Symmetries in physics

Stefan Weinzierl

June 27, 2013

Contents

1 Introduction

1.1 Literature

- Introductory texts:
  - J.P. Elliot and P.G. Dawber, Symmetry in Physics, Macmillan Press, 1979
  - W. Fulton and J. Harris, Representation theory, Springer, 1991
  - B. Hall, Lie groups, Lie Algebras and Representations, Springer, 2003
- Physics related:
  - D.B. Lichtenberg, Unitary symmetry and elementary particles, Academic Press, 1970
  - H. Georgi, Lie Algebras in Particle Physics, Benjamin/Cummings Publishing Company, 1982
  - M. Schottenloher, Geometrie und Symmetrie in der Physik, Vieweg, 1995
  - M. Nakahara, Geometry, Topology and Physics, IOP, 1990
- Classics:
  - N. Bourbaki, Groupes et algèbres de Lie, Hermann, 1972
  - H. Weyl, The classical groups, Princeton University Press, 1946
- Differential geometry:
  - S. Helgason, Differential Geometry, Lie Groups and Symmetric Spaces, AMS, 1978
- Hopf algebras:
  - Ch. Kassel, Quantum Groups, Springer, 1995
- Specialised topics:
  - Ch. Reutenauer, Free Lie Algebras, Clarendon Press, 1993
  - V. Kac, Infinite dimensional Lie algebras, Cambridge University Press, 1983

1.2 Motivation

Symmetries occur in many physical systems, from molecules, crystals, atoms, nuclei to elementary particles. Symmetries occur in classical physics as well as in quantum physics. **A symmetry is expressed by a transformation, which leave the physical system invariant**. Examples for such transformations are translations, rotations, inversions, particle interchanges. The symmetry transformations form a group and we are led to study group theory.

A few examples:

Example 1: One classical particle in one dimension.
Consider a particle of mass m moving in one dimension under the influence of a potential $V(x)$. Newton's law gives the equation of motion

$$m\ddot{x} = -\frac{d}{dx}V(x).$$

Suppose now, that $V(x)$ is constant, in other words that it is invariant under translations. Then we have

$$m\ddot{x} = 0$$

and integrating this equation we obtain

$$m\dot{x} = \text{const},$$

showing that the momentum $p = m\dot{x}$ is conserved.

Example 2: One classical particle in two dimensions.
In two dimensions the motion of the particle is governed by the two equations

$$\begin{aligned} m\ddot{x} &= -\frac{\partial}{\partial x}V(x,y), \\ m\ddot{y} &= -\frac{\partial}{\partial y}V(x,y). \end{aligned}$$

Suppose now that $V(x,y)$ is invariant with respect to rotations about the origin, in other words that V is independent of the angle ϕ if expressed in terms of the polar coordinates r, ϕ rather than the cartesian coordinates x and y. In this case we have

$$\frac{\partial}{\partial \phi}V = 0.$$

From

$$\begin{aligned} x &= r\cos\phi, \\ y &= r\sin\phi \end{aligned}$$

we obtain

$$\frac{\partial}{\partial \phi}V = \frac{\partial x}{\partial \phi}\frac{\partial V}{\partial x} + \frac{\partial y}{\partial \phi}\frac{\partial V}{\partial y} = -y\frac{\partial V}{\partial x} + x\frac{\partial V}{\partial y},$$

and therefore

$$0 = \frac{\partial}{\partial \phi}V = -y\frac{\partial V}{\partial x} + x\frac{\partial V}{\partial y} = -y(-m\ddot{x}) + x(-m\ddot{y}) = m\frac{d}{dt}(y\dot{x} - x\dot{y}).$$

This shows that the angular momentum is conserved:

$$m\left(y\dot{x}-x\dot{y}\right) \;=\; \text{const.}$$

Example 3: Noether theorem.
The two examples above are a special case of the Noether theorem. In classical mechanics we can describe a physical system by generalised coordinates q_i and a Lagrange function $L(\vec{q},\dot{\vec{q}},t)$. We consider the case of a family of coordinate transformations,

$$q_i' \;=\; f_i\left(\vec{q},t,\alpha\right), \qquad 1 \le i \le n,$$

depending on a real parameter α, such that all functions f_i are continously differentiable with respect to α and sucht that $\alpha = 0$ corresponds to the identity transformation

$$q_i \;=\; f_i\left(\vec{q},t,0\right), \qquad 1 \le i \le n.$$

If there exists an $\varepsilon > 0$, such that for all $|\alpha| < \varepsilon$ we have

$$L'\left(\vec{q}',\dot{\vec{q}}',t\right) \;=\; L\left(\vec{q},\dot{\vec{q}},t\right)+\frac{d}{dt}\Lambda\left(\vec{q},t,\alpha\right)+O\left(\alpha^2\right),$$

with

$$\Lambda\left(\vec{q},t,0\right) \;=\; 0,$$

then it follows, that the quantity

$$I \;=\; \left(\sum_{i=1}^{n} p_i \left.\frac{\partial f_i}{\partial\alpha}\right|_{\alpha=0}\right) - \left.\frac{\partial\Lambda}{\partial\alpha}\right|_{\alpha=0}$$

is conserved. This is Noether's theorem. It states the every continuously symmetry of the Lagrange function leads to a conserved quantity. Note that it is only required that the transformed Lagrange function agrees with the original one only up to a gauge transformation.

Example 4: Two particles connected by springs.
Consider two particles of equal mass m connected to each other and to fixed supports by springs with spring constant λ. The kinetic and potential energies are

$$\begin{aligned} T &= \frac{1}{2}m\left(\dot{x}_1^2+\dot{x}_2^2\right), \\ V &= \frac{1}{2}\lambda\left(x_1^2+x_2^2+\left(x_1+x_2\right)^2\right). \end{aligned}$$

This system is symmetric under the interchange $x_1 \leftrightarrow x_2$. The equations of motion are

$$\begin{aligned} m\ddot{x}_1 &= -\lambda x_1-\lambda\left(x_1+x_2\right), \\ m\ddot{x}_2 &= -\lambda x_2-\lambda\left(x_1+x_2\right). \end{aligned}$$

This suggests new coordinates

$$q_1 = x_1 + x_2,$$
$$q_2 = x_1 - x_2.$$

Adding and subtracting the two equations above we obtain

$$m\ddot{q}_1 = -3\lambda q_1,$$
$$m\ddot{q}_2 = -\lambda q_2.$$

In terms of the coordinates q_1 and q_2 the solutions are harmonic oscillations with frequencies

$$\omega_1 = \sqrt{\frac{3\lambda}{m}}, \qquad \omega_2 = \sqrt{\frac{\lambda}{m}}.$$

Let us denote the symmetry transformation $x_1 \leftrightarrow x_2$ by σ. The new coordinates q_1 and q_2 are even and odd, respectively, under the symmetry transformation:

$$\sigma q_1 = q_1,$$
$$\sigma q_2 = -q_2.$$

Example 5: Parity transformations in quantum mechanics.
Consider a quantum mechanical particle in an energy eigenstate with energy E. Let us assume that this eigenstate is non-degenerate. If the potential has the reflection symmetry

$$V(\vec{x}) = V(-\vec{x}),$$

it follows that also the Hamilton operator has this property: $H(\vec{x}) = H(-\vec{x})$. Now, if $\psi(\vec{x})$ is an eigenfunction of the Hamilton operator with eigenvalue E,

$$H\psi(\vec{x}) = E\psi(\vec{x}),$$

it follows that also $\psi(-\vec{x})$ is an eigenfunction with eigenvalue E:

$$H\psi(-\vec{x}) = E\psi(-\vec{x}).$$

Since we assumed that the energy state is non-degenerate we must have

$$\psi(-\vec{x}) = c\psi(\vec{x}).$$

Repeating the symmetry operation twice we obtain

$$\psi(\vec{x}) = c^2\psi(\vec{x}),$$

and thus

$$c = \pm 1.$$

Hence, $\psi(\vec{x})$ is either even or odd. This leads to selection rules in quantum mechanics: The transition probability for a decay from some initial state ψ_i to a final state ψ_f is proportional to the square of hte integral

$$I \;=\; \int d^3x\, \psi_f^*(\vec{x})\, O(\vec{x})\, \psi_i(\vec{x})$$

where $O(\vec{x})$ depends on the particular decay process. If $O(\vec{x})$ is an even function of $\vec{x}$, the integral is non-zero only if ψ_i and ψ_f are both even or both odd.

Example 6: Particle physics
In particle physics one often observes that certain particles form a pattern (mathematically we say they form a representation of a group). If some particles are already discovered and the pattern is known, one is able to predict the remaining particles of the pattern.

- The π^0 meson in the isospin triple (π^+, π^0, π^-).
- The Ω^- in the baryon decuplet.
- The charm quark as partner of the strange quark.
- The top quark as partner of the bottom quark.
- The Z-boson as a third mediator of the weak force.

Summary: Understanding the symmetry properties of a physical system is useful for the following reasons:

- Gives insight (origin of selection rules in quantum mechanics)
- Simplifies calculations (conserved quantities)
- Makes predictions (new particles)
- Gauge symmetries are the key ingredient for the understanding of the fundamental forces.

2 Basics of group theory

2.1 Definition of a group

A non-empty set G together with a composition $\cdot : G \times G \to G$ is called a group $(G, \cdot)$ if

G1: The composition $\cdot$ is associative : $a \cdot (b \cdot c) = (a \cdot b) \cdot c$

G2: There exists a neutral element : $e \cdot a = a \cdot e = a$ for all $a \in G$

G3: For all $a \in G$ there exists an inverse a^{-1} : $a^{-1} \cdot a = a \cdot a^{-1} = e$

One can actually use a weaker system of axioms:

G1': The composition $\cdot$ is associative : $a \cdot (b \cdot c) = (a \cdot b) \cdot c$

G2': There exists a left-neutral element : $e \cdot a = a$ for all $a \in G$

G3': For all $a \in G$ there exists an left-inverse a^{-1} : $a^{-1} \cdot a = e$

The first system of axioms clearly implies the second system of axioms. To show that the second system also implies the first one, we show the following:
a) If e is a left-neutral element, and e' is a right-neutral element, then $e = e'$.
Proof:

$$\begin{aligned} e' &= e \cdot e' && e \text{ is left-neutral} \\ &= e, && e' \text{ is right-neutral} \end{aligned}$$

b) If b is a left-inverse to a, and b' is a right-inverse to a, then $b = b'$.
Proof:

$$\begin{aligned} b &= b \cdot e && e \text{ is right-neutral} \\ &= b \cdot \left(a \cdot b'\right) && b' \text{ is right-inverse of } a \\ &= (b \cdot a) \cdot b' && \text{associativity} \\ &= e \cdot b' && b \text{ is left-inverse of } a \\ &= b' && e \text{ is left-neutral} \end{aligned}$$

c) If b is a left-inverse to a, i.e. $b \cdot a = e$, then b is also the right-inverse to a.
Proof:

$$\begin{aligned} (a \cdot b) \cdot (a \cdot b) &= a \cdot (b \cdot a) \cdot b \\ &= a \cdot e \cdot b \\ &= a \cdot b \end{aligned}$$

Therefore $a \cdot b = e$.
d) If e is the left-neutral element, then e is also right-neutral.
Proof:

$$\begin{aligned} a &= e \cdot a \\ &= \left(a^{-1} \cdot a\right) \cdot a \\ &= \left(a \cdot a^{-1}\right) \cdot a \\ &= a \cdot \left(a^{-1} \cdot a\right) \\ &= a \cdot e \end{aligned}$$

This completes the proof that the second system of axioms is equivalent to the first system of axioms. To verify that a given set together with a given composition forms a group it is therefore sufficient to verify axioms (G2') and (G3') instead of axioms (G2) and (G3).

More definitions:
A group $(G, \cdot)$ is called **Abelian** if the operation $\cdot$ is commutative : $a \cdot b = b \cdot a$.

The number of elements in the set G is called the **order of the group**. If this number is finite, we speak of a finite group. In the case where the order is infinite, we can further distinguish the case where the set is countable or not. For Lie groups we are in particular interested in the latter case. For finite groups we can write down all possible compositions in a **composition table**. In such a composition table each element occurs exactly once in each row and column.

Examples

a) The trivial example: Let $G = \{e\}$ and $e \cdot e = e$. This is a group with one element.

b) $\mathbb{Z}_2$: Let $G = \{0, 1\}$ and denote the composition by $+$. The composition is given by the following composition table:

$+$	0	1
0	0	1
1	1	0

$\mathbb{Z}_2$ is of order 2 and is Abelian.

c) $\mathbb{Z}_n$: We can generalise the above example and take $G = \{0, 1, 2, ..., n-1\}$. We define the addition by

$$a + b \quad = \quad a + b \bmod n,$$

where on the l.h.s. "+" denotes the composition in $\mathbb{Z}_n$, whereas on the r.h.s. "+" denotes the usual addition of integer numbers. $\mathbb{Z}_n$ is a group of order n and is Abelian.

d) The symmetric group S_n: Let X be a set with distinct n elements and set

$$G = \{\sigma | \sigma : X \to X \text{ permutation of } X\}$$

As composition law we take the composition of permutations. The symmetric group has order

$$|S_n| = n!$$

For $n \geq 3$ this group is non-Abelian:

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} &= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \\ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} &= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}. \end{aligned}$$

e) $(\mathbb{Z}, +)$: The integer numbers with addition form an Abelian group. The order of the group is infinite, but countable.

f) $(\mathbb{R}, +)$: The real numbers with addition form an Abelian group. The order of the group is not countable.

g) $(\mathbb{R}^*, \cdot)$: Denote by $\mathbb{R}^* = \mathbb{R} \backslash \{0\}$ the real numbers without zero. The set $\mathbb{R}^*$ with the multiplication as composition law forms an Abelian group.

h) Rotations in two dimensions: Consider the set of 2×2-matrixes

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix},$$

together with matrix multiplication as composition. To check this, one has to show that

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}$$

can again be written as

$$\begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix}.$$

Using the addition theorems of sin and cos one finds $\gamma = \alpha + \beta$. The elements of this group are not countable, but they form a compact set.

2.2 Group morphisms

Let a and b be elements of a group $(G,*)$ with composition $*$ and let a' and b' be elements of a group $(G',\circ)$ with composition $\circ$. We are interested in mappings between groups which preserve the structure of the compositions.

Homomorphism: We call a mapping $f: G \to G'$ a homomorphism, if

$$f(a*b) \;=\; f(a)\circ f(b).$$

Isomorphism: We call a mapping $f: G \to G'$ an isomorphism, if it is bijective and a homomorphism.

Automorphism: We call a mapping $f: G \to G$ from the group G into the group G itself an automorphism, if it is an isomorphism.

We consider an example for an isomorphism: We take

$$G \;=\; (\mathbb{Z}_n,+),$$

and

$$G' \;=\; \left(\left\{e^{2\pi i\frac{0}{n}}, e^{2\pi i\frac{1}{n}}, \dots, e^{2\pi i\frac{(n-1)}{n}}\right\}, \cdot\right).$$

This is the group of the n-th roots of unity. These two groups are isomorphic. The isomorphism is given by

$$\begin{aligned} f \;:\; & G \to G', \\ & k \to e^{2\pi i\frac{k}{n}}. \end{aligned}$$

2.3 Subgroups

A non-empty subset $H \subseteq G$ is said to be a **subgroup** of a group G, if H is itself a group under the same law of composition as that of G.

Every group has two trivial subgroups:

- The group consisting just of the identity element $\{e\}$.
- The whole group G.

A subgroup H is called a **proper subgroup** if $H \neq G$.

Example: Consider the group

$$\mathbb{Z}_6 = \{0,1,2,3,4,5\}.$$

The set

$$H = \{0, 2, 4\}$$

is a subgroup of $\mathbb{Z}_6$, isomorphic to $\mathbb{Z}_3$. The isomorphism is given by

$$\begin{aligned} f \quad : \quad & H \to \mathbb{Z}_3, \\ & 2n \to n. \end{aligned}$$

Generating set: Let G be a group and $S \subseteq G$. The group

$$\langle S \rangle = \cap \{U | U \text{ subgroup of } G \text{ with } S \subseteq U\}$$

is called the subgroup generated by S. We say that G is generated by S, if

$$G \quad = \quad \langle S \rangle.$$

We say that G is **finitely generated** if there is a finite set $S = \{a_1, ..., a_n\}$ such that

$$G \quad = \quad \langle a_1, ..., a_n \rangle.$$

Note that "finitely generated" does not imply that the group has a finite number of elements. The group $(\mathbb{Z}, +)$ is generated by the element 1, but has an infinite number of elements.

A group G is called **cyclic**, if it is generated by one element

$$G \quad = \quad \langle a \rangle$$

Remark: Cyclic groups are always Abelian.

We define the **order** of an element $a \in G$ as

$$\text{ord}\, a \quad = \quad |\langle a \rangle|.$$

The order of an element a is either infinity or equal to the smallest positive number s such that

$$a^s \quad = \quad e,$$

where

$$a^s \quad = \quad \underbrace{a \circ ... \circ a}_{s \text{ times}}$$

The orders of the elements of a group give us useful information about the group: Let us consider the following example. Assume that we are considering a group with four elements $\{e, a, b, c\}$. Assume further that the elements a and b are of order two:

$$a^2 = b^2 = e.$$

Based on this knowledge we know the following entries in the composition table:

	e	a	b	c
e	e	a	b	c
a	a	e	?	
b	b		e	
c	c			

In the place with the question mark we must put c, otherwise c would occur in the last column twice. All other entries can be figured out with similar considerations and we obtain the composition table

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

2.4 Cosets

Consider a subgroup H of a group G. Let $a \in G$. The set

$$aH = \{ah_1, ah_2, ah_3, \ldots\}$$

is called a **left coset** of H in G. The number of distinct left cosets of H in G is called the **index** of H in G and is denoted by

$$|G:H|$$

Theorem (Lagrange):

1. $$G = \bigcup_{a \in G} aH$$

2. Two left cosets of H in G are either identical or have no common element.

3. If two of the numbers $|G|$, $|H|$ and $|G:H|$ are finite, then also the third one is and we have the relation

$$|G| = |G:H|\,|H|\,.$$

Proof:

1. H contains the neutral element. We therefore have

$$\bigcup_{a \in G} aH \supseteq \bigcup_{a \in G} ae = \bigcup_{a \in G} a = G.$$

On the other hand it is obvious that

$$\bigcup_{a\in G} aH \subseteq G$$

and the claim follows.

2. Assume that a_1H and a_2H have one element in common. Then

$$\begin{aligned} a_1h_1 &= a_2h_2 \\ a_2^{-1}a_1 &= h_2h_1^{-1} \in H \\ a_2^{-1}a_1H &= H \\ a_1H &= a_2H, \end{aligned}$$

which shows that the two cosets are identical.

3. The proof follows from

$$|aH| = |H|.$$

The theorem of Lagrange has some important consequences for finite groups: Let G be a finite group.

- The order $|H|$ of each subgroup divides the order $|G|$ of the group.
- The order of each element a divides the order $|G|$ of the group.
- For each element a we have

$$a^{|G|} = e.$$

As a further consequence we have the following: If G is a finite group, where the order is a prime number, then G is cyclic.

2.5 Conjugacy classes

Let G be a group. An element b is conjugate to an element a if there is an element $g \in G$ sucht that

$$b = gag^{-1}.$$

Remark: $b = gag^{-1}$ is equivalent to $a = g^{-1}bg$.
This defines an **equivalence relation**:

- a is conjugated to itself.

- If a is conjugated to b, then b is conjugated to a.
- If a is conjugated to b and b is conjugated to c, then a is conjugated to c.

The set of all elements conjugate to a is called the **conjugacy class** of a. The set of all elements of a group can be decomposed into disjoint conjugacy classes.

It is often useful to consider conjugacy classes instead of all individual elements, because group elements of the same conjugacy class will behave similar.

2.6 Normal subgroups

Let G be a group and H a subgroup of G. We have already considered the left cosets aH, $a \in G$. We may ask under which conditions do the left cosets form again a group ?

We start with a definition: We call a subgroup N of G a **normal subgroup** if for all $a \in G$

$$aNa^{-1} \subseteq N.$$

Remark: This means that for all $n_1 \in N$ and for all $a \in G$ there exists an $n_2 \in N$ such that

$$an_1a^{-1} = n_2.$$

For a proper normal subgroup N of G one writes $N \lhd G$.

For a normal subgroup the left and right cosets are equal

$$aN = Na.$$

This is exactly the property we need such that the left cosets have a well-defined composition law:

$$(aN)(bN) = (ab)N,$$

or in more detail

$$(an_1)(bn_2) = a(n_1b)n_2 = a\left(bn_1'\right)n_2 = ab\underbrace{n_1'n_2}_{n_3} = (ab)n_3.$$

We summarise: If N is a normal subgroup of G, then

$$G/N = \{aN | a \in G\}$$

together with the composition law

$$(aN)(bN) = (ab)N$$

is a group, called the **factor group** of G by N.

Remark: We have

$$|G/N| \quad = \quad |G:N|\,.$$

Remark 2: Suppose N_1 is a normal subgroup of G_1 and suppose that N_2 is a normal subgroup of G_2. Let us further assume that N_1 and N_2 as well as the corresponding factor groups are isomorphic:

$$\begin{aligned} N_1 &\cong N_2, \\ G_1/N_1 &\cong G_2/N_2. \end{aligned}$$

This does **not** imply that G_1 and G_2 are isomorphic, as the counter-example $G_1 = \mathbb{Z}_4$ and $G_2 = \mathbb{Z}_2 \times \mathbb{Z}_2$ shows.

Example: Let us consider the dihedral group D_n for $n \geq 3$. This is the symmetry group of a regular polygon. The symmetry operations are rotations through an angle $2\pi/n$ and reflexions. In mathematical terms, this group is generated by two elements a and b with

$$a^n = b^2 = e, \qquad ab = ba^{n-1}.$$

The element a corresponds to a rotation through the angle $2\pi/n$, hence $a^n = e$, the element b corresponds to a reflexion, hence $b^2 = e$. This group is non-Abelian and one easily convinces oneself that $ab = ba^{n-1}$ (note that $a^{-1} = a^{n-1}$). The element a generates a cyclic subgroup $\langle a \rangle$ of order n. We can write D_n as left cosets of $\langle a \rangle$:

$$D_n \quad = \quad \langle a \rangle \;\cup\; b\langle a \rangle.$$

This shows that D_n has $(2n)$ elements. The subgroup $\langle a \rangle$ is a normal subgroup:

$$ba^j b^{-1} = ba^j b = ba^{j-2}ab = ba^{j-2}ba^{n-1} = bba^{j(n-1)} = a^{j(n-1)} \in \langle a \rangle.$$

The factor group $D_n/\langle a \rangle$ is of order 2 and hence

$$D_n/\langle a \rangle \quad \cong \quad \mathbb{Z}_2.$$

The element b generates a cyclic subgroup $\langle b \rangle$ of order 2. This subgroup is not a normal subgroup:

$$\left(ba^j\right) b \left(ba^j\right)^{-1} \quad = \quad ba^j ba^{n-j} b = ba^j b^2 a^{(n-j)(n-1)} = ba^{j+n^2-n-jn+j} = ba^{2j}.$$

Let us now look more specifically at D_4, the dihedral group with $n = 4$. We work out the conjugacy classes and find

$$\{e\}, \{a, a^3\}, \{a^2\}, \{b, ba^2\}, \{ba, ba^3\}.$$

2.7 Direct product

Let G_1, ..., G_n be groups. In the set

$$G = G_1 \times \ldots \times G_n = \{(a_1, \ldots, a_n) \,|a_i \in G_i, 1 \le i \le n\}$$

we define a composition by

$$(a_1, \ldots, a_n)(b_1, \ldots, b_n) = (a_1 b_1, \ldots, a_n b_n).$$

G is called the **direct product** of the groups G_1, ..., G_n. The neutral element in G is given by $(e_1, \ldots, e_n)$, the inverse element to $(a_1, \ldots, a_n)$ is given by $(a_1^{-1}, \ldots, a_n^{-1})$.

For the order of the group G we have

$$|G| = \prod_{i=1}^{n} |G_i|.$$

Let us denote the trivial group by $E = \{e\}$ and

$$\tilde{G}_i = \underbrace{E \times \ldots \times E}_{(i-1)\text{ times}} \times G_i \times \underbrace{E \times \ldots \times E}_{(n-i)\text{ times}}$$

We can show that $\tilde{G}_i$ is a normal subgroup in G and that

$$\tilde{G}_i \cap \left(\tilde{G}_1 \ldots \tilde{G}_{i-1} \tilde{G}_{i+1} \ldots \tilde{G}_n\right) = \{e\}$$

We call a group G the **inner direct product** of normal subgroups N_i, $1 \le i \le n$, if

$$G = N_1 N_2 \ldots N_n,$$
$$N_i \cap (N_1 \ldots N_{i-1} N_{i+1} \ldots N_n) = \{e\}.$$

Furthe properties of inner direct products: If $G = N_1 \ldots N_n$ is an inner direct product, then we have

- The elements of N_i commute with the elements of N_j for $i \neq j$:

$$a_i a_j = a_j a_i.$$

- Each element $a \in G$ can be represented uniquely (up to ordering) as

$$a = a_1 \ldots a_n,$$

  with $a_i \in N_i$.

The converse is also true: If G_i $(1 \leq i \leq n)$ are subgroups of G, such that the elements of G_i commute with the elements of G_j (for $i \neq j$) and each element $a \in G$ has a unique representation

$$a = a_1 ... a_n$$

with $a_i \in G_i$, then the G_i's are normal subgroups of G and G is an inner direct product of the G_i's.

Consequence: If n and m are two positive numbers, which share no common divisor, then

$$\mathbb{Z}_{nm} \cong \mathbb{Z}_n \times \mathbb{Z}_m.$$

Without a proof we state the following theorem, which characterises completely the finitely generated Abelian groups:

Theorem on finitely generated Abelian groups: Every finitely generated Abelian group G is a direct product of finitely many cyclic groups, in other words

$$G \cong \mathbb{Z}_{p_1^{k_1}} \times ... \times \mathbb{Z}_{p_r^{k_r}} \times \mathbb{Z} \times ... \times \mathbb{Z},$$

with (not necessarily distinct) prime numbers p_1, ..., p_r.

An example where direct products occur in physics is given by the gauge symmetry of the Standard Model of particle physics. The gauge group is given by

$$U(1) \times SU(2) \times SU(3),$$

where $U(1)$ is the gauge group corresponding to the hypercharge, $SU(2)$ is the gauge group corresponding to the weak isospin and $SU(3)$ is the gauge group corresponding to the colour charges.

2.8 The theorems of Sylow

In this section we state without proof the three theorems of Sylow. These theorems are very helpful to discuss the structure of finite groups. From the theorem of Lagrange we know, that the order of a subgroup must divide the order of the group. But the theorem of Lagrange does not make any statement on the existence of a subgroup for a given divisor. This situation is clarified with the theorems of Sylow.

First theorem of Sylow: Let G be a finite group of order $n = p^r m$, where p is a prime number and $\gcd(p,m) = 1$. Then there exists for each j with $1 \leq j \leq r$ a subgroup H of G with order p^j.

Corollary: Let G be a finite group and let p be a prime number, which divides the order of G. Then G contains an element of order p.

Definition: Let p be a prime number and let G be a group. We call G a p**-group**, if the order of every element is a power of p.

Definition: Let G be a group and let P be a subgroup of G. We call P a p**-Sylow group** of G, if

- P is a p-group
- P is maximal, i.e. if P' is another p-subgroup of G with $P \subseteq P'$, then $P = P'$.

It can be shown that if P is a p-Sylow subgroup of G, then also aPa^{-1} is a p-Sylow subgroup of G for all $a \in G$.

Second theorem of Sylow: Let G be a finite group of order $n = p^r m$, where p is a prime number and $\gcd(p,m) = 1$. Let P be a p-Sylow group of G and let H be a p-subgroup of G. Then there exists an element a such that

$$aHa^{-1} \quad \subseteq \quad P.$$

Corollary: Assume that for a given prime number p there is only one p-Sylow group P of G. Then P is a normal subgroup.

Third theorem of Sylow: Let G be a finite group and assume that the prime number p is a divisor of the order of the group. Then the number of the p-Sylow groups of G is also a divisor of the order of the group and of the form

$$1 + kp,$$

with $k \geq 0$.

2.9 The group rearrangement theorem

Let G be a group and $a \in G$. Consider the set

$$aG \quad = \quad \{ag | g \in G\}.$$

We have

$$aG \quad = \quad G.$$

Proof: We show $aG \subseteq G$ and $aG \supseteq G$.
Let us start with "$\subseteq$": Take an arbitrary $b \in G$. We have $ab \in G$ and therefore $aG \subseteq G$.
We then show "$\supseteq$": Take an arbitrary $c \in G$. The element $(a^{-1}c)$ is again in G and therefore $a\left(a^{-1}c\right) = c$. This shows $aG \supseteq G$.

An important application of the group rearrangement theorem is the following: Consider a function

$$f \quad : \quad G \to \mathbb{R}$$

defined on the group G. Then

$$\sum_{a \in G} f(a) \quad = \quad \sum_{a \in G} f(ab)$$

for all $b \in G$.

3 Lie groups

3.1 Manifolds

3.1.1 Definition

A **topological space** is a set M together with a family $\mathcal{T}$ of subsets of M satisfying the following properties:

1. $\emptyset \in \mathcal{T}, M \in \mathcal{T}$
2. $U_1, U_2 \in \mathcal{T} \Rightarrow U_1 \cap U_2 \in \mathcal{T}$
3. For any index set A we have $U_\alpha \in \mathcal{T}; \alpha \in A \Rightarrow \bigcup\limits_{\alpha \in A} U_\alpha \in \mathcal{T}$

The sets $U \in \mathcal{T}$ are called **open**.

A topological space is called **Hausdorff** if for any two distinct points $p_1, p_2 \in M$ there exists open sets $U_1, U_2 \in \mathcal{T}$ with

$$p_1 \in U_1, \quad p_2 \in U_2, \quad U_1 \cap U_2 = \emptyset.$$

A map between topological spaces is called **continous** if the preimage of any open set is again open.

A bijective map which is continous in both directions is called a **homeomorphism**.

An **open chart** on M is a pair (U, φ), where U is an open subset of M and φ is a homeomorphism of U onto an open subset of $\mathbb{R}^n$.

A **differentiable manifold** of dimension n is a Hausdorff space with a collection of open charts $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ such that

M1:

$$M = \bigcup_{\alpha \in A} U_\alpha.$$

M2: For each pair $\alpha, \beta \in A$ the mapping $\varphi_\beta \circ \varphi_\alpha^{-1}$ is an infinitely differentiable mapping of $\varphi_\alpha\left(U_\alpha \cap U_\beta\right)$ onto $\varphi_\beta\left(U_\alpha \cap U_\beta\right)$.

A differentiable manifold is also often denoted as a C^∞ manifold. As we will only be concerned with differentiable manifolds, we will often omitt the word “differentiable” and just speak about manifolds.

The collection of open charts $(U_\alpha, \varphi_\alpha)_{\alpha \in A}$ is called an **atlas**.

If $p \in U_\alpha$ and

$$\varphi_\alpha(p) \;=\; (x_1(p), ..., x_n(p)),$$

the set U_α is called the **coordinate neighbourhood** of p and the numbers $x_i(p)$ are called the **local coordinates** of p.

Note that in each coordinate neighbourhood M looks like an open subset of $\mathbb{R}^n$. But note that we do not require that M be $\mathbb{R}^n$ globally.

Consider two manifolds M and N with dimensions m and n. Let x_i be coordinates on M and y_j be coordinates on N. A mapping $f : M \to N$ between two manifolds is called **analytic**, if for each point $p \in M$ there exits a neighbourhood U of p and n power series P_j, $j = 1, ..., n$ such that

$$y_j(f(q)) \;=\; P_j\left(x_1(q) - x_1(p), ..., x_m(q) - x_m(p)\right)$$

for all $q \in U$.

An **analytic manifold** is a manifold where the mapping $\varphi_\beta \circ \varphi_\alpha^{-1}$ is analytic.

3.1.2 Examples

a) $\mathbb{R}^n$: The space $\mathbb{R}^n$ is a manifold. $\mathbb{R}^n$ can be covered with a single chart.

b) S^1: The circle

$$S^1 = \{\vec{x} \in \mathbb{R}^2 | |\vec{x}|^2 = 1\}$$

is a manifold. For an atlas we need at least two charts.

c) The set of rotation matrices in two dimensions:

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix},$$

The set of all these matrices forms a manifold homeomorphic to the circle S^1.

3.1.3 Morphisms

Homeomorphism: A map $f : M \to N$ between two manifolds M and N is called a homeomorphism if it is bijective and both the mapping $f : M \to N$ and the inverse $f^{-1} : N \to M$ are continous.

Diffeomorphism: A map $f : M \to N$ is called a diffeomorphism if it is a homeomorphism and both f and f^{-1} are infinitely differentiable.

Analytic diffeomorphism: The map $f : M \to N$ is a diffeomorphism and analytic.

3.2 Lie groups

3.2.1 Definition

A Lie group G is a group which is also an analytic manifold, such that the mappings

$$\begin{array}{rcl} G \times G & \to & G, \\ (a,b) & \to & a \cdot b, \end{array}$$

and

$$\begin{array}{rcl} G & \to & G, \\ a & \to & a^{-1} \end{array}$$

are analytic.

Remark: Instead of the two mappings above, it is sufficient to require that the mapping

$$\begin{array}{rcl} G \times G & \to & G, \\ (a,b) & \to & a \cdot b^{-1} \end{array}$$

is analytic.

3.2.2 Examples

The most important examples of Lie groups are matrix groups with matrix multiplication as composition. In order to have an inverse, the matrices must be non-singular.

a) $GL(n,\mathbb{R})$, $GL(n,\mathbb{C})$: The group of non-singular $n \times n$ matrices with real or complex entries. $GL(n,\mathbb{R})$ has n^2 real parameters, $GL(n,\mathbb{C})$ has $2n^2$ real parameters.

b) $SL(n,\mathbb{R})$, $SL(n,\mathbb{C})$: The group of non-singular $n \times n$ matrices with real or complex entries and

$$\det A = 1.$$

$SL(n,\mathbb{R})$ has $n^2 - 1$ real parameters, while $SL(n,\mathbb{C})$ has $2(n^2 - 1)$ real parameters.

c) $O(n)$: The group of orthogonal $n \times n$ matrices defined through

$$RR^T = 1.$$

The group $O(n)$ has $n(n-1)/2$ real parameters. The group $O(n)$ can also be defined as the transformation group of a real n-dimensional vector space, which preserves the inner product

$$\sum_{i=1}^{n} x_i^2$$

d) $SO(n)$: The group of special orthogonal $n \times n$ matrices defined through

$$RR^T = 1 \quad \text{and} \quad \det R = 1.$$

The group $SO(n)$ has $n(n-1)/2$ real parameters.

e) $U(n)$: The group of unitary $n \times n$ matrices defined through

$$UU^\dagger = 1.$$

The group $U(n)$ has n^2 real parameters. The group $U(n)$ can also be defined as the transformation group of a complex n-dimensional vector space, which preserves the inner product

$$\sum_{i=1}^{n} z_i^* z_i$$

f) $SU(n)$: The group of special unitary $n \times n$ matrices defined through

$$UU^\dagger = 1 \quad \text{and} \quad \det U = 1.$$

The group $SU(n)$ has $n^2 - 1$ real parameters.

g) $Sp(n, \mathbb{R})$: The symplectic group is the group of $2n \times 2n$ matrices satisfying

$$M^T \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} M = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

The group $Sp(n, \mathbb{R})$ has $(2n+1)n$ real parameters. The group $Sp(n, \mathbb{R})$ can also be defined as the transformation group of a real $2n$-dimensional vector space, which preserves the inner product

$$\sum_{j=1}^{n} \left(x_j y_{j+n} - x_{j+n} y_j \right).$$

3.3 Algebras

3.3.1 Definition

Let K be a field and A a vector space over the field K. A is called an algebra, if there is an additional composition

$$\begin{aligned} A \times A &\to A \\ (a_1, a_2) &\to a_1 a_2 \end{aligned}$$

such that the algebra multiplication is K-linear.

$$\begin{aligned} (r_1 a_1 + r_2 a_2) a_3 &= r_1 (a_1 a_3) + r_2 (a_2 a_3) \\ a_3 (r_1 a_1 + r_2 a_2) &= r_1 (a_3 a_1) + r_2 (a_3 a_2) \end{aligned}$$

Remark: It is not necessary to require that K is a field. It is sufficient to have a commutative ring R with 1. In this case one replaces the requirement for A to be a vector space by the requirement that A is an unital R-modul. The difference between a field K and a commutative ring R with 1 lies in the fact that in the ring R the multiplicative inverse might not exist.

An algebra is called associative if

$$(a_1 a_2) a_3 = a_1 (a_2 a_3)$$

An algebra is called commutative if

$$a_1 a_2 = a_2 a_1$$

An unit element $\mathbf{1}_A \in A$ satisfies

$$\mathbf{1}_A \, a = a.$$

Note that it is not required that A has a unit element. If there is one, note that difference between $\mathbf{1}_A \in A$ and $1_K \in K$: The latter always exists and we have the scalar multiplication with one:

$$1_K \, a = a.$$

3.3.2 Examples

a) Consider the set of $n \times n$ matrices over $\mathbb{R}$ with the composition given by matrix multiplication. This gives an associative, non-commutative algebra with a unit element given by the unit matrix.

b) Consider the set of $n \times n$ matrices over $\mathbb{R}$ where the composition is defined by

$$[a,b] = ab - ba.$$

This defines a non-associative, non-commutative algebra. There is no unit element.

3.4 Lie algebras

3.4.1 Definition

For a Lie algebra it is common practice to denote the composition of two elements a and b by $[a,b]$. An algebra is called a Lie-algebra if the composition satisfies

$$\begin{aligned} [a,a] &= 0, \\ [a,[b,c]] + [b,[c,a]] + [c,[a,b]] &= 0. \end{aligned}$$

Remark: Consider again the example above of the set of $n \times n$ matrices over $\mathbb{R}$ where the composition is defined by the commutator

$$[a,b] = ab - ba.$$

Clearly this definition satisfies $[a,a]=0$. It fullfills the Jacobi identity:

$$\begin{aligned}
& [a,[b,c]]+[b,[c,a]]+[c,[a,b]] = \\
&= abc-acb-bca+cba+bca-bac-cab+acb+cab-cba-abc+bac \\
&= 0.
\end{aligned}$$

Matrix algebras with the commutator as composition are therefore Lie algebras.

Let A be a Lie algebra and $X_1,...,X_n$ a basis of A as a vector space. $[X_i,X_j]$ is again in A and can be expressed as a linear combination of the basis vectors X_k:

$$\left[X_i,X_j\right] = \sum_{k=1}^{n} c_{ijk} X_k.$$

The coefficents c_{ijk} are called the structure constants of the Lie algebra. For matrix algebras the X_i's are anti-hermitian matrices.

The notation above is mainly used in the mathematical literature. In physics a slightly different convention is often used: Denote by $T_1,...,T_n$ a basis of A as a (complex) vector space. Then

$$[T_a,T_b] = i \sum_{c=1}^{n} f_{abc} T_c.$$

For matrix algebras the T_a's are hermitian matrices.

We can get from one convention to the other one by letting

$$T_a = iX_a.$$

In this case we have

$$f_{abc} = c_{abc}.$$

3.4.2 The exponential map

In this section we focus on matrix Lie groups. Let us first define the matrix exponential. For an $n \times n$ matrix X we define $\exp X$ by

$$\exp X = \sum_{n=0}^{\infty} \frac{X^n}{n!}.$$

Theorem: For any $n \times n$ real or complex matrix X the series converges.

A few properties:
1. We have

$$\exp(0) = 1.$$

2. $\exp X$ is invertible and

$$(\exp X)^{-1} = \exp(-X).$$

3. We have

$$\exp[(\alpha+\beta)X] = \exp(\alpha X)\exp(\beta X).$$

4. If $XY = YX$ then

$$\exp(X+Y) \quad = \quad \exp X \exp Y.$$

5. If A is invertible then

$$\exp\left(AXA^{-1}\right) \quad = \quad A\exp(X)A^{-1}.$$

6. We have

$$\frac{d}{dt}\exp(tX) \quad = \quad X\exp(tX) = \exp(tX)X.$$

In particular

$$\left.\frac{d}{dt}\exp(tX)\right|_{t=0} \quad = \quad X.$$

Point 1 is obvious. Points 2 and 3 are special cases of 4. To prove point 4 it is essential that X and Y commute:

$$\begin{aligned}
\exp X \exp Y &= \sum_{i=0}^{\infty}\frac{X^i}{i!}\sum_{j=0}^{\infty}\frac{Y^j}{j!} = \sum_{n=0}^{\infty}\sum_{i=0}^{n}\frac{X^i}{i!}\frac{Y^{n-i}}{(n-i)!} = \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{i=0}^{n}\binom{n}{i}X^iY^{n-i} \\
&= \sum_{n=0}^{\infty}\frac{1}{n!}(X+Y)^n = \exp(X+Y).
\end{aligned}$$

Proof of point 5:

$$\exp\left(AXA^{-1}\right) \quad = \quad \sum_{n=0}^{\infty}\frac{1}{n!}\left(AXA^{-1}\right)^n = \sum_{n=0}^{\infty}\frac{1}{n!}AX^nA^{-1} = A\exp(X)A^{-1}.$$

Proof of point 6:

$$\frac{d}{dt}\exp(tX) \quad = \quad \frac{d}{dt}\sum_{n=0}^{\infty}\frac{1}{n!}(tX)^n = \sum_{n=0}^{\infty}\frac{1}{(n-1)!}t^{n-1}X^n = X\exp(tX) = \exp(tX)X.$$

Computation of the exponential of a matrix:

Case 1: X is diagonalisable.

If $X = ADA^{-1}$ with $D = \mathrm{diag}\,(\lambda_1, \lambda_2, ...)$ we have

$$\exp X = \exp ADA^{-1} = A \exp(D) A^{-1} = A\,\mathrm{diag}\left(e^{\lambda_1}, e^{\lambda_2}, ...\right) A^{-1}.$$

Case 2: X is nilpotent.

A matrix X is called nilpotent, if $X^m = 0$ for some positive m. In this case the series terminates:

$$\exp X = \sum_{n=0}^{m-1} \frac{X^n}{n!}$$

Case 3: X is arbitrary.

A general matrix X may be neither diagonalisable nor nilpotent. However, any matrix X can uniquely be written as

$$X = S + N,$$

where S is diagonalisable and N is nilpotent and $SN = NS$. Then

$$\exp X = \exp S \exp N$$

and $\exp S$ and $\exp N$ can be computed as in the previous two cases.

3.4.3 Relation between Lie algebras and Lie groups

Let G be a Lie group. Assume that as a manifold it has dimension n. G is also a group. Choose a local coordinate system, such that the identity element e is given by

$$e = g(0, ..., 0).$$

A lot of information on G can be obtained from the study of G in the neighbourhood of e. Let

$$g(\theta_1, ..., \theta_n)$$

denote a general point in the local chart containing e. Let us write

$$\begin{aligned} g(0, ..., \theta_a, ..., 0) &= g(0, ..., 0, ..., 0) + \theta_a X^a + O(\theta^2) \\ &= g(0, ..., 0, ..., 0) - i\theta_a T^a + O(\theta^2). \end{aligned}$$

We also have

$$\begin{aligned} X^a &= \lim_{\theta_a \to 0} \frac{g(0, ..., \theta_a, ..., 0) - g(0, ..., 0, ..., 0)}{\theta_a}, \\ T^a &= i \lim_{\theta_a \to 0} \frac{g(0, ..., \theta_a, ..., 0) - g(0, ..., 0, ..., 0)}{\theta_a}. \end{aligned}$$

The T^a's (and the X^a's) are called the **generators** of the Lie group G.

Theorem: The commutators of the generators T^a of a Lie group are linear combinations of the generators and satisfy a Lie algebra.

$$\left[T^a, T^b\right] = i \sum_{c=1}^{n} f^{abc} T^c.$$

We will often use Einstein's summation convention and simply write

$$\left[T^a, T^b\right] = i f^{abc} T^c.$$

In order to proove this theorem we have to show that the commutator is again a linear combination of the generators. We start with the definition of a one-parameter subgroup of $GL(n,\mathbb{C})$: A map $g : \mathbb{R} \to GL(n,\mathbb{C})$ is called a one-parameter sub-group of $GL(n,\mathbb{C})$ if
1. $g(t)$ is continous.

2. $g(0) = 1$.

3. For $t_1, t_2 \in \mathbb{R}$ we have

$$g(t_1 + t_2) = g(t_1)\, g(t_2).$$

If $g(t)$ is a one-parameter sub-group of $GL(n,\mathbb{C})$ then there exists a unique $n \times n$ matrix X such that

$$g(t) = \exp(tX).$$

X is given by

$$X = \left. \frac{d}{dt} g(t) \right|_{t=0}.$$

There is a one-to-one correspondence between linear combinations of the generators

$$X = -i\theta_a T^a$$

and the one-parameter sub-groups

$$g(t) = \exp(tX) \quad \text{with} \quad X = \left. \frac{d}{dt} g(t) \right|_{t=0}.$$

If $A \in G$ and if Y defines a one-parameter sub-group of G, then also AYA^{-1} defines a one-parameter sub-group of G. The non-trivial point is to check that $\exp\left[t\left(AYA^{-1}\right)\right]$ is again in G. This follows from

$$\exp\left[t\left(AYA^{-1}\right)\right] = A \exp(tY)\, A^{-1}.$$

Therefore AYA^{-1} is a linear combination of the generators. Now we take for $A = \exp(\lambda X)$. This implies that

$$\exp(\lambda X)\, Y \exp(-\lambda X)$$

is a linear combination of the generators. Since the vector space spanned by the generators is topologically closed, also the derivative with respect to λ belongs to this vector space and we have shown that

$$\left.\frac{d}{d\lambda}\exp(\lambda X)\, Y \exp(-\lambda X)\right|_{\lambda=0} = XY - YX = [X,Y]$$

is again a linear combination of the generators.

We have seen that by studying a Lie group G in the neighbourhood of the identity we can obtain from the Lie group G the corresponding Lie algebra $\mathfrak{g}$. We can now ask if the converse is also true: Given the Lie algebra $\mathfrak{g}$, can we reconstruct the Lie group G ? The answer is that this can almost be done. Note that a Lie group need not be connected. The Lorentz group is an example of a Lie group which is not connected. Given a Lie algebra we have information about the connected component in which the idenity lies. The exponential map takes us from the Lie algebra into the group. In the neighbourhood of the identity we have

$$g(\theta_1, ..., \theta_n) = \exp\left(-i \sum_{a=1}^{n} \theta_a T^a\right).$$

3.4.4 Examples

As an example for the generators of a group let us study the cases of $SU(2)$ and $SU(3)$, as well as the groups $U(2)$ and $U(3)$. A common normalisation for the generators is

$$\mathrm{Tr}\, T^a T^b = \frac{1}{2}\delta^{ab}.$$

a) The group $SU(2)$ is a three-paramter group. The generators are proportional to the Pauli matrices:

$$T^1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad T^2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad T^3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

b) The group $SU(3)$ has eight parameters. The generators can be taken as the Gell-Mann matrices:

$$T^1 = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T^2 = \frac{1}{2}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad T^3 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$T^4 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad T^5 = \frac{1}{2}\begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad T^6 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$T^7 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad T^8 = \frac{1}{2\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

c) For the groups $U(2)$ and $U(3)$ add the generator

$$T^0 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

for $U(2)$, respectively the generator

$$T^0 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

for $U(3)$.

3.4.5 The Fierz identity

Problem: Denote by T^a the generators of $SU(n)$ or $U(n)$. Evaluate traces like

$$\mathrm{Tr}\, T^a T^b T^a T^b,$$

where a sum over a and b is implied.

We first consider the case for $SU(N)$. The Fierz identity reads for $SU(N)$:

$$T^a_{ij} T^a_{kl} = \frac{1}{2}\left(\delta_{il}\delta_{jk} - \frac{1}{N}\delta_{ij}\delta_{kl}\right).$$

Proof: T^a and the unit matrix form a basis of the $N \times N$ hermitian matrices, therefore any hermitian matrix A can be written as

$$A = c_0 1 + c_a T^a.$$

The constants c_0 and c_a are determined using the normalization condition and the fact that the T^a are traceless. We first take the trace on both sides:

$$\mathrm{Tr}\,(A) \;=\; c_0\,\mathrm{Tr}\,1 + c_a\,\mathrm{Tr}\,T^a = c_0 N,$$

therefore

$$c_0 \;=\; \frac{1}{N}\mathrm{Tr}\,(A)\,.$$

Now we multiply first both sides with T^b and take then the trace:

$$\mathrm{Tr}\left(AT^b\right) \;=\; c_0\,\mathrm{Tr}\,T^b + c_a\,\mathrm{Tr}\,T^a T^b = c_a \frac{1}{2}\delta^{ab},$$

therefore

$$c_a \;=\; 2\mathrm{Tr}\,(T^a A)\,.$$

Putting both results together we obtain

$$A \;=\; \frac{1}{N}\mathrm{Tr}\,(A)\,1 + 2\mathrm{Tr}\,(AT^a)\,T^a$$

Let us now write this equation in components

$$\begin{aligned} A_{ij} &= \frac{1}{N}\mathrm{Tr}\,(A)\,1_{ij} + 2\mathrm{Tr}\,(AT^a)\,T^a_{ij}, \\ A_{ij} &= \frac{1}{N}A_{ll}1_{ij} + 2A_{lk}T^a_{kl}T^a_{ij}, \end{aligned}$$

Therefore

$$A_{lk}\left(2T^a_{ij}T^a_{kl} + \frac{1}{N}\delta_{ij}\delta_{kl} - \delta_{il}\delta_{jk}\right) \;=\; 0.$$

This has to hold for an arbitrary A, therefore the Fierz identity follows. Useful formulae involving traces:

$$\begin{aligned} \mathrm{Tr}\,(T^a X)\,\mathrm{Tr}\,(T^a Y) &= \frac{1}{2}\left[\mathrm{Tr}\,(XY) - \frac{1}{N}\mathrm{Tr}\,(X)\,\mathrm{Tr}\,(Y)\right], \\ \mathrm{Tr}\,(T^a X T^a Y) &= \frac{1}{2}\left[\mathrm{Tr}\,(X)\,\mathrm{Tr}\,(Y) - \frac{1}{N}\mathrm{Tr}\,(XY)\right]. \end{aligned}$$

In the case of a $U(N)$-group the identity matrix is part of the generators and the Fierz identity takes the simpler form

$$T^a_{ij}T^a_{kl} \;=\; \frac{1}{2}\delta_{il}\delta_{jk}.$$

As a consequence we have for a $U(N)$-group for the traces

$$\begin{aligned}\mathrm{Tr}\left(T^a X\right)\mathrm{Tr}\left(T^a Y\right) &= \frac{1}{2}\mathrm{Tr}\left(XY\right),\\ \mathrm{Tr}\left(T^a X T^a Y\right) &= \frac{1}{2}\mathrm{Tr}\left(X\right)\mathrm{Tr}\left(Y\right).\end{aligned}$$

It is also useful to know, that the structure constants f^{abc} can expressed in terms of traces over the generators: From

$$\left[T^a, T^b\right] = i f^{abc} T^c$$

one derives by multiplying with T^d and taking the trace:

$$i f^{abc} = 2\left[\mathrm{Tr}\left(T^a T^b T^c\right) - \mathrm{Tr}\left(T^b T^a T^c\right)\right]$$

This yields an expression of the structure constants in terms of the matrices of the fundamental representation. We can now calculate for the group $SU(N)$ the fundamental and the adjoint Casimirs:

$$\begin{aligned}\left(T^a T^a\right)_{ij} &= C_F \delta_{ij} = \frac{N^2-1}{2N}\delta_{ij},\\ f^{abc} f^{dbc} &= C_A \delta^{ad} = N\delta^{ad}.\end{aligned}$$

For the group $SU(N)$ we define the symmetric tensor d^{abc} through

$$\left\{T^a, T^b\right\} = d^{abc} T^c + \frac{1}{N}\delta^{ab}.$$

Here, $\{...,...\}$ denotes the anti-commutator

$$\{A, B\} = AB + BA.$$

With the same steps as above one finds that

$$d^{abc} = 2\left[\mathrm{Tr}\left(T^a T^b T^c\right) + \mathrm{Tr}\left(T^b T^a T^c\right)\right].$$

From this expression and the fact that the trace is cyclic we see explicitly that d^{abc} is symmetric in all indices.

3.5 Gauge symmetries

Lie groups play an essential role in describing internal symmetries in physics. The simplest example is given by electrodynamics. We denote by $A_\mu(x)$ the gauge potential of electrodynamics and by

$$F_{\mu\nu} = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)$$

the field strength tensor. We denote further by $U(x)$ an element of a $U(1)$-Lie group, smoothly varying with x. We may write

$$U(x) = \exp(-i\theta(x)),$$

where $\theta(x)$ is a real smooth function. Recall that the group $U(1)$ has one generator, which is the 1×1 unit matrix. This is simply one and we don't write it explicitly. We then consider the transformation

$$A'_\mu(x) = U(x)\left(A_\mu(x) + \frac{i}{e}\partial_\mu\right)U(x)^\dagger.$$

Here, e denotes the electric charge. Working out the expression for $A'_\mu(x)$ in terms of $\theta(x)$ we find

$$A'_\mu(x) = A_\mu(x) - \frac{1}{e}\partial_\mu\theta(x).$$

This is nothing else than a gauge transformation $A'_\mu(x) = A_\mu(x) - \partial_\mu\chi(x)$ with $\chi(x) = \theta(x)/e$. The field strength transforms as

$$F'_{\mu\nu} = UF_{\mu\nu}U^\dagger.$$

For a $U(1)$-group U and $U^\dagger$ commute with $F_{\mu\nu}$ and we get

$$F'_{\mu\nu} = F_{\mu\nu}.$$

Therefore for a $U(1)$-transformation the field strength is invariant. As a consequence, also the Lagrange density

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$$

is invariant:

$$\mathcal{L}' = \mathcal{L}.$$

Now let us see if this can be generalised to $SU(N)$. In electrodynamics we can interpret the gauge potential as a quantity, which takes values in the Lie algebra of $U(1)$. There is only one generator for the $U(1)$-group. In the group $SU(N)$ there are more generators (N^2-1 to be precise), and we start from the following generalisation of the gauge potential:

$$T^a A^a_\mu(x),$$

where the T^a's are the generators of the Lie algebra of $SU(N)$ and a ranges from 1 to N^2-1. We consider again a transformation of the form

$$T^a A^{a\prime}_\mu(x) = U(x)\left(T^a A^a_\mu(x) + \frac{i}{g}\partial_\mu\right)U(x)^\dagger,$$

where $U(x)$ is now an element of $SU(N)$. (g is a coupling constant replacing the electromagnetic coupling e). $U(x)$ can be written in terms of the generators as

$$U(x) = \exp\left(-i\theta_a(x)T^a\right),$$

The group $SU(N)$ is non-Abelian and the generators T^a do not commute. As a consequence the expression

$$\partial_\mu\left(T^a A_\nu^a\right) - \partial_\nu\left(T^a A_\mu^a\right)$$

does not transform nicely. As a short-hand notation we write in the following

$$A_\mu = T^a A_\mu^a.$$

We have

$$\begin{aligned} \partial_\mu A_\nu{}' &= \partial_\mu\left[U\left(A_\nu + \frac{i}{g}\partial_\nu\right)U^\dagger\right] = \partial_\mu\left[UA_\nu U^\dagger + \frac{i}{g}U\left(\partial_\nu U^\dagger\right)\right] \\ &= U\left(\partial_\mu A_\nu\right)U^\dagger + \left(\partial_\mu U\right)A_\nu U^\dagger + UA_\nu\left(\partial_\mu U^\dagger\right) + \frac{i}{g}\left(\partial_\mu U\right)\left(\partial_\nu U^\dagger\right) + \frac{i}{g}U\left(\partial_\mu\partial_\nu U^\dagger\right) \end{aligned}$$

Since $UU^\dagger = 1$ we have

$$0 = \partial_\mu\left(UU^\dagger\right) = \left(\partial_\mu U\right)U^\dagger + U\left(\partial_\mu U^\dagger\right)$$

and hence

$$\left(\partial_\mu U\right)U^\dagger = -U\left(\partial_\mu U^\dagger\right).$$

We can use this relation to rewrite

$$\begin{aligned} \partial_\mu A_\nu{}' &= U\left(\partial_\mu A_\nu\right)U^\dagger - \left(U\partial_\mu U^\dagger\right)\left(UA_\nu U^\dagger\right) + \left(UA_\nu U^\dagger\right)\left(U\partial_\mu U^\dagger\right) \\ &\quad -\frac{i}{g}\left(U\partial_\mu U^\dagger\right)\left(U\partial_\nu U^\dagger\right) + \frac{i}{g}U\left(\partial_\mu\partial_\nu U^\dagger\right) \end{aligned}$$

We then find that

$$\begin{aligned} \partial_\mu A_\nu{}' - \partial_\nu A_\mu{}' &= U\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right)U^\dagger \\ &\quad -\left[UA_\mu U^\dagger, U\partial_\nu U^\dagger\right] - \left[U\partial_\mu U^\dagger, UA_\nu U^\dagger\right] - \frac{i}{g}\left[U\partial_\mu U^\dagger, U\partial_\nu U^\dagger\right]. \end{aligned}$$

If we now define the field strength in the non-Abelian case by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig\left[A_\mu, A_\nu\right]$$

we obtain the transformation law

$$F'_{\mu\nu} = UF_{\mu\nu}U^\dagger.$$

We now define the Lagrange density as

$$\mathcal{L} \;=\; -\frac{1}{2}\,\mathrm{Tr}\,F_{\mu\nu}F^{\mu\nu}.$$

This Lagrange density is invariant under $SU(N)$-gauge transformations:

$$\mathcal{L}' \;=\; -\frac{1}{2}\,\mathrm{Tr}\,F'_{\mu\nu}F^{\mu\nu\prime} = -\frac{1}{2}\,\mathrm{Tr}\,\left(UF_{\mu\nu}U^\dagger\right)\left(UF^{\mu\nu}U^\dagger\right) = -\frac{1}{2}\,\mathrm{Tr}\,F_{\mu\nu}F^{\mu\nu} = \mathcal{L}.$$

Going back from our short-notation to the more detailed notation

$$A_\mu \;=\; T^a A^a_\mu,$$

one denotes the field strength also by

$$F^a_{\mu\nu} \;=\; \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu.$$

Obviously, we have the relation

$$F_{\mu\nu} \;=\; T^a F^a_{\mu\nu}.$$

In terms of $F^a_{\mu\nu}$ the Lagrange density reads

$$\mathcal{L} \;=\; -\frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu}.$$

4 Representation theory

4.1 Group actions

An action of a group G on a set X is a correspondence that associates to each element $g \in G$ a map $\phi_g : X \to X$ in such a way that

$$\begin{array}{l} \phi_{g_1 g_2} = \phi_{g_1}\phi_{g_2}, \\ \phi_e \;\text{ is the identity map on } X, \end{array}$$

where e denotes the neutral element of the group. Instead of $\phi_g(x)$ one often writes gx.

A group action of G on X gives rise to a natural equivalence relation on X: $x_1 \in X$ and $x_2 \in X$ are equivalent, if they can be obtained from one another by the action of some group element $g \in G$. The equivalence class of a point $x \in X$ is called the **orbit** of x.

G is said to act **effectively** on X, if the homomorphism from G into the group of transformations of X is injective.

G is said to act **transitively** on X, if there is only one orbit. A set X where a group G acts transitively is called a homogeneous space. Every orbit of a (not necessarily transitive) group action is a homogeneous space.

The stabilizer (or the isotropy subgroup or the little group) H_x of a point $x \in X$ is the subgroup of G that leave x fixed, e.g. $h \in H_x$ if $hx = x$. When H_x is the trivial subgroup for all $x \in X$, we say that the action of G on X is **free**.

If G acts on X and on Y, then a map $\psi : X \to Y$ is said to be G**-equivariant** if $\psi \circ g = g \circ \psi$ for all $g \in G$.

4.2 Representations

Let V be a finite-dimensional vector space and $GL(V)$ the group of automorphisms of V. Typically $V = \mathbb{R}^n$ or $V = \mathbb{C}^n$ and $GL(V) = GL(n,\mathbb{R})$ or $GL(V) = GL(n,\mathbb{C})$.

Definition: A representation of a goup G is a homomorphism ρ from G to $GL(V)$

$$g \;\to\; \rho(g).$$

The composition in $GL(V)$ is given by matrix multiplication. Since ρ is a homomorphism we have

$$\rho(g_1 g_2) \;=\; \rho(g_1)\rho(g_2).$$

This implies

$$\begin{aligned} \rho(e) &= 1, \\ \rho\left(g^{-1}\right) &= [\rho(g)]^{-1}. \end{aligned}$$

The trivial representation:

$$\rho(g) = 1, \quad \forall g.$$

Remark: In general more than one group element can be mapped on the identity. If the mapping $\rho : G \to GL(V)$ is one-to-one, i.e.

$$\rho(g) = \rho(g') \quad \Rightarrow \quad g = g'$$

then the representation is called **faithful**.

Strictly speaking a representation is a set of (non-singular) matrices, e.g. a sub-set of $GL(V)$. Very often we will also speak about the vector space V, on which these matrices act, as a representation of G.

In this sense a **sub-representation** of a representation V is a vector sub-space W of V, which is invariant under G:

$$\rho(g)w \in W \ \ \forall g \in G \text{ and } w \in W.$$

A representation V is called **irreducible** if there is no proper non-zero invariant sub-space W of V. (This excludes the trivial invariant sub-spaces $W = \{0\}$ and $W = V$.)

If V_1 and V_2 are representations of G, the **direct sum** $V_1 \oplus V_2$ and the **tensor product** $V_1 \otimes V_2$ are again representations:

$$\begin{aligned} g\left(v_1 \oplus v_2\right) &= \left(gv_1\right) \oplus \left(gv_2\right), \\ g\left(v_1 \otimes v_2\right) &= \left(gv_1\right) \otimes \left(gv_2\right), \end{aligned}$$

Two representations ρ_1 and ρ_2 of the same dimension are called **equivalent**, if there exists a non-singular matrix S such that

$$\rho_1(g) = S\rho_2(g)S^{-1}, \quad \forall g \in G.$$

For finite groups and compact Lie groups it can be shown that any representation is equivalent to a unitary representation.
For finite groups the proof goes as follows: Suppose we start with an arbitrary (i.e. not necessarily unitary) representation $\rho_2(g)$. We would like to find a S such that $\rho_1(g) = S\rho_2(g)S^{-1}$ is a unitary matrix for all g. We set S to be a hermitian matrix which satisfies

$$S^2 = \sum_{g' \in G} \rho_2\left(g'\right)^{\dagger} \rho_2\left(g'\right).$$

Then

$$\begin{aligned}\rho_2(g)^\dagger S^2 \rho_2(g) &= \rho_2(g)^\dagger \sum_{g' \in G} \rho_2(g')^\dagger \rho_2(g') \rho_2(g) = \sum_{g' \in G} \rho_2(g'g)^\dagger \rho_2(g'g) \\ &= \sum_{g'' \in G} \rho_2(g'')^\dagger \rho_2(g'') = S^2.\end{aligned}$$

We therefore have

$$\begin{aligned}\rho_2(g)^\dagger S^2 \rho_2(g) &= S^2, \\ \rho_2(g)^\dagger S^2 &= S^2 \rho_2(g)^{-1}, \\ S^{-1} \rho_2(g)^\dagger S &= S \rho_2(g)^{-1} S^{-1}, \\ \left(S \rho_2(g) S^{-1}\right)^\dagger &= \left(S \rho_2(g) S^{-1}\right)^{-1},\end{aligned}$$

This shows that $\rho_1(g) = S\rho_2(g)S^{-1}$ satisfies

$$\rho_1(g)^\dagger = \rho_1(g)^{-1},$$

in other words, $\rho_1(g)$ is a unitary matrix.

This proof carries over to the case of compact Lie groups by replacing the sum in the definition of S by an integration over all group elements.

The goal of representation theory: Classify and study all representations of a group G up to equivalence. This will be done by decomposing an arbitrary representation into direct sums of irreducible representations.

4.3 Schur's lemmas

Lemma 1: Any matrix M which commutes with all the matrices $\rho(g)$ of an irreducible representation of a group G must be a multiple of the unit matrix:

$$M = c\mathbf{1}.$$

Proof: We have

$$\rho(g)M = M\rho(g) \quad \forall g \in G.$$

If $\rho(g)$ is of dimension n, then M must be square of dimension n. Let us assume that $\rho(g)$ is unitary. Then

$$M^\dagger \rho(g)^\dagger = \rho(g)^\dagger M^\dagger.$$

Multiply by $\rho(g)$ from left and right:

$$\rho(g)M^\dagger = M^\dagger \rho(g).$$

Therefore also $M^\dagger$ commutes with all $\rho(g)$, and so do the hermitian matrices

$$\begin{aligned} H_1 &= M + M^\dagger, \\ H_2 &= i\left(M - M^\dagger\right). \end{aligned}$$

Any hermitian matrix may be diagonalised by a unitary transformation:

$$D = U^{-1} H U.$$

If we define now

$$\rho'(g) = U^{-1} \rho(g) U,$$

we have

$$\rho'(g) D = D \rho'(g).$$

Let $D = \text{diag}(\lambda_1, ..., \lambda_n)$ and consider now the i, j element of this matrix equation:

$$\begin{aligned} \left[\rho'(g)\right]_{ij} \lambda_j &= \lambda_i \left[\rho'(g)\right]_{ij}, \\ \left(\lambda_i - \lambda_j\right) \left[\rho'(g)\right]_{ij} &= 0. \end{aligned}$$

Suppose that a certain eigenvalue λ of D occurs k times and that, by a suitable ordering the first k positions of D are occupied by λ. Then

$$\lambda_1 = ... = \lambda_k \neq \lambda_l, \;\; k+1 \leq l \leq n.$$

This implies that

$$\left[\rho'(g)\right]_{ij} = 0 \qquad \begin{array}{l} \text{for } 1 \leq i \leq k, \;\; k+1 \leq j \leq n, \\ \text{or } 1 \leq j \leq k, \;\; k+1 \leq i \leq n. \end{array}$$

Hence $\rho'(g)$ is of the form

$$\left(\begin{array}{cc} ... & 0 \\ 0 & ... \end{array} \right)$$

and is thus reducible, contrary to the initial assumption. Thus if and only if all the eigenvalues of D are the same $\rho'(g)$ will be irreducible. In other words, D and hence M must be a multiple of the unit matrix.

Lemma 2: If $\rho_1(g)$ and $\rho_2(g)$ are two irreducible representations of a group G of dimensions n_1 and n_2 respectively and if a rectangular matrix M of dimension $n_1 \times n_2$ exists sucht that

$$\rho_1(g) M = M \rho_2(g), \;\; \forall g \in G$$

then either
(a) $M = 0$ or
(b) $n_1 = n_2$ and $\det M \neq 0$, in which case $\rho_1(g)$ and $\rho_2(g)$ are equivalent.

Proof: Let us assume without loss of generality that $\rho_1(g)$ and $\rho_2(g)$ are unitary representations.

$$\begin{aligned} M^\dagger \rho_1(g)^\dagger &= \rho_2(g)^\dagger M^\dagger, \\ M^\dagger \rho_1(g^{-1}) &= \rho_2(g^{-1}) M^\dagger. \end{aligned}$$

Multiply by M from the right:

$$M^\dagger \rho_1(g^{-1}) M = \rho_2(g^{-1}) M^\dagger M.$$

By assumption $\rho_1(g^{-1}) M = M \rho_2(g^{-1})$ and therefore

$$M^\dagger M \rho_2(g^{-1}) = \rho_2(g^{-1}) M^\dagger M.$$

By lemma 1 we conclude

$$M^\dagger M = \lambda \mathbf{1}.$$

Consider the case $n_1 = n_2 = n$:

$$\det M^\dagger M = \det M^\dagger \det M = \lambda^n.$$

If $\lambda \neq 0$ then $\det M \neq 0$ and therefore M^{-1} exists. From $\rho_1(g) M = M \rho_2(g)$ it follows that

$$\rho_1(g) = M \rho_2(g) M^{-1}$$

and $\rho_1(g)$ and $\rho_2(g)$ are equivalent.

If on the other hand $\lambda = 0$ we have

$$\begin{aligned} \sum_k M^\dagger_{ik} M_{ki} &= 0, \\ \sum_k |M_{ki}|^2 &= 0. \end{aligned}$$

This is only possible for $M_{ki} = 0$ and hence

$$M = 0.$$

To complete the proof we consider the case $n_1 \neq n_2$. Let us assume $n_1 < n_2$. Construct M' from M by adding $n_2 - n_1$ rows of zeros:

$$M' = \begin{pmatrix} M \\ 0 \end{pmatrix}$$

$$M'^{\dagger} = \left(M^{\dagger} \quad 0 \right)$$

We have

$$M'^{\dagger} M' = M^{\dagger} M$$

and thus

$$\det M^{\dagger} M = \det M'^{\dagger} M' = \det M'^{\dagger} \det M' = 0.$$

Hence $\lambda = 0$ and $M^{\dagger} M = 0$. It follows $M = 0$ as before.

Application: Orthogonality theorem for finite groups. Let G be a finite group and let ρ_1 and ρ_2 be irreducible representations of dimension n_1 and n_2. Then

$$\sum_{g \in G} \rho_1(g)_{ij} \rho_2(g^{-1})_{kl} = \begin{cases} 0 & \rho_1 \text{ and } \rho_2 \text{ are inequivalent,} \\ \frac{|G|}{n_1} \delta_{il} \delta_{kj} & \rho_1 \text{ and } \rho_2 \text{ are identical,} \\ \dots & \rho_1 \text{ and } \rho_2 \text{ are equivalent, but not identical.} \end{cases}$$

Proof: Assume that ρ_1 and ρ_2 are inequivalent. Consider

$$M = \frac{1}{|G|} \sum_{g \in G} \rho_1(g) X \rho_2(g^{-1}),$$

where X is an arbitrary $n_1 \times n_2$ matrix. Then

$$\begin{aligned} \rho_1(g')M &= \rho_1(g') \frac{1}{|G|} \sum_{g \in G} \rho_1(g) X \rho_2(g^{-1}) = \frac{1}{|G|} \sum_{g \in G} \rho_1(g'g) X \rho_2(g^{-1}) \\ &= \frac{1}{|G|} \sum_{g \in G} \rho_1(g) X \rho_2(g^{-1}g') = \frac{1}{|G|} \sum_{g \in G} \rho_1(g) X \rho_2(g^{-1}) \rho_2(g') = M \rho_2(g'). \end{aligned}$$

By Schur's second lemma we have $M = 0$, therefore

$$\frac{1}{|G|} \sum_{g \in G} \rho_1(g)_{ij'} X_{j'k'} \rho_2(g^{-1})_{k'l} = 0.$$

Since X was arbitrary we can take $X = \delta_{jj'} \delta_{kk'}$ and we have

$$\sum_{g \in G} \rho_1(g)_{ij} \rho_2(g^{-1})_{kl} = 0.$$

Now consider the case where ρ_1 and ρ_2 are identical: $\rho_1 = \rho_2 = \rho$. Take again

$$M = \frac{1}{|G|} \sum_{g \in G} \rho(g) X \rho(g^{-1}).$$

One shows again

$$\rho(g)M = M\rho(g).$$

Therefore by Schur's first lemma

$$\frac{1}{|G|}\sum_{g\in G}\rho(g)_{ij'}X_{j'k'}\rho(g^{-1})_{k'l} = c\delta_{il}.$$

Again take $X = \delta_{jj'}\delta_{kk'}$:

$$\frac{1}{|G|}\sum_{g\in G}\rho(g)_{ij}\rho(g^{-1})_{kl} = c\delta_{il}.$$

To find c take the trace on both sides:

$$\delta_{kj} = cn_1,$$

and therefore

$$\sum_{g\in G}\rho(g)_{ij}\rho(g^{-1})_{kl} = \frac{|G|}{n_1}\delta_{kj}\delta_{il}.$$

Another consequence of Schur's first lemma: All irreducible representation of an Abelian group are one-dimensional.

4.4 Representation theory for finite groups

A finite group G admits only finitely many irreducible representations V_i up to isomorphism.

Example: Consider the symmetric group S_3, the permutation group of three elements, which is the simplest non-abelian group. This group has two one-dimensional representations: The trivial one I and the alternating representation A defined by

$$gv = \text{sign}(g)v.$$

There is a natural representation, in which S_3 acts on $\mathbb{C}^3$ by

$$g\cdot(z_1,z_2,z_3) = \left(z_{g^{-1}(1)},z_{g^{-1}(2)},z_{g^{-1}(3)}\right)$$

This representation is reducible: The line spanned by the sum

$$e_1+e_2+e_3$$

is an invariant sub-space. The complementary sub-space

$$V = \{(z_1,z_2,z_3)|z_1+z_2+z_3=0\}$$

defines an irreducible representation. This representation is called the standard representation.
It can be shown that any representation of S_3 can be decomposed into these three irreducible representations

$$W = I^{\oplus n_1}\oplus A^{\oplus n_2}\oplus V^{\oplus n_3}.$$

4.4.1 Characters

Definition:If V is a representation of G, its character χ_V is the complex-valued function on the group defined by

$$\chi_V(g) = \mathrm{Tr}(\rho(g)).$$

In particular we have

$$\chi_V\left(hgh^{-1}\right) = \chi_V(g).$$

Let V and W be representation of G. Then

$$\begin{aligned}
\chi_{V\oplus W} &= \chi_V + \chi_W, \\
\chi_{V\otimes W} &= \chi_V \cdot \chi_W, \\
\chi_{V^*} &= (\chi_V)^*, \\
\chi_{\wedge^2 V}(g) &= \frac{1}{2}\left[\chi_V(g)^2 - \chi_V(g^2)\right], \\
\chi_{\mathrm{Sym}^2 V}(g) &= \frac{1}{2}\left[\chi_V(g)^2 + \chi_V(g^2)\right]
\end{aligned}$$

Orthogonality theorem for characters: For finite groups we had the orthogonality theorem. If we consider unitary representations and if we make the agreement that if two representations are equivalent, we take them to be identical, the orthogonality theorem can be written as

$$\sum_{g\in G} \rho_\alpha(g)_{ij}\rho_\beta(g)^*_{lk} = \frac{|G|}{n_1}\delta_{il}\delta_{kj}\delta_{\alpha\beta}$$

(Note that for an unitary representation we have $\rho(g^{-1})_{kl} = \rho(g)^*_{lk}$.) Now we set $i = j$ and sum, and we set $l = k$ and sum:

$$\sum_{g\in G} \chi_\alpha(g)\chi_\beta(g)^* = |G|\,\delta_{\alpha\beta}$$

Since the character is a class function we can write

$$\sum_{g\in G} = \sum_{\text{classes }\kappa} n_\kappa,$$

where n_κ denotes the number of elements in the class C_κ. Therefore

$$\sum_\kappa n_\kappa \chi_\alpha(C_\kappa)\chi_\beta(C_\kappa)^* = |G|\,\delta_{\alpha\beta}$$

Character table:

	$n_1 C_1$	$n_2 C_2$	$n_2 C_2$	...
ρ_1	$\chi_1(C_1)$	$\chi_1(C_2)$	$\chi_1(C_3)$	...
ρ_2	$\chi_2(C_1)$	$\chi_2(C_2)$	$\chi_2(C_3)$	...
ρ_3	$\chi_3(C_1)$	$\chi_3(C_2)$	$\chi_3(C_3)$	...
...	...	...	...	...

The number of orthogonal vectors corresponds to the number of inequivalent representations. The dimension of the space is given by the number of classes. Therefore the number of inequivalent representations is smaller or equal to the number of classes. In fact equality holds. To show this, we consider one specific (reducible) representation, called the **regular representation**. The regular representation is defined by

$$g_a g_b = \sum_c \rho^R_{cb}(g_a)\, g_c$$

Note that the matrix $\rho^R_{cb}(g_a)$ has in each column exactly one 1 and all other entries in this column are zero. This defines a representation: We have

$$\begin{aligned} g_d g_a g_b &= g_d \sum_c \rho^R_{cb}(g_a)\, g_c = \sum_e \rho^R_{ec}(g_d) \sum_c \rho^R_{cb}(g_a)\, g_e \\ &= \sum_e \left(\sum_c \rho^R_{ec}(g_d)\, \rho^R_{cb}(g_a) \right) g_e. \end{aligned}$$

On the other hand we have

$$g_d g_a g_b = (g_d g_a)\, g_b = \sum_e \rho^R_{eb}(g_d g_a)\, g_e.$$

Therefore it follows

$$\rho^R_{eb}(g_d g_a) = \sum_c \rho^R_{ec}(g_d)\, \rho^R_{cb}(g_a),$$

or in matrix notation

$$\rho^R(g_d g_a) = \rho^R(g_d) \cdot \rho^R(g_a).$$

We have already seen that the 1 appears in each column of the matrix $\rho^R(g)$ exactly once. The 1 appears always on the diagonal if $g = e$, otherwise it appears never on the diagonal. This implies for the character

$$\chi^R(g) = \begin{cases} 0 & g \neq e, \\ |G| & g = e. \end{cases}$$

In general, the regular representation is reducible. We write

$$\rho^R(g) = \bigoplus_\alpha a_\alpha \rho_\alpha(g),$$

where the sum is over all irreducible representations and a_α gives the number of times the irreducible representation ρ_α is contained in ρ^R. For the characters we have then

$$\chi^R(g) \quad = \quad \sum_\alpha a_\alpha \chi_\alpha(g).$$

The number a_α we can obtain from the orthogonality theorem as follows

$$\frac{1}{|G|}\sum_\kappa n_\kappa \chi_\alpha \chi^R(g)^* \quad = \quad \frac{1}{|G|}\sum_\kappa n_\kappa \chi_\alpha \sum_\beta a_\beta \chi_\beta(g)^* = \sum_\beta a_\beta \delta_{\alpha\beta} = a_\alpha.$$

Using the fact that $\chi^R(g) = 0$ for all $g \neq e$ we have on the other side

$$\frac{1}{|G|}\sum_\kappa n_\kappa \chi_\alpha \chi^R(g)^* \quad = \quad \frac{1}{|G|}\chi_\alpha(e)|G| = \chi_\alpha(e) = n_\alpha,$$

where n_α is the dimension of the irreducible representation ρ_α. Thus

$$a_\alpha \quad = \quad n_\alpha.$$

The irreducible representation ρ_α appears exactly n_α times in the decomposition of ρ^R. In particular, each irreducible representation appears in the decomposition of the regular representation. We also have

$$|G| \quad = \quad \sum_\alpha a_\alpha n_\alpha = \sum_\alpha n_\alpha^2.$$

From the orthogonality theorem for finite groups it follows that we can view $\rho_\alpha(g)_{ij}$ for fixed α and fixed i, j as a vector in a $|G|$-dimensional space. The orthogonality theorem tells us, that there are

$$\sum_\alpha n_\alpha^2$$

orthogonal vectors. On the other hand we have just shown that this number equals $|G|$, therefore the vectors $\rho_\alpha(g)_{ij}$ span the full space. Therefore any vector in this space can be written as a linear combination of these basis vectors. In particular the component v_a with respect to the standard bases is given by

$$v_a \quad = \quad \sum_{\alpha,i,j} c(\alpha,i,j)\rho_\alpha(g_a)_{ij}.$$

Let us now focus on vectors which are constant on classes. For those vectors we have with $g_c = g_b^{-1} g_a g_b$

$$\begin{aligned}
v_a &= \frac{1}{|G|} \sum_{b=1}^{|G|} v_c \\
&= \frac{1}{|G|} \sum_{b=1}^{|G|} \sum_{\alpha,i,j} c(\alpha,i,j) \rho_\alpha \left(g_b^{-1} g_a g_b \right)_{ij} \\
&= \frac{1}{|G|} \sum_{b=1}^{|G|} \sum_{\alpha,i,j} c(\alpha,i,j) \sum_{k,l} \rho_\alpha \left(g_b^{-1} \right)_{ik} \rho_\alpha (g_a)_{kl} \, \rho_\alpha (g_b)_{lj} \\
&= \sum_{\alpha,i,j} c(\alpha,i,j) \frac{1}{n_\alpha} \sum_{k,l} \delta_{ij} \delta_{lk} \rho_\alpha (g_a)_{kl} \\
&= \sum_{\alpha,i} \frac{c(\alpha,i,i)}{n_\alpha} \chi_\alpha (g_a)
\end{aligned}$$

These vectors span a subspace of dimension equal to the number of classes n_{class}. The above equations show that the characters of the irreducible representations span this subspace. Hence there must be exactly n_{class} such characters, which is the desired result.

Criteria for reducibility: Assume that

$$\rho(g) = \bigoplus_\alpha a_\alpha \rho_\alpha(g)$$

Then

$$\chi(g) = \sum_\alpha a_\alpha \chi_\alpha(g).$$

Conisder now

$$\frac{1}{|G|} \sum_g \chi(g) \chi(g)^* = \sum_\alpha |a_\alpha|^2 \quad \begin{cases} = & 1 \quad \rho \text{ irreducible} \\ > & 1 \quad \rho \text{ reducible} \end{cases}$$

This gives a criteria to check if a representation is irreducible.

Let us consider again the orthogonality theorem for characters:

$$\sum_\kappa n_\kappa \chi_\alpha(C_\kappa) \chi_\beta(C_\kappa)^* = |G| \, \delta_{\alpha\beta}$$

If we define

$$\zeta_\kappa^\alpha = \sqrt{\frac{n_\kappa}{|G|}} \chi_\alpha(C_\kappa)$$

we have

$$\sum_{\kappa} \zeta_{\kappa}^{\alpha} \zeta_{\kappa}^{\beta *} = \delta^{\alpha\beta}.$$

We can view ζ_{κ}^{α} as an entry of a $n_{\mathrm{class}} \times n_{\mathrm{class}}$-matrix. In matrix notation we can write

$$\zeta \cdot \zeta^{\dagger} = \mathbf{1}.$$

It follows that $\zeta^{-1} = \zeta^{\dagger}$ and also

$$\zeta^{\dagger} \cdot \zeta = \mathbf{1}.$$

In other words

$$\sum_{\alpha} \zeta_{\kappa}^{\alpha *} \zeta_{\lambda}^{\alpha} = \delta_{\kappa\lambda}.$$

Converting back to our original notation we have

$$\sum_{\alpha} \chi_{\alpha}(C_{\kappa})^{*} \chi_{\alpha}(C_{\lambda}) = \frac{|G|}{n_{\kappa}} \delta_{\kappa\lambda}.$$

This defines an orthogonality relation between the columns of the character table.

4.4.2 Application: Molecular vibration

As an application of representation theory of finite groups we will study the vibration modes of molecules. We will treat the atoms of the molecules classically as point particles moving in a potential which has a minimum when they are in their equilibrium position. We will consider small displacements from the equilibrium position.

Let us assume that the molecules has n atoms. We will denote the positions of the atoms by

$$(\vec{x}_1, \ldots, \vec{x}_n)$$

and the equilibrium position by $(\vec{x}_1^{(0)}, \ldots, \vec{x}_n^{(0)})$. It is convenient to introduce $3n$ coordinates η_i, $i = 1, \ldots, 3n$, describing the displacement from the equilibrium position. The Lagrange function of the system reads

$$\begin{aligned} L &= T - V, \\ T &= \frac{1}{2} \sum_{i=1}^{3n} T_{ij} \dot{\eta}_i \dot{\eta}_j \\ V &= \frac{1}{2} \sum_{i=1}^{3n} V_{ij} \eta_i \eta_j. \end{aligned}$$

From classical mechanics we know that by suitable change of coordinates we can achive that the Lagrange function reads

$$L = \frac{1}{2}\sum_{i=1}^{3n}\dot{\xi}_i^2 - \frac{1}{2}\sum_{i}^{3n}\omega_i^2\xi_i^2.$$

The ω_i's are called the frequencies of the normal modes. The change of coordinates involves the diagonalisation of $(3n)\times(3n)$ matrices.

We will now discuss how the task of obtaining the normal modes can be simplified using group theory. As an example we will discuss the water molecule H_2O. Let us agree that the equilibrium position of the water molecule is in the $x-z$ plane, with the O-atom along the z-axis and the two H-atoms along the x-axis. We assume that the two H-atoms cannot be distinguished. We first determine the symmetry group, which leaves the equilibrium position invariant. We can rotate the system by 180° along the z-axis, since the two H-atoms are not distinguished, we obtain the same configuration again. This defines a $\mathbb{Z}_2$-group. We denote the element, which generates the group by a. Secondly, we have a reflection in the $x-z$ plane: Changing $y\to -y$ will not affect the equilibrium position. This defines another $\mathbb{Z}_2$-group. We denote the element, which generates the group by b. In summary we find the symmetry group $\mathbb{Z}_2\times\mathbb{Z}_2$, with the generators a and b. This group has four elements $\{e,a,b,ab\}$ and is Abelian. We therefore have four classes (each element is in a class of its own) and as a consequence four irreducible representations ρ_1, ..., ρ_4. The character table is easily obtained:

	e	a	b	ab
ρ_1	1	1	1	1
ρ_2	1	-1	1	-1
ρ_3	1	1	-1	-1
ρ_4	1	-1	-1	1

The $(3n)$ coordinates η_i define a $(3n)$-dimensional representation ρ of this group. This representation is in general reducible. We will now discuss how often a given irreducible representation occurs in the $(3n)$-dimensional one, i.e. we look for the decomposition

$$\rho = \bigoplus_{\alpha} a_\alpha\rho_\alpha.$$

To this aim let us group the $(3n)$-coordinates in tuples of 3:

$$(\eta_1,...,\eta_{3n}) = (\eta_{1,1},\eta_{1,2},\eta_{1,3},\eta_{2,1},...,\eta_{n,3}).$$

ρ acts on this representation as

$$\eta_i' = \sum_{j=1}^{3n}\rho(g)_{ij}\eta_j,$$

or writing this alternatively

$$\eta'_{i,j} = \sum_{k=1}^{n}\sum_{l=1}^{3} \rho(g)_{i,k;j,l}\eta_{k,l}.$$

We can think of $\rho(g)_{i,k;j,l}$ as a $n \times n$-matrix, whose entries are 3×3-matrices. For the character of this representation we have

$$\chi(g) = \sum_{i=1}^{n}\sum_{j=1}^{3} \rho(g)_{i,i;j,j}.$$

In particular we observe that for the trace only the 3×3-matrices on the diagonal of the $n \times n$-matrix contribute. In other words: Only the displacements of the atoms which are left unmoved by the symmetry operation are relevant. Let us now consider an atom i which is not moved by a symmetry operation. The effect of a rotation by an angle θ through the z-axis on the three displacements $\eta_{i,1}$, $\eta_{i,2}$, $\eta_{i,3}$ is given by

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The trace of this matrix is

$$\chi(R(\theta)) = 2\cos\theta + 1.$$

For $\theta = 180^\circ$ we find

$$\chi = -1.$$

The reflection in the $x - z$ plane is described by

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

its character is

$$\chi = 1.$$

Finally, the combined operation of a rotation by $\theta = 180^\circ$ and a reflection is given by

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

with character

$$\chi = 1.$$

We can now obtain the characters of the representation ρ:

	e	a	b	ab
ρ	9	-1	3	1

Not all irreducible representations in the decomposition of ρ correspond to true vibrations: The 9 generalised coordinates η_i contain 6 coordinates desribing the centre-of-mass motion and the rigid rotation of the molecule. The centre-of-mass motion is described by the vector

$$\vec{X} = \frac{1}{M}\sum_{i=1}^{n} m_i \vec{x}_i,$$

with $M = \sum\limits_{i=1}^{n} m_i$. This is a three-dimensional representation of the symmetry group with character

	e	a	b	ab
ρ_{trans}	3	-1	1	1

The rigid rotation can be described by the three quantities

$$\frac{1}{M}\sum_{i=1}^{n} m_i \vec{x}_i^{(0)} \times \vec{\eta}_i = \frac{1}{M}\sum_{i=1}^{n} m_i \left(\begin{array}{c} x_{i,1}^{(0)}\eta_{i,2} - x_{i,2}^{(0)}\eta_{i,1} \\ x_{i,2}^{(0)}\eta_{i,3} - x_{i,3}^{(0)}\eta_{i,2} \\ x_{i,3}^{(0)}\eta_{i,1} - x_{i,1}^{(0)}\eta_{i,3} \end{array} \right).$$

This transforms as a three-dimensional vector under rotations, but as a pseudo-vector under reflections. Indeed, for the symmetry transformation $y \to -y$ we have the transformation matrix

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

and the character

$$\chi = -1.$$

We therefore find the character of rigid rotations as

	e	a	b	ab
ρ_{rot}	3	-1	-1	-1

We are not so much interested in the centre-of-mass motion and the rigid rotation. We subtract these characters from the character of ρ and obtain

	e	a	b	ab
$\rho - \rho_{\text{trans}} - \rho_{\text{rot}}$	3	1	3	1

We can now find the decomposition of $\rho - \rho_{\text{trans}} - \rho_{\text{rot}}$ in terms of the irreducible representations. We make the ansatz

$$\rho - \rho_{\text{trans}} - \rho_{\text{rot}} = \bigoplus_{\alpha} a_\alpha \rho_\alpha.$$

The multiplicity a_α we obtain from the orthogonality theorem for characters

$$a_\alpha = \frac{1}{|G|} \sum_{\kappa} n_\kappa \chi_\alpha \chi^*_{3n-\text{trans}-\text{rot}}.$$

For the water molecule $|G| = 4$ and all classes have exactly one element: $n_\kappa = 1$. We find

$$\begin{aligned}
a_1 &= \frac{1}{4}(1\cdot 3 + 1\cdot 1 + 1\cdot 3 + 1\cdot 1) = 2,\\
a_2 &= \frac{1}{4}(1\cdot 3 + (-1)\cdot 1 + 1\cdot 3 + (-1)\cdot 1) = 1,\\
a_3 &= \frac{1}{4}(1\cdot 3 + 1\cdot 1 + (-1)\cdot 3 + (-1)\cdot 1) = 0,\\
a_4 &= \frac{1}{4}(1\cdot 3 + (-1)\cdot 1 + (-1)\cdot 3 + 1\cdot 1) = 0.
\end{aligned}$$

Therefore

$$\rho - \rho_{\text{trans}} - \rho_{\text{rot}} = 2\rho_1 + \rho_2.$$

We therefore find three vibrational modes. Two transform as the trivial representation ρ_1. The displacements in these modes are left invariant under the symmetry group $\mathbb{Z}_2 \times \mathbb{Z}_2$. One of these two modes is given by the vibration, where both H-atoms move along the z-axis in the same direction, while the O-atom moves along the z-axis in the opposite direction. The other mode is given by an oscillation, where the O-atom is at rest, and the H-atoms move along the x-axis in opposite directions.

The third mode transforms as the representation ρ_2. It transforms trivially under reflections, therefore the motion is in the $x-z$ plane. It has however a non-trivial transformation under the rotation of 180° around the z-axis. This is oscillation is given by a motion, where the O-atom moves along the x-axis in one direction, the two H-atoms have both a component along the x-axis in the opposite direction. In addition the two H-atoms have opposite components along the z-axis.

4.4.3 Application: Quantum mechanics

A quantum mechanical system is described by the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{x},t) = \hat{H} \psi(\vec{x},t).$$

If the Hamilton operator $\hat{H}$ is time-independent we can make the ansatz

$$\psi(\vec{x},t) = \psi(\vec{x}) \exp\left(-\frac{i}{\hbar}Et\right)$$

and we obtain the time-independent Schrödinger equation

$$\hat{H}\psi(\vec{x}) = E\psi(\vec{x}).$$

Consider now a group of transformations acting on the coordinates $\vec{x}$:

$$\vec{x}' = g\vec{x}, \quad g \in G.$$

This induces a transformation on the wave function by

$$\psi'(\vec{x}) = \rho(g)\psi(\vec{x}) = \psi\left(g^{-1}\vec{x}\right).$$

An operator transforms as

$$\hat{O}' = \rho(g)\hat{O}\rho(g^{-1}).$$

We are in particular interested in transformations, which leave the Hamilton operator invariant:

$$\hat{H} = \rho(g)\hat{H}\rho(g^{-1}).$$

Multiplying this equation by $\rho(g)$ from the right we obtain $\hat{H}\rho(g) = \rho(g)\hat{H}$, and we see that this is equivalent to the statement that $\rho(g)$ commutes with the Hamilton operator

$$\left[\hat{H},\rho(g)\right] = 0.$$

Remark: Usually the Hamiltonian of a quantum mechanical system is given by the sum of the kinetic and potential energy operator. As symmetry transformations we will usually consider translations, rotations and reflections. The kinetic energy operator is invariant under these transformations, therefore the full Hamiltonian is invariant if the potential energy operator is:

$$\hat{V}(\vec{x}) = \hat{V}(g\vec{x}).$$

Let us now consider a quantum mechanical system with a Hamilton operator $\hat{H}$, which is invariant under a finite symmetry group G. In this case :

- The eigenfunctions for a given eigenvalue E form a representation of the symmetry group G.
- The energy E_α corresponding to an irreducible representation ρ_α will be at least n_α-fold degenerate, where n_α is the dimension of the irreducible representation ρ_α.

Proof: The set of all degenerate eigenfunctions for the eigenvalue E form a vectorspace V. If ψ and ϕ are two eigenfunctions with the eigenvalue E, so is any linear combination of them. This vector space defines a representation of the symmetry group G. If $\psi \in V$ and $\psi' = \rho(g)\psi$ then

$$\hat{H}\psi' = \hat{H}\rho(g)\psi = \rho(g)\hat{H}\psi = \rho(g)E\psi = E\left(\rho(g)\psi\right) = E\psi'.$$

Therefore $\psi' \in V$. V is either irreducible or reducible. In the latter case V may be decomposed into irreducible components. In both cases, if V contains the irreducible representation V_α, it follows that

$$\dim V \geq \dim V_\alpha = n_\alpha.$$

Example: We consider the quantum mechanical harmonic oscillator in one dimension. The Hamilton operator is

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2}{2}\hat{x}^2.$$

We define the characteristic length

$$x_0 = \sqrt{\frac{\hbar}{m\omega}}.$$

Obviously, the Hamilton operator is invariant under the reflection $x \to -x$. The reflection a generates a symmetry group $\mathbb{Z}_2$. The character table of this group is

	e	a
ρ_1	1	1
ρ_2	1	-1

It is well-known that the eigenvalues are given by

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right)$$

and the eigenfunctions are given by

$$\psi_n(x) = \frac{1}{\sqrt{2^n n! \sqrt{\pi} x_0}} \exp\left(-\frac{1}{2}\left(\frac{x}{x_0}\right)^2\right) H_n\left(\frac{x}{x_0}\right),$$

with the Hermite polynomials

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}.$$

We see that the eigenfunctions ψ_n for even n transform as the trivial representation ρ_1 of $\mathbb{Z}_2$, while the eigenfunctions for odd n transform as the representation ρ_2 of $\mathbb{Z}_2$.

4.5 Representation theory for Lie groups

4.5.1 Irreducible representation of $SU(2)$ and $SO(3)$

The groups $SU(2)$ and $SO(3)$ have the same Lie algebra:

$$[I_a, I_b] = i\varepsilon_{abc} I_c.$$

For $SU(2)$ we can take the I^a's proportional to the Pauli matrices

$$I_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad I_2 = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad I_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

This defines a representation of $SU(2)$ which is called the fundamental representation. (It is not a representation of $SO(3)$.)
Quite generally the structure constants provide a representation known as the adjoint or vector representation:

$$(M_b)_{ac} = i f_{abc}.$$

For $SU(2)$ and $SO(3)$:

$$M_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The dimension of the adjoint representation equals the dimension of the parameter space of the group and the numbers of generators.

Let us now discus more systematically all irreducible representations.
Definition: A **Casimir operator** is an operator, which commutes with all the generators of the group.
Example: For $SU(2)$

$$I^2 = I_1^2 + I_2^2 + I_3^2$$

is a Casimir operator:

$$\left[I^2, I_a\right] = 0.$$

Example 2: Let us consider $SU(3)$, with the generators T^a, $a = 1, ..., 8$. Here we find two independent Casimir operators, which we call C_2 and C_3. The first one is given by

$$C_2 = T^a T^a$$

and is called the quadratic Casimir operator. The symmetric tensor d^{abc} is defined for $SU(3)$ through

$$\left\{T^a, T^b\right\} = d^{abc} T^c + \frac{1}{3}\delta^{ab}.$$

Here, $\{\ldots,\ldots\}$ denotes the anti-commutator

$$\{A,B\} = AB+BA.$$

We then define C_3 by

$$C_3 = d^{abc}T^aT^bT^c.$$

C_3 is called the cubic Casimir operator. Since C_2 and C_3 are Casimir operators, we have

$$[C_2,T^a]=0, \qquad [C_3,T^a]=0.$$

Continuing in this line, it can be shown that $SU(n)$ has $(n-1)$ independent Casimir operators C_2, C_3, ..., C_{n-1}. The group $U(n)$ has n independent Casimir operators C_1, C_2, ..., C_{n-1}.

Definition: The **rank of a Lie algebra** is the number of simultaneously diagonalisable generators.
Example 1: $SU(2)$ has rank one, the convention is to take I_3 diagonal.
Example 2: $SU(3)$ has rank two, in the Gell-Mann representation T_3 and T_8 are diagonal.

Theorem: The number of independent Casimir operators is equal to the rank of the Lie algebra. The proof can be found in many textbooks.

The eigenvalues of the Casimir operators may be used to label the irreducible representations. The eigenvalues of the diagonal generators can be used to label the basis vectors within a given irreducible representation.

Example $SU(2)$:

$$\begin{aligned} I^2\left|\lambda,m\right\rangle &= \lambda\left|\lambda,m\right\rangle, \\ I_3\left|\lambda,m\right\rangle &= m\left|\lambda,m\right\rangle. \end{aligned}$$

Consider

$$\left(I_1^2+I_2^2\right)\left|\lambda,m\right\rangle = \left(I^2-I_3^2\right)\left|\lambda,m\right\rangle = \left(\lambda-m^2\right)\left|\lambda,m\right\rangle.$$

Further

$$\left\langle\lambda,m\right|I_1^2\left|\lambda,m\right\rangle = \left\langle\lambda,m\right|I_1^\dagger I_1\left|\lambda,m\right\rangle = \left|I_1\left|\lambda,m\right\rangle\right|^2 \geq 0.$$

A similar consideration applies to $\left\langle\lambda,m\right|I_2^2\left|\lambda,m\right\rangle$. Therefore

$$\lambda-m^2 \geq 0.$$

For a given λ the possible values of m are bounded:

$$-\sqrt{\lambda}\leq m\leq\sqrt{\lambda}$$

Define

$$I_\pm = \frac{1}{\sqrt{2}}\left(I_1 \pm i I_2\right)$$

$$[I_3, I_\pm] = \pm I_\pm, \qquad \left[I^2, I_\pm\right] = 0.$$

The last relation implies

$$\begin{aligned} \left(I^2 I_\pm - I_\pm I^2\right)|\lambda, m\rangle &= 0, \\ I^2\left(I_\pm |\lambda, m\rangle\right) &= \lambda\left(I_\pm |\lambda, m\rangle\right). \end{aligned}$$

Therefore the operators $I_\pm$ don't change λ. From the commutation relation with I_3 we obtain

$$\begin{aligned} \left(I_3 I_\pm - I_\pm I_3\right)|\lambda, m\rangle &= \pm I_\pm |\lambda, m\rangle, \\ I_3\left(I_\pm |\lambda, m\rangle\right) &= (m \pm 1)\left(I_\pm |\lambda, m\rangle\right). \end{aligned}$$

Therefore $I_\pm |\lambda, m\rangle$ is proportional to $|\lambda, m \pm 1\rangle$ unless zero. Recall that the values of m are bounded, therefore there is a maximal value m_{max} and a minimal value m_{min}:

$$\begin{aligned} I_+ |\lambda, m_{max}\rangle &= 0, \\ I_- |\lambda, m_{min}\rangle &= 0. \end{aligned}$$

Now

$$\begin{aligned} I^2 = I_1^2 + I_2^2 + I_3^2 &= 2I_+ I_- + I_3^2 - I_3 \\ &= 2I_- I_+ + I_3^2 + I_3 \end{aligned}$$

Therefore

$$\begin{aligned} I^2 |\lambda, m_{max}\rangle &= \left(2I_- I_+ + I_3^2 + I_3\right)|\lambda, m_{max}\rangle, \\ \lambda |\lambda, m_{max}\rangle &= m_{max}\left(m_{max} + 1\right)|\lambda, m_{max}\rangle, \end{aligned}$$

and

$$\lambda = m_{max}\left(m_{max} + 1\right).$$

Similar:

$$\begin{aligned} I^2 |\lambda, m_{min}\rangle &= \left(2I_+ I_- + I_3^2 - I_3\right)|\lambda, m_{min}\rangle, \\ \lambda |\lambda, m_{min}\rangle &= m_{min}\left(m_{min} - 1\right)|\lambda, m_{min}\rangle, \end{aligned}$$

and

$$\lambda = m_{min}\left(m_{min} - 1\right).$$

From

$$
\begin{aligned}
m_{max}^2 + m_{max} &= m_{min}^2 - m_{min}, \\
(m_{max} + m_{min}) \underbrace{(m_{max} - m_{min} + 1)}_{>0} &= 0
\end{aligned}
$$

it follows

$$m_{min} = -m_{max}.$$

Since the ladder operators raise or lower m by one unit we must have that m_{max} and m_{min} differ by an integer, therefore

$$2m_{max} = \text{integer}.$$

Let us write $m_{max} = j$. Then $2j$ is an integer and

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

$$\lambda = j(j+1)$$

Normalisation:

$$I_\pm |\lambda, m\rangle = A_\pm |\lambda, m \pm 1\rangle$$

With $I_\pm^\dagger = I_\mp$ we have

$$|A_\pm|^2 = \langle \lambda, m| I_\pm^\dagger I_\pm |\lambda, m\rangle = \langle \lambda, m| I_\mp I_\pm |\lambda, m\rangle = \frac{1}{2} \langle \lambda, m| I^2 - I_3 (I_3 \pm 1) |\lambda, m\rangle$$

and therefore

$$|A_\pm|^2 = \frac{1}{2} (j(j+1) - m(m \pm 1))$$

Condon-Shortley convention:

$$A_\pm = \sqrt{\frac{j(j+1) - m(m \pm 1)}{2}}.$$

The representation of $SU(2)$ corresponding to $j = 0, 1, 2, ...$ are also representations of $SO(3)$, but the one corresponding to $j = 1/2, 3/2, ...$ are not.

4.5.2 The Cartan basis

Definition: Suppose a Lie algebra A has a sub-algebra B such that the commutator of any element of A (T^a say) with any element of B (T^b say) always lies in B, then B is said to be an **ideal** of A:

$$\left[T^a, T^b\right] \quad \in \quad B.$$

Every Lie algebra has two trivial ideals: A and $\{0\}$.

A Lie algebra is called **simple** if it is non-Abelian and has no non-trivial ideals.

A Lie algebra is called **semi-simple** if it has no non-trivial Abelian ideals.

A Lie algebra is called **reductive** if it is the sum of a semi-simple and an abelian Lie algebra.

A simple Lie algebra is also semi-simple and a semi-simple Lie algebra is also reductive.

Examples: The Lie algebras

$$su(n), so(n), sp(n)$$

are simple.
Semi-simple Lie algebras are sums of simple Lie algebras:

$$su(n_1) \oplus su(n_2).$$

Reductive Lie algebras may have in addition an abelian part:

$$u(1) \oplus su(2) \oplus su(3).$$

From Schur's lemma we know that abelian Lie groups have only one-dimensional irreducible representations. Therefore let us focus on Lie groups corresponding to semi-simple Lie algebras. A Lie group, which has a semi-simple Lie algebra, is for obvious reasons called semi-simple. We first would like to have a criterion to decide, whether a Lie algebra is semi-simple or not: If

$$\left[T^a, T^b\right] \quad = \quad i f^{abc} T^c,$$

define

$$g^{ab} \quad = \quad f^{acd} f^{bcd}.$$

A criterion due to Cartan say that a Lie algebra is semi-simple if and only if

$$\det g \quad \neq \quad 0.$$

For $SU(n)$ we find

$$g^{ab} = C_A \delta^{ab}.$$

Let us now define the Cartan standard form of a Lie algebra. To motivate the Cartan standard form let us as an example suppose that we have

$$\left[T^1, T^2\right] = 0, \quad \left[T^1, T^3\right] \neq 0, \quad \left[T^2, T^3\right] \neq 0.$$

If we now make a change of basis

$$T^{1'} = T^1 + T^3, \quad T^{2'} = T^2, \quad T^{3'} = T^3,$$

none of the new commutators vanishes. It is certainly desirable to pick a basis, where a maximum number of commutators vanish and the non-vanishing ones are rather simple. This will bring us to the Cartan standard form. Let us assume that

$$\begin{aligned} A &= \sum_{a=1}^{n} c_a T^a, \\ X &= \sum_{a=1}^{n} x_a T^a, \end{aligned}$$

such that

$$[A, X] = \rho X.$$

ρ is called a **root** of the Lie algebra. We then have

$$\begin{aligned} [A, X] &= i c_a x_b f^{abc} T^c = \rho x_c T^c, \\ \left(c_a x_b i f^{abc} - \rho x_c\right) &= 0, \\ \left(c_a i f^{abc} - \rho \delta^{bc}\right) x_b &= 0. \end{aligned}$$

For a non-trivial solution we must have

$$\det\left(c_a i f^{abc} - \rho \delta^{bc}\right) = 0.$$

In general the secular equation will give a n-th order polynomial is ρ. Solving for ρ one obtains n roots. One root may occur more than once. The degree of degeneracy is called the multiplicity of the root.

Theorem (Cartan): If A is chosen sucht that the secular equation has the maximum number of distince roots, then only the root $\rho = 0$ is degenerate. Further if r is the multiplicity of that root, there exist r linearly independent generators H_i, which mutually commute

$$\left[H_i, H_j\right] = 0, \quad i, j = 1, ..., r.$$

r is the rank of the Lie algebra.

Notation: Latin indices for $1,\dots,r$, e.g. H_i and greek indices for the remaining $(n-r)$ generators E_α $(\alpha = 1,\dots,n-r)$.

Example $SU(2)$:

$$\left[I^a, I^b\right] = i\varepsilon^{abc} I^c.$$

Take $A = I^3$:

$$\left[I^3, X\right] = \rho X.$$

Secular equation:

$$\begin{aligned}
\det\left(i\varepsilon^{3bc} - \rho\delta^{bc}\right) &= 0, \\
\begin{vmatrix} -\rho & i & 0 \\ -i & -\rho & 0 \\ 0 & 0 & -\rho \end{vmatrix} &= 0, \\
-\rho^3 + \rho &= 0, \\
\rho\left(\rho^2 - 1\right) &= 0.
\end{aligned}$$

Therefore the roots are $0, \pm 1$. We have

$$\begin{aligned}
\rho = 0 \qquad & \left[I^3, X\right] = 0 \quad \Rightarrow X = I^3 = H_1, \\
\rho = 1 \qquad & \left[I^3, X\right] = X \quad \Rightarrow X = \frac{1}{\sqrt{2}}\left(I^1 + iI^2\right) = E_1, \\
\rho = -1 \qquad & \left[I^3, X\right] = -X \quad \Rightarrow X = \frac{1}{\sqrt{2}}\left(I^1 - iI^2\right) = E_2.
\end{aligned}$$

Theorem: For any compact semi-simple Lie group, non-zero roots occur in pairs of opposite sign and are denoted E_α and $E_{-\alpha}$ $(\alpha = 1,\dots,(n-r)/2)$.

We thus have the Cartan standard form:

$$\begin{aligned}
\left[H_i, H_j\right] &= 0, \\
\left[H_i, E_\alpha\right] &= \rho(\alpha, i) E_\alpha.
\end{aligned}$$

As a short-hand notation the last equation is also often written as

$$\left[H_i, E_\alpha\right] = \alpha_i E_\alpha.$$

The standard normalisation for the Cartan basis is

$$\sum_{\alpha=1}^{(n-r)/2} \rho(\alpha, i)\, \rho(\alpha, j) = \delta_{ij}.$$

Cartan standard form of $SU(2)$:

$$H_1 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad E_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

The roots are

$$[H, E_1] = E_1, \qquad [H, E_{-1}] = -E_1.$$

Cartan standard form of $SU(3)$:

$$H_1 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad H_2 = \frac{1}{3\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix},$$
$$E_1 = \frac{1}{\sqrt{3}}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E_2 = \frac{1}{\sqrt{3}}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E_3 = \frac{1}{\sqrt{3}}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$
$$E_{-1} = \frac{1}{\sqrt{3}}\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E_{-2} = \frac{1}{\sqrt{3}}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad E_{-3} = \frac{1}{\sqrt{3}}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

The roots are

$$\begin{aligned} [H_1, E_1] &= \frac{1}{3}\sqrt{6}E_1, \qquad & [H_2, E_1] &= 0, \\ [H_1, E_2] &= \frac{1}{6}\sqrt{6}E_2, \qquad & [H_2, E_2] &= \frac{1}{2}\sqrt{2}E_2 \\ [H_1, E_3] &= -\frac{1}{6}\sqrt{6}E_3, \qquad & [H_2, E_3] &= \frac{1}{2}\sqrt{2}E_3. \end{aligned}$$

The r numbers α_i, $i = 1, \ldots, r$ can be regarded as the components of a root vector $\vec{\alpha} = (\alpha_1, \ldots, \alpha_r)$ of dimension r.

Example: For $SU(3)$, the root vectors corresponding to E_1 E_2 and E_3 are

$$\vec{\alpha}(E_1) = \begin{pmatrix} \frac{1}{3}\sqrt{6} \\ 0 \end{pmatrix}, \qquad \vec{\alpha}(E_2) = \begin{pmatrix} \frac{1}{6}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \end{pmatrix}, \qquad \vec{\alpha}(E_3) = \begin{pmatrix} -\frac{1}{6}\sqrt{6} \\ \frac{1}{2}\sqrt{2} \end{pmatrix},$$
$$\vec{\alpha}(E_{-1}) = \begin{pmatrix} -\frac{1}{3}\sqrt{6} \\ 0 \end{pmatrix}, \qquad \vec{\alpha}(E_{-2}) = \begin{pmatrix} -\frac{1}{6}\sqrt{6} \\ -\frac{1}{2}\sqrt{2} \end{pmatrix}, \qquad \vec{\alpha}(E_{-3}) = \begin{pmatrix} \frac{1}{6}\sqrt{6} \\ -\frac{1}{2}\sqrt{2} \end{pmatrix}.$$

Theorem: If $\vec{\alpha}$ is a root vector, so is $-\vec{\alpha}$, (since roots always occur in pairs of opposite sign).

Theorem: If $\vec{\alpha}$ and $\vec{\beta}$ are root vectors then

$$\frac{2\vec{\alpha}\cdot\vec{\beta}}{|\alpha|^2} \quad \text{and} \quad \frac{2\vec{\alpha}\cdot\vec{\beta}}{|\beta|^2}$$

are integers. Suppose these integers are p and q. Then

$$\frac{\left(\vec{\alpha}\cdot\vec{\beta}\right)^2}{|\alpha|^2\,|\beta|^2} \;=\; \frac{pq}{4} = \cos^2\theta \le 1.$$

Therefore

$$pq \;\le\; 4.$$

It follows that

$$\cos^2\theta \;=\; 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1.$$

Case $\theta = 0^\circ$: This is the trivial case $\vec{\alpha} = \vec{\beta}$.
Case $\theta = 30^\circ$: We have $pq = 3$ and $p = 1, q = 3$ or $p = 3, q = 1$. Let us first discuss $p = 1, q = 3$. This means

$$\frac{2\vec{\alpha}\cdot\vec{\beta}}{|\alpha|^2} = 1, \qquad \frac{2\vec{\alpha}\cdot\vec{\beta}}{|\beta|^2} = 3.$$

Therefore

$$\frac{|\alpha|^2}{|\beta|^2} \;=\; 3.$$

The case $p = 3, q = 1$ is similar and in summary we obtain

$$\frac{|\alpha|^2}{|\beta|^2} \;=\; 3 \text{ or } \frac{1}{3}.$$

Case $\theta = 45^\circ$: We have $pq = 2$ and $p = 1, q = 2$ or $p = 2, q = 1$. It follows

$$\frac{|\alpha|^2}{|\beta|^2} \;=\; 2 \text{ or } \frac{1}{2}.$$

Case $\theta = 60^\circ$: We have $pq = 1$ and $p = 1, q = 1$. It follows

$$\frac{|\alpha|^2}{|\beta|^2} \;=\; 1.$$

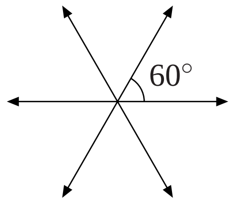
Case $\theta = 90^\circ$: In this case $p = 0$ and $q = 0$. This leaves the ratio $|\alpha|^2/|\beta|^2$ undetermined.

The cases $\theta = 120^\circ$, $\theta = 135^\circ$, $\theta = 150^\circ$ and $\theta = 180^\circ$ are analogous to the ones discussed above.

If $\vec{\alpha}$ and $\vec{\beta}$ are root vectors so is

$$\vec{\gamma} \;=\; \vec{\beta} - \frac{2\vec{\alpha}\cdot\vec{\beta}}{\alpha^2}\vec{\alpha}$$

Example: The root diagram of $SU(3)$:



4.5.3 Weights

Let us first recall some basic facts: The rank of a Lie algebra is the number of simultaneously diagonalizable generators. In the following we will denote the rank of a Lie algebra by r.

We already mentioned the following theorem : The rank of a Lie algebra is equal to the number of independent Casimir operators. (A Casimir operator is an operator, which commutes with all the generators.)

For a Lie algebra of rank r we therefore have r Casimir operators and r simultaneously diagonalizable generators H_i.

The eigenvalues of the Casimir operators may be used to label the irreducible representations. The eigenvalues of the diagonal generators H_i may be used to label the states within a given irreducible representation.

Let $\vec{\lambda}$ be a shorthand notation for $\vec{\lambda} = (\lambda_1, ..., \lambda_r)$, a set of eigenvalues of Casimir operators and let $\vec{m}$ be a shorthand notation for $\vec{m} = (m_1, ..., m_r)$, a set of eigenvalues of the diagonal generators:

$$H_i |\vec{\lambda}, \vec{m}\rangle = m_i |\vec{\lambda}, \vec{m}\rangle$$

The vector $\vec{m}$ is called the weight vector.

Example $SU(3)$: Let us consider the fundamental representation. The vector space is spanned by the three vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

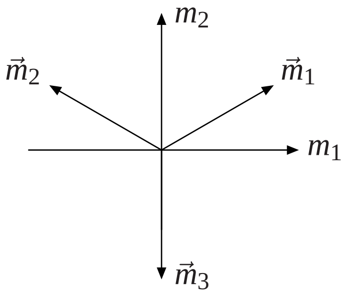
We have

$$\begin{aligned} (H_1, H_2)\, e_1 &= \left(\frac{1}{\sqrt{6}}, \frac{1}{3\sqrt{2}} \right) e_1, \\ (H_1, H_2)\, e_2 &= \left(-\frac{1}{\sqrt{6}}, \frac{1}{3\sqrt{2}} \right) e_2, \\ (H_1, H_2)\, e_3 &= \left(0, -\frac{2}{3\sqrt{2}} \right) e_3. \end{aligned}$$

This gives the weight vectors

$$\vec{m}_1 = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{3\sqrt{2}} \end{pmatrix}, \quad \vec{m}_2 = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ \frac{1}{3\sqrt{2}} \end{pmatrix}, \quad \vec{m}_3 = \begin{pmatrix} 0 \\ -\frac{\sqrt{2}}{3} \end{pmatrix}.$$

and the weight diagram



Consider now the complex conjugate representation of the fundamental representation: If

$$\rho = \exp\left(-i\theta_a T^a\right)$$

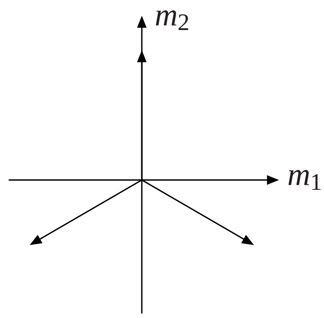
is a representation, then also

$$\rho^* = \exp\left(i\theta_a T^{a*}\right) = \exp\left(-i\theta_a T^{a\prime}\right)$$

is a representation and we have

$$T^{a\prime} = -T^{a*}.$$

It follows that the weights of the complex conjugate representation are negatives of those of the fundamental representation:



Note that in general the complex conjugate representation ρ^* is inequivalent to ρ. This is in contrast to $SU(2)$, where one can find a S, such that

$$SI^aS^{-1} = -I^{a*}, \quad S\rho S^{-1} = \rho^*.$$

The generators $E_{\pm\alpha}$ are generalisations of the raising and lowering operators $I_\pm$ of $SU(2)$. Suppose

$$H_i|\vec{\lambda},\vec{m}\rangle = m_i|\vec{\lambda},\vec{m}\rangle$$

and

$$[H_i, E_\alpha] = \alpha_i E_\alpha.$$

Then

$$\begin{aligned}
[H_i, E_\alpha]\, |\vec{\lambda}, \vec{m}\rangle &= \alpha_i E_\alpha |\vec{\lambda}, \vec{m}\rangle, \\
H_i E_\alpha |\vec{\lambda}, \vec{m}\rangle - E_\alpha H_i |\vec{\lambda}, \vec{m}\rangle &= \alpha_i E_\alpha |\vec{\lambda}, \vec{m}\rangle, \\
H_i \left(E_\alpha |\vec{\lambda}, \vec{m}\rangle \right) &= (m_i + \alpha_i) \left(E_\alpha |\vec{\lambda}, \vec{m}\rangle \right).
\end{aligned}$$

Therefore $E_\alpha|\vec{\lambda}, \vec{m}\rangle$ is proportional to $|\vec{\lambda}, \vec{m} + \vec{\alpha}\rangle$ unless zero. Therefore the weight vectors within an irreducible representation differ by a linear combination of root vectors with integer coefficients.
Example: In the $SU(2)$ case the weight vectors were one-dimensional. Within one irreducible representation all weights could be obtained from m_{max} by applying the lowering operator I_-. The action of I_- corresponds to a shift in the weight proportional to a root vector.

In the $SU(2)$ case we also found that within an irreducible representations the weights are bounded, i.e. there is a maximal weight, for which $I_+ |\lambda, m_{max}\rangle = 0$. We now look how this fact generalises: We start with the definition of the multiplicity of a a weight: The number of different eigenstates with the same weight is called the **multiplicity of the weight**. A weight is said to be **simple** if the multiplicity is 1.
For Lie algebras with $r \geq 2$, weights are not necessarily simple.

Theorem : Given a weight $\vec{m}$ and a root vector $\vec{\alpha}$ then

$$\frac{2\vec{\alpha} \cdot \vec{m}}{\alpha^2}$$

is an integer and

$$\vec{m}' = \vec{m} - \frac{2\vec{\alpha} \cdot \vec{m}}{\alpha^2} \vec{\alpha}$$

is also a weight vector. $\vec{m}$ and $\vec{m}'$ are called **equivalent weights**.
Geometrically, $\vec{m}'$ is obtained from $\vec{m}$ by a reflection in the plane perpendicular to $\vec{\alpha}$. For $SU(2)$ the weights m and $-m$ within an irreducible representation are equivalent.

Ordering of weights: The convention for $SU(n)$ is the following: $\vec{m}$ is said to be higher than $\vec{m}'$ if the r^{th} component of $(\vec{m} - \vec{m}')$ is positive (if zero look at the $(r-1)^{\text{th}}$ component, if this one is also zero, look at the $(r-2)^{\text{th}}$ component, etc.).

The highest weight of a set of equivalent weights is said to be **dominant**.
(In the case of an irreducible representation of $SU(2)$ the dominant weight is the one with m_{max}.)

Theorem: For any compact semi-simple Lie algebra there exists for any irreducible representation a highest weight. Furthermore this highest weight is also simple.
(Recall: In the $SU(2)$ case we first showed that the values of m are bounded, and then obtained all other states in the irreducible representation by applying the lowering operator to the state with m_{max}.)

Theorem: For every simple Lie algebra of rank r there are r dominant weights $\vec{M}^{(i)}$, called fundamental dominant weights, such that any other dominant weight $\vec{M}$ is a linear combination of the $\vec{M}^{(i)}$

$$\vec{M} = \sum_{i=1}^{r} n_i \vec{M}^{(i)}$$

where the n_i are non-negative integers.

Note that there exists r fundamental irreducible representations, which have the r different $\vec{M}^{(i)}$'s as their highest weight. We can label the irreducible representations by $(n_1, n_2, ..., n_r)$ instead of the eigenvalues of the Casimirs.

Example: For $SU(2)$ we can label the irreducible representations either by the eigenvalue λ of the Casimir operator I^2, or by a number n with the relation

$$\lambda = j(j+1), \quad j = \frac{n}{2}.$$

Note that there is one fundamental representation, which is two-dimensional. The dominant weight of the fundamental representation is $1/2$. All other dominant weights j are non-negative integer multiples of this fundamental dominant weight:

$$j = n \cdot \frac{1}{2}$$

4.6 Tensor methods

We have already seen how to construct new representation out of given ones through the operations of the direct sum and the tensor product: If V_1 and V_2 are representations of G, the direct sum $V_1 \oplus V_2$ and the tensor product $V_1 \otimes V_2$ are again representations:

$$\begin{aligned} g(v_1 \oplus v_2) &= (gv_1) \oplus (gv_2), \\ g(v_1 \otimes v_2) &= (gv_1) \otimes (gv_2), \end{aligned}$$

We now turn to the question how to construct new irreducible representations out of given irreducible ones. If V_1 and V_2 are irreducible representations, the direct sum is $V_1 \oplus V_2$ is reducible and decomposes into the irreducible representations V_1 and V_2. Nothing new here. More interesting is the tensor product, which we will study in the following.

4.6.1 Clebsch-Gordan series

To motivate the discussion of tensor methods we start again from the $SU(2)$ example and its relation to the spin of a physical system. Suppose we have to independent spin operators $\vec{J}_1$ and $\vec{J}_2$, describing the spin of particle 1 and 2, respectively.

$$\left[J_{1\,i}, J_{2\,j}\right] \;=\; 0 \quad \forall i, j$$

Let us now define the total spin as

$$\begin{aligned}\vec{J} &= \vec{J}_1 + \vec{J}_2, \\ J_z &= J_{1\,z} + J_{2\,z}.\end{aligned}$$

We use the following notation:

$$\begin{array}{ll}|j_1, m_1\rangle & \text{eigenstate of } J_1^2 \text{ and } J_{1\,z} \\ |j_2, m_2\rangle & \text{eigenstate of } J_2^2 \text{ and } J_{2\,z}\end{array}$$

We define

$$|j_1, j_2, m_1, m_2\rangle \;=\; |j_1, m_1\rangle \otimes |j_2, m_2\rangle\,.$$

The set

$$\{|j_1, j_2, m_1, m_2\rangle\}$$

are eigenvectors of

$$\left\{J_1^2, J_2^2, J_{1\,z}, J_{2\,z}\right\}$$

and is referred to as the uncoupled basis. In general these states are not eigenstates of J^2 and the basis is reducible. This can be seen easily:

$$J^2 \;=\; \left(\vec{J}_1 + \vec{J}_2\right)\left(\vec{J}_1 + \vec{J}_2\right) = J_1^2 + J_2^2 + 2\vec{J}_1\vec{J}_2,$$

and $2\vec{J}_1\vec{J}_2$ fails to commute with $J_{1\,z}$ and $J_{2\,z}$. To find a better basis, we look for a set of mutually commuting operators. The set

$$\left\{J^2, J_z, J_1^2, J_2^2\right\}$$

is such a set and an eigenbasis for this set is labelled by

$$\{|j, m, j_1, j_2\rangle\}\,.$$

This basis is called the coupled basis and carries and irreducible representation of dimension $2j+1$. Of course we can express each vector in the coupled basis through a linear combination of the uncoupled basis:

$$|j, m, j_1, j_2\rangle \;=\; \sum_{m_1, m_2; m_1+m_2=m} C^{jm}_{j_1 j_2 m_1 m_2}\, |j_1, j_2, m_1, m_2\rangle$$

The coefficients $C^{jm}_{j_1 j_2 m_1 m_2}$ are called the Clebsch-Gordan coefficients. The Clebsch-Gordan coefficients are tabulated in the particle data group tables.

Example: We take $j_1 = j_2 = 1/2$ and use the short-hand notation

$$\begin{aligned}
\left|\uparrow\uparrow\right\rangle &= \left| j_1 = \frac{1}{2}, j_2 = \frac{1}{2}, m_1 = \frac{1}{2}, m_2 = \frac{1}{2} \right\rangle, \\
\left|\uparrow\downarrow\right\rangle &= \left| j_1 = \frac{1}{2}, j_2 = \frac{1}{2}, m_1 = \frac{1}{2}, m_2 = -\frac{1}{2} \right\rangle, \\
\left|\downarrow\uparrow\right\rangle &= \left| j_1 = \frac{1}{2}, j_2 = \frac{1}{2}, m_1 = -\frac{1}{2}, m_2 = \frac{1}{2} \right\rangle, \\
\left|\downarrow\downarrow\right\rangle &= \left| j_1 = \frac{1}{2}, j_2 = \frac{1}{2}, m_1 = -\frac{1}{2}, m_2 = -\frac{1}{2} \right\rangle,
\end{aligned}$$

For the coupled basis we have $j \in \{0,1\}$ and we find

$$\begin{aligned}
\left| j = 1, m = 1, j_1 = \frac{1}{2}, j_2 = \frac{1}{2} \right\rangle &= \left|\uparrow\uparrow\right\rangle, \\
\left| j = 1, m = 0, j_1 = \frac{1}{2}, j_2 = \frac{1}{2} \right\rangle &= \frac{1}{\sqrt{2}} \left(\left|\uparrow\downarrow\right\rangle + \left|\downarrow\uparrow\right\rangle \right), \\
\left| j = 1, m = -1, j_1 = \frac{1}{2}, j_2 = \frac{1}{2} \right\rangle &= \left|\downarrow\downarrow\right\rangle, \\
\left| j = 0, m = 0, j_1 = \frac{1}{2}, j_2 = \frac{1}{2} \right\rangle &= \frac{1}{\sqrt{2}} \left(\left|\uparrow\downarrow\right\rangle - \left|\downarrow\uparrow\right\rangle \right),
\end{aligned}$$

Note that the three states with $j = 1$ form an irreducible representation, as does the state with $j = 0$. The tensor product of two spin $1/2$ states decomposed therefore as

$$\mathbf{2} \otimes \mathbf{2} = \mathbf{3} \oplus \mathbf{1},$$

where $\mathbf{n}$ denotes an irreducible representation of dimension n.

4.6.2 The Wigner-Eckart theorem

Let us make a small detour and discuss the Wigner-Eckart theorem. Consider first in quantum mechanics the matrix element of an operator O between two states $|\phi\rangle$ and $|\psi\rangle$:

$$M = \left\langle \phi \middle| \hat{O} \middle| \psi \right\rangle.$$

Suppose that a unitary group transformation ($U^{-1} = U^\dagger$) acts on the states as

$$\left|\psi'\right\rangle = U \left|\psi\right\rangle$$

and on operators as

$$\hat{O}' = U \hat{O} U^\dagger.$$

Then

$$M' = \langle\phi'|\hat{O}'|\psi'\rangle = \left\langle\phi\left|U^\dagger\left(U\hat{O}U^\dagger\right)U\right|\psi\right\rangle = \langle\phi|\hat{O}|\psi\rangle = M.$$

If

$$\hat{O}' = \hat{O},$$

or equivalently

$$\left[\hat{O},U\right] = 0$$

we say that the operator $\hat{O}$ transforms as a scalar (or as a singlet). This is the simplest case. We discuss now more general cases. We first fix the group to be $SU(2)$. Suppose we are given a set of $(2k+1)$ operators T_q^k, $-k \leq q \leq k$, which transform irreducible under the group $SU(2)$. That is to say that

$$\left(T_q^k\right)' = UT_q^kU^\dagger$$

can be expressed as

$$\left(T_q^k\right)' = \mathcal{D}_{qq'}^{(k)}T_{q'}^k,$$

where we denote the $(2k+1)\times(2k+1)$ matrix representation of the transformation U by

$$\mathcal{D}_{qq'}^{(k)},$$

with $q,q' = -k,...,k$. We call T_q^k a **tensor operator of rank** k.

A tensor operator of rank zero is a scalar. An example of a tensor operator of rank 1 is given by the three generators of $SU(2)$:

$$\begin{aligned} J_1 &= \frac{1}{\sqrt{2}}\left(I_1+iI_2\right), \\ J_0 &= I_3, \\ J_{-1} &= \frac{1}{\sqrt{2}}\left(I_1-iI_2\right). \end{aligned}$$

$\{J_{-1},J_0,J_1\}$ define the spherical basis. The generators transform under $SU(2)$ as the adjoint representation. For $SU(2)$ the adjoint representation is the **3** representation. The spherical basis transforms in this notation as

$$\left(J_q\right)' = \mathcal{D}_{qq'}^{(1)}J_{q'}.$$

An equivalent definition for a tensor operator is a set of $(2k+1)$ operators satisfying

$$\begin{aligned} \left[I_3,T_q^k\right] &= qT_q^k, \\ \left[I_\pm,T_q^k\right] &= \sqrt{(k\mp q)(k\pm q+1)}T_{q\pm1}^k = \sqrt{k(k+1)-q(q\pm1)}T_{q\pm1}^k. \end{aligned}$$

We can now state the Wigner-Eckart theorem:

$$\left\langle j'm' \left| T_q^k \right| jm \right\rangle = \frac{1}{\sqrt{2j'+1}} C_{jkmq}^{j'm'} \left\langle j' \left|\left| T^k \right|\right| j \right\rangle .$$

The important point is that the double bar matrix element $\langle j'||T^k||j\rangle$ is independent of m, m' and q. The dependence on m, m' and q is entirely given by the Clebsch-Gordan coefficients $C_{jkmq}^{j'm'}$.

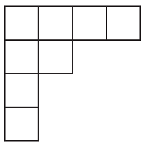
4.6.3 Young diagrams

We have seen that the tensor product of two fundamental representations of $SU(2)$ decomposes as

$$\mathbf{2} \otimes \mathbf{2} = \mathbf{3} \oplus \mathbf{1},$$

into a direct sum of irreducible representations. We generalise this now to general irreducible representations of $SU(N)$.

Definition: A Young diagram is a collection of m boxes $\Box$ arranged in rows and left-justified. To be a legal Young diagram, the number of boxes in a row must not increase from top to bottom. An example for a Young diagram is



Let us denote the number of boxes in row j by λ_j. Then a Young diagram is a partition of m defined by the numbers $(\lambda_1, \lambda_2, ..., \lambda_n)$ subject to

$$\lambda_1 + \lambda_2 + ... + \lambda_n = m,$$
$$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n.$$

The example diagram above therefore corresponds to

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (4, 2, 1, 1)$$

The number of rows is denoted by n. For $SU(N)$ we consider only Young diagrams with $n \leq N$.

Let us further define $(n-1)$ numbers p_j by

$$\begin{aligned} p_1 &= \lambda_1 - \lambda_2, \\ p_2 &= \lambda_2 - \lambda_3, \\ &\;\;... \\ p_{n-1} &= \lambda_{n-1} - \lambda_{n-2}. \end{aligned}$$

The example above has

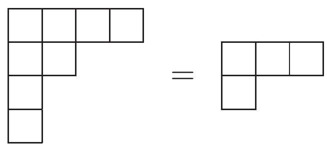
$$(p_1, p_2, p_3) \quad = \quad (2, 1, 0)$$

Correspondence between Young diagrams and irreducible representations: Recall from the last lecture that we could label any irreducible representation of a simple Lie algebra of rank r by either the r eigenvalues of the Casimir operators or by the r numbers $(p_1, ..., p_r)$ appearing when expressing the dominant weight of the representation in terms of the fundamental dominant weights:

$$\vec{M} \quad = \quad \sum_{i=1}^{r} p_i \, \vec{M}^{(i)}$$

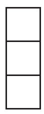
The group $SU(N)$ has rank $N-1$ and we associate to an irreducible representation of $SU(N)$ given through $(p_1, ..., p_{N-1})$ the Young diagram corresponding to $(p_1, ..., p_{N-1})$.

As only differences in the number boxes between succesive rows matter, we are allowed to add any completed column of N boxes from the left. Therefore in $SU(4)$ we have

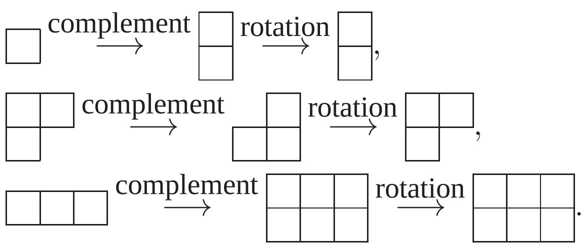


The fundamental representation of $SU(N)$ is always represented by a single box

The trivial (or singlet) representation is alway associated with a column of N boxes. For $SU(3)$:



The complex conjugate representation of a given representation is associated with the conjugate Young diagram. This diagram is obtained by taking the complement with respect to complete columns of N boxes and rotate through 180° to obtain a legal Young diagram.
Examples for $SU(3)$:



The hook rule for the dimensionality of an irreducible representation:

i) Place integers in the boxes, starting with N in the top left box, increase in steps of 1 across rows, decrease in steps of 1 down columns:

$$\begin{array}{|c|c|c|}\hline N & N+1 & N+2 \\ \hline N-1 & N & \multicolumn{1}{c}{} \\ \cline{1-2} N-2 & N-1 & \multicolumn{1}{c}{} \\ \cline{1-2}\end{array}$$

ii) Compute the numerator as the product of all integers.

iii) The denominator is given by multiplying all hooks of a Young diagram. A hook is the number of boxes that one passes through on entering the tableau along a row from the right hand side and leaving down a column.

Some examples for $SU(3)$:

$$\begin{array}{|c|c|}\hline 3 & 4 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}: \qquad \dim = \frac{2\cdot 3\cdot 4}{1\cdot 3\cdot 1} = 8,$$

$$\begin{array}{|c|c|c|}\hline 3 & 4 & 5 \\ \hline\end{array}: \qquad \dim = \frac{3\cdot 4\cdot 5}{1\cdot 2\cdot 3} = 10,$$

$$\begin{array}{|c|c|c|}\hline 3 & 4 & 5 \\ \hline 2 & 3 & 4 \\ \hline\end{array}: \qquad \dim = \frac{2\cdot 3\cdot 4\cdot 3\cdot 4\cdot 5}{1\cdot 2\cdot 3\cdot 2\cdot 3\cdot 4} = 10.$$

Rules for tensor products: We now give rules for tensor products of irreducible representations represented by Young diagrams. As an example we take in $SU(3)$

$$\begin{array}{|c|}\hline \\ \hline \\ \hline\end{array} \otimes \begin{array}{|c|c|}\hline & \\ \hline & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$$

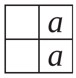
i) Label the boxes of the second factor by row, e.g. a, b, c, ...:

$$\begin{array}{|c|c|}\hline & \\ \hline & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \longrightarrow \begin{array}{|c|c|}\hline a & a \\ \hline b & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}.$$

ii) Add the boxes with the a's from the lettered diagram to the right-hand ends of the rows of the unlettered diagram to form all possible legitimate Young diagrams that have no more than one a per column.

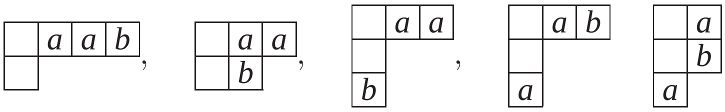
$$\begin{array}{|c|c|c|}\hline & a & a \\ \hline & \multicolumn{2}{c}{} \\ \cline{1-1}\end{array}, \quad \begin{array}{|c|c|}\hline & a \\ \hline & \multicolumn{1}{c}{} \\ \cline{1-1} a & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$$

Note that the diagram

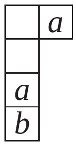


is not allowed since it has one column with two *a*'s.

iii) Repeat the same with the *b*'s, then with the *c*'s, etc.

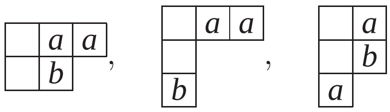


Note that the diagram

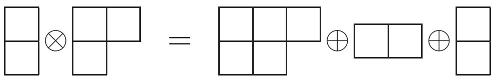


is not allowed for $SU(3)$, since it has more than 3 rows.

iv) A sequence of letters *a*, *b*, *c*, ... is admissible if at any point in the sequence at least as many *a*'s have occured as *b*'s, at least as many *b*'s have occured as *c*'s, etc. Thus *abcd* and *aabcb* are admissible sequences, while *abb* and *acb* are not. From the diagrams in step iii) throw away all diagrams in which the sequence of letters formed by reading right to left in the first row, then in the second row, etc., is not admissible. This leaves



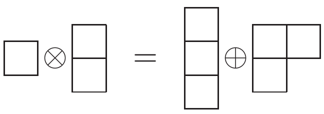
Removing complete columns of 3 boxes, we finally obtain



For the dimensions we have

$$\mathbf{3} \otimes \mathbf{8} \quad = \quad \mathbf{15} \oplus \mathbf{6} \oplus \mathbf{3}.$$

As a further example let us calculate in $SU(3)$ the tensor product of the fundamental representation with its complex conjugate representation:



For the dimensions we have

$$\mathbf{3} \otimes \bar{\mathbf{3}} \quad = \quad \mathbf{1} \oplus \mathbf{8}.$$

As a final example let us consider

$$\ydiagram{1} \otimes \ydiagram{1} \otimes \ydiagram{1} \quad = \quad \ydiagram{1,1,1} \oplus \ydiagram{2,1} \oplus \ydiagram{2,1} \oplus \ydiagram{3}$$

For the dimensions we have

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} \quad = \quad \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}.$$

4.7 Applications

4.7.1 Selection rules in quantum mechanics

We consider a time-independent quantum mechanical system

$$\hat{H}\psi(\vec{x}) \quad = \quad E\psi(\vec{x}),$$

together with a group G acting on the states by

$$\psi'(\vec{x}) \quad = \quad \rho(g)\psi(\vec{x}),$$

and on operators by

$$\hat{O}' \quad = \quad \rho(g)\hat{O}\rho(g^{-1}).$$

If the group G leaves the Hamilton operator invariant

$$\hat{H} \quad = \quad \rho(g)\hat{H}\rho(g^{-1}),$$

we may label the states by the irreducible representations of the group G. Likewise, we focus now on operators $\hat{O}$, which transform as an irreducible representation of the group G. We denote by $\psi^{(\alpha)}(\vec{x})$ a set of states transforming like the irreducible representation ρ_α:

$$\psi^{(\alpha)'}(\vec{x}) \quad = \quad \rho_\alpha(g)\psi^{(\alpha)}(\vec{x}),$$

Further we denote by $\hat{O}^{(\beta)}$ a set of operators transforming like like the irreducible representation ρ_β

$$\hat{O}^{(\beta)\prime} \quad = \quad \rho_\beta(g)\hat{O}^{(\beta)},$$

and finally by $\phi^{(\gamma)}$ a set of states transforming like the irreducible representation ρ_γ

$$\phi^{(\gamma)'}(\vec{x}) \quad = \quad \rho_\gamma(g)\phi^{(\gamma)}(\vec{x}).$$

Recall that we defined for $SU(2)$ a tensor operator as a set of $(2k+1)$ operators T^k_q ($-k \leq q \leq k$), which transform irreducible as

$$\left(T^k_q\right)' \quad = \quad \mathcal{D}^{(k)}_{qq'} T^k_{q'}$$

under $SU(2)$. The transformation law $\hat{O}^{(\beta)\prime} = \rho_\beta(g)\hat{O}^{(\beta)}$ is nothing else than the straightforward generalisation to an arbitrary group.

We are interested in the matrix elements

$$\left\langle \phi^{(\gamma)} \middle| \hat{O}^{(\beta)} \middle| \psi^{(\alpha)} \right\rangle.$$

Suppose the decomposition of the tensor product $\rho_\alpha \otimes \rho_\beta$ into irreducible representations reads

$$\rho_\alpha \otimes \rho_\beta = \sum_\delta n_\delta \rho_\delta,$$

where the number n_δ indicates, how often the irreducible representation ρ_δ occurs in the decomposition.

We then have the following theorem: If ρ_γ does not occur in the decomposition $\sum\limits_\delta n_\delta \rho_\delta$, then the matrix

$$\left\langle \phi^{(\gamma)} \middle| \hat{O}^{(\beta)} \middle| \psi^{(\alpha)} \right\rangle$$

vanishes. This is called a selection rule.

As an example we consider the case $G = SU(2)$. The irreducible representations are labelled by $j = 0, 1/2, 1, 3/2, ...$ and the states within an irreducible representation by m with $-j \le m \le j$. The Wigner-Eckard theorem states

$$\left\langle j'm' \middle| T^k_q \middle| jm \right\rangle = \frac{1}{\sqrt{2j'+1}} C^{j'm'}_{jkmq} \left\langle j' \middle|\middle| T^k \middle|\middle| j \right\rangle.$$

The Clebsch-Gordon coefficients $C^{j'm'}_{jkmq}$ vanish whenever the irreducible representation $\rho_{j'}$ is not contained in the decomposition of the tensor product $\rho_j \otimes \rho_k$.

4.7.2 Gauge symmetries and the Standard Model of particle physics

The Standard Model of elementary particle physics is based on a gauge theory with gauge group

$$SU(3)_{\text{colour}} \times SU(2)_{\text{weak isospin}} \times U(1)_Y,$$

where $SU(3)_{\text{colour}}$ corresponds to the strong interactions, $SU(2)_{\text{weak isospin}}$ to the weak interaction and $U(1)_Y$ to the hypercharge. The gauge symmetry of $SU(2)_{\text{weak isospin}} \times U(1)_Y$ is spontaneously broken to a subgroup $U(1)_{\text{Q}}$, where $U(1)_{\text{Q}}$ corresponds to the electroc charge. We will not go here into the details of the mechanism of spontaneous symmetry breaking, but focus on an unbroken gauge theory. The strong interactions with gauge group SU_{colour} provide an example.

All particles are classified according to representations of the gauge group. Fermions transform as the fundamental representation of the gauge group. For $SU(3)_{\text{colour}}$ the quark fields form a three-dimensional representation $q_i(x)$, $i = 1, 2, 3$. Let us denote a gauge transformation by

$$U(x) = \exp\left(-i\theta_a(x)T^a\right),$$

the dependence on the space-time coordinates x indicates that the gauge transformation may vary from one space-time point to another. The hermitian matrices T^a are the generators of the gauge group in the fundamental representation. Under this gauge transformation, the quarks transform as follows:

$$q'_i(x) \;=\; U(x)_{ij} q_j(x),$$

or in vector/matrix-notation without indices

$$q'(x) \;=\; U(x) q(x).$$

The anti-fermions transform as the complex conjugate of the fundamental representation. In physics we usually take the anti-quark fields $\bar{q}_i(x)$, $i = 1,2,3$ as the components of a bra-vector and write the transformation law as

$$\bar{q}'(x) \;=\; \bar{q}(x) U^\dagger(x).$$

Taking the transpose of this equation we get

$$\left[\bar{q}'(x)\right]^T \;=\; U^*(x) \bar{q}(x)^T.$$

We note that the combination $\bar{q}(x) \cdot q(x)$ is gauge-invariant:

$$\bar{q}'(x) \cdot q'(x) \;=\; \left(\bar{q}(x) U^\dagger(x)\right) \cdot (U(x) q(x)) = \bar{q}(x) \cdot q(x).$$

The gauge boson fields are given by the gauge potentials $A^a_\mu(x)$, where the index a runs from 1 to the number of generators of the Lie group. For $SU(N)$ this number is given by $N^2 - 1$. For $SU(3)_{\text{colour}}$ we have eight gauge boson fields, which are called gluon fields. We have already seen that the gauge potential transforms as

$$T^a A^{a\prime}_\mu(x) \;=\; U(x) \left(T^a A^a_\mu(x) + \frac{i}{g} \partial_\mu \right) U(x)^\dagger.$$

This transformation law ensures that the expression

$$\mathcal{L} \;=\; -\frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu}, \quad \text{with} \quad F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu$$

is invariant. For space-time independent transformations $U(x) = U$ the transformation law reduces to

$$T^a A^{a\prime}_\mu(x) \;=\; U T^a A^a_\mu(x) U^\dagger.$$

For infinitessimal transformations we have

$$\begin{aligned}
T^a A^{a\prime}_\mu(x) &= \left(1 - i\theta_b T^b + \ldots\right) T^a A^a_\mu(x) \left(1 + i\theta_c T^c + \ldots\right) \\
&= T^a A^a_\mu(x) + i \left[T^a, T^b\right] \theta_b A^a_\mu + \ldots = T^a A^a_\mu(x) - f^{abc} T^c \theta_b A^a_\mu + \ldots \\
&= T^a A^a_\mu(x) + f^{cba} T^c \theta_b A^a_\mu + \ldots = T^a A^a_\mu(x) + f^{abc} T^a \theta_b A^c_\mu + \ldots
\end{aligned}$$

We therefore find

$$A^{a\,\prime}_{\mu}(x) \;=\; \left(\delta^{ac} + f^{abc}\theta_b\right) A^c_{\mu} = \exp\left(-i\theta_b M^b_{ac}\right) A^c_{\mu},$$

where

$$M^b_{ac} \;=\; i f^{abc}$$

are hermitian $(N^2-1)\times(N^2-1)$ matrices defining the adjoint representation. Therefore the gauge bosons belong to the adjoint representation.

4.7.3 The naive parton model

In the early days of particle physics only three quarks (up, down and strange) were known, together with the corresponding anti-quarks (anti-up, anti-down and anti-strange). Further it was observed that there is an approximate $SU(3)_{\text{flavour}}$-symmetry, called flavour symmetry. Under flavour symmetry the quarks (u,d,s) transform as the fundamental representation of $SU(3)_{\text{flavour}}$, while the anti-quarks $(\bar{u},\bar{d},\bar{s})$ transform as the complex conjugate of the fundamental representation. In the naive parton model mesons consist of a quark and an anti-quark. As a short hand notation we write

$$q\bar{q}' = q \otimes \bar{q}'.$$

The tensor product forms a nine-dimensional representation. As a basis we can take

$$u\bar{u}, u\bar{d}, u\bar{s}, d\bar{u}, d\bar{d}, d\bar{s}, s\bar{u}, s\bar{d}, s\bar{s}.$$

This representation is reducubible. Using Young diagrams we find

$$\yng(1) \otimes \yng(1,1) \;=\; \yng(1,1,1) \oplus \yng(2,1)$$

For the dimensions we have

$$\mathbf{3} \otimes \bar{\mathbf{3}} \;=\; \mathbf{1} \oplus \mathbf{8}.$$

Therefore the tensor representation reduces to a one-dimensional (singlet) representation and an eight-dimensional (octet) representation. Let us first discuss the singlet representation: The linear combination

$$\eta' \;=\; \frac{1}{\sqrt{3}}\left(u\bar{u} + d\bar{d} + s\bar{s}\right)$$

transforms as a singlet under $SU(3)_{\text{flavour}}$. This can be seen as follows: We write

$$\vec{q} = \begin{pmatrix} u \\ d \\ s \end{pmatrix}, \qquad \vec{\bar{q}} = \begin{pmatrix} \bar{u} \\ \bar{d} \\ \bar{s} \end{pmatrix}.$$

Under a $SU(3)_{\text{flavour}}$ transformation, the quarks and the antiquarks transform as

$$q_i' = U_{ij} q_j, \qquad \bar{q}_i' = U_{ij}^* \bar{q}_j.$$

For the η' we can equally well write

$$\eta' = \frac{1}{\sqrt{3}} \vec{\bar{q}}^T \cdot \vec{q}.$$

This linear combination transforms under a $SU(3)_{\text{flavour}}$ transformation as follows:

$$(\eta')' = \frac{1}{\sqrt{3}} \left(\vec{\bar{q}}^T\right)' \cdot \vec{q}' = \frac{1}{\sqrt{3}} \left(U^* \vec{\bar{q}}\right)^T U \vec{q} = \frac{1}{\sqrt{3}} \vec{\bar{q}}^T U^\dagger U \vec{q} = \eta'.$$

(On the left-hand side of this equation the first prime is part of the name, the second prime denotes the transformed quantity.) Therefore the η' transforms into itself and is a singlet. (The factor $1/\sqrt{3}$ is only included for the normalisation. If the states $q\bar{q}$ have norm 1, so does η'.)

Before discussing the octet representation we first look how the operators H_1, H_2, $E_{\pm 1}$, $E_{\pm 2}$ and $E_{\pm 3}$ act on the states $q\bar{q}$ of the nine-dimensional representation. A finite $SU(3)_{\text{flavour}}$ transformation acts on such a state as

$$U\left(q \otimes \bar{q}\right) = (Uq) \otimes \left(U^* \bar{q}\right).$$

We can always write

$$U = \exp\left(-i\theta_a T^a\right).$$

In order to find the action of the generators on the tensor representation we expand to first order in θ_a:

$$(1 - i\theta_a T^a + \ldots)\left(q \otimes \bar{q}\right) = q \otimes \bar{q} - (i\theta_a T^a q) \otimes \bar{q} + q \otimes (i\theta_a T^{a*})\bar{q} + \ldots$$

Here we assumed that the parameters θ_a are real. Therefore

$$T^a\left(q \otimes \bar{q}\right) = (T^a q) \otimes \bar{q} - q \otimes \left(T^{a*} \bar{q}\right).$$

The generators H_1 and H_2 are diagonal and real ($H_1^* = H_1$, $H_2^* = H_2$). We can use this formula to obtain the action of H_1 and H_2 on the states $q\bar{q}$. For example H_1 acts on $u\bar{u}$ as

$$H_1(u\bar{u}) = \frac{1}{\sqrt{6}} u\bar{u} - \frac{1}{\sqrt{6}} u\bar{u} = 0.$$

Doing this for all other basis vectors $q\bar{q}$ of the nine-dimensional representation, we find that H_1 is given in this representation by

$$H_1 = \text{diag}\left(0, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, 0, -\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, 0\right).$$

Similar, H_2 is given by

$$H_2 \quad = \quad \mathrm{diag}\left(0,0,\frac{1}{\sqrt{2}},0,0,\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}},0\right).$$

For the non-diagonal generators we have to be a little bit more carefully. Let us first discuss the simpler $SU(2)$ case. We can write

$$i\theta_1 I_1 + i\theta_2 I_2 \quad = \quad i\theta_+ I_+ + i\theta_- I_-$$

with

$$\theta_\pm = \frac{1}{\sqrt{2}}(\theta_1 \mp i\theta_2), \qquad I_\pm = \frac{1}{\sqrt{2}}(I_1 \pm iI_2).$$

Note that the coefficients $\theta_\pm$ are now complex. Therefore complex conjugation gives

$$(i\theta_+ I_+ + i\theta_- I_-)^* \quad = \quad -i\theta_+^* I_+^* - i\theta_-^* I_-^* = -i\theta_- I_+ - i\theta_+ I_-$$

Therefore

$$I'_+ = -I_-, \qquad I'_- = -I_+.$$

This is true in general: E_α acts on the complex conjugate representation as $-E_{-\alpha}$. We have

$$E_1(u) = 0, \quad E_1(d) = \frac{1}{\sqrt{3}}u, \quad E_1(s) = 0,$$

and with the explanations above

$$E_1(\bar{u}) = -\frac{1}{\sqrt{3}}\bar{d}, \quad E_1\left(\bar{d}\right) = 0, \quad E_1(\bar{s}) = 0.$$

This allows us to write down the action of E_1 on the nine-dimensional basis $q\bar{q}'$:

$$E_1 \quad = \quad \frac{1}{\sqrt{3}}\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

In a similar way we can obtain the matrix representations of E_{-1}, $E_{\pm 2}$ and $E_{\pm 3}$.

Note that we can represent η' in our basis as

$$\eta' \quad = \quad (1,0,0,0,1,0,0,0,1)^T.$$

We have

$$H_1\eta' = 0, \qquad H_2\eta' = 0,$$

therefore η' corresponds to the weight vector $(0,0)$. We further have

$$E_{\pm\alpha}\eta' \;=\; 0,$$

i.e. η' is annihilated by all operators $E_{\pm\alpha}$.

We are now in a position to tackle the octet representation: We first recall that the fundamental representation has the weight vectors

$$\vec{m}_1 = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{3\sqrt{2}} \end{pmatrix}, \quad \vec{m}_2 = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ \frac{1}{3\sqrt{2}} \end{pmatrix}, \quad \vec{m}_3 = \begin{pmatrix} 0 \\ -\frac{\sqrt{2}}{3} \end{pmatrix}.$$

The highest weight vector of the fundamental representation is

$$\vec{M}_1 = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{3\sqrt{2}} \end{pmatrix},$$

The complex conjugate representation has the weight vectors

$$\vec{m}'_1 = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{3\sqrt{2}} \end{pmatrix}, \quad \vec{m}'_2 = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{3\sqrt{2}} \end{pmatrix}, \quad \vec{m}'_3 = \begin{pmatrix} 0 \\ \frac{\sqrt{2}}{3} \end{pmatrix}.$$

The highest weight vector is here

$$\vec{M}_2 = \vec{m}'_3 = \begin{pmatrix} 0 \\ \frac{\sqrt{2}}{3} \end{pmatrix}.$$

$\vec{M}_1$ and $\vec{M}_2$ are the two fundamental dominant weights. The octet representation has the dominant weight

$$\vec{M} \;=\; \vec{M}_1 + \vec{M}_2 = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

To which state does this weight vector correspond ? Let us call the state K^+. The weights $1/\sqrt{6}$ and $1/\sqrt{2}$ are the eigenvalues of H_1 and H_2 applied to K^+. In other words, we must have

$$\begin{aligned} H_1K^+ &= \frac{1}{\sqrt{6}}K^+, \\ H_2K^+ &= \frac{1}{\sqrt{2}}K^+. \end{aligned}$$

These equations are easily solved and one finds $K^+ = cu\bar{s}$, where c is some constant. Requiring that K^+ has unit norm leads to

$$K^+ \;=\; u\bar{s}.$$

There are two ways to obtain the other states in this representation. The first possibility constructs first all possible weights of the representation. For the octet representation the occuring weights are

$$\vec{m}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \vec{m}_1 = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \quad \vec{m}_2 = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \quad \vec{m}_3 = \begin{pmatrix} -\frac{\sqrt{6}}{3} \\ 0 \end{pmatrix},$$
$$\vec{m}_4 = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}, \quad \vec{m}_5 = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}, \quad \vec{m}_6 = \begin{pmatrix} \frac{\sqrt{6}}{3} \\ 0 \end{pmatrix}.$$

For each weight we can then repeat the exercise and solve

$$H_1\phi = m_1^{(i)}\phi, \qquad H_2\phi = m_2^{(i)}\phi,$$

where $m_1^{(i)}$ and $m_2^{(i)}$ are the components of $\vec{m}_i = (m_1^{(i)}, m_2^{(i)})$ and ϕ is the state which we would like to solve for. For the weight vectors $\vec{m}_1$ to $\vec{m}_6$ we find

$$\begin{array}{lcl}
\vec{m}_1 & : & K^+ = u\bar{s}, \\
\vec{m}_2 & : & K^0 = d\bar{s}, \\
\vec{m}_3 & : & \pi^- = d\bar{u}, \\
\vec{m}_4 & : & K^- = s\bar{u}, \\
\vec{m}_5 & : & \bar{K}^0 = s\bar{d}, \\
\vec{m}_6 & : & \pi^+ = u\bar{d}.
\end{array}$$

The weight $\vec{m}_0$ is degenerate. Solving

$$H_1\phi = 0, \qquad H_2\phi = 0$$

yields

$$\phi = c_1 u\bar{u} + c_2 d\bar{d} + c_3 s\bar{s}$$

with arbitrary constants c_1, c_2 and c_3. We recall that $c_1 = c_2 = c_3$ corresponds to the singlet representation, therefore for the octet representation we are only interested in the vector space orthogonal to η'. This gives a two-dimensional vector space. A convenient basis is given by

$$\begin{aligned}
\pi^0 &= \frac{1}{\sqrt{2}}\left(u\bar{u} - d\bar{d}\right), \\
\eta &= \frac{1}{\sqrt{6}}\left(u\bar{u} + d\bar{d} - 2s\bar{s}\right).
\end{aligned}$$

Both states belong to the octet representation: We know that the octet representation is eight-dimensional and that the weight spaces corresponding to $\vec{m}_1$ to $\vec{m}_6$ are one-dimensional. Therefore the weight space corresponding to $\vec{m}_0$ must be two-dimensional.

The second possibility of finding the remaining states in the octet representation starting from the state K^+ with the dominant weight is given by repeatidly applying the operators $E_{\pm\alpha}$ to the state K^+. We then obtain multiples of the other states. For example:

$$\begin{aligned} E_{-1}K^+ &= \frac{1}{\sqrt{3}} d\bar{s} = \frac{1}{\sqrt{3}} K^0, \\ E_{-2}K^+ &= \frac{1}{\sqrt{3}} (s\bar{s} - u\bar{u}) = -\frac{1}{\sqrt{6}}\pi^0 - \frac{1}{\sqrt{2}}\eta, \\ E_{-3}K^+ &= -\frac{1}{\sqrt{3}} u\bar{d} = -\frac{1}{\sqrt{3}}\pi^+. \end{aligned}$$

We can repeat this procedure with the newly found states K^0, π^+ and the linear combination of π^0 and η, until we have found all states in the representation.

The classification of the pseudo-scalar meson π^0, $\pi^\pm$, K^0, $\bar{K}^0$, $K^\pm$, η and η' according to the representations of $SU(3)_{\text{flavour}}$ was very important in early days of particle physics. With four quark flavours (up, down, strange, charm) the symmetry group can be extended to $SU(4)_{\text{flavour}}$. Adding a fifth quark (bottom quark) would bring us to $SU(5)_{\text{flavour}}$. In principle one could also thing about $SU(6)_{\text{flavour}}$ by adding the top quark. However, the classification of pseudo-scalar mesons according to $SU(6)_{\text{flavour}}$ is not useful, since the top quark is so heavy and does therefore not form composite bound states like pseudo-scalar mesons.

It should be added that the naive parton model has some short-comings. The most important ones are:

- $SU(3)_{\text{flavour}}$ is only an approximate symmetry. The flavour symmetry is explicitly broken by mass terms. As the strange quark mass differs the most from the quark masses of the up- and the down-quark, corrections due to the strange quark mass give the dominant contribution to $SU(3)_{\text{flavour}}$-breaking terms.

- In the modern understanding, a meson does not consist of a quark and an antiquark alone, but in addition contains an indefinite number of gluons and quark-antiquark pairs.

- The physical particles η and η' do not correspond exactly to the pure octet state and the pure singlet state, but are mixtures of both:

$$\begin{aligned} \eta &= \frac{\cos\varphi}{\sqrt{6}} \left(u\bar{u} + d\bar{d} - 2s\bar{s} \right) - \frac{\sin\varphi}{\sqrt{3}} \left(u\bar{u} + d\bar{d} + s\bar{s} \right). \\ \eta' &= \frac{\sin\varphi}{\sqrt{6}} \left(u\bar{u} + d\bar{d} - 2s\bar{s} \right) + \frac{\cos\varphi}{\sqrt{3}} \left(u\bar{u} + d\bar{d} + s\bar{s} \right). \end{aligned}$$

5 The classification of semi-simple Lie algebras

Recall: For a semi-simple Lie algebra $\mathfrak{g}$ of dimension n snd r we had the Cartan standard form

$$\begin{aligned} \left[H_i, H_j\right] &= 0, \\ \left[H_i, E_\alpha\right] &= \alpha_i E_\alpha, \end{aligned}$$

with generators H_i, $i = 1, \ldots, r$ as well as the generators E_α and $E_{-\alpha}$ with $\alpha = 1, \ldots, (n-r)/2$.

The generators H_i generate an Abelian sub-algebra of $\mathfrak{g}$. This sub-algebra is called the **Cartan sub-algebra** of $\mathfrak{g}$.

The r numbers α_i, $i = 1, \ldots, r$ are the components of the root vector $\vec{\alpha} = (\alpha_1, \ldots, \alpha_r)$.

We have already seen that if if $\vec{\alpha}$ and $\vec{\beta}$ are root vectors so is

$$\vec{\gamma} = \vec{\beta} - \frac{2\vec{\alpha}\cdot\vec{\beta}}{\alpha^2}\vec{\alpha}$$

Let us now put this a little bit more formally. For any root vector α we define a mapping W_α from the set of root vectors to the set of root vectors by

$$W_\alpha(\beta) = \vec{\beta} - \frac{2\vec{\alpha}\cdot\vec{\beta}}{\alpha^2}\vec{\alpha}$$

W_α can be described as the reflection by the plane Ω_α perpendicular to α. It is clear that this mapping is an involution: After two reflections one obtains the original root vector again. The set of all these mappings W_α generates a group, which is called the **Weyl group**.

Since W_α maps a root vector to another root vector, we have the following theorem:

Theorem: The set of root vectors is invariant under the Weyl group.

Actually, a more general result holds: We have seen that if $\vec{m}$ is a weight and if $\vec{\alpha}$ is a root vector then

$$W_\alpha(\vec{m}) = \vec{m} - \frac{2\vec{\alpha}\cdot\vec{m}}{\alpha^2}\vec{\alpha}$$

is again a weight vector. Therefore we can state that the following theorem:

Theorem: The set of weights of any representation of $\mathfrak{g}$ is invariant under the Weyl group.

The previous theorem is a special case of this one, as the root vectors are just the weights of the adjoint representation.

For the weights we defined an ordering. $\vec{m}$ is said to be higher than $\vec{m}'$ if the r^{th} component of $(\vec{m}-\vec{m}')$ is positive (if zero look at the $(r-1)^{\text{th}}$ component). This applies equally well to roots.

Definition: A root vectors $\vec{\alpha}$ is called positive, if $\vec{\alpha} > \vec{0}$.

Therefore the set of non-zero root vectors R decomposes into

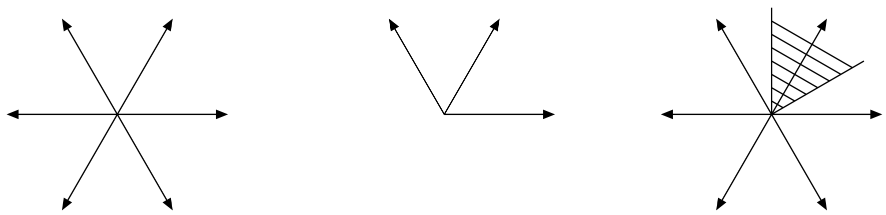
$$R = R^+ \cup R^-,$$

where R^+ denotes the positive roots and R^- denotes the negative roots.

Definition: The (closed) **Weyl chamber** relative to a given ordering is the set of points $\vec{x}$ in the r-dimensional space of root vectors, such that

$$2\frac{\vec{x}\cdot\vec{\alpha}}{\alpha^2} \geq 0 \qquad \forall\vec{\alpha}\in R^+.$$

Example: The Weyl chamber for $SU(3)$:



The root system   The positive roots   The Weyl chamber

Let us further recall that if $\vec{\alpha}$ and $\vec{\beta}$ are root vectors then

$$\frac{2\vec{\alpha}\cdot\vec{\beta}}{|\alpha|^2} \quad \text{and} \quad \frac{2\vec{\alpha}\cdot\vec{\beta}}{|\beta|^2}$$

are integers. This restricts the angle between two root vectors to

$$0^\circ, 30^\circ, 45^\circ, 60^\circ, 90^\circ, 120^\circ, 135^\circ, 150^\circ, 180^\circ$$

For $\theta = 30^\circ$ or $\theta = 150^\circ$ the ratio of the length of the two root vectors is

$$\frac{|\alpha|^2}{|\beta|^2} = 3 \text{ or } \frac{1}{3}.$$

For $\theta = 45^\circ$ or $\theta = 135^\circ$ the ratio of the length of the two root vectors is

$$\frac{|\alpha|^2}{|\beta|^2} = 2 \text{ or } \frac{1}{2}.$$

For $\theta = 60^\circ$ or $\theta = 120^\circ$ the ratio of the length of the two root vectors is

$$\frac{|\alpha|^2}{|\beta|^2} = 1.$$

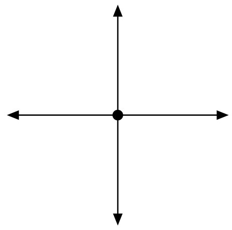
Let us summarise: The root system R of a Lie algebra has the following properties:

1. R is a finite set.
2. If $\vec{\alpha} \in R$, then also $-\vec{\alpha} \in R$.
3. For any $\vec{\alpha} \in R$ the reflection W_α maps R to itself.
4. If $\vec{\alpha}$ and $\vec{\beta}$ are root vectors then $2\vec{\alpha} \cdot \vec{\beta} / |\alpha|^2$ is an integer.

This puts strong constraints on the geometry of a root system. Let us now try to find all possible root systems of rank 1 and 2. For rank 1 the root vectors are one-dimensional and the only possibility is

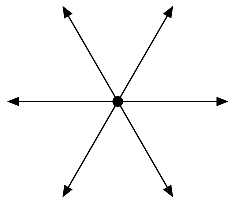
A_1:

This is the root system of $SU(2)$. For rank 2 we first note that due to property (3) the angle between two roots must be the same for any pair of adjacent roots. It will turn out that any of the four angles 90°, 60°, 45° and 30° can occur. Once this angle is specified, the relative lengths of the roots are fixed except for the case of right angles. Let us start with the case $\theta = 90^\circ$. Up to rescaling the root system is

$A_1 \times A_1$: 

This corresponds to $SU(2) \times SU(2)$. This group is semi-simple, but not simple. In general, the direct sum of two root systems is again a root system. A root system which is not a direct sum is called irreducible. An irreducible root system corresponds to a simple group. We would like to classify the irreducible root systems.

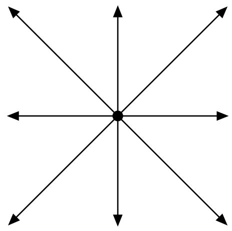
For the angle $\theta = 60^\circ$ we have

A_2 

This is the root system of $SU(3)$.

For the angle $\theta = 45^\circ$ we have

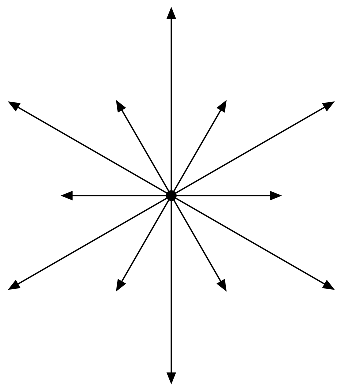
B_2:



This is the root system of $SO(5)$.

Finally, for $\theta = 30^\circ$ we have

G_2:



This is the root system of the exceptional Lie group G_2.

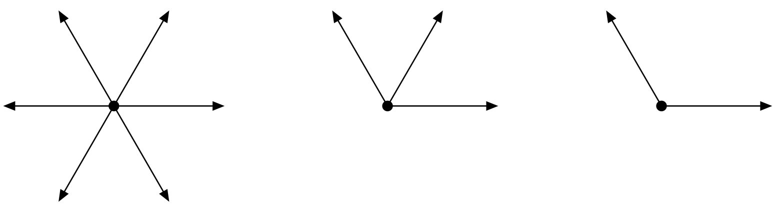
5.1 Dynkin diagrams

Let us try to reduce further the data of a root system. We already learned that with the help of an ordering we can divide the root vectors into a disjoint union of positive and negative roots:

$$R = R^+ \cup R^-.$$

Definition: A positive root vector is called **simple** if it is not the sum of two other positive roots.

Example: For $SU(3)$ we have



root system positive roots simple roots

The angle between the two simple roots is $\theta = 120°$.

The Dynkin diagram of the root system is constructed by drawing one node $\circ$ for each simple root and joining two nodes by a number of lines depending on the angle θ between the two roots:

no lines	○ ○	if $\theta = 90°$
one line	○—○	if $\theta = 120°$
two lines	○⇒○	if $\theta = 135°$
three lines	○⇛○	if $\theta = 150°$

When there is one line, the roots have the same length. If two roots are connected by two or three lines, an arrow is drawn pointing from the longer to the shorter root.

Example: The Dynkin diagram of $SU(3)$ is

○—○

5.2 The classification

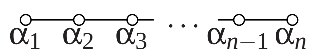
Semi-simple groups are a direct product of simple groups. For a compact group, all unitary representations are finite dimensional.
Real compact semi-simple Lie algebras $\mathfrak{g}$ are in one-to-one correspondence (up to isomorphisms) with complex semi-simple Lie algebras $\mathfrak{g}^{\mathbb{C}}$ obtained as the complexification of $\mathfrak{g}$. Therefore the classification of real compact semi-simple Lie algebras reduces to the classification of complex semi-simple Lie algebras.

Theorem: Two complex semi-simple Lie algebras are isomorphic if and only if they have the same Dynkin diagram.

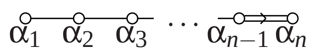
Theorem: A complex semi-simple Lie algebra is simple if and only if its Dynkin diagram is connected.

We have the following classification:

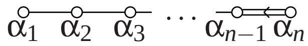
- $A_n \cong SL(n+1,\mathbb{C})$



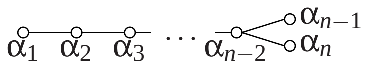
- $B_n \cong SO(2n+1,\mathbb{C})$



- $C_n \cong Sp(n,\mathbb{C})$

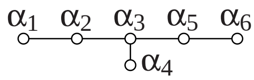


- $D_n \cong SO(2n,\mathbb{C})$

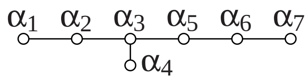


The exceptional groups are

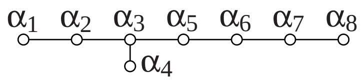
- E_6



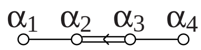
- E_7



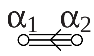
- E_8



- F_4



- G_2



Summary: The classical real compact simple Lie algebras are

$$\begin{aligned} A_n &= SU(n+1) \\ B_n &= SO(2n+1) \\ C_n &= Sp(n) \\ D_n &= SO(2n) \end{aligned}$$

The exceptional groups are

$$E_6, E_7, E_8, F_4, G_2$$

A semi-simple Lie algebra is determined up to isomorphism by specifying which simple summands occur and how many times each one occurs.

5.3 Proof of the classification

Recall: The root system R of a Lie algebra has the following properties:

1. R is a finite set.
2. If $\vec{\alpha} \in R$, then also $-\vec{\alpha} \in R$.
3. For any $\vec{\alpha} \in R$ the reflection W_α maps R to itself.
4. If $\vec{\alpha}$ and $\vec{\beta}$ are root vectors then $2\vec{\alpha} \cdot \vec{\beta} / |\alpha|^2$ is an integer.

With the help of an ordering we can divide the root vectors into a disjoint union of positive and negative roots:

$$R = R^+ \cup R^- .$$

A positive root vector is called simple if it is not the sum of two other positive roots.

The angle between two simple roots is either 90°, 120°, 135° or 150°

The Dynkin diagram of the root system is constructed by drawing one node $\circ$ for each simple root and joining two nodes by a number of lines depending on the angle θ between the two roots:

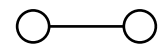
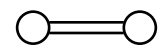
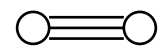
no lines	$\circ \quad \circ$	if $\theta = 90^\circ$
one line	$\circ\!-\!\!-\!\circ$	if $\theta = 120^\circ$
two lines	$\circ\!\Rightarrow\!\circ$	if $\theta = 135^\circ$
three lines	$\circ\!\Rrightarrow\!\circ$	if $\theta = 150^\circ$

When there is one line, the roots have the same length. If two roots are connected by two or three lines, an arrow is drawn pointing from the longer to the shorter root.

Theorem: The only possible connected Dynkin diagrams are the ones listed in the previous section.

To prove this theorem it is sufficient to consider only the angles between the simple roots, the relative length do not enter the proof.

Such diagrams, without the arrows to indicate the relative lengths, are called **Coxeter diagrams**. Define a diagram of n nodes, with each pair connected by 0, 1, 2 or 3 lines, to be **admissible** if there are n independent unit vectors $\vec{e}_1$, ..., $\vec{e}_n$ in a Euclidean space with the angle between $\vec{e}_i$ and $\vec{e}_j$ as follows:

no lines		if $\theta = 90^\circ$
one line		if $\theta = 120^\circ$
two lines		if $\theta = 135^\circ$
three lines		if $\theta = 150^\circ$

Theorem: The only connected admissible Coxeter graphs are the ones of the previous section (without the arrows).

To prove this theorem, we will first prove the following lemmata:

(i) Any sub-diagram of an admissible diagram, obtained by removing some nodes and all lines to them, will also be admissible.

(ii) There are at most $(n-1)$ pairs of nodes that are connected by lines. The diagram has no loops.

(iii) No node has more than three lines to it.

(iv) In an admissible diagram, any string of nodes connected to each other by one line, with none but the ends of the string connected to any other nodes, can be collapsed to one node, and the resulting diagram remains admissible.

Proof of (i): Suppose we have an admissible diagram with n nodes. By definition there are n vectors $\vec{e}_j$, such that the angle between a pair of vectors is in the set

$$\{90^\circ, 120^\circ, 135^\circ, 150^\circ\}$$

Removing some of the vectors $\vec{e}_j$ does not change the angles between the remaining ones. Therefore any sub-diagram of an admissible diagram is again admissible.

Proof of (ii): We have

$$2\vec{e}_i \cdot \vec{e}_j \in \{0, -1, -\sqrt{2}, -\sqrt{3}\}$$

Therefore if $\vec{e}_i$ and $\vec{e}_j$ are connected we have $\theta > 90^\circ$ and

$$2\vec{e}_i \cdot \vec{e}_j \quad \leq \quad -1.$$

Now

$$0 \quad < \quad \left(\sum_i \vec{e}_i\right) \cdot \left(\sum_j \vec{e}_i\right) \quad = \quad n + 2\sum_{i<j} \vec{e}_i \cdot \vec{e}_j \quad < \quad n - \#\,\text{connected pairs}$$

Therefore

$$\#\,\text{connected pairs} \quad < \quad n.$$

Connecting n nodes with $(n-1)$ connections (of either 1, 2 or 3 lines) implies that there are no loops.

Proof of (iii): We first note that

$$\left(2\vec{e}_i \cdot \vec{e}_j\right)^2 \quad = \quad \#\,\text{number of lines between } \vec{e}_i \text{ and } \vec{e}_j.$$

Consider the node $\vec{e}_1$ and let $\vec{e}_i$, $i = 2, \ldots, j$ bet the nodes connected to $\vec{e}_1$. We want to show

$$\sum_{i=2}^{j} (2\vec{e}_1 \cdot \vec{e}_i)^2 \quad < \quad 4.$$

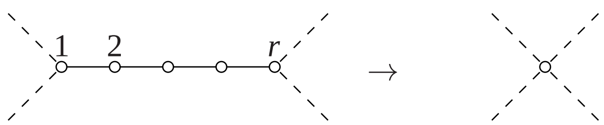
Since there are no loops, no pair of $\vec{e}_2$,...,$\vec{e}_j$ is connected. Therefore $\vec{e}_2$, ..., $\vec{e}_j$ are perpendicular unit vectors. Further, by assumption $\vec{e}_1$, $\vec{e}_2$,...,$\vec{e}_j$ are linearly independent vectors. Therefore $\vec{e}_1$ is not in the span of $\vec{e}_2$,...,$\vec{e}_j$. It follows

$$1 \quad = \quad (\vec{e}_1 \cdot \vec{e}_1)^2 \quad > \quad \sum_{i=2}^{j} (\vec{e}_1 \cdot \vec{e}_i)^2$$

and therefore

$$\sum_{i=2}^{j} (\vec{e}_1 \cdot \vec{e}_i)^2 < 1.$$

Proof of (iv):



If $\vec{e}_1$, ..., $\vec{e}_r$ are the unit vectors corresponding to the string of nodes as indicated above, then

$$\vec{e}\,' \quad = \quad \vec{e}_1 + \ldots + \vec{e}_r$$

is a unit vector since

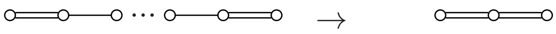
$$\begin{aligned} \vec{e}\,' \cdot \vec{e}\,' &= (\vec{e}_1 + \ldots + \vec{e}_r)^2 = r + 2\vec{e}_1 \cdot \vec{e}_2 + +2\vec{e}_2 \cdot \vec{e}_3 + \ldots + +2\vec{e}_{r-1} \cdot \vec{e}_r \\ &= r - (r-1) = 1. \end{aligned}$$

Further $\vec{e}\,'$ satisfies the same conditions with respect to the other vectors since $\vec{e}\,' \cdot \vec{e}_j$ is either $\vec{e}_1 \cdot \vec{e}_j$ or $\vec{e}_r \cdot \vec{e}_j$.

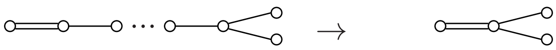
With the help of these lemmata we can now prove the original theorem:

From (iii) it follows that the only connected diagram with a triple line is G_2.

Further we cannot have a diagram with two double lines, otherwise we would have a sub-diagram, which we could contract as

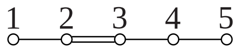


contradicting again (iii). By the same reasoning we cannot have a diagram with a double line and a triple node:



Again this contradicts (iii).

To finish the case with double lines, we rule out the diagram



Consider the vectors

$$\vec{v} = \vec{e}_1 + 2\vec{e}_2, \qquad \vec{w} = 3\vec{e}_3 + 2\vec{e}_4 + \vec{e}_5.$$

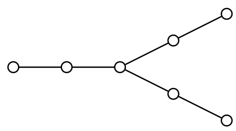
We find

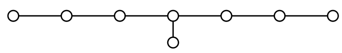
$$(\vec{v} \cdot \vec{w})^2 = 18, \quad |\vec{v}|^2 = 3, \quad |\vec{w}|^2 = 6.$$

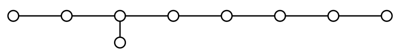
This violates the Cauchy-Schwarz inequality

$$(\vec{v} \cdot \vec{w})^2 \quad < \quad |\vec{v}|^2 \cdot |\vec{w}|^2 \,.$$

By a similar reasoning one rules out the following (sub-) graphs with single lines:







These sub-diagrams rules out all graphs not in the list of the previous section. To finish the proof of the theorem it remains to show that all graphs in the list are admissible. This is equivalent to show that for each Dynkin diagram in the list there exists a corresponding Lie algebra. (The simple root vectors of such a Lie algebra will then have automatically the corresponding angles of the Coxeter diagram.)

To prove the existence it is sufficient to give for each Dynkin diagram an example of a Lie algebra corresponding to it. For the four families A_n, B_n, C_n and D_n we have already seen that they correspond to the Lie algebras of $SU(n+1)$, $SO(2n+1)$, $Sp(n)$ and $SO(2n)$ (or $SL(n+1,\mathbb{C})$, $SO(2n+1,\mathbb{C})$, $Sp(n,\mathbb{C})$ and $SO(2n,\mathbb{C})$ in the complex case). In addition one can write down explicit matrix representations for the Lie algebras corresponding to the five exceptional groups E_6, E_7, E_8, F_4 and G_2.

Finally for the uniqueness let us recall the following theorem: Two complex semi-simple Lie algebras are isomorphic if and only if they have the same Dynkin diagram.