, \quad a = 0, \dots, N_f^2 - 1. \end{aligned} \quad (5.23)$$

Since one can transform right- and left-handed fields **separately**, the full symmetry group is $(U(N_f)_r \times U(N_f)_\ell, \cdot)$, the so-called **chiral symmetry** of QCD.

A mass term **explicitly** breaks this symmetry, as one readily confirms,

$$\bar{\psi} m \psi = \bar{\psi}_\ell m \psi_r + \bar{\psi}_r m \psi_\ell. \quad (5.24)$$

This term is only symmetric under those chiral transformations which fulfill $\hat{U}_\ell^\dagger = \hat{U}_\ell^{-1} = \hat{U}_r^{-1}$, or $\alpha_r^a = \alpha_\ell^a$. This is the symmetry of **vector transformations** $V = r + \ell$,

$$\hat{U}_V = \exp\left(-i\alpha_V^a \hat{T}_a\right) \in U(N_f)_V, \quad \alpha_V^a \equiv \alpha_{r+\ell}^a = \alpha_r^a + \alpha_\ell^a, \quad (5.25)$$

which forms the diagonal subgroup $(U(N_f)_V, \cdot)$ of the chiral symmetry group $(U(N_f)_r \times U(N_f)_\ell, \cdot)$. On the other hand, the symmetry of **axial-vector transformations** $A = r - \ell$,

$$\hat{U}_A = \exp\left(-i\gamma_5 \alpha_A^a \hat{T}_a\right) \in U(N_f)_A, \quad \alpha_A^a \equiv \alpha_{r-\ell}^a = \alpha_r^a - \alpha_\ell^a, \quad (5.26)$$

is **explicitly** broken. (There is no such symmetry, since the mass term requires $\alpha_r^a = \alpha_\ell^a$, which means that $\alpha_A^a \equiv 0$.)

The individual quark masses are non-zero,

$$\begin{aligned} m_u &\simeq 2.3 \text{ MeV}, & m_d &\simeq 4.8 \text{ MeV}, \\ m_s &\simeq 95 \text{ MeV}, & m_c &\simeq 1.275 \text{ GeV}, \\ m_b &\simeq 4.18 \text{ GeV}, & m_t &\simeq 173.21 \text{ GeV}, \end{aligned} \quad (5.27)$$

consequently, the chiral symmetry of QCD is explicitly broken. If all quark masses were equal, then the residual symmetry of QCD would be that of the vector transformations, $(U(N_f)_V, \cdot)$. Since $U(N) = SU(N) \otimes U(1)$, cf. Exercise 2, and the $U(1)_V$ symmetry represents according to Noether's theorem simply the (trivial) conservation of quark number,

one usually considers the group $(SU(N_f), \cdot)$ of **special unitary vector transformations**, or short, the $(SU(N_f), \cdot)$ **flavor symmetry**.

This symmetry is broken because the quark masses (5.27) are all of different value. However, for some quark flavors the breaking is less strong than for others. For instance, the mass difference between up and down quark is, in comparison to a typical hadronic mass scale of $M_h \sim 1$ GeV, vanishingly small. This results in an approximate $(SU(2), \cdot)$ flavor symmetry, the so-called **isospin symmetry** of QCD and thus of the strong interaction, which will be discussed in the next section. Moreover, also the mass difference between strange and up or down quark is, on a hadronic mass scale, small, so that one can assume to good approximation also an approximate $(SU(3), \cdot)$ flavor symmetry of the strong interaction. This will be discussed in the next-to-next section. More strongly broken is the $(SU(4), \cdot)$ flavor symmetry, which also takes into account the charm quark. It is nevertheless reasonable to also consider this symmetry in order to classify hadrons with the quantum number charm, which will be done at the end of this chapter.

5.4 Isospin symmetry

The masses of up and down quark are, in comparison to a hadronic mass scale $M_h \sim 1$ GeV, almost of equal magnitude (and also vanishingly small),

$$m_u \simeq 2.3 \text{ MeV} < m_d \simeq 4.8 \text{ MeV} \ll M_h \sim 1 \text{ GeV} .$$

Therefore, one can to good approximation assume that QCD possesses an $(SU(2), \cdot)$ flavor symmetry. (Since the quark masses are vanishingly small compared to M_h , one may think that the symmetry group is actually larger, namely that of chiral symmetry for $N_f = 2$ massless flavors, $U(2)_r \times U(2)_\ell$. This is, however, not true because of **spontaneous breaking** of chiral symmetry by a non-vanishing quark condensate $\langle \bar{\Psi}\Psi \rangle \neq 0$.)

We write

$$\bar{m} \equiv \frac{1}{2} (m_u + m_d) , \quad \delta m \equiv m_d - m_u , \quad (5.28)$$

so that

$$m_u = \bar{m} - \frac{\delta m}{2} , \quad m_d = \bar{m} + \frac{\delta m}{2} , \quad (5.29)$$

and neglect the mass difference δm in the following. Then, the $(SU(2), \cdot)$ flavor symmetry of QCD is even an **exact** symmetry. Up and down quark form an $(SU(2), \cdot)$ **doublet**, [2],

$$\begin{array}{ccc} |d\rangle & & |u\rangle \\ \bullet & & \bullet \\ \hline -1/2 & 0 & 1/2 \end{array} \quad \longrightarrow \quad I_3[\hbar]$$

If we denote the states with $|I I_3\rangle$, we have

$$\begin{aligned} |u\rangle &\equiv | \text{“up”} \rangle = \left| \frac{1}{2} \frac{1}{2} \right\rangle , \\ |d\rangle &\equiv | \text{“down”} \rangle = \left| \frac{1}{2} - \frac{1}{2} \right\rangle . \end{aligned}$$

The corresponding **anti-doublet** $[\bar{2}]$ consists of the corresponding antiparticles, the anti-up \bar{u} and anti-down quark \bar{d} . Formally, it has the same shape as the doublet,

$$\begin{array}{c} | \bar{u} \rangle \qquad | \bar{d} \rangle \\ \bullet \qquad \bullet \\ \hline -1/2 \quad 0 \quad 1/2 \end{array} \longrightarrow I_3[\hbar]$$

The assignment of states reads

$$\begin{aligned} |\bar{u}\rangle &\equiv | \text{“anti-up”} \rangle = \left| \frac{1}{2} \ - \frac{1}{2} \right\rangle, \\ |\bar{d}\rangle &\equiv | \text{“anti-down”} \rangle = \left| \frac{1}{2} \ \frac{1}{2} \right\rangle. \end{aligned}$$

Because of the formal similarity with [2] the coupling rules for the anti-doublet are identical with those of the doublet, i.e., in analogy to Eq. (4.46) we have

$$[2] \otimes [\bar{2}] = [1] \oplus [3]. \quad (5.30)$$

Because of the $(SU(2), \cdot)$ symmetry, all hadrons which are composed of up and down quarks (or the respective anti-quarks) must also fall into $(SU(2), \cdot)$ multiplets. The assignment follows the coupling rules for $(SU(2), \cdot)$, which we have discussed in Sec. 4.5. Since the eigenvalues of the Hamilton operator $\hat{H}(\hat{C}_1)$ are degenerate on a multiplet, all hadrons of a given multiplet must possess the **same mass**, if the $(SU(2), \cdot)$ flavor symmetry is exact. Therefore, the $(SU(2), \cdot)$ flavor symmetry is also called **isobaric spin symmetry**, or short **isospin symmetry** (greek: ἴσος βαρύς = equally heavy). This is the origin of the labels I and I_3 for magnitude and 3–component of the isospin.

Mesons are colorless **quark–anti-quark states**. The possible $(SU(2), \cdot)$ flavor multiplets can be determined from Eq. (5.30), i.e., there is a **flavor singlet** and a **flavor triplet**. Considering the respective Clebsch-Gordan coefficients one obtains

$$\begin{array}{c} (|u\bar{u}\rangle - |d\bar{d}\rangle)/\sqrt{2} \\ \bullet \\ \hline 0 \end{array} \longrightarrow I_3[\hbar] \qquad \begin{array}{c} |d\bar{u}\rangle \quad (|u\bar{u}\rangle + |d\bar{d}\rangle)/\sqrt{2} \quad |u\bar{d}\rangle \\ \bullet \qquad \bullet \qquad \bullet \\ \hline -1 \qquad 0 \qquad 1 \end{array} \longrightarrow I_3[\hbar]$$

There are, however, different kinds of mesons with the same flavor content, but different behavior under Lorentz transformations, i.e., they differ in their spin J . The usual classification follows according to the isospin I of the meson, and then according to spin J , parity P , and charge conjugation C , with the notation J^{PC} . Some of the $N_f = 2$ mesons are listed in Tab. 5.1.

The state of the triplet differ in their electric charge, which can be calculated according to the formula

$$Q = e I_3 \quad (\text{mesons}). \quad (5.31)$$

As one observes, the isospin symmetry is nearly exact, only the different states of the pion triplet differ slightly in their masses,

$$\delta m_\pi = m_{\pi^\pm} - m_{\pi^0} = 4.59358 \text{ MeV} \ll m_{\pi^\pm}, m_{\pi^0}.$$

5 Unitary symmetries of the strong interaction

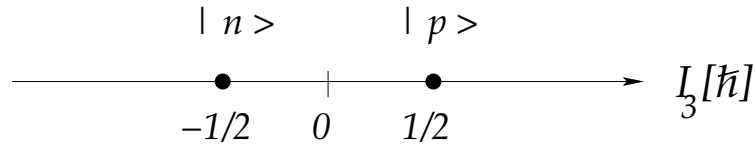
Mesons	[1] ($I = 0$)	[3] ($I = 1$)	
Scalars ($J^{PC} = 0^{++}$)	$f_0(1370)$	$a_0^\pm(1450)$	$a_0^0(1450)$
Mass [MeV]	1350 ± 150	1474 ± 19	
Pseudoscalars ($J^{PC} = 0^{-+}$)	η	π^\pm	π^0
Mass [MeV]	547.862 ± 0.018	139.57018 ± 0.00035	134.9766 ± 0.0006
Vectors ($J^{PC} = 1^{--}$)	ω	ρ^\pm	ρ^0
Mass [MeV]	782.65 ± 0.12	775.26 ± 0.25	
Axial-vectors ($J^{PC} = 1^{++}$)	$f_1(1285)$	$a_1^\pm(1260)$	$a_1^0(1260)$
Masse [MeV]	1281.9 ± 0.5	1230 ± 40	

Table 5.1: Mesons which consist of a quark and an anti-quark.

Baryons are colorless states composed of **three quarks**. The possible flavor multiplets can be determined from

$$[2] \otimes [2] \otimes [2] = [2] \otimes ([1] \oplus [3]) = [2] \oplus [2] \oplus [4]. \quad (5.32)$$

One of the doublets is the **nucleon doublet**, consisting of **proton** and **neutron**:



Taking into account the Clebsch-Gordan coefficients, the flavor content of proton and neutron is

$$|p\rangle \equiv |\text{“proton”}\rangle = \left| \frac{1}{2} \frac{1}{2} \right\rangle = \frac{1}{\sqrt{6}} (|uud\rangle + |udu\rangle - 2|duu\rangle),$$

$$|n\rangle \equiv |\text{“neutron”}\rangle = \left| \frac{1}{2} -\frac{1}{2} \right\rangle = \frac{1}{\sqrt{6}} (2|udd\rangle - |dud\rangle - |ddu\rangle),$$

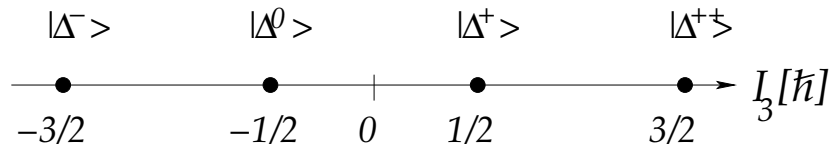
The quantum numbers of the nucleon doublet read $I(J^P) = \frac{1}{2}(\frac{1}{2}^+)$ and the masses are

$$m_n = 939.565379 \pm 0.000021 \text{ MeV}, \quad m_p = 938.272046 \pm 0.000021 \text{ MeV},$$

so that the mass difference (and thus isospin violation) is again small compared to the absolute magnitude of the nucleon masses,

$$\delta m_N = m_n - m_p = 1.293333 \text{ MeV} \ll m_n, m_p.$$

The quadruplet in Eq. (5.32) is that of the so-called **Delta resonances**:



Taking into account the Clebsch-Gordan coefficients, the flavor content of the Δ resonances reads

$$\begin{aligned} |\Delta^{++}\rangle &= \left| \frac{3}{2} \frac{3}{2} \right\rangle = |uuu\rangle, \\ |\Delta^+\rangle &= \left| \frac{3}{2} \frac{1}{2} \right\rangle = \frac{1}{\sqrt{3}} (|uud\rangle + |udu\rangle + |duu\rangle), \\ |\Delta^0\rangle &= \left| \frac{3}{2} - \frac{1}{2} \right\rangle = \frac{1}{\sqrt{3}} (|udd\rangle + |dud\rangle + |ddu\rangle), \\ |\Delta^-\rangle &= \left| \frac{3}{2} - \frac{3}{2} \right\rangle = |ddd\rangle. \end{aligned}$$

The quantum numbers of the Δ quadruplet are $I(J^P) = \frac{3}{2}(\frac{3}{2}^+)$ and the mass is $m_\Delta = 1232$ MeV.

In order to compute the electric charge of the baryons, Eq. (5.31) needs to be modified:

$$Q = e \left(I_3 + \frac{1}{2} \right) \quad (\text{baryons}). \quad (5.33)$$

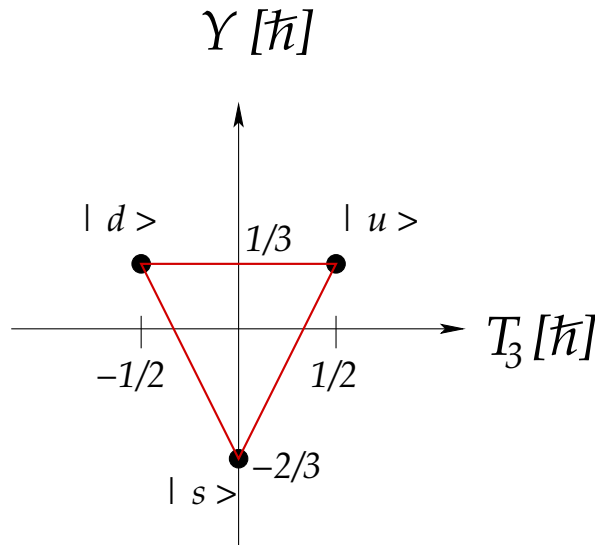
5.5 Strangeness and $(SU(3), \cdot)$ flavor symmetry

The difference between the masses of up, down, and strange quarks is, in comparison to a typical hadronic mass scale $M_h \sim 1$ GeV, still relatively small,

$$\begin{aligned} \delta m_{ud} &= m_d - m_u \simeq 2.5 \text{ MeV} \\ &< \delta m_{ds} = m_s - m_d \simeq 90 \text{ MeV} \\ &\simeq \delta m_{us} = m_s - m_u \simeq 92.7 \text{ MeV} \\ &\ll M_h \sim 1 \text{ GeV}, \end{aligned}$$

so that one can speak of an approximate $(SU(3), \cdot)$ flavor symmetry of the strong interaction. In the following, we first neglect the mass differences δm_{ud} , δm_{ds} , and δm_{us} and assume an **exact** $(SU(3), \cdot)$ flavor symmetry.

Then, the three quark flavors form an $(SU(3), \cdot)$ triplet:



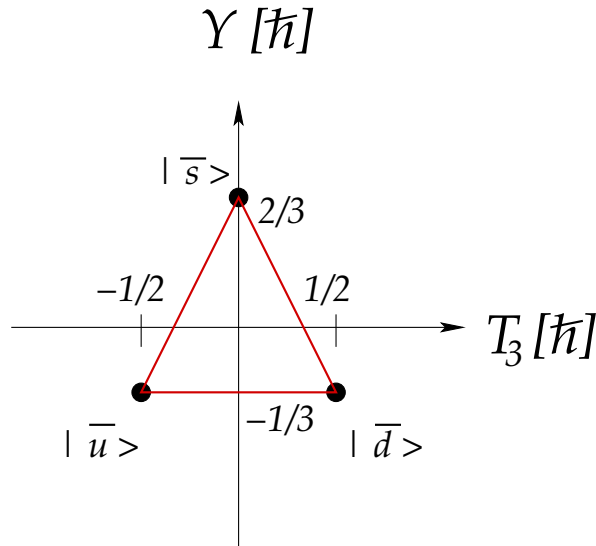
5 Unitary symmetries of the strong interaction

The role of the 3–component of isospin from the preceding section is now assumed by T_3 . The assignment of states is

$$\begin{aligned} |u\rangle \equiv |\text{“up”}\rangle &\equiv \left| \frac{1}{2} \frac{1}{3} \right\rangle, \\ |d\rangle \equiv |\text{“down”}\rangle &\equiv \left| -\frac{1}{2} \frac{1}{3} \right\rangle, \\ |s\rangle \equiv |\text{“strange”}\rangle &\equiv \left| 0 -\frac{2}{3} \right\rangle. \end{aligned}$$

Note the analogy to the assignment of the three colors red, green, and blue in the case of the $(SU(3), \cdot)$ color symmetry of QCD.

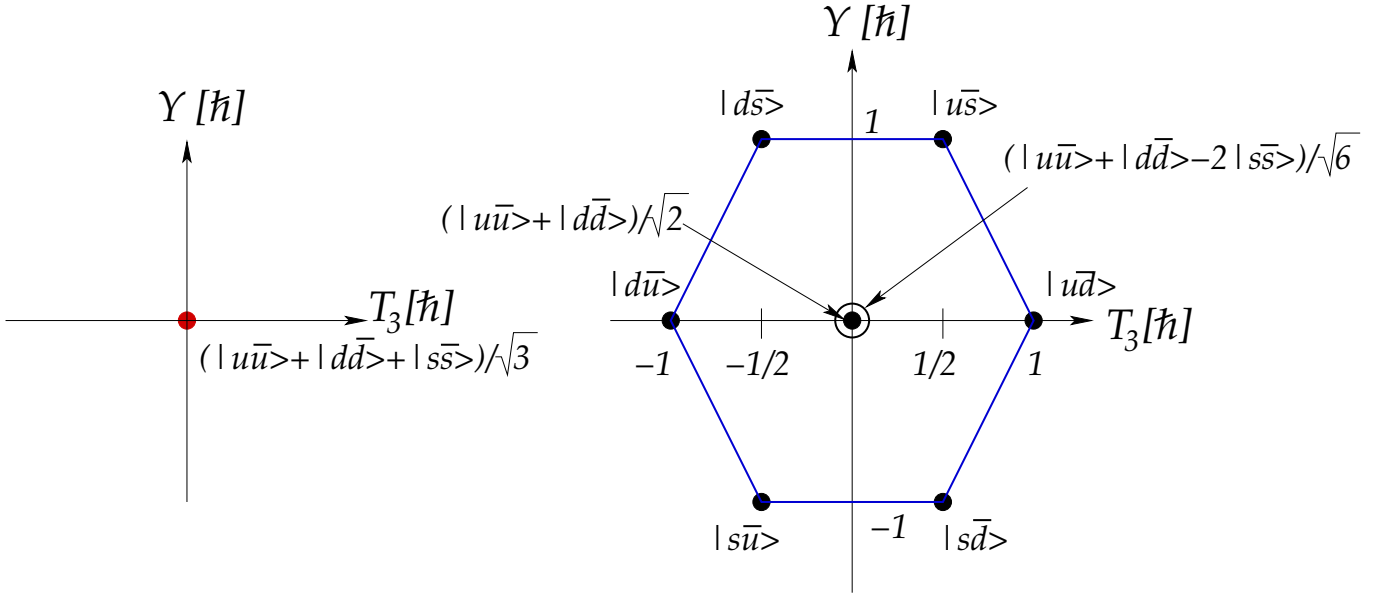
Anti-quarks form an anti-triplet:



The assignment of states is

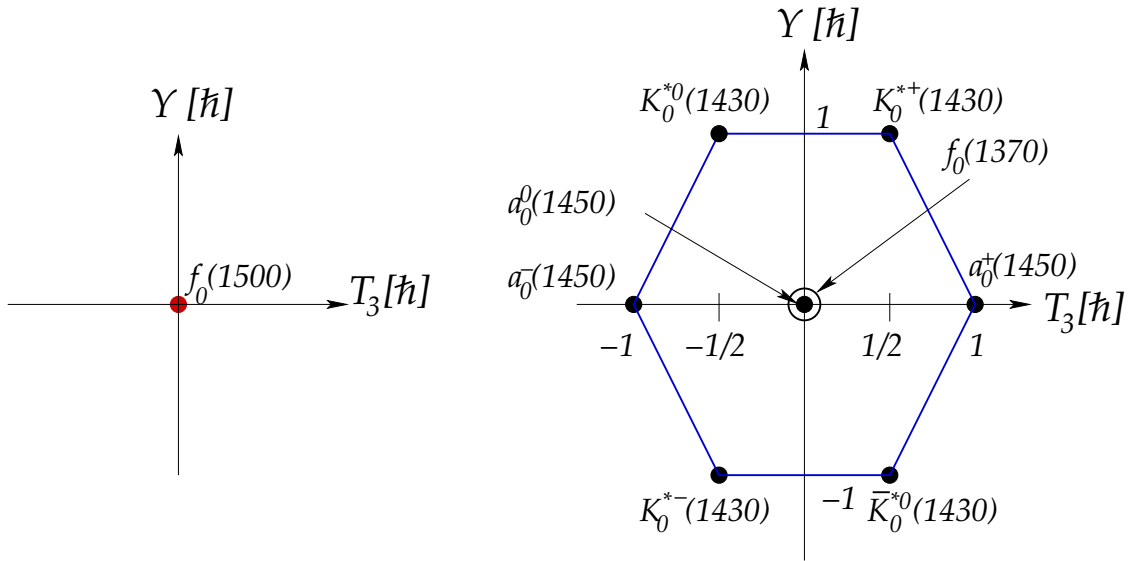
$$\begin{aligned} |\bar{u}\rangle \equiv |\text{“anti-up”}\rangle &\equiv \left| -\frac{1}{2} -\frac{1}{3} \right\rangle, \\ |\bar{d}\rangle \equiv |\text{“anti-down”}\rangle &\equiv \left| \frac{1}{2} -\frac{1}{3} \right\rangle, \\ |\bar{s}\rangle \equiv |\text{“anti-strange”}\rangle &\equiv \left| 0 \frac{2}{3} \right\rangle. \end{aligned}$$

According to the coupling rule (4.50), **mesons** form a **singlet** and an **octet**:



We distinguish

- (i) **Scalar mesons**, $J^{PC} = 0^{++}$:

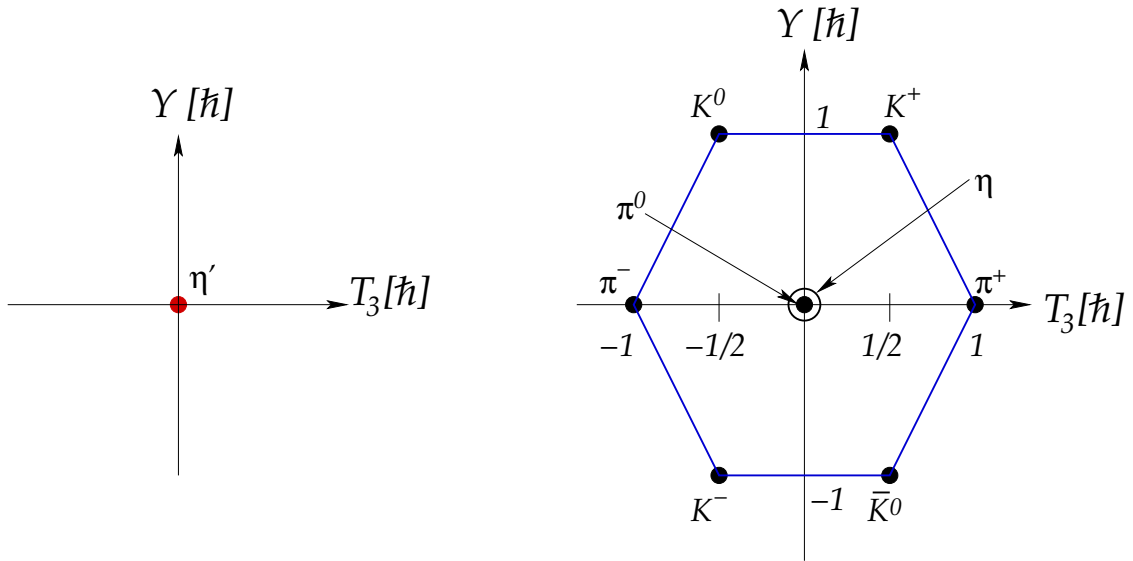


As we observe, the isosinglet $f_0(1370)$ and the isotriplet of $a_0(1260)$ mesons from the $N_f = 2$ case are parts of the octet. New as compared to the latter case are the four scalar K_0^* mesons, which form two isospin doublets inside the octet. The new isoscalar $f_0(1500)$ meson constitutes the singlet. The masses and isospin quantum numbers of these mesons are:

$$\begin{aligned}
 f_0(1500) &: I = 0, & m_{f_0(1500)} &= 1505 \pm 6 \text{ MeV}, \\
 K_0^*(1430) &: I = \frac{1}{2}, & m_{K_0^*} &= 1425 \pm 50 \text{ MeV}.
 \end{aligned}$$

5 Unitary symmetries of the strong interaction

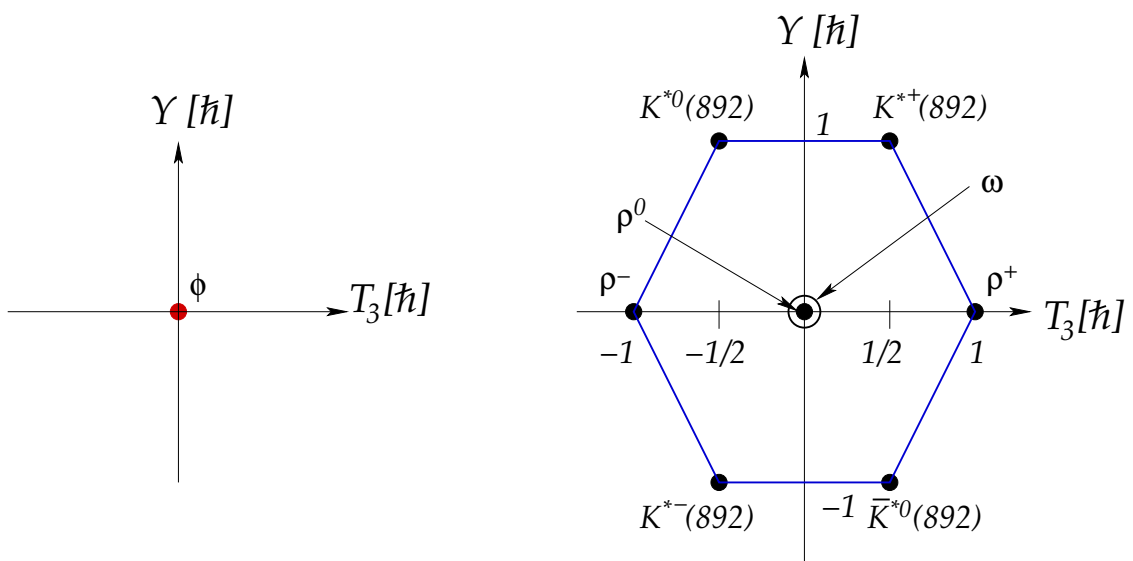
(ii) **Pseudoscalar mesons**, $J^{PC} = 0^{-+}$:



As we observe, the isosinglet η and the pion isotriplet from the $N_f = 2$ case are parts of the octet. New as compared to the latter case are the four pseudoscalar K mesons, or short **kaons**, which form two isospin doublets inside the octet. The new isoscalar η' meson constitutes the singlet. The masses and isospin quantum numbers of these mesons are:

$$\begin{aligned} \eta' &: I = 0, & m_{\eta'} &= 957.78 \pm 0.06 \text{ MeV}, \\ K &: I = \frac{1}{2}, & m_{K^\pm} &= 493.677 \pm 0.016 \text{ MeV}, \\ & & m_{K^0} &= 497.614 \pm 0.024 \text{ MeV}. \end{aligned}$$

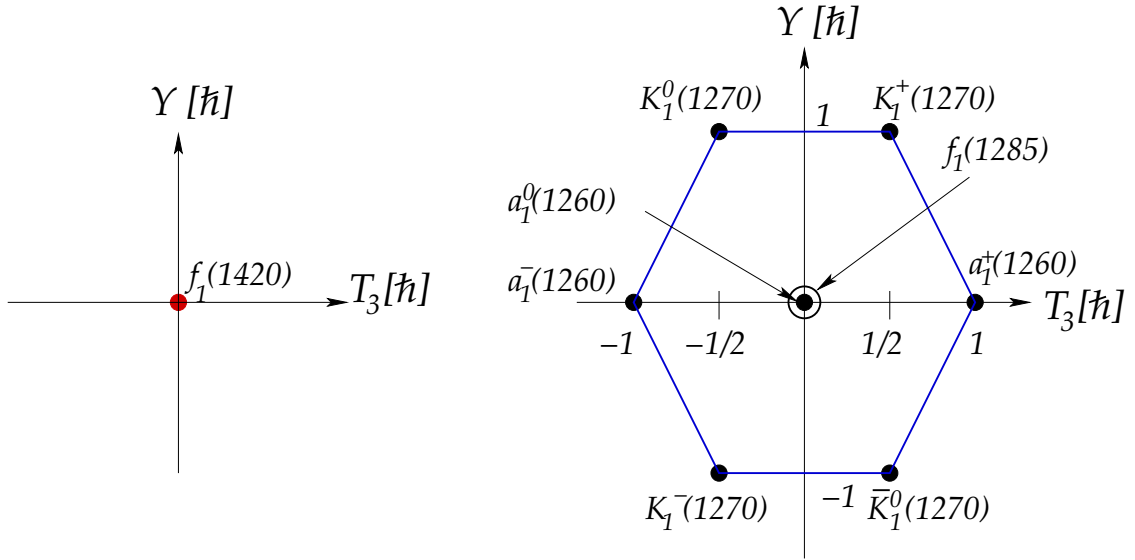
(iii) **Vector mesons**, $J^{PC} = 1^{--}$:



As we observe, the isosinglet ω and the isotriplet of the ρ mesons from the $N_f = 2$ case are parts of the octet. New as compared to the latter case are the four vector K^* mesons, which form two isospin doublets inside the octet. The new isoscalar ϕ meson constitutes the singlet. The masses and isospin quantum numbers of these mesons are:

$$\begin{aligned} \phi & : I = 0 , & m_\phi & = 1019.461 \pm 0.019 \text{ MeV} , \\ K^* & : I = \frac{1}{2} , & m_{K^*} & = 891.66 \pm 0.26 \text{ MeV} . \end{aligned}$$

(iv) **Axial-vector mesons**, $J^{PC} = 1^{++}$:

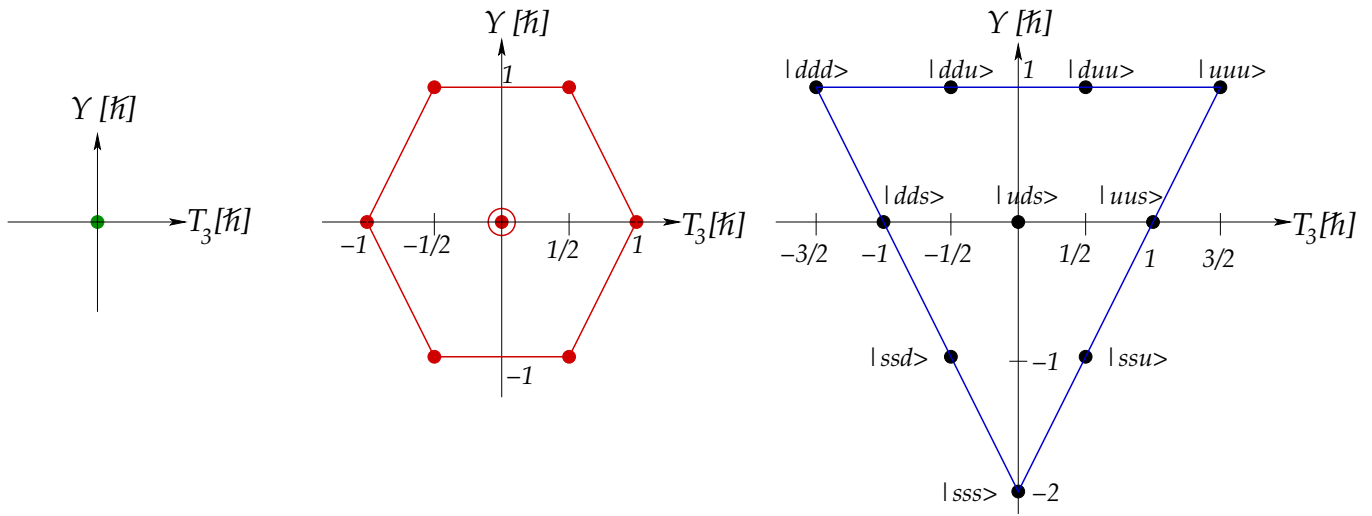


The isosinglet $f_1(1285)$ and the isotriplet of the a_1 mesons from the case $N_f = 2$ are parts of the octet. New as compared to the case $N_f = 2$ are the four axial-vector K_1 mesons, which form two isospin doublets inside the octet. The new isoscalar $f_1(1420)$ meson constitutes the singlet. The masses and isospin quantum numbers of these mesons are:

$$\begin{aligned} f_1(1420) & : I = 0 , & m_{f_1(1420)} & = 1426.4 \pm 0.9 \text{ MeV} , \\ K_1 & : I = \frac{1}{2} , & m_{K_1} & = 1272 \pm 7 \text{ MeV} . \end{aligned}$$

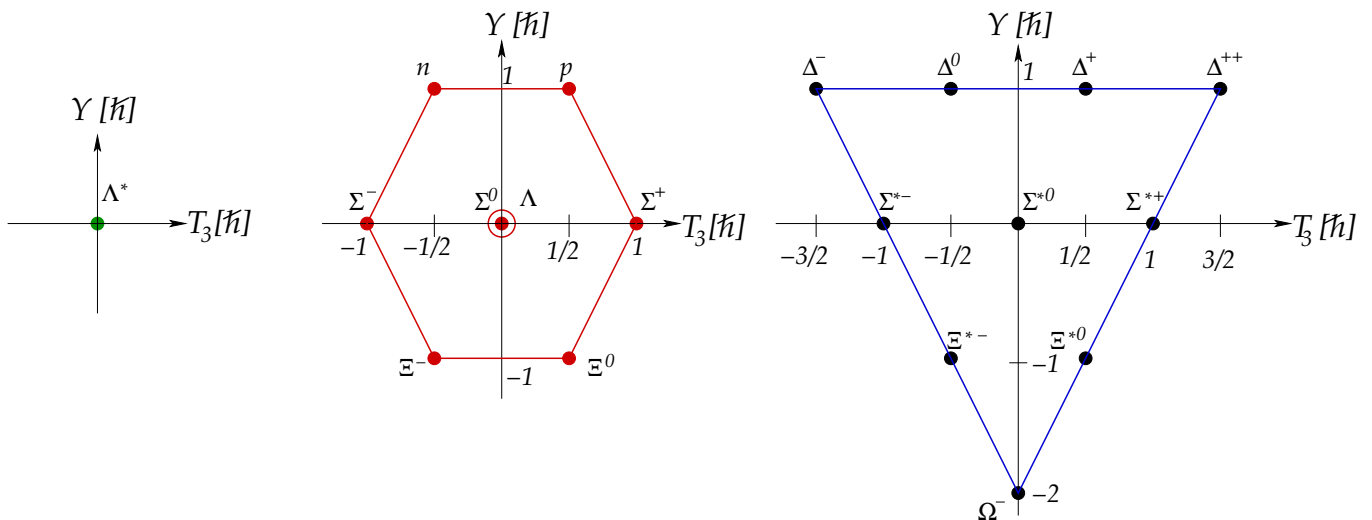
5 Unitary symmetries of the strong interaction

According to the coupling rule (4.52), **baryons** form a **singlet**, two **octets**, and a **decuplet**:



We only show one of the two octets, and we just indicate the flavor content of the states of the decuplet, without respecting proper symmetrization. The flavor content of the states of the singlet and of the two octets is in principle the same as that of the corresponding states of the decuplet, but the quarks appear in different combinations, corresponding to the respective symmetry of the state under exchange of particles. Due to lack of time, we cannot discuss this further at this point.

We just mention the assignment of physical baryon states to these multiplets:



We observe that the nucleon isodoublet is part of the octet and the Δ isoquadruplet is

part of the decuplet. New baryon states as compared to the case $N_f = 2$ are:

$$\begin{aligned}
 \Lambda & : I(J^P) = 0(\frac{1}{2}^+) , & m_\Lambda & = 1115.683 \pm 0.006 \text{ MeV} , \\
 \Sigma^+ & : I(J^P) = 1(\frac{1}{2}^+) , & m_{\Sigma^+} & = 1189.37 \pm 0.07 \text{ MeV} , \\
 \Sigma^0 & : I(J^P) = 1(\frac{1}{2}^+) , & m_{\Sigma^0} & = 1192.642 \pm 0.024 \text{ MeV} , \\
 \Sigma^- & : I(J^P) = 1(\frac{1}{2}^+) , & m_{\Sigma^-} & = 1197.449 \pm 0.030 \text{ MeV} , \\
 \Xi^0 & : I(J^P) = \frac{1}{2}(\frac{1}{2}^+) , & m_{\Xi^0} & = 1314.86 \pm 0.20 \text{ MeV} , \\
 \Xi^- & : I(J^P) = \frac{1}{2}(\frac{1}{2}^+) , & m_{\Xi^-} & = 1321.71 \pm 0.07 \text{ MeV} , \\
 \Lambda^* & : I(J^P) = 0(\frac{1}{2}^+) , & m_{\Lambda^*} & = 1630 \pm 70 \text{ MeV} , \\
 \Sigma^{*+} & : I(J^P) = 1(\frac{3}{2}^+) , & m_{\Sigma^{*+}} & = 1382.80 \pm 0.35 \text{ MeV} , \\
 \Sigma^{*0} & : I(J^P) = 1(\frac{3}{2}^+) , & m_{\Sigma^{*0}} & = 1383.7 \pm 1.0 \text{ MeV} , \\
 \Sigma^{*-} & : I(J^P) = 1(\frac{3}{2}^+) , & m_{\Sigma^{*-}} & = 1387.2 \pm 0.5 \text{ MeV} , \\
 \Xi^{*0} & : I(J^P) = \frac{1}{2}(\frac{3}{2}^+) , & m_{\Xi^{*0}} & = 1531.80 \pm 0.32 \text{ MeV} , \\
 \Xi^{*-} & : I(J^P) = \frac{1}{2}(\frac{3}{2}^+) , & m_{\Xi^{*-}} & = 1535.0 \pm 0.6 \text{ MeV} , \\
 \Omega^- & : I(J^P) = 0(\frac{3}{2}^+) , & m_{\Omega^-} & = 1672.45 \pm 0.29 \text{ MeV} .
 \end{aligned}$$

The corresponding **anti-baryons** are members of a singlet, two anti-octets (which have, however, the same shape as the octets), and an anti-decuplet. Besides these baryons, there is a multitude of further excited states, as well as such of negative parity. Because of identical flavor content, also these states must be members of a singlet, two octets, and a decuplet. Further details can be found in the compilation of the Particle Data Group [5].

With the help of the hypercharge, the empirically found charge formulas (5.31) and (5.33) can now be generalized to **all** $(SU(3), \cdot)$ multiplets,

$$Q = e \left(T_3 + \frac{1}{2} Y \right) . \quad (5.34)$$

This is the so-called **Gell-Mann–Nishijima formula**. One readily checks that it holds **for meson as well as for baryon multiplets**. If it is valid for all $(SU(3), \cdot)$ multiplets, it should also apply to the triplet. Thus it follows that

$$\begin{aligned}
 Q_u & = e \left(\frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3} \right) = \frac{2}{3} e , \\
 Q_d & = e \left(-\frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3} \right) = -\frac{1}{3} e , \\
 Q_s & = e \left(0 - \frac{1}{2} \cdot \frac{2}{3} \right) = -\frac{1}{3} e .
 \end{aligned} \quad (5.35)$$

Murray Gell-Mann was the first to systematically assign hadrons to $(SU(3), \cdot)$ flavor multiplets and interpret them as quark–anti-quark (mesons) or three-quark states (baryons),

respectively. His realization culminated in postulating the existence of quarks as “building blocks” of the hadrons.

If the $(SU(3), \cdot)$ flavor symmetry were **exact**, all hadrons of a given multiplet should be **degenerate** in mass. This is obviously not the case. However, we observe that the **isospin symmetry**, i.e., the mass degeneracy inside an isospin multiplet as part of an $(SU(3), \cdot)$ flavor multiplet still holds to very good approximation. Deviations are of the order of the mass difference between up and down quark. Quite analogously, the violation of $(SU(3), \cdot)$ flavor symmetry, i.e., the **mass difference** between various isospin multiplets inside an $(SU(3), \cdot)$ flavor multiplet, is of the order of the mass difference between up or down and strange quark. For instance, for the baryons we have

$$\begin{aligned} m_\Lambda - m_N &\simeq 176.76 \text{ MeV} , \\ m_\Xi - m_\Lambda &\simeq 202.60 \text{ MeV} , \\ m_{\Sigma^*} - m_\Delta &\simeq 152.57 \text{ MeV} , \\ m_{\Xi^*} - m_{\Sigma^*} &\simeq 148.83 \text{ MeV} , \\ m_\Omega - m_{\Xi^*} &\simeq 139.05 \text{ MeV} , \end{aligned}$$

where we have assumed averaged values inside an isospin multiplet. Obviously, the mass of the baryons grows proportional to the number of strange quarks contained in them, by an amount of the order of the strange quark mass.

This observation can also be expressed in terms of an empirical formula. To this end, we write the Hamilton operator of QCD symbolically in the form

$$\hat{H}_{\text{QCD}} = \hat{H}_{SU(3)} + \hat{H}_{\text{FSB}} , \quad (5.36)$$

where the first part, $\hat{H}_{SU(3)}$, is assumed to be **invariant** under $(SU(3), \cdot)$ flavor transformations, while the second part, \hat{H}_{FSB} , breaks $(SU(3), \cdot)$ flavor symmetry **explicitly**. If $\hat{H}_{\text{FSB}} = 0$, all hadrons of an $(SU(3), \cdot)$ flavor multiplet must have **identical** masses, since $\hat{H}_{SU(3)}$ can only depend on the Casimir operators \hat{C}_1, \hat{C}_2 of $(SU(3), \cdot)$, which assume the same eigenvalues on any given multiplet. In the following, the eigenvalue of $\hat{H}_{SU(3)}$ on a given multiplet will be denoted by a .

The observed violation of mass degeneracy inside an $(SU(3), \cdot)$ flavor multiplet is proportional to the number of strange quarks or to the **hypercharge** of the respective isospin multiplet. On the other hand, isospin invariance is (approximately) fulfilled on isospin multiplets. Therefore, the symmetry-breaking part \hat{H}_{FSB} must be proportional to \hat{Y} (or higher powers of \hat{Y}), and may depend on the Casimir operator \hat{T}^2 of the isospin group. We thus make the Ansatz

$$\hat{H}_{\text{FSB}} = b\hat{Y} + c\left(\hat{T}^2 - \frac{1}{4}\hat{Y}^2\right) . \quad (5.37)$$

If we are interested in accounting for isospin violation as well, we would have to add a part $\sim \hat{T}_3$. Equations (5.36) and (5.37) imply that the mass of a hadron on a given $(SU(3), \cdot)$ flavor multiplet can be calculated from the formula

$$m = a + bY + c\left[T(T+1) - \frac{1}{4}Y^2\right] , \quad (5.38)$$

where the constants a, b, c have dimension energy (in natural units) and assume different values on different multiplets. Equation (5.38) is the so-called **Gell-Mann–Okubo mass formula**. Let us check its validity at hand of the baryon octet. For the various isospin multiplets inside the baryon octet we have:

$$\begin{aligned} m_N &= a + b + c \left(\frac{3}{4} - \frac{1}{4} \right) = a + b + \frac{c}{2}, \\ m_\Lambda &= a, \\ m_\Sigma &= a + c(2 - 0) = a + 2c, \\ m_\Xi &= a - b + c \left(\frac{3}{4} - \frac{1}{4} \right) = a - b + \frac{c}{2}. \end{aligned}$$

Forming linear combinations, one can eliminate b ,

$$\frac{1}{2}(m_N + m_\Xi) = a + \frac{c}{2} = \frac{3}{4}m_\Lambda + \frac{1}{4}m_\Sigma.$$

This equation is fulfilled to good approximation, because

$$\frac{1}{2}(m_N + m_\Xi) \simeq 1128.60 \text{ MeV} \simeq 1135.05 \text{ MeV} \simeq \frac{3}{4}m_\Lambda + \frac{1}{4}m_\Sigma.$$

The deviation is only 6.5 MeV, i.e., of the order of magnitude of isospin violation, which is not taken into account in the Gell-Mann–Okubo formula (5.38).

For the decuplet we compute

$$\begin{aligned} m_\Delta &= a + b + c \left(\frac{3}{2} \cdot \frac{5}{2} - \frac{1}{4} \right) = a + b + \frac{7}{2}c, \\ m_{\Sigma^*} &= a + c(2 - 0) = a + 2c, \\ m_{\Xi^*} &= a - b + c \left(\frac{3}{4} - \frac{1}{4} \right) = a - b + \frac{c}{2}, \\ m_\Omega &= a - 2b + c(0 - 1) = a - 2b - c. \end{aligned}$$

This leads to the equations

$$\frac{1}{2}(m_\Delta + m_{\Xi^*}) = a + 2c \equiv m_{\Sigma^*} \equiv 2m_{\Xi^*} - m_\Omega,$$

which are also fulfilled to very good approximation:

$$\begin{aligned} \frac{1}{2}(m_\Delta + m_{\Xi^*}) &\simeq 1382.70 \text{ MeV}, \\ m_{\Sigma^*} &\simeq 1384.57 \text{ MeV}, \\ 2m_{\Xi^*} - m_\Omega &\simeq 1394.35 \text{ MeV}. \end{aligned}$$

5.6 Charm and ($SU(4), \cdot$) flavor symmetry

Lecture 9 and 10

The considerations regarding the ($SU(3), \cdot$) flavor symmetry of the strong interaction can also be extended to the charm quark (and in principle also to bottom and top quark). However, because of the large mass of the charm quark,

$$m_c \simeq 1.266 \text{ GeV} \gg m_s \simeq 95 \text{ MeV},$$

5 Unitary symmetries of the strong interaction

the $(SU(4), \cdot)$ flavor symmetry is much stronger explicitly broken as the $(SU(3), \cdot)$ flavor symmetry. The mass difference $\delta m_{cs} = m_c - m_s \simeq 1.161 \text{ GeV} \sim M_h$ is now even of the order of the hadronic mass scale.

The generators of $(SU(4), \cdot)$ are (as direct generalization of the generators of $(SU(3), \cdot)$ and in natural units, $\hbar = 1$)

$$\hat{T}_a = \frac{1}{2} \hat{\lambda}_a, \quad a = 1, \dots, 15,$$

mit

$$\begin{aligned} \hat{\lambda}_1 &= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & \hat{\lambda}_2 &= \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & \hat{\lambda}_3 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ \hat{\lambda}_4 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & \hat{\lambda}_5 &= \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & \hat{\lambda}_6 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \\ \hat{\lambda}_7 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & \hat{\lambda}_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & & (5.39) \\ \hat{\lambda}_9 &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, & \hat{\lambda}_{10} &= \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, & \hat{\lambda}_{11} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \\ \hat{\lambda}_{12} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix}, & \hat{\lambda}_{13} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, & \hat{\lambda}_{14} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}, \end{aligned}$$

as well as

$$\hat{\lambda}_{15} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}. \quad (5.40)$$

The structure constants of $(SU(4), \cdot)$ can again be computed from Eqs. (4.13) and (4.16); we leave this as homework exercise.

The group $(SU(4), \cdot)$ has three Cartan generators,

$$\{\hat{T}_3, \hat{T}_8, \hat{T}_{15}\},$$

and, since it is a semi-simple Lie group, also three Casimir operators,

$$\{\hat{C}_1, \hat{C}_2, \hat{C}_3\}.$$

States of ($SU(4), \cdot$) multiplets are therefore characterized by the eigenvalues of the Casimir operators and Cartan generators,

$$|C_1 C_2 C_3 T_3 T_8 T_{15}\rangle . \quad (5.41)$$

Due to historical reasons, however, one classifies states a little bit differently. To this end we go back one step and consider the group ($SU(3), \cdot$). One defines the **strangeness operator** via the relation

$$\hat{Y} \equiv \hat{B} + \hat{S} , \quad (5.42)$$

where \hat{Y} is the hypercharge operator and \hat{B} the **baryon-number operator**. For mesons one has $B = 0$ and for baryons $B = 1$ (anti-baryons $B = -1$). Quarks carry therefore $B = 1/3$ and anti-quarks $B = -1/3$. The strangeness operator $\hat{S} = \hat{Y} - \hat{B}$ then reads in fundamental representation (and in natural units)

$$\hat{S} = \frac{1}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} .$$

Here we have used the fact that the baryon number in the fundamental representation (i.e., for quarks) is $B = 1/3$ and therefore $\hat{B} = \frac{1}{3} \mathbb{1}_3$. Then we have

$$\begin{aligned} \hat{S} |u\rangle &= \hat{S} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 0 , \\ \hat{S} |d\rangle &= \hat{S} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0 , \\ \hat{S} |s\rangle &= \hat{S} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = -1 |s\rangle . \end{aligned}$$

Up and down quarks thus carry strangeness $S = 0$, while strange quarks carry strangeness $S = -1$. The minus sign is simply historical convention. The strangeness operator measures exclusively the number of strange quarks inside a hadronic state. Therefore, ($SU(3), \cdot$) flavor states can also be characterized by their strangeness quantum number S instead of Y , if we also specify their baryon quantum number B ,

$$|C_1 C_2 T_3 Y\rangle \longrightarrow |C_1 C_2 B T_3 S\rangle .$$

For ($SU(4), \cdot$) one proceeds accordingly. We generalize the relation (5.42) by introducing the so-called **charm operator** \hat{C} ,

$$\hat{Y} = \hat{B} + \hat{S} + \hat{C} , \quad (5.43)$$

where

$$\hat{C} = \frac{3}{4} \hat{B} - \sqrt{\frac{3}{2}} \hat{T}_{15} . \quad (5.44)$$

In fundamental representation this operator reads

$$\hat{C} = \frac{1}{4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} - \frac{1}{4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5.45)$$

The charm operator thus measures the number of charm quarks in a given state. Due to

$$\hat{C} |c\rangle = \hat{C} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = 1 |c\rangle$$

charm quarks carry the charm quantum number $C = +1$. (Note the opposite sign as compared to the strangeness quantum number.)

If one specifies B , instead of T_8 and T_{15} one can also use S and C in Eq. (5.41) in order to classify $(SU(4), \cdot)$ states,

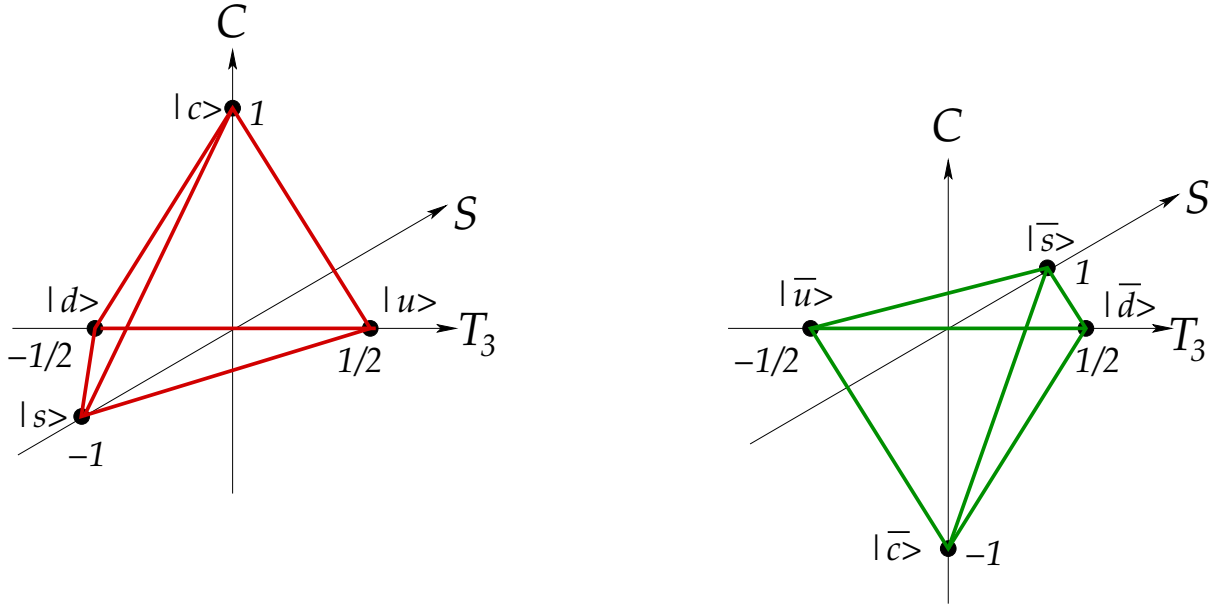
$$|C_1 C_2 C_3 T_3 T_8 T_{15}\rangle \longrightarrow |C_1 C_2 C_3 B T_3 S C\rangle.$$

The Gell-Mann–Nishijima charge formula is still given by Eq. (5.34), if we use Eq. (5.43) for the hypercharge. For the fundamental $(SU(4), \cdot)$ quadruplet all relevant quantum numbers are listed in Tab. 5.2.

	B	T_3	S	C	Y	$Q[e]$
u	$\frac{1}{3}$	$\frac{1}{2}$	0	0	$\frac{1}{3}$	$\frac{2}{3}$
d	$\frac{1}{3}$	$-\frac{1}{2}$	0	0	$\frac{1}{3}$	$-\frac{1}{3}$
s	$\frac{1}{3}$	0	-1	0	$-\frac{2}{3}$	$-\frac{1}{3}$
c	$\frac{1}{3}$	0	0	1	$\frac{4}{3}$	$\frac{2}{3}$

Table 5.2: Quantum numbers of the fundamental $(SU(4), \cdot)$ quadruplet.

$(SU(4), \cdot)$ flavor multiplets can be represented in the **three-dimensional** $(T_3 - S - C)$ space, e.g. the **quadruplet** [4] and the **anti-quadruplet** [$\bar{4}$]:

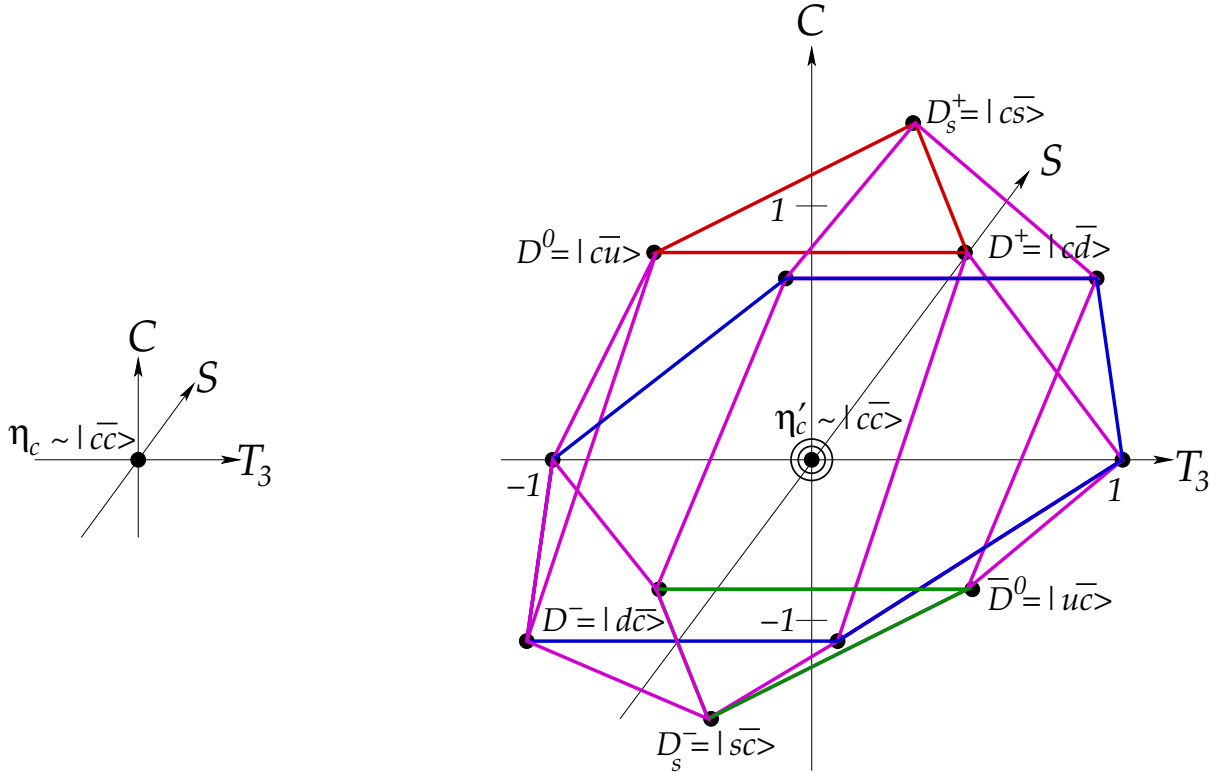


In the $(T_3 - S)$ plane one observes the $(SU(3), \cdot)$ sub-multiplets consisting of up, down, and strange quark (triplet) or of the corresponding anti-quarks (anti-triplet).

Higher multiplets can again be generated by coupling fundamental quadruplets. We list the most important coupling rules (without proof):

$$\begin{aligned}
 [4] \otimes [\bar{4}] &= [1] \oplus [15], \\
 [4] \otimes [4] \otimes [4] &= [\bar{4}] \oplus [20] \oplus [20] \oplus [20'].
 \end{aligned}$$

Accordingly, one finds mesons in $(SU(4), \cdot)$ singlets or octets and baryons in $(SU(4), \cdot)$ anti-quadruplets as well as three 20-plets. For the pseudoscalar mesons this is exemplified in the following figure:



In the 15-plet one recognizes one octet (blue) at $C = 0$. This is the pseudoscalar-meson octet we have already encountered when discussing the $(SU(3), \cdot)$ flavor symmetry. The individual states are no longer explicitly labelled. New states are an $(SU(3), \cdot)$ triplet (green) at $C = -1$ and an $(SU(3), \cdot)$ anti-triplet (red) at $C = 1$. The corresponding states are

$$\begin{aligned}
 D^\pm & : I(J^{PC}) = \frac{1}{2}(0^{-+}) , \quad m_{D^\pm} = 1869.5 \pm 0.4 \text{ MeV} , \\
 D^0 & : I(J^{PC}) = \frac{1}{2}(0^{-+}) , \quad m_{D^0} = 1864.91 \pm 0.17 \text{ MeV} , \\
 D_s^\pm & : I(J^{PC}) = 0(0^{-+}) , \quad m_{D_s^\pm} = 1969.0 \pm 1.4 \text{ MeV} , \\
 \eta_c & : I(J^{PC}) = 0(0^{-+}) , \quad m_{\eta_c} = 2981.0 \pm 1.1 \text{ MeV} , \\
 \eta_c' & : I(J^{PC}) = 0(0^{-+}) , \quad m_{\eta_c'} = 3638.9 \pm 1.3 \text{ MeV} .
 \end{aligned}$$

The vector meson corresponding to the $(SU(4), \cdot)$ singlet η_c , the so-called J/Ψ meson ($m_{J/\Psi} = 3096.916 \pm 0.011 \text{ MeV}$) was the first particle carrying charm which was experimentally detected, in the year 1974 by Samuel Ting at the Alternating Gradient Synchrotron (AGS) of Brookhaven National Laboratory (BNL) on Long Island. For this achievement he received (together with Burton Richter of Stanford Linear Accelerator Center, SLAC) the Nobel prize 1976.

6 The Poincaré group

In Chapter 2 we had considered **rotations in space**. Together with the so-called **Lorentz boosts**, i.e., transformations between inertial frames encountered in special relativity, they form the so-called **Lorentz group** \mathbb{L} (or more precisely, the so-called proper orthochronous Lorentz group \mathbb{L}_+^\uparrow , since space-reflection and time-reversal transformations will not be considered). Moreover, we had considered **translations in space and time** in Chapter 2. Taking these together with the Lorentz group, we obtain the so-called **Poincaré group** \mathbb{P} .

6.1 Characterization of the Poincaré group

Let us consider a **Poincaré transformation** of the space-time position vector,

$$x^\mu \longrightarrow x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu, \quad (6.1)$$

i.e., the **Lorentz transformation**

$$x^\mu \longrightarrow \bar{x}^\mu = \Lambda^\mu{}_\nu x^\nu \quad (6.2)$$

of the position vector, followed by a **space-time translation** by the constant 4-vector a^μ ,

$$\bar{x}^\mu \longrightarrow x'^\mu = \bar{x}^\mu + a^\mu. \quad (6.3)$$

Remark: Of course, one can also first perform the space-time translation,

$$x^\mu \longrightarrow \hat{x}^\mu = x^\mu + \hat{a}^\mu, \quad (6.4)$$

and then the Lorentz transformation,

$$\hat{x}^\mu \longrightarrow x'^\mu = \Lambda^\mu{}_\nu \hat{x}^\nu = \Lambda^\mu{}_\nu x^\nu + \Lambda^\mu{}_\nu \hat{a}^\nu. \quad (6.5)$$

If we denote $\Lambda^\mu{}_\nu \hat{a}^\nu \equiv a^\mu$, we again obtain Eq. (6.1).

Two successive Poincaré transformations can therefore be written as follows:

$$\begin{aligned} x^\mu \longrightarrow x''^\mu &= \bar{\Lambda}^\mu{}_\nu x'^\nu + \bar{a}^\mu \\ &= \bar{\Lambda}^\mu{}_\nu (\Lambda^\nu{}_\rho x^\rho + a^\nu) + \bar{a}^\mu \\ &= \bar{\Lambda}^\mu{}_\nu \Lambda^\nu{}_\rho x^\rho + \bar{\Lambda}^\mu{}_\nu a^\nu + \bar{a}^\mu. \end{aligned} \quad (6.6)$$

Two successive Lorentz transformations yield (because of the closure of the Lorentz group) again a Lorentz transformation, i.e.,

$$\bar{\Lambda}^\mu{}_\nu \Lambda^\nu{}_\rho \equiv \hat{\Lambda}^\mu{}_\rho \in \mathbb{L}. \quad (6.7)$$

6 The Poincaré group

If we denote

$$\hat{a}^\mu \equiv \bar{\Lambda}^\mu{}_\nu a^\nu + \bar{a}^\mu, \quad (6.8)$$

we realize that two successive Poincaré transformations, Eq. (6.6), again yield a Poincaré transformation,

$$x^\mu \longrightarrow x''^\mu = \hat{\Lambda}^\mu{}_\rho x^\rho + \hat{a}^\mu. \quad (6.9)$$

This shows the **closure** of the Poincaré group. If we consider the transformation of an object in an **arbitrary representation of the group** instead of the transformation of a 4-vector x^μ , and the corresponding representation of an element of the group, $U(\Lambda, a) \in D(\mathbb{P})$, then Eq. (6.6) reads as follows:

$$U(\bar{\Lambda}, \bar{a})U(\Lambda, a) = U(\bar{\Lambda}\Lambda, \bar{\Lambda}a + \bar{a}) \equiv U(\hat{\Lambda}, \hat{a}). \quad (6.10)$$

The **identity element** of the Poincaré group has the form $U(1, 0)$, so that

$$U(1, 0)U(\Lambda, a) = U(\Lambda, a)U(1, 0) = U(\Lambda, a). \quad (6.11)$$

The **inverse element** is

$$U^{-1}(\Lambda, a) \equiv U(\Lambda^{-1}, -\Lambda^{-1}a) \in D(\mathbb{P}), \quad (6.12)$$

since according to Eq. (6.10) we have

$$U^{-1}(\Lambda, a)U(\Lambda, a) = U(\Lambda^{-1}, -\Lambda^{-1}a)U(\Lambda, a) = U(\Lambda^{-1}\Lambda, \Lambda^{-1}a - \Lambda^{-1}a) \equiv U(1, 0). \quad (6.13)$$

With Eq. (6.10) we compute on the one hand

$$U(\hat{\Lambda}, \hat{a})U(\bar{\Lambda}, \bar{a})U(\Lambda, a) = U(\hat{\Lambda}, \hat{a})U(\bar{\Lambda}\Lambda, \bar{\Lambda}a + \bar{a}) = U(\hat{\Lambda}\bar{\Lambda}\Lambda, \hat{\Lambda}\bar{\Lambda}a + \hat{\Lambda}\bar{a} + \hat{a}), \quad (6.14)$$

and on the other hand

$$U(\hat{\Lambda}, \hat{a})U(\bar{\Lambda}, \bar{a})U(\Lambda, a) = U(\hat{\Lambda}\bar{\Lambda}, \hat{\Lambda}\bar{a} + \hat{a})U(\Lambda, a) = U(\hat{\Lambda}\bar{\Lambda}\Lambda, \hat{\Lambda}\bar{\Lambda}a + \hat{\Lambda}\bar{a} + \hat{a}). \quad (6.15)$$

Comparison of Eq. (6.14) with (6.15) proves the **law of associativity** for the Poincaré group. With Eqs. (6.10), (6.11), (6.12), (6.14), and (6.15) we have proven that the Poincaré group fulfills the **group axioms**.

6.2 Generators

The Poincaré group is a **Lie group**, i.e., each element $U(\Lambda, a)$ has a representation as a linear operator,

$$\hat{U}(\vec{\alpha}) = \exp \left(-i \sum_{j=1}^N \alpha_j \hat{X}_j \right), \quad (6.16)$$

cf. Eq. (3.10) (using natural units $\hbar = c = 1$). How many generators (or parameters, respectively) are there? **Four** generators enact translations in the three spatial directions

and in time. We know them already from the discussion in Chapter 2; they can be compactly written as a 4-vector:

$$\hat{P}^\mu = i\partial^\mu = i \left(\frac{\partial}{\partial t}, -\vec{\nabla} \right)^T, \quad (6.17)$$

where the right-hand side of the equation corresponds to the **representation** of \hat{P}^μ as linear differential operator. On the other hand, the Lorentz group has **six** generators, **three** for rotations around the three spatial axes and **three** for boosts in each spatial direction. Thus, the Poincaré group has **ten generators**.

The four generators (6.17) of space-time translations we have already encountered. But what are the six generators of the Lorentz group? We first write these generators compactly in the form of an antisymmetric Lorentz tensor of rank 2, $\hat{J}^{\mu\nu} = -\hat{J}^{\nu\mu}$. The relationship to the generators of rotations (which of course have to correspond to the angular momentum operator) and to that of the boosts will be discussed later. One readily convinces oneself of the fact that due to the antisymmetry of $\hat{J}^{\mu\nu}$, only six of the 16 possible components are independent. Then we can write Eq. (6.16) in the form

$$\hat{U}(\omega, a) = \exp \left(-\frac{i}{2} \omega_{\mu\nu} \hat{J}^{\mu\nu} + i a_\mu \hat{P}^\mu \right). \quad (6.18)$$

The four parameters of space-time translations are summarized in the constant 4-vector a^μ . The choice of sign in front of the translation part corresponds to the convention used in Chapter 2. The six parameters of Lorentz transformations we have written as components of a rank-2 tensor $\omega^{\mu\nu}$. This tensor can also be chosen to be antisymmetric, $\omega^{\mu\nu} = -\omega^{\nu\mu}$, since any possibly existing symmetric part would vanish anyway after complete contraction with the antisymmetric tensor $\hat{J}^{\mu\nu}$, thus does not need to be considered. The six parameters are on the one hand the three angles for rotations around the x , y , and z directions, and on the other hand the three rapidities for boosts along these spatial directions.

In order to see this, let us look at the explicit form of the Lorentz-transformation matrix. A **spatial rotation**, e.g. around the angle ϕ around the z axis, is enacted by

$$\Lambda_{D\nu}^\mu(0, 0, \phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \phi & \sin \phi & 0 \\ 0 & -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (6.19)$$

For an infinitesimal angle $\phi \ll 1$ this reads

$$\Lambda_{D\nu}^\mu(0, 0, \phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \phi & 0 \\ 0 & -\phi & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \equiv g^\mu{}_\nu + \phi (g^{\mu 2} g^1{}_\nu - g^{\mu 1} g^2{}_\nu). \quad (6.20)$$

A **boost**, e.g. with rapidity χ in z direction, is enacted by

$$\Lambda_{B\nu}^\mu(0, 0, \chi) = \begin{pmatrix} \cosh \chi & 0 & 0 & -\sinh \chi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh \chi & 0 & 0 & \cosh \chi \end{pmatrix}. \quad (6.21)$$

For infinitesimal rapidities $\chi \ll 1$ this reads

$$\Lambda_{B\nu}^\mu(0, 0, \chi) = \begin{pmatrix} 1 & 0 & 0 & -\chi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\chi & 0 & 0 & 1 \end{pmatrix} = g^\mu{}_\nu + \chi (g^{\mu 3} g^0{}_\nu - g^{\mu 0} g^3{}_\nu) . \quad (6.22)$$

For arbitrary infinitesimal Lorentz transformations, Eqs. (6.20) and (6.22) can be generalized as

$$\Lambda^\mu{}_\nu = g^\mu{}_\nu + \omega^\mu{}_\nu , \quad (6.23)$$

with the antisymmetric tensor $\omega^\mu{}_\nu$ of the parameters of Lorentz transformations.

Let us remark that the representation (6.18) of an element of the Poincaré group as linear operator is not yet completely determined. It depends on which **irreducible representation of the group** this linear operator is supposed to act. This then influences the form of the generators.

6.3 The Poincaré algebra

We now determine the **Poincaré algebra** fulfilled by the generators $\hat{J}^{\mu\nu}$ and \hat{P}^μ . To this end we first consider how the generators behave under Poincaré transformations. An arbitrary infinitesimal Poincaré transformation can be written as $U(1 + \omega, \epsilon)$, where on account of Eq. (6.23) we have written an infinitesimal Lorentz transformation compactly as $\Lambda \equiv 1 + \omega$, with the tensor $\omega^{\mu\nu}$ of the infinitesimal parameters of rotations and boosts, and we have introduced an infinitesimal space-time translation vector ϵ^μ . We now compute using Eq. (6.10)

$$\begin{aligned} U(\Lambda, a)U(1 + \omega, \epsilon)U^{-1}(\Lambda, a) &= U(\Lambda(1 + \omega), \Lambda\epsilon + a)U(\Lambda^{-1}, -\Lambda^{-1}a) \\ &= U(\Lambda(1 + \omega)\Lambda^{-1}, -\Lambda(1 + \omega)\Lambda^{-1}a + \Lambda\epsilon + a) \\ &= U(1 + \Lambda\omega\Lambda^{-1}, \Lambda\epsilon - \Lambda\omega\Lambda^{-1}a) . \end{aligned} \quad (6.24)$$

Using Eq. (6.18) and expanding the left-hand and right-hand sides to leading order in the infinitesimal parameters,

$$\begin{aligned} U(\Lambda, a) \left(1 - \frac{i}{2} \omega_{\mu\nu} \hat{J}^{\mu\nu} + i \epsilon_\mu \hat{P}^\mu \right) U^{-1}(\Lambda, a) & \quad (6.25) \\ &= 1 - \frac{i}{2} \omega_{\mu\nu} U(\Lambda, a) \hat{J}^{\mu\nu} U^{-1}(\Lambda, a) + i \epsilon_\mu U(\Lambda, a) \hat{P}^\mu U^{-1}(\Lambda, a) \\ &= 1 - \frac{i}{2} \Lambda_{\mu\rho} \omega^\rho{}_\sigma (\Lambda^{-1})^\sigma{}_\nu \hat{J}^{\mu\nu} + i (\Lambda_{\mu\rho} \epsilon^\rho - \Lambda_{\mu\rho} \omega^\rho{}_\sigma (\Lambda^{-1})^\sigma{}_\nu a^\nu) \hat{P}^\mu \\ &= 1 - \frac{i}{2} \left(\Lambda_{\mu\rho} \omega^\rho{}_\sigma \Lambda_\nu{}^\sigma \hat{J}^{\mu\nu} + \Lambda_{\mu\rho} \omega^\rho{}_\sigma \Lambda_\nu{}^\sigma a^\nu \hat{P}^\mu - \Lambda_{\mu\rho} \omega_\sigma{}^\rho \Lambda_\nu{}^\sigma a^\nu \hat{P}^\mu \right) + i \Lambda_{\mu\rho} \epsilon^\rho \hat{P}^\mu , \end{aligned}$$

where we have employed the property $\Lambda^{-1} = \Lambda^T$ of the Lorentz-transformation matrices and the antisymmetry of $\omega^{\mu\nu}$. We now relabel the indices $\mu \leftrightarrow \nu$ and $\rho \leftrightarrow \sigma$ in the last

term in parentheses and obtain after subtracting one and multiplying both sides with $2i$

$$U(\Lambda, a) \left(\omega_{\mu\nu} \hat{J}^{\mu\nu} - 2\epsilon_\mu \hat{P}^\mu \right) U^{-1}(\Lambda, a) = \Lambda_{\mu\rho} \omega^\rho{}_\sigma \Lambda_\nu{}^\sigma \left(\hat{J}^{\mu\nu} + a^\nu \hat{P}^\mu - a^\mu \hat{P}^\nu \right) - 2 \Lambda_{\mu\rho} \epsilon^\rho \hat{P}^\mu . \quad (6.26)$$

A further relabelling of indices on the right-hand side ($\mu \leftrightarrow \rho$, $\nu \leftrightarrow \sigma$) leads to

$$U(\Lambda, a) \left(\omega_{\mu\nu} \hat{J}^{\mu\nu} - 2\epsilon_\mu \hat{P}^\mu \right) U^{-1}(\Lambda, a) = \Lambda_{\rho\mu} \omega^\mu{}_\nu \Lambda_\sigma{}^\nu \left(\hat{J}^{\rho\sigma} + a^\sigma \hat{P}^\rho - a^\rho \hat{P}^\sigma \right) - 2 \Lambda_{\rho\mu} \epsilon^\mu \hat{P}^\rho . \quad (6.27)$$

The parameters $\omega_{\mu\nu}$ and ϵ_μ can be chosen arbitrarily, so that one can compare their coefficients on the left- and right-hand side to obtain the transformation behavior of the generators:

$$U(\Lambda, a) \hat{J}^{\mu\nu} U^{-1}(\Lambda, a) = \Lambda_\rho{}^\mu \Lambda_\sigma{}^\nu \left(\hat{J}^{\rho\sigma} + a^\sigma \hat{P}^\rho - a^\rho \hat{P}^\sigma \right) , \quad (6.28)$$

$$U(\Lambda, a) \hat{P}^\mu U^{-1}(\Lambda, a) = \Lambda_\rho{}^\mu \hat{P}^\rho . \quad (6.29)$$

We now consider the infinitesimal form of Eqs. (6.28) and (6.29), i.e., for $\Lambda \equiv 1 + \omega$ and $a \equiv \epsilon$. To first order in infinitesimal parameters we obtain from Eq. (6.28)

$$\begin{aligned} & \left(1 - \frac{i}{2} \omega_{\rho\sigma} \hat{J}^{\rho\sigma} + i\epsilon_\rho \hat{P}^\rho \right) \hat{J}^{\mu\nu} \left(1 + \frac{i}{2} \omega_{\alpha\beta} \hat{J}^{\alpha\beta} - i\epsilon_\alpha \hat{P}^\alpha \right) \\ &= \hat{J}^{\mu\nu} - \frac{i}{2} \omega_{\rho\sigma} \left[\hat{J}^{\rho\sigma}, \hat{J}^{\mu\nu} \right] + i\epsilon_\rho \left[\hat{P}^\rho, \hat{J}^{\mu\nu} \right] \\ &= (g_\rho{}^\mu + \omega_\rho{}^\mu) (g_\sigma{}^\nu + \omega_\sigma{}^\nu) \left(\hat{J}^{\rho\sigma} + \epsilon^\sigma \hat{P}^\rho - \epsilon^\rho \hat{P}^\sigma \right) \\ &= \hat{J}^{\mu\nu} + \epsilon^\nu \hat{P}^\mu - \epsilon^\mu \hat{P}^\nu + \omega_\rho{}^\mu \hat{J}^{\rho\nu} + \omega_\sigma{}^\nu \hat{J}^{\mu\sigma} \\ &= \hat{J}^{\mu\nu} + \epsilon_\rho \left(g^{\rho\nu} \hat{P}^\mu - g^{\rho\mu} \hat{P}^\nu \right) + \omega_{\rho\sigma} \left(g^{\sigma\mu} \hat{J}^{\rho\nu} - g^{\rho\nu} \hat{J}^{\mu\sigma} \right) \\ &= \hat{J}^{\mu\nu} + \epsilon_\rho \left(g^{\rho\nu} \hat{P}^\mu - g^{\rho\mu} \hat{P}^\nu \right) + \frac{1}{2} \omega_{\rho\sigma} \left(g^{\sigma\mu} \hat{J}^{\rho\nu} - g^{\rho\mu} \hat{J}^{\sigma\nu} - g^{\rho\nu} \hat{J}^{\mu\sigma} + g^{\sigma\nu} \hat{J}^{\mu\rho} \right) , \end{aligned} \quad (6.30)$$

where in the last step we have doubled the last term, accounting for that by an additional factor of $1/2$, relabelled the indices $\rho \leftrightarrow \sigma$, and finally used the antisymmetry of $\omega_{\rho\sigma}$. Comparing the coefficients of $\omega_{\rho\sigma}$ and ϵ_ρ in the second and sixth line, we obtain the commutation relations

$$\left[\hat{J}^{\rho\sigma}, \hat{J}^{\mu\nu} \right] = i \left(g^{\sigma\mu} \hat{J}^{\rho\nu} - g^{\rho\mu} \hat{J}^{\sigma\nu} - g^{\rho\nu} \hat{J}^{\mu\sigma} + g^{\sigma\nu} \hat{J}^{\mu\rho} \right) , \quad (6.31)$$

$$\left[\hat{P}^\rho, \hat{J}^{\mu\nu} \right] = i \left(g^{\rho\mu} \hat{P}^\nu - g^{\rho\nu} \hat{P}^\mu \right) . \quad (6.32)$$

We perform a completely analogous calculation for Eq. (6.29):

$$\begin{aligned} & \left(1 - \frac{i}{2} \omega_{\rho\sigma} \hat{J}^{\rho\sigma} + i\epsilon_\rho \hat{P}^\rho \right) \hat{P}^\mu \left(1 + \frac{i}{2} \omega_{\alpha\beta} \hat{J}^{\alpha\beta} - i\epsilon_\alpha \hat{P}^\alpha \right) \\ &= \hat{P}^\mu - \frac{i}{2} \omega_{\rho\sigma} \left[\hat{J}^{\rho\sigma}, \hat{P}^\mu \right] + i\epsilon_\rho \left[\hat{P}^\rho, \hat{P}^\mu \right] \\ &= (g_\rho{}^\mu + \omega_\rho{}^\mu) \hat{P}^\rho \\ &= \hat{P}^\mu + \frac{1}{2} \omega_{\rho\sigma} \left(g^{\sigma\mu} \hat{P}^\rho - g^{\rho\mu} \hat{P}^\sigma \right) . \end{aligned} \quad (6.33)$$

Comparison of coefficients in the second and fourth line leads on the hand again to Eq. (6.32) (after a suitable relabelling of indices), and on the other hand to

$$\left[\hat{P}^\rho, \hat{P}^\mu \right] = 0 . \quad (6.34)$$

Equations (6.31), (6.32), and (6.34) represent the **Poincaré algebra**. From Eqs. (6.32) and (6.34) follows that there is an (abelian) ideal, thus the Poincaré group is **not a (semi-)simple group**. Since the group of space-time translations is not compact, the Poincaré is also **not a compact group**.

If we identify

$$\hat{P}^0 \equiv \hat{H} , \quad (6.35)$$

$$\hat{\vec{J}} \equiv \left(\hat{J}^{23}, \hat{J}^{31}, \hat{J}^{12} \right)^T , \quad (6.36)$$

$$\hat{\vec{K}} \equiv \left(\hat{J}^{01}, \hat{J}^{02}, \hat{J}^{03} \right)^T , \quad (6.37)$$

where $\hat{\vec{J}}$ is the **operator of total angular momentum** (i.e., the sum of orbital angular momentum and spin) and $\hat{\vec{K}}$ is the **boost operator**, the Poincaré algebra (6.31), (6.32), and (6.34) can be brought into the following form:

$$\begin{aligned} \left[\hat{J}^i, \hat{J}^j \right] &= i\epsilon^{ijk} \hat{J}^k , & \left[\hat{J}^i, \hat{K}^j \right] &= i\epsilon^{ijk} \hat{K}^k , \\ \left[\hat{K}^i, \hat{K}^j \right] &= -i\epsilon^{ijk} \hat{J}^k , & \left[\hat{J}^i, \hat{P}^j \right] &= i\epsilon^{ijk} \hat{P}^k , \\ \left[\hat{K}^i, \hat{P}^j \right] &= -i\hat{H}\delta^{ij} & \left[\hat{K}^i, \hat{H} \right] &= -i\hat{P}^i , \\ \left[\hat{J}^i, \hat{H} \right] &= \left[\hat{P}^i, \hat{H} \right] = 0 . \end{aligned} \quad (6.38)$$

We leave the proof as homework exercise. The first commutation relation is, of course, obvious, since it represents the angular-momentum algebra. This holds for orbital angular momentum as well as for spin, and thus also for total angular momentum.

6.4 Casimir operators

We define the so-called **Pauli-Lubanski pseudovector**

$$\hat{W}_\mu \equiv \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} \hat{J}^{\nu\rho} \hat{P}^\sigma . \quad (6.39)$$

Here, $\epsilon_{\mu\nu\rho\sigma}$ is the completely antisymmetric tensor of rank 4, with $\epsilon^{0123} = +1$. It fulfills the contraction formula

$$\epsilon^{\mu\nu\alpha\beta} \epsilon_{\mu\tau\rho\sigma} = g^\nu_\tau (g^\alpha_\rho g^\beta_\sigma - g^\alpha_\sigma g^\beta_\rho) + g^\nu_\rho (g^\alpha_\tau g^\beta_\sigma - g^\alpha_\sigma g^\beta_\tau) + g^\nu_\sigma (g^\alpha_\rho g^\beta_\tau - g^\alpha_\tau g^\beta_\rho) . \quad (6.40)$$

A further contraction over the second index yields

$$\begin{aligned} \epsilon^{\mu\nu\alpha\beta} \epsilon_{\mu\nu\rho\sigma} &= 4(g^\alpha_\sigma g^\beta_\rho - g^\alpha_\rho g^\beta_\sigma) + g^\alpha_\rho g^\beta_\sigma - g^\alpha_\sigma g^\beta_\rho + g^\alpha_\rho g^\beta_\sigma - g^\alpha_\sigma g^\beta_\rho \\ &= 2(g^\alpha_\sigma g^\beta_\rho - g^\alpha_\rho g^\beta_\sigma) . \end{aligned} \quad (6.41)$$

The Pauli-Lubanski pseudovector is **orthogonal** to the 4-momentum vector. Namely, using Eq. (6.34), we obtain

$$\begin{aligned}\hat{W}_\mu \hat{P}^\mu &= \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} \hat{J}^{\nu\rho} \hat{P}^\sigma \hat{P}^\mu = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} \hat{J}^{\nu\rho} \hat{P}^\mu \hat{P}^\sigma = \frac{1}{2} \epsilon_{\sigma\nu\rho\mu} \hat{J}^{\nu\rho} \hat{P}^\sigma \hat{P}^\mu \\ &= -\frac{1}{2} \epsilon_{\mu\nu\rho\sigma} \hat{J}^{\nu\rho} \hat{P}^\sigma \hat{P}^\mu = -\hat{W}_\mu \hat{P}^\mu \equiv 0,\end{aligned}\quad (6.42)$$

where in the last step of the first line we relabelled indices $\sigma \leftrightarrow \mu$ and to the second line we have used the antisymmetry of the tensor $\epsilon_{\mu\nu\rho\sigma}$.

Proposition: The following two operators are the **two Casimir operators** of the Poincaré group:

$$\hat{P}^2 = \hat{P}_\mu \hat{P}^\mu \quad \text{und} \quad \hat{W}^2 = \hat{W}_\mu \hat{W}^\mu. \quad (6.43)$$

Remark: According to Eqs. (6.31), (6.32), and (6.34) the Poincaré algebra contains an **abelian ideal**, consisting of the generators of space-time translations. Thus, it is **not a semi-simple group**, therefore Racah's theorem does not apply to draw a conclusion about the number of generators from the rank of the group. However, it is in fact true that the two operators given in Eq. (6.43) are the only Casimir operators of the Poincaré group.

Proof: We first show that \hat{P}^2 commutes with all generators of the group. Using the Poincaré algebra (6.31), (6.32), and (6.34) we obtain

$$\begin{aligned}[\hat{P}^2, \hat{P}^\rho] &= g_{\mu\nu} [\hat{P}^\mu \hat{P}^\nu, \hat{P}^\rho] = g_{\mu\nu} \left\{ \hat{P}^\mu [\hat{P}^\nu, \hat{P}^\rho] + [\hat{P}^\mu, \hat{P}^\rho] \hat{P}^\nu \right\} = 0, \\ [\hat{P}^2, \hat{J}^{\rho\sigma}] &= g_{\mu\nu} [\hat{P}^\mu \hat{P}^\nu, \hat{J}^{\rho\sigma}] = g_{\mu\nu} \left\{ \hat{P}^\mu [\hat{P}^\nu, \hat{J}^{\rho\sigma}] + [\hat{P}^\mu, \hat{J}^{\rho\sigma}] \hat{P}^\nu \right\} \\ &= ig_{\mu\nu} \left\{ \hat{P}^\mu (g^{\nu\rho} \hat{P}^\sigma - g^{\nu\sigma} \hat{P}^\rho) + (g^{\mu\rho} \hat{P}^\sigma - g^{\mu\sigma} \hat{P}^\rho) \hat{P}^\nu \right\} \\ &= i (\hat{P}^\rho \hat{P}^\sigma - \hat{P}^\sigma \hat{P}^\rho + \hat{P}^\sigma \hat{P}^\rho - \hat{P}^\rho \hat{P}^\sigma) = 0.\end{aligned}\quad (6.44)$$

Now we show that \hat{W}^2 commutes with all generators of the group. First, on account of the commutation relations (6.32) and (6.34) we conclude that

$$\begin{aligned}[\hat{W}^\alpha, \hat{P}^\mu] &= \frac{1}{2} \epsilon^\alpha{}_{\nu\rho\sigma} [\hat{J}^{\nu\rho} \hat{P}^\sigma, \hat{P}^\mu] \\ &= \frac{1}{2} \epsilon^\alpha{}_{\nu\rho\sigma} \left\{ \hat{J}^{\nu\rho} [\hat{P}^\sigma, \hat{P}^\mu] + [\hat{J}^{\nu\rho}, \hat{P}^\mu] \hat{P}^\sigma \right\} \\ &= -\frac{i}{2} \epsilon^\alpha{}_{\nu\rho\sigma} (g^{\mu\nu} \hat{P}^\rho - g^{\mu\rho} \hat{P}^\nu) \hat{P}^\sigma \\ &= -\frac{i}{2} (\epsilon^{\alpha\mu\nu\sigma} - \epsilon^{\alpha\nu\mu\sigma}) \hat{P}_\nu \hat{P}_\sigma = -\frac{i}{2} \epsilon^{\alpha\mu\nu\sigma} [\hat{P}_\nu, \hat{P}_\sigma] \equiv 0,\end{aligned}\quad (6.45)$$

where to the last line we have relabelled the indices $\rho \leftrightarrow \nu$ in the first term and then in the next-to-last step used the antisymmetry of $\epsilon^{\mu\alpha\nu\sigma}$.

Furthermore, we require

$$[\hat{W}^\alpha, \hat{J}^{\mu\nu}] = i (g^{\alpha\mu} \hat{W}^\nu - g^{\alpha\nu} \hat{W}^\mu). \quad (6.46)$$

6 The Poincaré group

The proof of this relation utilizes the definition of a new tensor,

$$\tilde{J}_{\mu\nu} \equiv \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} \hat{J}^{\rho\sigma} . \quad (6.47)$$

Contracting from the left with $\epsilon^{\alpha\beta\mu\nu}$ and multiplying by $-1/2$ we obtain using Eq. (6.41) and the antisymmetry of $\hat{J}^{\rho\sigma}$

$$-\frac{1}{2} \epsilon^{\alpha\beta\mu\nu} \tilde{J}_{\mu\nu} = -\frac{1}{4} \epsilon^{\alpha\beta\mu\nu} \epsilon_{\mu\nu\rho\sigma} \hat{J}^{\rho\sigma} = \frac{1}{2} (g^\alpha_\rho g^\beta_\sigma - g^\alpha_\sigma g^\beta_\rho) \hat{J}^{\rho\sigma} \equiv \hat{J}^{\alpha\beta} . \quad (6.48)$$

In order to prove Eq. (6.46), we need the following commutation relation:

$$\begin{aligned} [\tilde{J}^{\rho\sigma}, \hat{J}^{\mu\nu}] &= \frac{1}{2} \epsilon^{\rho\sigma}{}_{\alpha\beta} [\hat{J}^{\alpha\beta}, \hat{J}^{\mu\nu}] \\ &= \frac{i}{2} \epsilon^{\rho\sigma}{}_{\alpha\beta} (g^{\beta\mu} \hat{J}^{\alpha\nu} - g^{\alpha\mu} \hat{J}^{\beta\nu} - g^{\alpha\nu} \hat{J}^{\mu\beta} + g^{\beta\nu} \hat{J}^{\mu\alpha}) \\ &= \frac{i}{2} (\epsilon^{\rho\sigma\alpha\mu} \hat{J}_\alpha{}^\nu - \epsilon^{\rho\sigma\mu\beta} \hat{J}_\beta{}^\nu - \epsilon^{\rho\sigma\nu\beta} \hat{J}_\beta{}^\mu + \epsilon^{\rho\sigma\alpha\nu} \hat{J}_\alpha{}^\mu) \\ &= -i (\epsilon^{\rho\sigma\mu\alpha} \hat{J}_\alpha{}^\nu - \epsilon^{\rho\sigma\nu\alpha} \hat{J}_\alpha{}^\mu) . \end{aligned} \quad (6.49)$$

Here we have used Eq. (6.31) for the second step and relabelled the indices in the second and third term ($\beta \rightarrow \alpha$) for the fourth step. Furthermore, we exploited the antisymmetry of $\epsilon_{\alpha\beta\mu\rho}$ and $\hat{J}^{\mu\nu}$. Inserting Eq. (6.48), we obtain using Eq. (6.40)

$$\begin{aligned} [\tilde{J}^{\rho\sigma}, \hat{J}^{\mu\nu}] &= \frac{i}{2} (\epsilon^{\rho\sigma\mu\alpha} \epsilon_\alpha{}^{\nu\beta\gamma} - \epsilon^{\rho\sigma\nu\alpha} \epsilon_\alpha{}^{\mu\beta\gamma}) \tilde{J}_{\beta\gamma} \\ &= -\frac{i}{2} [g^{\mu\nu} (g^{\rho\gamma} g^{\sigma\beta} - g^{\rho\beta} g^{\sigma\gamma}) + g^{\mu\beta} (g^{\rho\nu} g^{\sigma\gamma} - g^{\rho\gamma} g^{\sigma\nu}) + g^{\mu\gamma} (g^{\rho\beta} g^{\sigma\nu} - g^{\rho\nu} g^{\sigma\beta}) \\ &\quad - (\mu \leftrightarrow \nu)] \tilde{J}_{\beta\gamma} \\ &= -i (g^{\rho\nu} \tilde{J}^{\mu\sigma} - g^{\rho\mu} \tilde{J}^{\nu\sigma} - g^{\sigma\nu} \tilde{J}^{\mu\rho} + g^{\sigma\mu} \tilde{J}^{\nu\rho}) \\ &= i (g^{\sigma\mu} \tilde{J}^{\rho\nu} - g^{\rho\mu} \tilde{J}^{\sigma\nu} - g^{\rho\nu} \tilde{J}^{\mu\sigma} + g^{\sigma\nu} \tilde{J}^{\mu\rho}) , \end{aligned} \quad (6.50)$$

where to the next-to-last step we have used the antisymmetry of $\tilde{J}^{\mu\nu}$ and to the last step reordered the terms somewhat (again using the antisymmetry of $\tilde{J}^{\mu\nu}$). Comparison with Eq. (6.31), reveals that $\tilde{J}^{\rho\sigma}$ fulfills **the same** commutation relation with $\hat{J}^{\mu\nu}$ as $\hat{J}^{\rho\sigma}$ with itself.

An important property of the tensor $\tilde{J}^{\mu\nu}$ is

$$\begin{aligned} \hat{P}^\nu \tilde{J}_{\mu\nu} &= \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} \hat{P}^\nu \hat{J}^{\rho\sigma} = \frac{1}{2} \epsilon_{\mu\rho\sigma\nu} \hat{P}^\nu \hat{J}^{\rho\sigma} \\ &= \frac{1}{2} \epsilon_{\mu\rho\sigma\nu} \left\{ \hat{J}^{\rho\sigma} \hat{P}^\nu + [\hat{P}^\nu, \hat{J}^{\rho\sigma}] \right\} \\ &= \hat{W}_\mu + \frac{i}{2} \epsilon_{\mu\rho\sigma\nu} (g^{\nu\rho} \hat{P}^\sigma - g^{\nu\sigma} \hat{P}^\rho) \equiv \hat{W}_\mu . \end{aligned} \quad (6.51)$$

Here we have made use of the definition (6.39) of the Pauli-Lubanski pseudovector, of the commutation relation (6.32), as well as of the antisymmetry of $\epsilon_{\mu\rho\sigma\nu}$.

Using the relations (6.50) and (6.51), as well as Eq. (6.32), we prove Eq. (6.46):

$$\begin{aligned}
[\hat{W}^\alpha, \hat{J}^{\mu\nu}] &= [\hat{P}_\beta \tilde{J}^{\alpha\beta}, \hat{J}^{\mu\nu}] = \hat{P}_\beta [\tilde{J}^{\alpha\beta}, \hat{J}^{\mu\nu}] + g_{\beta\rho} [\hat{P}^\rho, \hat{J}^{\mu\nu}] \tilde{J}^{\alpha\beta} \\
&= i\hat{P}_\beta \left(g^{\beta\mu} \tilde{J}^{\alpha\nu} - g^{\alpha\mu} \tilde{J}^{\beta\nu} - g^{\alpha\nu} \tilde{J}^{\mu\beta} + g^{\beta\nu} \tilde{J}^{\mu\alpha} \right) + ig_{\beta\rho} \left(g^{\rho\mu} \hat{P}^\nu - g^{\rho\nu} \hat{P}^\mu \right) \tilde{J}^{\alpha\beta} \\
&= i \left(\hat{P}^\mu \tilde{J}^{\alpha\nu} + g^{\alpha\mu} \hat{W}^\nu - g^{\alpha\nu} \hat{W}^\mu + \hat{P}^\nu \tilde{J}^{\mu\alpha} + \hat{P}^\nu \tilde{J}^{\alpha\mu} - \hat{P}^\mu \tilde{J}^{\alpha\nu} \right) \\
&= i \left(g^{\alpha\mu} \hat{W}^\nu - g^{\alpha\nu} \hat{W}^\mu \right), \quad \text{q.e.d.} \tag{6.52}
\end{aligned}$$

Now it is no longer difficult to show that \hat{W}^2 is a Casimir-Operator. On the one hand we have on account of Eq. (6.45)

$$[\hat{W}^2, \hat{P}^\rho] = g_{\mu\nu} [\hat{W}^\mu \hat{W}^\nu, \hat{P}^\rho] = g_{\mu\nu} \left\{ \hat{W}^\mu [\hat{W}^\nu, \hat{P}^\rho] + [\hat{W}^\mu, \hat{P}^\rho] \hat{W}^\nu \right\} = 0, \tag{6.53}$$

and on the other hand we compute with the help of Eq. (6.46)

$$\begin{aligned}
[\hat{W}^2, \hat{J}^{\rho\sigma}] &= g_{\mu\nu} [\hat{W}^\mu \hat{W}^\nu, \hat{J}^{\rho\sigma}] = g_{\mu\nu} \left\{ \hat{W}^\mu [\hat{W}^\nu, \hat{J}^{\rho\sigma}] + [\hat{W}^\mu, \hat{J}^{\rho\sigma}] \hat{W}^\nu \right\} \\
&= i \left\{ \hat{W}_\nu \left(g^{\nu\rho} \hat{W}^\sigma - g^{\nu\sigma} \hat{W}^\rho \right) + \left(g^{\mu\rho} \hat{W}^\sigma - g^{\mu\sigma} \hat{W}^\rho \right) \hat{W}_\mu \right\} \\
&= i \left(\hat{W}^\rho \hat{W}^\sigma - \hat{W}^\sigma \hat{W}^\rho + \hat{W}^\sigma \hat{W}^\rho - \hat{W}^\rho \hat{W}^\sigma \right) = 0, \quad \text{q.e.d.} \tag{6.54}
\end{aligned}$$

Finally, we give an explicit representation of \hat{W}^2 as function of the generators of the Poincaré group:

$$\begin{aligned}
\hat{W}^2 &= \hat{W}^\mu \hat{W}_\mu = \frac{1}{4} \epsilon^{\mu\nu\alpha\beta} \epsilon_{\mu}{}^{\tau\rho\sigma} \hat{J}_{\nu\alpha} \hat{P}_\beta \hat{J}_{\tau\rho} \hat{P}_\sigma \\
&= \frac{1}{4} [g^{\nu\tau} (g^{\alpha\sigma} g^{\beta\rho} - g^{\alpha\rho} g^{\beta\sigma}) + g^{\nu\rho} (g^{\alpha\tau} g^{\beta\sigma} - g^{\alpha\sigma} g^{\beta\tau}) + g^{\nu\sigma} (g^{\alpha\rho} g^{\beta\tau} - g^{\alpha\tau} g^{\beta\rho})] \hat{J}_{\nu\alpha} \hat{P}_\beta \hat{J}_{\tau\rho} \hat{P}_\sigma \\
&= \frac{1}{4} \left(\hat{J}^{\tau\sigma} \hat{P}^\rho - \hat{J}^{\tau\rho} \hat{P}^\sigma + \hat{J}^{\rho\tau} \hat{P}^\sigma - \hat{J}^{\rho\sigma} \hat{P}^\tau + \hat{J}^{\sigma\rho} \hat{P}^\tau - \hat{J}^{\sigma\tau} \hat{P}^\rho \right) \hat{J}_{\tau\rho} \hat{P}_\sigma \\
&= -\frac{1}{2} \left(\hat{J}^{\tau\rho} \hat{P}^\sigma - \hat{J}^{\tau\sigma} \hat{P}^\rho + \hat{J}^{\rho\sigma} \hat{P}^\tau \right) \hat{J}_{\tau\rho} \hat{P}_\sigma = -\left(\frac{1}{2} \hat{J}^{\tau\rho} \hat{P}^\sigma + \hat{J}^{\rho\sigma} \hat{P}^\tau \right) \hat{J}_{\tau\rho} \hat{P}_\sigma, \tag{6.55}
\end{aligned}$$

where we have frequently used the antisymmetry of $\hat{J}^{\mu\nu}$. Now we employ the commutation relation (6.32) and obtain

$$\begin{aligned}
\hat{W}^2 &= -\left(\frac{1}{2} \hat{J}^{\tau\rho} g^{\sigma\alpha} + \hat{J}^{\rho\sigma} g^{\tau\alpha} \right) \left[\hat{J}_{\tau\rho} \hat{P}_\alpha + i \left(g_{\alpha\tau} \hat{P}_\rho - g_{\alpha\rho} \hat{P}_\tau \right) \right] \hat{P}_\sigma \\
&= -\frac{1}{2} \hat{J}^{\tau\rho} \hat{J}_{\tau\rho} \hat{P}^2 - \frac{i}{2} \hat{J}^{\tau\rho} [\hat{P}_\rho, \hat{P}_\tau] - \hat{J}^{\rho\sigma} \hat{J}_{\tau\rho} \hat{P}^\tau \hat{P}_\sigma - 4i \hat{J}^{\rho\sigma} \hat{P}_\rho \hat{P}_\sigma + i \hat{J}^{\rho\sigma} \hat{P}_\rho \hat{P}_\sigma \\
&= -\frac{1}{2} \hat{J}^{\tau\rho} \hat{J}_{\tau\rho} \hat{P}^2 + \hat{J}^{\rho\sigma} \hat{J}_{\rho\tau} \hat{P}_\sigma \hat{P}^\tau, \tag{6.56}
\end{aligned}$$

where we utilized Eq. (6.34) and the antisymmetry of $\hat{J}^{\mu\nu}$.

What are the eigenvalues of \hat{W}^2 ? Choosing eigenstates of the momentum operator \hat{P}^μ , we may replace the momentum operator in Eq. (6.56) by its eigenvalue, $\hat{P}^\mu \equiv P^\mu$. Since

6 The Poincaré group

\hat{W}^2 is Lorentz-invariant, we may evaluate it in the rest frame of the respective particle, $P^\mu = (m, \vec{0})^T$. Then we obviously have

$$\hat{W}^2 = -\frac{m^2}{2} \hat{J}^{\tau\rho} \hat{J}_{\tau\rho} + m^2 \hat{J}^{\rho 0} \hat{J}_{\rho 0} = -\frac{m^2}{2} \left(\hat{J}^{\tau 0} \hat{J}_{\tau 0} + \hat{J}^{0\rho} \hat{J}_{0\rho} + \hat{J}^{ij} \hat{J}_{ij} - 2\hat{J}^{\rho 0} \hat{J}_{\rho 0} \right) = -m^2 \hat{\vec{J}}^2, \quad (6.57)$$

where we have replaced $\hat{J}^{ij} \equiv \epsilon^{ijk} \hat{J}^k$, cf. Eq. (6.36). Here, \hat{J}^k is the k th component of the total angular momentum operator, i.e., the sum of spin and orbital angular momentum. In the rest frame of the particle, the latter vanishes, and $\hat{J}^k \equiv \hat{S}^k$. The eigenvalues W^2 of the Casimir operators are therefore

$$W^2 = -m^2 S(S+1), \quad (6.58)$$

where S is the spin of the particle. For an explicit discussion for particles with spin zero, one, $3/2$, and two, see Ref. [6].

Bibliography

- [1] H. Georgi, *Lie Algebras in Particle Physics* (Frontiers in Physics 54, Westview Press, 1982)
- [2] M. Hamermesh, *Group Theory and Its Application to Physical Problems* (Addison–Wesley Publishing Co., Reading, 1962)
- [3] H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover Publications, New York, 1950)
- [4] W. Greiner, B. Müller, *Quantum Mechanics: Symmetries* (Springer, Berlin, 1994)
- [5] <http://pdg.lbl.gov>
- [6] A. Königstein, F. Giacosa, and D.H. Rischke, “Classical and quantum theory of the massive spin-two field,” *Annals Phys.* **368**, 16 (2016)