Lie Groups and Lie Algebras for Physicists

Harold Steinacker

Lecture Notes[1]

spring 2023
University of Vienna

Fakultät für Physik
Universität Wien
Boltzmanngasse 5, A-1090 Wien, Austria

Email: `harold.steinacker@univie.ac.at`

[1]These notes are un-finished and undoubtedly contain many mistakes. They are are not intended for publication in its present form, but can hopefully serve as a useful resouce, keeping in mind these limitations. If you find a typo or mistake, please let me know!

Contents

1 Introduction

Lie groups are of great importance in modern theoretical physics. Their main application is in the context of *symmetries*. Symmetries are typically certain transformations (rotations, ...) of a physical system

$$\Phi : \mathcal{M} \to \mathcal{M} \tag{1}$$

which map allowed configurations (such as solutions of some equation of motion) into other allowed configurations. It turns out that this is an extremely powerful concept, because it restricts the dynamics of a system, and allows to use the powerful mathematical tools of group theory.

Since symmetry transformations are specific *maps* of some configuration space, they are *associative*, they can be *iterated*, and hopefully *reversed*. This leads immediately to the concept of a group. *Lie groups* are *continuous groups*, i.e. they contain infinitely many (more precisely a continuum of) different transformations which are related in a differentiable way. It turns out that their structure is essentially encoded in their associated *Lie algebras*, which are very useful for explicit calculation. In fact, pretty much everything in the context of group theory can in principle be calculated. If applicable, group theory provides a natural description and organization of a physical system. For example, in the context of Lagrangian systems (= almost all systems), symmetries arising from Lie groups lead to *conservation laws* via Noethers Theorem.

Some of the applications of Lie groups in physics are as follows:

- *translations*, leading to plane waves, Fourier transforms, the concepts of energy and momentum, and most of your homework problems so far
- *rotations in* $\mathbb{R}^3$ (i.e. $SO(3)$), which leads to the concept of angular momentum
- In Quantum Mechanics, rotations are generalized to $SU(2)$, leading to the concept of *spin* (and precise calculations of Hydrogen atoms etc. etc.)
- Einstein understood that the rotations in $\mathbb{R}^3$ should be extended to *rotations in Minkowski space*, which are described by $SO(1,3)$ leading e.g. to $E = mc^2$.
- Wigner realized that $SO(1,3)$ should be extended to the *Poincaré group*, leading to the correct ("kinematical") description of elementary particles: they are *irreducible unitary representations of the Poincaré group.*
- Modern theories of the dynamics of elementary particles are based on the concept of *gauge groups*, which are infinite-dimensional Lie groups based on classical Lie groups. For the standard model it is $SU(3) \times SU(2) \times U(1)$, and people try to extend it to groups like $SU(5), SO(8), E_6, ...$

  The concept of a quark is entirely based on the group theory of $SU(3)$, and will be explained later.

  At least sometimes gauge groups can be considered as something like $SU(\infty)$.

  There are further "approximate" symmetries, broken symmetries, ... which are very useful in elementary particle theory.
- In string theory, the whole zoo of Lie groups and -algebras occurs including infinite-dimensional ones like the Virasoro algebra, affine Lie algebras, etc.

The examples above are Lie groups. Some interesting discrete groups are:

- crystallographic groups, leading to a classification of crystals
- lattice translations, leading to Bloch waves etc. in solid state physics
- the symmetric group (permutation group), leading e.g. to the concept of Fermions and Bosons

Notice that all of these are *transformation groups*, i.e. they act on some space of states via invertible transformations.

2 Groups

Definition 1 *A **group** is a set G, together with a map*

$$\begin{array}{rccc} \mu: & G \times G & \to & G, \\ & (g_1, g_2) & \mapsto & g_1 \cdot g_2 \end{array} \tag{2}$$

with the following properties:

1. *Associativity: for all $g_1, g_2 \in G$,*

$$g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3. \tag{3}$$

2. *There exists an element e (the **identity element**) in G such that for all $g \in G$,*

$$g \cdot e = e \cdot g = g. \tag{4}$$

3. *For all $g \in G$, there exists an element $g^{-1} \in G$ (the **inverse element**) with*

$$g \cdot g^{-1} = g^{-1} \cdot g = e. \tag{5}$$

If $g \cdot h = h \cdot g$ for all $g, h \in G$, then the group is said to be **commutative** (or **abelian**).

It is easy to show that the identity element e is unique, and so is the inverse for each $g \in G$.

Examples of groups are the integers $\mathbb{Z}$ with the group law being addition, the permutation group (symmetric group) of n elements, and the integers $\mathbb{Z}_n$ modulo n with addition.

A **Lie group** is a group which is also a differentiable manifold; the precise definition will be given later.

Typical examples of Lie groups are the reals $\mathbb{R}$ with the group law being addition, $\mathbb{R} - \{0\}$ and $\mathbb{C} - \{0\}$ with the group law being multiplication, the complex numbers with unit modulus S^1 and multiplication, and matrix groups such as $SU(n), SO(n), GL(n)$,... .

Definition 2 *A **subgroup** of a group G is a subset H of G with the following properties:*

1. *The identity is an element of H.*

2. If $h \in H$, then $h^{-1} \in H$.

3. If $h_1, h_2 \in H$, then $h_1 h_2 \in H$.

It follows that H is a group, with the same product operation as G (but restricted to H). A typical example of a subgroup is the group of orthogonal matrices $SO(n) \subset GL(n)$.

Definition 3 *Let G and H be groups. A map $\phi : G \to H$ is called a* ***homomorphism*** *if $\phi(g_1 g_2) = \phi(g_1)\phi(g_2)$ for all $g_1, g_2 \in G$. If in addition, ϕ is bijective, then ϕ is called an* ***isomorphism****.*

It is easy to see that if e_G the identity element of G, and e_H the identity element of H and $\phi : G \to H$ is a homomorphism, then $\phi(e_G) = e_H$, and $\phi(g^{-1}) = \phi(g)^{-1}$ for all $g \in G$.

The main use of groups in physics is as *transformation groups*, which means that a (Lie) group G acts on some space $\mathcal{M}$ of states of a physical system. This is formalized as follows:

Definition 4 *A* left action *of a Lie group G on a space $\mathcal{M}$ is a map*

$$\begin{aligned} G \times \mathcal{M} &\to \mathcal{M}, \\ (g, \psi) &\mapsto g \triangleright \psi \end{aligned} \tag{6}$$

which respects the group law, $(g_1 g_2) \triangleright \psi = (g_1) \triangleright (g_2 \triangleright \psi)$ and $e \triangleright \psi = \psi$. Equivalently, it is a group homomorphism

$$\pi : G \to Map(\mathcal{M}, \mathcal{M}) \tag{7}$$

from G into the invertible maps from $\mathcal{M}$ to itself, given by $(\pi(g)) \triangleright \psi = g \triangleright \psi$ ("transformation group").

Usually one only needs linear transformations, i.e. maps $\pi : G \to GL(V)$ on some vector space V. Because this is so important, one attaches a name to that concept:

Definition 5 *Let G be a group. Then a* ***(real, complex) representation*** *of G is a group homomorphism*

$$\pi : G \to GL(V)$$

where V is a (real, complex) vector space (i.e. $V = \mathbb{R}^n$ resp. $V = \mathbb{C}^n$ essentially). Equivalently, it is given by a map $G \times V \to V$ as above.

One of the main results of the theory of Lie groups is the classification and description of such "linear" representations. The principal tool is to reduce this problem to an analogous problem for Lie algebras. The goal of this lecture is to explain these things.

3 Examples of Lie groups in physics

3.1 The rotation group $SO(3)$ and its universal covering group $SU(2)$

$SO(3)$ is the rotation group of $\mathbb{R}^3$ which is relevant in classical Mechanics. It acts on the space $\mathbb{R}^3$ as

$$\begin{aligned} SO(3) \times \mathbb{R}^3 &\rightarrow \mathbb{R}^3, \\ (g, \vec{x}) &\mapsto g \cdot \vec{x} \end{aligned} \tag{8}$$

In particular, this is the simplest of all representations of $SO(3)$, denoted by $\pi_3 : SO(3) \rightarrow GL(\mathbb{R}^3)$.

If a physical system is isolated, one should be able to rotate it, i.e. there there should be an action of $SO(3)$ on the space of states $\mathcal{M}$ (=configuration space). In Quantum Mechanics, the space of states is described by a vector space V (the Hilbert space), which therefore should be a representation of $SO(3)$.

It turns out that sometimes (if we deal with spin), $SO(3)$ should be "replaced" by the "spin group" $SU(2)$. In fact, $SU(2)$ and $SO(3)$ are almost (but not quite!) isomorphic. More precisely, there exists a Lie group homomorphism $\phi : SU(2) \rightarrow SO(3)$ which maps $SU(2)$ onto $SO(3)$, and which is *two-to-one*. This is a nice illustration of the importance of global aspects of Lie groups.

To understand this, consider the space V of all 2×2 complex matrices which are hermitean and have trace zero,

$$V = \{\text{hermitean traceless } 2 \times 2 \text{ matrices}\} = \{\begin{pmatrix} x^3, & x^1 - ix^2 \\ x^1 + ix^2, & -x^3 \end{pmatrix}, x^i \in \mathbb{R}\} \tag{9}$$

This is a three-dimensional *real* vector space with the following basis

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

(the Pauli matrices), hence any $x \in V$ can be written uniquely as $x = x^i \sigma_i$. We may define an inner product on V by the formula

$$\langle x, y \rangle = \frac{1}{2} \text{trace}(xy)$$

(Exercise: check that this is an inner product.) A direct computation shows that $\{\sigma_1, \sigma_2, \sigma_3\}$ is an orthonormal basis for V. Hence we can identify V with $\mathbb{R}^3$. Now if U is an element of $SU(2)$ and x is an element of V, then it is easy to see that UxU^{-1} is in V. Thus for each $U \in SU(2)$, we can define a linear map ϕ_U of V to itself by the formula

$$\begin{aligned} \phi: \quad SU(2) \times V &\rightarrow V \\ (U,x) &\rightarrow \phi_U(x) = UxU^{-1} \end{aligned} \tag{10}$$

Moreover, given $U \in \mathsf{SU}(2)$, and $x, y \in V$, note that

$$\langle \phi_U(x), \phi_U(y) \rangle = \frac{1}{2}\text{trace}(UxU^{-1}UyU^{-1}) = \frac{1}{2}\text{trace}(xy) = \langle x, y \rangle$$

Thus ϕ_U is an orthogonal transformation of $V \cong \mathbb{R}^3$, which we can think of as an element of $O(3)$. It follows that the map $U \rightarrow \phi_U$ is a map of $SU(2)$ into $O(3)$. It is very easy to check that this map is a homomorphism (i.e., $\phi_{U_1U_2} = \phi_{U_1}\phi_{U_2}$), and that it is continuous.

Now recall that every element of $O(3)$ has determinant ± 1. Since $SU(2)$ is connected (Exercise), and the map $U \rightarrow \phi_U$ is continuous, ϕ_U actually maps into $SO(3)$. Thus

$$\begin{aligned} \phi: \quad SU(2) &\rightarrow SO(3) \\ U &\rightarrow \phi_U \end{aligned} \tag{11}$$

is a Lie group homomorphism of $SU(2)$ into $SO(3)$. In particular, every representation of $SO(3)$ is automatically a representation of $SU(2)$, but the converse is not true. The map $U \rightarrow \phi_U$ is not one-to-one, since for any $U \in SU(2)$, $\phi_U = \phi_{-U}$. (Observe that if U is in $SU(2)$, then so is $-U$.) In particular, only rotations around 720 degree lead back to the identity in $SU(2)$. This happens for spin in Q.M.

This was illustrated by Dirac as follows: ...

It is now easy to show that ϕ_U is a two-to-one map of $SU(2)$ onto $SO(3)$. Moreover, $SU(2)$ is simply connected, and one can show that it is in a sense the "universal cover" of $SO(3)$, i.e. the "universal rotation group" (i.e. there is no other covering-group of $SU(2)$).

3.1.1 Finite and "infinitesimal" rotations

The rotation operators (or rotation matrices) of vectors in $\mathbb{R}^3$ are well-known to be

$$R(\phi\vec{e_x}) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & \sin(\phi) \\ 0 & -\sin(\phi) & \cos(\phi) \end{pmatrix}, \quad R(\phi\vec{e_y}) := \begin{pmatrix} \cos(\phi) & 0 & -\sin(\phi) \\ 0 & 1 & 0 \\ \sin(\phi) & 0 & \cos(\phi) \end{pmatrix}, \quad \text{und}$$

$$R(\phi\vec{e_z}) := \begin{pmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

One can show (exercise!) that rotations around the axis $\frac{\vec{\phi}}{|\vec{\phi}|}$ by the angle $|\vec{\phi}|$ take the form

$$R(\vec{\phi}) = e^{i\vec{\phi}\cdot\vec{J}} \qquad \in SO(3)$$

for $\vec{\phi} \in \mathbb{R}^3$ and

$$J_x := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad J_y := \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \quad \text{und}$$

$$J_z := \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Note that

$$R((\varphi_1 + \varphi_2)\vec{e}) = R(\varphi_1\vec{e})R(\varphi_2\vec{e}).$$

for any $\vec{e}$. "Infinitesimal" rotations therefore have the form

$$R(\varepsilon\vec{\phi}) = \mathbb{1} + i\varepsilon\vec{\phi}\cdot\vec{J}$$

Therefore the J_i are called "generators" of rotations, and one can check that they satisfy

$$[J_i, J_j] = i\epsilon_{ijk}J_k. \tag{12}$$

This is the "rotation algebra", i.e. the Lie algebra $so(3)$ of $SO(3)$. In general, any (linear) operators $J_i \in GL(V)$ satisfying (12) are called "angular momentum generators", and $R(\vec{\phi}) = e^{i\vec{\phi}\cdot\vec{J}}$ für $\vec{\phi} \in \mathbb{R}^3$ are called rotation operators (in mathematics usually iJ_i is used). The fact that the generators don't commute reflects the fact that the group is non-abelian. One can show that the group structure (i.e. the "table of multiplication") of $SO(3)$ is (almost) uniquely determined by these commutation relations. The precise statement will be given later.

There are many non-equivalent "realizations" (i.e. representations) of (12), one for each half-integer spin. The "simplest" (smallest, fundamental) one is the spin $\frac{1}{2}$ representation, given by the Pauli-matrices: ${J^{(1/2)}}_i = \frac{1}{2}\sigma_i$, where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} o & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Finite rotations of spin $\frac{1}{2}$ objects are obtained as

$$R^{(1/2)}(\vec{\phi}) = e^{i\vec{\phi}\cdot\vec{J}^{(1/2)}} \qquad \in SU(2)$$

They act on "spinors", i.e. elements of $\mathbb{C}^2$. One can easily verify that the spin $\frac{1}{2}$ representation of a rotation around 2π is equal to $-\mathbb{1}$, and rotations around 4π give the identity. One can now show that every representation of the rotation algebra $so(3)$

induces automatically a representation of $SU(2)$ on the representation space V of the generators J_i using the formula

$$\begin{array}{rccc} \pi: & SU(2) & \rightarrow & GL(V) \\ & U = e^{i\vec{\phi}\cdot\vec{J}^{(1/2)}} & \mapsto & e^{i\vec{\phi}\cdot\vec{J}} \end{array} \tag{13}$$

This is a group homomorphism! For example, the group homomorphism (11) can be written as

$$\Phi(e^{i\vec{\phi}\cdot\vec{J}^{(1/2)}}) = e^{i\vec{\phi}\cdot\vec{J}^{(1)}}$$

However, not every representation of $so(3)$ induces a representation of $SO(3)$: this is prevented by global subtleties (related to rotations around 2π). This relation between Lie groups and Lie algebras is very general, and constitutes the core of the theory of Lie groups.

There are also some less obvious applications of these groups in physics. For example, we briefly discuss isospin in nuclear physics.

Isospin Nuclear physics studies how protons p and neutrons n bind together to form a nucleus. The dominating force is the *strong* force, which is much stronger that the electromagnetic and weak forces (not to mention gravity).

A lot of nuclear physics can be explained by the simple assumption that the strong force is independent of the particle type ("flavor") - that is, it is the same for protons and neutrons.

Based on previous experience with QM, one is led to the idea that the neutron and the proton form a doublet $\begin{pmatrix} p \\ n \end{pmatrix}$, which transforms like a spin 1/2 representation of an "isospin" group $SU(2)$. (This is the most interesting group which has 2-dimensional representations). The symmetries are generated by $I_{1,2,3}$ which satisfies the usual $su(2)$ algebra $[I_i, I_j] = i\epsilon_{ijk}I_k$ (hence the name) and act via Pauli-matrices on the isospin doublets:

$$I_i \begin{pmatrix} p \\ n \end{pmatrix} = \frac{1}{2}\sigma_i \begin{pmatrix} p \\ n \end{pmatrix}$$

etc. That is, a proton is represented by

$$|p\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \mathbb{C}^2,$$

and a neutron by

$$|n\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \mathbb{C}^2.$$

Invariance of the strong (nuclear) force under this isospin $SU(2)$ would mean that the Hamiltonian which describes this system commutes with I_i,

$$[H, I_i] = 0.$$

Therefore the eigenstates will be isospin multiplets, and the energy (mass!) should depend only on the total isospin I and not on I_3. In practice, this is approximately correct.

For example, consider a system of 2 nucleons, which according to this idea can form the following states

$$\begin{aligned}
|I=1, I_3=1\rangle &= |p\rangle|p\rangle, \\
|I=1, I_3=0\rangle &= \frac{1}{\sqrt{2}}(|p\rangle|n\rangle + |n\rangle|p\rangle), \\
|I=1, I_3=-1\rangle &= |n\rangle|n\rangle, \\
|I=0, I_3=0\rangle &= \frac{1}{\sqrt{2}}(|p\rangle|n\rangle - |n\rangle|p\rangle)
\end{aligned}$$

as in systems of 2 spin $\frac{1}{2}$ particles. Now consider the three nuclei $^6He, ^6Li$ and ^6Be, which can be regarded respectively as nn, np, and pp system attached to a ^4He core (which has $I = 0$). After correcting for the Coulomb interaction and the neutron-proton mass difference, the observed nuclear masses are as follows

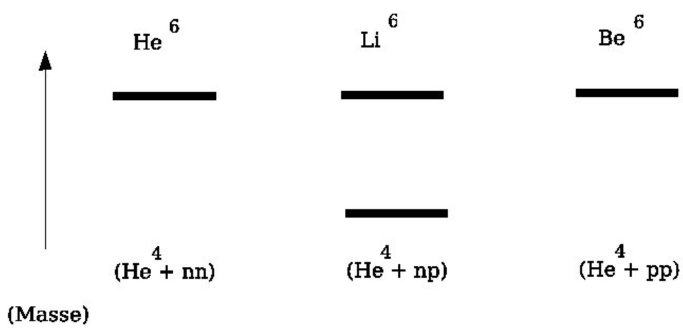


Figure 1: fig:nuclear-mass

This idea of isospin is a precursor of the current understanding that $|p\rangle = |uud\rangle$ and $|n\rangle = |udd\rangle$, where the up and down quarks form an isospin doublet $\begin{pmatrix} u \\ d \end{pmatrix}$. Later, a third quark flavor ("strange quarks") was discovered, leading to te extension of the $SU(2)$ isospin to $SU(3)$, the famous "eight-fold way" of Gell-Mann etal. We will consider this later.

Another important application of $SU(3)$ in physics is "color $SU(3)$", which is an exact symmetry (as opposed to the above "flavor symmetry", which is only approximate) of

QCD, the theory of strong interactions. Larger Lie group such as $SU(5), SO(10)$, and even exceptional groups such as E_8 (see later) play a central role in modern quantum field theory and string theory.

Lie groups and -algebras are also essential in many other branches of physics.

3.2 The Lorentz group $SO(3,1)$ and its universal cover $SL(2,\mathbb{C})$

This section explains the relativistic concept of spin, more precisely spin $\frac{1}{2}$. The existence of spin $\frac{1}{2}$ objects in physics implies that there should be a representation of the Lorentz group (or a suitable generalization of it) on 2-component objects. It is easy to extend the argument in Sec. 3.1 to show that $SL(2,\mathbb{C})$ is the universal covering group of the Lorentz group $SO(3,1)$. This provides the relativistic concept of spin.

Consider the (Lie) group

$$SL(2,\mathbb{C}) = \{M \in Mat(2,\mathbb{C});\ det(M) = 1\}, \tag{14}$$

and the following (real) vector space

$$X = \{\text{hermitean } 2\times 2 \text{ matrices}\} = \{\begin{pmatrix} x^0+x^3, & x^1-ix^2 \\ x^1+ix^2, & x^0-x^3 \end{pmatrix}, x^\mu \in \mathbb{R}\} \tag{15}$$

Hence any $x \in X$ can be written uniquely as $x = x^\mu \sigma_\mu$, where

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};\quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix};\quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix};\quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Observe that

$$\det(x) = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \tag{16}$$

is just the Minkowski metric on $X \cong M^4$.

Now consider a fixed $U \in SL(2,\mathbb{C})$. Using it, we define a linear map

$$\begin{aligned} \phi_U : X &\to X \\ x &\to \phi_U(x) := UxU^\dagger \end{aligned} \tag{17}$$

check: rhs $\in X$ for any $U \in SL(2,\mathbb{C})$. Moreover, given any $U \in SL(2,\mathbb{C})$ and $x \in X$, we have again

$$\det(\phi_U(x)) = \det(UxU^\dagger) = \det(x)$$

because $\det(U) = 1$. Thus ϕ_U preserves the Minkowski metric on X, and because it is linear it must be an element of the pseudo-orthogonal group $O(3,1) \subset GL(X)$. Hence we have defined a map

$$\begin{aligned} \phi : SL(2,\mathbb{C}) &\to O(3,1) \\ U &\to \phi_U \end{aligned} \tag{18}$$

check: this map is a group homomorphism, and continuous.

Since $\phi(\mathbb{1}) = \mathbb{1}$ and $SL(2,\mathbb{C})$ is connected, it follows that $\phi(SL(2,\mathbb{C}))$ is contained in the component of $SO(3,1)$ connected to the identity, hence $\phi : SL(2,\mathbb{C}) \to L_+^\uparrow$, the proper orthochronous Lorentz group (i.e. preserves sign of time, and $\in SO(3,1)$).

Again, $\phi : SL(2,\mathbb{C}) \to SO(3,1)$ is two-to-one, since for any $U \in SL(2,\mathbb{C})$, $\phi_U = \phi_{-U}$. It is again the "universal covering group" of the Lorentz group.

Due to the map $SL(2,\mathbb{C}) \to SO(3,1)$, every action (representation) of $SO(3,1)$ yields also an action (representation) of $SL(2,\mathbb{C})$, but the converse is not true. Indeed there are objects in relativistic physics which transform under $SL(2,\mathbb{C})$ but not under $SO(3,1)$. The basic such objects are columns or 2-component spinors

$$\psi = \begin{pmatrix} \psi^+ \\ \psi^- \end{pmatrix} \in C^2$$

with the obvious action

$$\begin{aligned} SL(2,\mathbb{C}) \times \mathbb{C}^2 &\to \mathbb{C}^2, \\ (M,\psi) &\to M \cdot \psi \end{aligned}$$

These are spinors (Weyl spinors; a Dirac spinor consists of 2 Weyl spinors), which are the "most basic" non-trivial objects which are consistent with special relativity. They describe e.g. neutrinos[2].

Finally, the Poincare group is defined as a combination of Lorentz transformations with translations. It consists of pairs $(\Lambda, a) \in SO(3,1) \times \mathbb{R}^4$ which act on Minkowski space as

$$x^\mu \to \Lambda^\mu_\nu x^\nu + a^\mu. \tag{19}$$

Accordingly, the group law is given by

$$(\Lambda, a) \cdot (\Lambda', a') = (\Lambda\Lambda', \Lambda a' + a) \tag{20}$$

It plays a fundamental role in quantum field theory, but since it structure is somewhat outside of the main focus of these lectures (i.e. semi-simple Lie algebras), we will not discuss it any further here.

4 Basic definitions and theorems on Lie groups

We now give the general theory of Lie groups. Because they are manifolds, this requires some background in differentiable manifolds.

[2]Strictly speaking the Lorentz group (resp. $SL(2,\mathbb{C})$) should be augmented to the Poincare group in this context, so that the spinors can depend on spacetime.

4.1 Differentiable manifolds

Here is a very short summary of the definitions and main concepts on differentiable manifolds. This is not entirely precise. The proofs can be found e.g. in [Warner].

Definition 6 *A topological space M is a m-**dimensional differentiable manifold** if it comes with a family $\{(U_i, \varphi_i)\}$ of coordinate systems ("charts") such that*

1. *$U_i \subset M$ open, $\cup_i U_i = M$ and*

   $$\varphi_i : U_i \to V_i \subset \mathbb{R}^m \qquad \text{is a homeomorphism (=continuous and invertible)}$$

2. *$\varphi_i \circ \varphi_j^{-1}$ is smooth ($\mathcal{C}^\infty$) where defined.*

(Picture)

Notation:

$$\varphi(p) = \begin{pmatrix} x^1(p) \\ \vdots \\ x^m(p) \end{pmatrix}, \qquad p \in M$$

Definition: *smooth maps* are

$$\begin{aligned} \mathcal{C}^\infty(M) &= \{f : M \to \mathbb{R}, \quad \mathcal{C}^\infty\} \\ \mathcal{C}^\infty(M,N) &= \{\Phi : M \to N, \quad \mathcal{C}^\infty\} \end{aligned}$$

the latter means that $\varphi_N \circ f \circ \varphi_M^{-1}$ is smooth for all coordinate systems if defined.

A smooth invertible map between manifolds $\Phi : M \to N$ is called a *diffeomorphism.*

Tangential space: let $p \in M$. The tangential space of M at p is defined as the space of all derivations (="directional derivatives") of functions at p, i.e.

$$T_p(M) = \{X : \mathcal{C}^\infty(M) \to \mathbb{R} \text{ derivation}\}$$

which means that

$$\begin{aligned} X[\lambda f + \mu g] &= \lambda X[f] + \mu X[g], \qquad\qquad f, g \in \mathcal{C}^\infty(M), \quad \lambda, \mu \in \mathbb{R} \\ X[fg] &= f(p)X[g] + g(p)X[f] \end{aligned}$$

In particular, this implies

$$X[c] = 0$$

for any constant function c.

example:

let $\varphi = \begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix}$ be a coordinate system containing the point p. Then

$$X_i := \frac{\partial}{\partial x^i}_{|p} : \quad f \to \mathbb{R} \qquad \text{partial derivative at } p$$

i.e.

$$X_i[f] = \frac{\partial}{\partial x^i}_{|p}[f] = \frac{\partial (f \circ \varphi^{-1})}{\partial x^i}_{|p}$$

Theorem 7

$$T_p(M) = \langle \frac{\partial}{\partial x^i} \rangle_{\mathbb{R}} \cong \mathbb{R}^m \qquad \textit{is a m-dimensional vector space}$$

i.e. a general tangent vector at p *has the form*

$$X_p = \sum a_i \frac{\partial}{\partial x^i}_{|p}, \qquad a_i \in \mathbb{R}$$

The point is that these are *first-order differential operators*, not higher-order ones. This is encoded in the coordinate-independent concept of a derivation. The theorem can be proved easily using a Taylor expansion of f near p.

A *vector field* $X \in T(M)$ is an assignment of a tangent vector for every $p \in M$. It has the form

$$X = \sum a_i(x) \frac{\partial}{\partial x^i}$$

(in some local coordinate system), and is parametrized by m "component functions" $a_i : M \to \mathbb{R}$. They depend of course on the coordinate system, and transform as follows under change of coordinates:

$$\frac{\partial}{\partial x^i} = \left(\frac{\partial y^j}{\partial x^i} \right) \frac{\partial}{\partial y^j}.$$

(chain rule).

The differential of a map or "tangential map": let

$$\Phi : M \to N$$

be a smooth map. Then one defines the "push-forward" map

$$\begin{aligned} d\Phi : T_p(M) &\rightarrow T_{\Phi(p)}(N), \\ X &\rightarrow (d\Phi(X))[f] := X[f \circ \Phi] \end{aligned}$$

where $f : N \rightarrow \mathbb{R}$. Sometimes this is also written as $d\Phi = T\Phi = \Phi_*$. Notice that this is indeed a derivation (Exercise)!

For example, consider a (smooth) curve in $M \subset \mathbb{R}^n$,

$$\gamma : \mathbb{R} \rightarrow M, \qquad \gamma(0) = p \in M.$$

Denote with $V_0 = \frac{d}{dt}$ the unit tangential vector at $0 \in \mathbb{R}$, which means that $V_0[g] = \frac{d}{dt}[g]$. Then the tangential vector along γ at p is obtained by

$$X_p = d\gamma(V_0)$$

i.e. for $f \in \mathcal{C}(\mathbb{R}^n)$ we have

$$X_p[f] = d\gamma(V_0)[f] = V_0[f \circ \gamma] = \frac{d}{dt}(f \circ \gamma) = \frac{\partial f}{\partial x^i}\frac{d\gamma^i}{dt} = \frac{d\vec{\gamma}}{dt} \cdot \vec{\nabla}\ [f]$$

which is indeed the "directional derivative" along γ. Hence

$$X_p = d\gamma(\frac{d}{dt}) = \frac{d\gamma^i}{dt}\frac{\partial}{\partial x^i},$$

and the components are just the components of $\frac{d\vec{\gamma}}{dt}$

<u>Examples:</u>

1. If (x^i) are coordinates on M and (y^i) coordinates on N, then the chain rule on $\mathbb{R}^n$ gives
$$d\Phi(\frac{\partial}{\partial x^i}|_p) = \sum \frac{\partial(y^j \circ \Phi)}{\partial x^i}\frac{\partial}{\partial y^j}|_{\Phi(p)} \tag{21}$$
2. Consider a map $\varphi : M \rightarrow \mathbb{R}_t$. Specializing the above to $\Phi = \varphi$, we obtain
$$d\varphi(\frac{\partial}{\partial x^k}|_p) = \sum \frac{\partial \varphi}{\partial x^k}\frac{d}{dt}$$
3. Consider some individual coordinate function $x^i : M \rightarrow \mathbb{R}_t$ of a chart containing p. Specializing the above to $\varphi = x^i$, we obtain
$$dx^i(\frac{\partial}{\partial x^k}|_p) = \sum \frac{\partial x^i}{\partial x^k}\frac{d}{dt} = \delta^i_k\,\frac{d}{dt}\ .$$
This means that one can identify dx^i as the dual of $\frac{\partial}{\partial x^k}$.

4. if $\varphi : M \to \mathbb{R}$, then we obtain by comparing the last two relations we obtain

$$d\varphi = \frac{\partial \varphi}{\partial x^i} dx^i$$

The tangential map satisfies the *chain rule*: if $\Phi : M \to N$ and $\Psi : N \to P$, then

Theorem 8

$$d(\Psi \circ \Phi) = d\Psi \circ d\Phi$$

more precisely

$$d(\Psi \circ \Phi)_m = d\Psi_{\Phi(m)} \circ d\Phi_m$$

Note that the proof is trivial in this framework, and reduces to the usual chain rule in coordinates (Exercise!).

Lie brackets of vector fields: Let X, Y vector fields on M. Then one can define a new vector field $[X, Y]$ on M via

$$[X, Y]_p(f) = X_p[Y[f]] - Y_p[X[f]].$$

One then easily shows

Theorem 9

- $[X, Y]$ *is indeed a vector field on* M *(derivation!)*
- $[X, X] = 0$*, hence* $[X, Y] = -[Y, X]$
- $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0.$
- $[fX, gY] = fg[X, Y] + fX[g]Y - gY[f]X.$
- *Given a map* $\phi : M \to N$*, let* $\tilde{X}, \tilde{Y}$ *be vector fields on* N *such that* $d\phi(X) = \tilde{X}$ *and* $d\phi(Y) = \tilde{Y}$ *(i.e.* X *and* $\tilde{X}$ *are "*ϕ*-related", etc.). Then* $[X, Y]$ *is* ϕ*-related to* $[\tilde{X}, \tilde{Y}]$*, i.e.*
  $d\Phi([X, Y]) = [\tilde{X}, \tilde{Y}] = [d\Phi(X), d\Phi(Y)]$

In particular, the space of all vector fields is a (infinite-dimensional) Lie algebra*! (see later...)*

Proof: easy verification.

In a coordinate system, we can write the vector fields as

$$X = X^i(x)\frac{\partial}{\partial x^i}$$

and similar Y. Then

$$\begin{aligned}
[X,Y] &= X^i(x)\frac{\partial}{\partial x^i}\,Y^j(x)\frac{\partial}{\partial x^j} - Y^j(x)\frac{\partial}{\partial x^j}\,X^i(x)\frac{\partial}{\partial x^i} \\
&= X^i(x)\frac{\partial Y^j(x)}{\partial x^i}\,\frac{\partial}{\partial x^j} - Y^j(x)\frac{\partial X^i(x)}{\partial x^j}\,\frac{\partial}{\partial x^i}
\end{aligned} \tag{22}$$

because the partial derivatives commute. This shows explicitly that the rhs is indeed again a vector field (as opposed to e.g. XY!!)

4.2 Lie groups

A *Lie group* G is a group which is also a differentiable manifold, such that the maps

$$\begin{aligned}
\mu : G\times G &\rightarrow G \\
(g_1, g_2) &\mapsto g_1\cdot g_2
\end{aligned}$$

and

$$\begin{aligned}
\nu : G &\rightarrow G \\
g &\mapsto g^{-1}
\end{aligned}$$

are smooth.

A *Lie subgroup* H of G is a (topological) subgroup which is also a (smooth) submanifold.

The *left translations* on G are the diffeomorphisms of G labeled by the elements $g \in G$ and defined by

$$\begin{aligned}
L_g : G &\rightarrow G \\
g' &\mapsto g\cdot g'
\end{aligned}$$

(similarly the right translations). They satisfy

$$L_g L_{g'} = L_{gg'}.$$

A *homomorphism* between Lie groups is a smooth map $\phi : G \rightarrow H$ which is a group homomorphism.

If $H = GL(V)$ for some vector space V, this is called a *representation* of G. One considers in particular the following types of representations:

$$\pi : G \to GL(n,\mathbb{R}) \qquad n - \text{dimensional “real” representation}$$
$$\pi : G \to GL(n,\mathbb{C}) \qquad n - \text{dimensional “complex” representation}$$
$$\pi : G \to U(n) \qquad n - \text{dimensional “unitary” representation}$$

An important problem both in physics and math is to find (ideally all) representations of G. This can be solved completely for a large class of Lie groups, and will be explained later in this course.

Examples for Lie groups:

- $(\mathbb{R}^n, +)$. The left-translation is $L_x(y) = x + y$, i.e. indeed translations of y by x.
- $\mathbb{C}^* = (\mathbb{C} - \{0\}, \cdot)$
- the complex numbers with unit modulus $U(1) = S^1$ and multiplication
- matrix groups:
  $$GL(n,\mathbb{R}) := \{A \in Mat(n,\mathbb{R});\ det(A) \neq 0\}$$
  similarly $GL(n,\mathbb{C})$, and
  $$\begin{aligned}
  SL(n,\mathbb{R}) &:= \{A \in Mat(n,\mathbb{R});\ \det(A) = 1\!\!1\ \},\\
  O(n) &:= \{A \in Mat(n,\mathbb{R});\ AA^T = 1\!\!1\ \},\\
  SO(n) &:= \{A \in Mat(n,\mathbb{R});\ AA^T = 1\!\!1\ , \det(A) = 1\},\\
  U(n) &:= \{A \in Mat(n,\mathbb{C});\ AA^\dagger = 1\!\!1\ \},\\
  SU(n) &:= \{A \in Mat(n,\mathbb{C});\ AA^\dagger = 1\!\!1\ , \det(A) = 1\},\\
  SP(n,\mathbb{R}) &:= \{A \in Mat(2n,\mathbb{R});\ A^T J A = J, \quad J = \begin{pmatrix} 0 & 1\!\!1_n \\ -1\!\!1_n & 0 \end{pmatrix}\}
  \end{aligned}$$
  (symplectic group)

  the Lorentz group
  $$O(3,1) = \{A \in Mat(n,\mathbb{R});\ A\eta A^T = \eta\}, \quad \eta = (1,-1,-1,-1),$$
  etc.
- the Poincare group (=Lorentz plus translations)

there are many more! exceptional groups, ...

4.3 Lie algebras

A *Lie algebra* $\mathfrak{g}$ over $\mathbb{R}$ (resp. $\mathbb{C}$ etc... any field) is a vector space over $\mathbb{R}$ resp. $\mathbb{C}$ and an operation (a *Lie bracket*)

$$[.,.] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

which is bilinear over $\mathbb{R}$ (resp. $\mathbb{C}$) and satisfies

$$[X, X] = 0$$

and the Jacobi identity

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

The first property implies

$$[X, Y] = -[Y, X] \qquad \text{“antisymmetry”}$$

Note: for any associative algebra $\mathcal{A}$, there is an associated Lie algebra $\mathfrak{g}$, which is $\mathcal{A}$ as a vector space and

$$[X, Y] := X \cdot Y - Y \cdot X \qquad \text{“commutator”}$$

The Jacobi identity is then trivial.

Examples: let

$$gl(n, \mathbb{R}) := Mat(n, \mathbb{R}) = Mat(n \times n, \mathbb{R})$$

with $[x, y] = xy - yx$.

The following Lie algebras are particularly important:

$$\begin{aligned} sl(n, \mathbb{R}) &:= \{A \in gl(n, \mathbb{R});\ \mathrm{Tr}(A) = 0\}, \\ so(n) &:= \{A \in gl(n, \mathbb{R});\ A^T = -A, \mathrm{Tr}(A) = 0\}, \\ u(n) &:= \{A \in gl(n, \mathbb{C});\ A^\dagger = -A\}, \\ su(n) &:= \{A \in gl(n, \mathbb{C});\ A^\dagger = -A, \mathrm{Tr}(A) = 0\}, \\ sp(n) &:= \{A \in gl(2n, \mathbb{R});\ A^T = JAJ\}, \end{aligned}$$

where the Lie algebra is again defined by the commutator.

Also, the space of vector fields on a manifold $\mathcal{M}$ together withe the Lie bracket forms an infinitesimal Lie algebra.

Further definitions:

A *subalgebra of a Lie algebra* is a subspace $\mathfrak{h} \subset \mathfrak{g}$ such that $[H_1, H_2] \in \mathfrak{h}$ whenever $H_1, H_2 \in \mathfrak{h}$. It is easy to check that the above Lie algebras are indeed Lie subalgebras of $gl(n, \mathbb{R})$.

A *Lie algebra homomorphism* is a linear map $\varphi : \mathfrak{g} \to \mathfrak{h}$ such that

$$\varphi([X, Y]) = [\varphi(X), \varphi(Y)] \qquad \forall X, Y \in \mathfrak{g}.$$

One can show that essentially all (finite-dimensional) Lie algebras are subalgebras of $gl(n)$ (Varadarajan, see e.g. [Hall]).

A *representation* of a Lie algebra $\mathfrak{g}$ is a Lie-algebra homomorphism

$$\pi : \mathfrak{g} \to gl(n, \mathbb{R}) \qquad \text{"real" representation}$$

$$\pi : \mathfrak{g} \to gl(n, \mathbb{C}) \qquad \text{"complex" representation}$$

Structure constants. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra, and let $X_1, \cdots, X_n$ be a basis for $\mathfrak{g}$ (as a vector space). Then for each i, j, $[X_i, X_j]$ can be written uniquely in the form

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

The constants c_{ij}^k are called the **structure constants** of $\mathfrak{g}$ (with respect to the chosen basis). Clearly, the structure constants determine the bracket operation on $\mathfrak{g}$. (Often in physics one uses $i\mathfrak{g}$ in order to have hermitian generators, which leads to $[X_i, X_j] = i \sum_k c_{ij}^k X_k$.)

The structure constants satisfy the following two conditions,

$$c_{ij}^k + c_{ji}^k = 0 \qquad \text{(antisymmetry)}$$

$$\sum_m (c_{ij}^m c_{mk}^l + c_{jk}^m c_{mi}^l + c_{ki}^m c_{mj}^l) = 0 \qquad \text{(Jacobi identity)}$$

4.4 The Lie algebra of a Lie group

Let G be a Lie group. Recall left translations on G, defined by $L_g : G \to G,\ g' \mapsto g \cdot g'$. Define

$$\mathfrak{g} := \{\text{left-invariant vector fields } X \text{ on G}\}$$

i.e.

$$dL_g(X) = X \quad \forall g \in G$$

or more precisely, $dL_g(X_{g'}) = X_{gg'}$.

Example: consider $G = (\mathbb{R}^n, +)$

We have $L_{\vec{a}}(\vec{x}) = \vec{x} + \vec{a}$. Then (Exercise) $dL_a(f^i(x)\frac{\partial}{\partial x^i}|_x) = f^i(x)\frac{\partial(L_a x)^j}{\partial x^i}\frac{\partial}{\partial x^j}|_{x+a} = f^i(x)\frac{\partial}{\partial x^i}|_{x+a}$, hence $dL_{\vec{a}}(X) = X$ for all $\vec{a} \in \mathbb{R}^n$ implies $f^i(\vec{x}) = f^i(\vec{x}+\vec{a}) \quad \forall \vec{a}$, hence $f^i = const$ and

$$\mathfrak{g} = \{f^i \frac{\partial}{\partial x^i}\} \cong \mathbb{R}^n.$$

Observe:

- $\mathfrak{g} \cong \mathbb{R}^n \cong T_e G$
- in general: given X_e, define $X_g := dL_g(X_e)$

  can show (easy): is left-invariant V.F.

  (Proof: $dL_{g'}(X_g) = dL_{g'}(dL_g(X_e)) = d(L_{g'}L_g)(X_e) = dL_{g'g}(X_e) = X_{g'g}$)
- $\mathfrak{g}$ is a *Lie algebra*: for $X, Y \in \mathfrak{g}$; define $[X, Y]$ ... Lie-bracket of left-invariant V.F.

  Lemma: $[X, Y]$ is again left-invariant V.F., because

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

  by theorem 9.

The relation between Lie groups and their Lie algebras is contained in the following central theorem:

Theorem 10 *Let G and H be Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively. Then:*

1. *If $\phi : G \to H$ is a homomorphism of Lie groups, then $d\phi : \mathfrak{g} \to \mathfrak{h}$ is a homomorphism of Lie algebras.*
2. *If $\varphi : \mathfrak{g} \to \mathfrak{h}$ is a homomorphism of Lie algebras and G is simply connected (and connected), then there exists a unique homomorphism of Lie groups $\phi : G \to H$ such that $\varphi = d\phi$.*
3. *If $\mathfrak{h} \subset \mathfrak{g}$ is a Lie subalgebra, then there exists a (connected) Lie subgroup $H \subset G$ such that $\mathfrak{h}$ is the Lie algebra of H.*

Proof:

1. let $X, Y \in \mathfrak{g}$. Define $\tilde{X}, \tilde{Y}$ to be those left-invariant vector fields on H such that $\tilde{X}_e = d\phi(X_e)$ and $\tilde{Y}_e = d\phi(Y_e)$.

   We observe that $\tilde{X}$ is related to X through $d\phi$, i.e. $\tilde{X} = d\phi(X)$ and $\tilde{Y} = d\phi(Y)$ on $\phi(G) \subset H$ (not only at the unit element). To see this, let $h = \phi(g)$, then $L_h \circ \phi = \phi \circ L_g$ since ϕ is a group homomorphism. Therefore

   $$dL_h(d\phi(X)) = d(L_h \circ \phi)(X) = d(\phi \circ L_g)(X) = d\phi dL_g(X) = d\phi(X),$$

   using the chain rule. Therefore $d\phi(X)$ must agree with $\tilde{X}$ on $\phi(G) \subset H$. Using Theorem 9, we now get

   $$[d\phi(X), d\phi(Y)] = [\tilde{X}, \tilde{Y}] = d\phi[X, Y]$$

   which means that $d\phi : \mathfrak{g} \to \mathfrak{h}$ is a homomorphism of Lie algebras.

2. very nontrivial, see e.g. [Warner].

3. also nontrivial! (existence of a smooth submanifold etc.), see e.g. [Warner].

qed

For example, consider the homomorphism $\phi : SU(2) \to SO(3)$. Because this is invertible near e, theorem 10 implies that their Lie algebras are isomorphic, $su(2) \cong so(3)$. Moreover $SU(2)$ is simply connected, hence the statement 2) applies: As soon as we know that $su(2) \cong so(3)$ (by simply checking it, see later!) it follows that there exists a group homomorphism ϕ as above. This is obviously a strong statement!

This example generalizes as follows: One can show that for every Lie group G, there exists a so-called "universal covering (Lie) group" $\tilde{G}$, which means that $\tilde{G}$ is a simply connected Lie group and that there exists a surjective group-homomorphism

$$\phi : \tilde{G} \to G$$

which is locally an isomorphism (i.e. in a neighborhood of the identity), but not globally. Globally, the map Φ is such that the inverse image of each $g \in G$ consists of k points in $\tilde{G}$ for some integer k (more precisely, the inverse image of a small $U \subset G$ consists of k homeomorphic copies of U). In particular, the Lie algebras of $\tilde{G}$ and G coincide by the above theorem, $\tilde{\mathfrak{g}} = \mathfrak{g}$, and $dim(G) = dim(\tilde{G})$. For example, $SU(2)$ is the universal cover of $SO(3)$.

This implies that whenever we have a homomorphism of Lie algebras $\varphi : \mathfrak{g} \to \mathfrak{h}$, there exists a homomorphism of Lie groups $\phi : \tilde{G} \to H$. This is the reason why

1. it is “better” (i.e. more general) to use $SU(2)$ rather than $SO(3)$
2. it is essentially enough to consider representations of Lie algebras, which is a “linear” problem and can be handled. The theorem then guarantees the existence of the representation of the Lie group $\tilde{G}$, and one can then decide if this also gives a rep. of G.
3. there is a one-to-one correspondence between representations of a (simply connected) Lie group $\tilde{G}$ and its Lie algebra $\mathfrak{g}$. The latter is much easier to handle. Later.

This extremely important result really depends on the full machinery of Lie groups, hence this lengthy preparation. But from now on, we will get more down to earth.

The most important examples of Lie groups (but not all!) are matrix groups, i.e. subgroups of $GL(N, \mathbb{R})$ (or $GL(N, \mathbb{C})$). In this case, the above general concepts become more transparent.

4.4.1 The Lie algebra of $GL(n, \mathbb{R})$

Recall that $GL(n, \mathbb{R})$ is an open subset of $\mathbb{R}^{n^2}$. A natural coordinate system on $GL(n, \mathbb{R})$ near the unit element $e = \mathbb{1}$ is given by the “Cartesian matrix coordinates”,

$$x^{ij}(g) := g^{ij} \qquad (i.e.\ \ x : GL(n) \hookrightarrow \mathbb{R}^{n^2}!)$$

where $g = \left(g^{ij}\right)$. A basis of tangent vectors $T_e(GL(n))$ is then given by the partial derivatives $\frac{\partial}{\partial x^{ij}}|_e$, i.e. a general tangent vector at e has the form

$$X_e^A = A_{ij} \frac{\partial}{\partial x^{ij}}|_e, \qquad A_{ij} \in \mathbb{R}$$

(sum convention). Hence $T_e(GL(n, \mathbb{R})) = Mat(n, \mathbb{R}) = gl(n, \mathbb{R})$ as vector space.

Denote with $gl(n) = Mat(n)$ the space of $n \times n$ matrices. We want to show that

$$Lie(GL(n)) = gl(n)$$

as *Lie algebras* (with the commutator for $gl(n)$), not just as vector spaces; the latter is evident.

Let us calculate the corresponding left-invariant vector field $X_g^A = dL_g(X_e^A)$. We can use the same coordinates near e and g, so that the map L_g has the “coordinate expression”

$$(L_g x)^{ij} = (gx)^{ij} = g^{ik} x^{kj}.$$

Then using (21), we have

$$dL_g(\frac{\partial}{\partial x^{ij}}|_e) = \frac{\partial (L_g x)^{kl}}{\partial x^{ij}}|_e \frac{\partial}{\partial x^{kl}}|_g = \frac{\partial (g^{km} x^{ml})}{\partial x^{ij}}|_e \frac{\partial}{\partial x^{kl}}|_g = \delta^{lj} g^{ki} \frac{\partial}{\partial x^{kl}}|_g = g^{ki} \frac{\partial}{\partial x^{kj}}|_g$$

Therefore for general $X_e = A_{ij} \frac{\partial}{\partial x^{ij}}|_e$, we have

$$X_g^A = dL_g(X_e^A) = g^{ki} A_{ij} \frac{\partial}{\partial g^{kj}}$$

where we write $g^{ij} = x^{ij}(g)$ for the (Cartesian) coordinate functions on $GL(n)$.

Now we can calculate the commutator: Let X^A, X^B be left-invariant vector fields as above. Noting that $\frac{\partial}{\partial g^{ij}}|_g g^{kl} = \delta^{ik}\delta^{jl}$, and using (22) we have

$$\begin{aligned}
[X^A, X^B] &= g^{ki} A_{ij} \frac{\partial}{\partial g^{kj}} g^{k'i'} B_{i'j'} \frac{\partial}{\partial g^{k'j'}} - g^{ki} B_{ij} \frac{\partial}{\partial g^{kj}} g^{k'i'} A_{i'j'} \frac{\partial}{\partial g^{k'j'}} \\
&= g^{ki} A_{ij}\, B_{jj'} \frac{\partial}{\partial g^{kj'}} - g^{ki} B_{ij}\, A_{jj'} \frac{\partial}{\partial g^{kj'}} \\
&= g^{ki}(A_{ij}\, B_{jj'} - B_{ij}\, A_{jj'}) \frac{\partial}{\partial g^{kj'}} = g^{ki}[A,B]_{ij'} \frac{\partial}{\partial g^{kj'}} = X^{[A,B]} \qquad (23)
\end{aligned}$$

But this is precisely the left-invariant vector-field associated to the commutator of the *matrices* A, B. Therefore we can identify

$$\boxed{gl(n,\mathbb{R}) \equiv Lie(GL(n,\mathbb{R})) \cong \{Mat(n,\mathbb{R});\ [A,B] = AB - BA\}}$$

which we considered before. Similarly one obtains

$$gl(n,\mathbb{C}) \equiv Lie(GL(n,\mathbb{C})) \cong \{Mat(n,\mathbb{C});\ [A,B] = AB - BA\}$$

4.4.2 Subgroups of $GL(n)$

Now one can obtain the Lie algebras corresponding to the other matrix Lie groups such as $SO(n)$ considered before: because they are subgroups of $GL(n)$, there is a (trivial) Lie group homomorphism

$$\phi : SO(n) \to GL(n)$$

etc., which by differentiating induces a Lie algebra homomorphism

$$d\phi : so(n) \to gl(n)$$

which is in fact injective. This means that we can consider e.g. $so(n)$ as a subalgebra of $gl(n,\mathbb{R})$, i.e.

$$so(n) \subset gl(n,\mathbb{R})\,.$$

In particular, we can just as well work with $gl(n, \mathbb{R})$, where the Lie algebra is given by the commutator $[A, B]$ for elements of $gl(n)$. We will show below that this gives precisely the matrix Lie algebras defined in section 4.3.

A similar observation applies to representations: any representation defines a group homomorphism $\pi : G \to GL(V)$, which means that $d\pi([X, Y]) = [d\pi(X), d\pi(Y)]$ is the commutator of the matrices $d\pi(X)$ and $d\pi(Y)$. This means again that we can work with $gl(V)$ and use the explicit matrix commutators.

It is quite easy to work with matrix groups. In particular, the exponential mapping (which exists for any Lie group, see later) is very transparent and useful here:

5 Matrix Lie groups and the exponential map

5.1 The Matrix Exponential

The exponential map plays a crucial role in the theory of Lie groups. This is the tool for passing explicitly from the Lie algebra to the Lie group.

Let X be a $n \times n$ real or complex matrix. We wish to define the exponential of X, e^X or $\exp X$, by the usual power series

$$e^X = \sum_{m=0}^{\infty} \frac{X^m}{m!}. \tag{24}$$

It is easy to show that for any $n \times n$ real or complex matrix X, this series (24) converges, and that the matrix exponential e^X is a smooth function of X.

Proposition 11 *Let X, Y be arbitrary $n \times n$ matrices. Then*

1. $e^0 = I$.
2. e^X *is invertible, and* $\left(e^X\right)^{-1} = e^{-X}$. *In particular,* $e^X \in GL(n)$.
3. $e^{(\alpha+\beta)X} = e^{\alpha X} e^{\beta X}$ *for all real or complex numbers* α, β.
4. *If* $[X, Y] = 0$, *then* $e^{X+Y} = e^X e^Y = e^Y e^X$.
5. *If* C *is invertible, then* $e^{CXC^{-1}} = Ce^X C^{-1}$.

The proof is elementary (Analysis lecture).

Note that in general $e^{X+Y} \neq e^X e^Y$, although equality holds by 4) if X and Y commute. This is a crucial point which one should never forget. There exists a formula – the Baker-Campbell-Hausdorff formula (52) – which allows to calculate products of the form (4) in terms of X and Y.

For example, consider

$$J_x := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}$$

which is a "rotation generator", i.e. $iJ_x \in so(3)$. We claimed previously that

$$R(\phi\vec{e_x}) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & \sin(\phi) \\ 0 & -\sin(\phi) & \cos(\phi) \end{pmatrix} = e^{i\phi\vec{e_x}\cdot\vec{J}} = e^{i\phi J_x}$$

Lets see if this is true: using $J_x^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, we get

$$\begin{aligned} e^{i\phi J_x} &= \sum_{m=0}^{\infty} \frac{(i\phi J_x)^m}{m!} = 1 + J_x^2 \sum_{n=1}^{\infty} \frac{(i\phi)^{2n}}{(2n)!} + J_x \sum_{n=0}^{\infty} \frac{(i\phi)^{2n+1}}{(2n+1)!} \\ &= \mathbb{1} + J_x^2(\cos(\phi) - 1) + iJ_x \sin(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & \sin(\phi) \\ 0 & -\sin(\phi) & \cos(\phi) \end{pmatrix} \end{aligned} \tag{25}$$

as desired.

Remark: One good way to calculate the exponential of a Matrix is to diagonalize it if possible: if $X = UDU^{-1}$, then $e^X = Ue^D U^{-1} = U\text{diag}(e^{d_i})U^{-1}$ by (5). Otherwise, one can bring X to Jordan normal form.

Further important formulas for the matrix exponential are as follows:

Theorem 12 *Let X be an $n \times n$ real or complex matrix. Then*

$$\det\left(e^X\right) = e^{\text{trace}(X)}.$$

<u>Proof:</u> *Case 1: X is diagonalizable.* Suppose there is a complex invertible matrix C such that

$$X = C\text{diag}(x_i)C^{-1}.$$

Then

$$e^X = C\text{diag}(e^{x_i})C^{-1}.$$

Thus $\text{trace}(X) = \sum \lambda_i$, and $\det(e^X) = \prod e^{\lambda_i} = e^{\sum \lambda_i}$. (Recall that $\text{trace}(CDC^{-1}) = \text{trace}(D)$.)

Case 2: X arbitrary. Not difficult, use e.g. Jordan normal form, or analytic continuation. q.e.d

Also, check it in the explicit example above.

5.2 One-parameter subgroups and the exponential map

Fix an axis $\vec{v}$ with $\|\vec{v}\| = 1$. Lets consider again rotations: we wrote finite rotations in the form

$$R(\phi) := R(\phi\vec{v}) := e^{i\phi\vec{v}\cdot\vec{J}} \tag{26}$$

We claim that these are rotations around the axis $\vec{v}$ and angle ϕ. How do we know this?

The rotations around $\vec{v}$ are clearly an abelian (1-parameter) subgroup of $SO(3)$, labeled by the angle $\phi \in \mathbb{R}$. This means that $R(\phi)$ is a group homomorphism from $\mathbb{R}$ to G,

$$R(\phi + \psi) = R(\phi)R(\psi).$$

Clearly the axis is fixed once we know that this is true for "infinitesimal" ϕ, and this must define a rotation around the angle ϕ. "Infinitesimal rotations" are given by $1 + i\phi\vec{v}\cdot\vec{J} + O(\phi^2)$. Note that

$$\frac{d}{d\phi}\bigg|\, R(\phi)_{\phi=0} = i\vec{v}\cdot\vec{J}$$

is a tangential vector to $SO(3)$ at the origin, i.e. it is an element of the Lie algebra $so(3)$.

Lets see what this means explicitly: Using

$$iJ_x e_y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} = -e_z$$

etc, we see that

$$iJ_j e_k = -\varepsilon_{jkl} e_l.$$

Hence "infinitesimal rotations" are given by

$$(1 + i\phi\vec{v}\cdot\vec{J})\vec{x} = 1 - \phi\vec{v}\times\vec{x}.$$

This really is an 'infinitesimal rotation" for "infinitesimal" ϕ, hence the claim is justified. Note that an infinitesimal rotation has the form $(\mathbb{1} + \varepsilon J)$ for $J \in so(3)$.

More generally, consider a $n \times n$ matrix $X \in gl(n)$. Recall that the Lie algebra $gl(n) = Lie(GL(n))$ is just the set of tangent vectors at $\mathbb{1}$. Hence to every $X \in gl(n)$ we can associate a curve $\gamma(t) = e^{tX}$, which satisfies

$$\gamma(t+s) = \gamma(t)\gamma(s) \qquad \forall t, s \in \mathbb{R}$$

by the properties of exp. This means that we have a Lie group homomorphism

$$\gamma : \mathbb{R} \to GL(n)$$

which satisfies

$$\left.\frac{d}{dt}\right|_{t=0} \gamma(t) = X \tag{27}$$

More generally for any Lie group G, each $X \in Lie(G)$ determines uniquely a 1-parameter subgroup in G. This is defined as follows: any given $X_e \in \mathfrak{g} = Lie(G)$ defines a (trivial) Lie algebra homomorphism or $\mathbb{R}$ to $\mathfrak{g}$. Then by by theorem 10, there exists a *unique* Lie group homomorphism $\gamma : \mathbb{R} \to G$ such that $d\gamma[\frac{d}{dt}]_{t=0} = X_e \in \mathfrak{g}$. Such a Lie group homomorphism $\gamma : \mathbb{R} \to G$ is called a *one-parameter subgroup* of G. (One can show that $\gamma(t)$ is the integral curve of the left-invariant vector field determined by X, which is e^{tX} for $GL(n)$). This leads to the general definition of the exponential map, which works for any Lie group:

Definition 13 *Let G be a Lie group, and $\mathfrak{g}$ its Lie algebra. Let $X \in \mathfrak{g}$. Let*

$$\exp_X : \mathbb{R} \to G$$

be the unique (by theorem 10) homomorphism of Lie groups such that

$$d\exp_X(\frac{d}{dt}) = X.$$

Then define

$$\exp : \mathfrak{g} \to G,$$

by setting

$$\exp(X) = \exp_X(1)$$

In the case $G = GL(n)$, this reduces to the matrix exponential as we've seen above. One can now show that all statements on proposition 11 remain true, and we will use the notation $\exp(X) = e^X$ interchangably. (The last property of proposition 11 leads to the adjoint representation.)

exp defines a diffeomorphism of a neighborhood of $0 \in \mathfrak{g}$ onto a neighborhood of $e \in G$ (picture). In the $GL(n)$ case, this can be seen easily since the local inverse is given by the matrix logarithm:

Theorem 14 *The function*

$$\log A = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A-I)^m}{m} \tag{28}$$

is well-defined and smooth on the set of all $n \times n$ complex matrices A with $\|A - I\| < 1$, and $\log A$ is real if A is real.

For all A with $\|A - I\| < 1$,

$$e^{\log A} = A.$$

For all X with $\|X\| < \log 2$, $\left\|e^X - 1\right\| < 1$ we have

$$\log e^X = X.$$

Moreover, one can show that

$$\exp : \mathfrak{g} \to G$$

is surjective for compact G. However, it is usually not injective. Furthermore, it is easy to show that if $\phi : G \to H$ is a Lie group homomorphism, then the diagram

$$\begin{array}{ccc} G & \xrightarrow{\phi} & H \\ {\scriptstyle \exp}\uparrow & & \uparrow{\scriptstyle \exp} \\ \mathfrak{g} & \xrightarrow{d\phi} & \mathfrak{h} \end{array} \tag{29}$$

commutes (using the uniqueness of the one-parameter subgroups).

This explains the Physicist's notion of "infinitesimal group elements": near the unit element, any group element can be written as $g = e^X$, and "infinitesimal" group elements are those of the form

$$e^{\epsilon X} = 1 + \epsilon X + o(\epsilon^2) \approx 1 + \epsilon X$$

for $X \in Lie(G)$ and "infinitesimal" ϵ.

The elements $X \in Lie(G)$ are called "generators" in physics.

For many arguments it is enough to consider these "infinitesimal group elements", which essentially amounts to working with the Lie algebra.

As a nice application of (29), we can obtain the following useful identity

$$\det(e^A) = e^{trA}$$

This follows from the following diagram

$$\begin{array}{ccc} Gl(V) & \xrightarrow{\det} & \mathbb{R}_+ \\ {\scriptstyle \exp}\uparrow & & \uparrow{\scriptstyle e} \\ gl(V) & \xrightarrow{tr} & \mathbb{R} \end{array} \tag{30}$$

noting that $d(\det)|_{\mathbb{1}} = tr$, which is easy to check.

5.2.1 The classical subgroups of $GL(n)$ and their Lie algebras.

One can use exp to calculate explicitly the most important Lie subgroups of $GL(n)$ and their Lie algebras. Recall the definitions of section 4.2, 4.3. Start with

$SO(n)$ and $so(n)$:

Recall that

$$so(n) := \{A \in Mat(n, \mathbb{R});\ A^T = -A\} \subset gl(n).$$

(this coincides with $o(n)$!)

Let $A \in o(n)$. Then $(e^A)^T = e^{-A} = (e^A)^{-1}$, which means that $e^A \in O(n)$ is an orthogonal matrix. Conversely, consider $g \in O(n)$ near $\mathbb{1}$, so that $g = e^A$ for some $A \in gl(n)$ (by theorem 14). Then $e^{A^T} = g^T = g^{-1} = e^{-A}$. Because exp is a local diffeomorphism (resp. by taking the matrix log), this implies that

$$A^T = -A.$$

This means that

$$\exp(so(n)) = SO(n) \subset GL(n),$$

therefore $so(n)$ is the Lie algebra of $SO(n)$. (recall theorem 10 which states that there exists a Lie subgroup of $GL(n)$ whose Lie algebra is $so(n)$, and the commutative diagram (29) which states that exp for $so(n)$ is really obtained by restriction of $gl(n)$ to $so(n)$).

The explicit form of the Lie algebra (the commutation relations) depends on the choice of basis. One useful basis for $so(n)$ is the following: Let

$$(M_{ab})_{jk} = \delta_{aj}\delta_{bk} - \delta_{bj}\delta_{ak},$$

which are antisymmetric $M_{ab} = -M_{ba}$. One can easily check that they satisfy the commutation relations

$$[M_{ab}, M_{cd}] = \delta_{bc}M_{ad} - \delta_{ac}M_{bd} - \delta_{bd}M_{ac} + \delta_{ad}M_{bc}.$$

$\underline{SL(n) \text{ and } sl(n)}$:

Recall that

$$sl(n, \mathbb{R}) := \{A \in Mat(n, \mathbb{R});\ \mathrm{Tr}(A) = 0\}.$$

Let $A \in sl(n)$. Then $\det(e^A) = e^{Tr(A)} = 1$, which means that $e^A \in SL(n)$. Conversely, consider $g \in SL(n)$ near $\mathbb{1}$, so that $g = e^A$ for some $A \in gl(n)$ (by theorem 14). Then $1 = dete^A = e^{Tr(A)}$. This implies that $Tr(A) = 0$, hence

$$\exp(sl(n)) = SL(n) \subset GL(n),$$

therefore $sl(n) = Lie(SL(n))$.

$\underline{U(n) \text{ and } u(n)}$:

Recall that

$$u(n) := \{A \in Mat(n, \mathbb{C});\ A^\dagger = -A\}.$$

Let $A \in u(n)$. Then $(e^A)^\dagger = e^{-A} = (e^A)^{-1}$, which means that $e^A \in U(n)$ is a unitary matrix. Conversely, consider $g \in U(n)$ near $\mathbb{1}$, so that $g = e^A$ for some $A \in gl(n)$ (by theorem 14). Then $e^{A^\dagger} = g^\dagger = g^{-1} = e^{-A}$. Because exp is a local diffeomorphism (resp. by taking the matrix log), this implies that $A^\dagger = -A$, hence

$$\exp(u(n)) = U(n) \subset GL(n),$$

therefore $u(n) = Lie(U(n))$.

Similarly, $su(n) = Lie(SU(n)) = \{A \in Mat(n, \mathbb{C});\ A^\dagger = -A, Tr(A) = 0\}$.

We can now easily compute the dimensions of these Lie groups, simply by computing the dimension of their Lie algebras. One finds that

$U(n)$ has dimension n^2 (as real manifold!!),
$SU(n)$ has dimension $n^2 - 1$ (as real manifold!!),
$SL(n, \mathbb{C})$ has dimension $2n^2 - 2$,
$SL(n, \mathbb{R})$ has dimension $n^2 - 1$,
$O(n, \mathbb{R})$ and $SO(n, \mathbb{R})$ have dimension $n(n - 1)/2$,

There are various "real sectors" of these classical Lie groups resp. algebras. A typical example is the Lorentz group:

5.3 Example: Lie algebra and exponential map for $SO(3)$ and $SU(2)$.

To illustrate this, reconsider $SO(3)$ in detail. According to the above, its Lie algebra is

$$so(3) := \{A \in Mat(n, \mathbb{R});\ A^T = -A, Tr(A) = 0\}.$$

A convenient basis of $so(3)$ is given by

$$X_1 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad X_2 := \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad X_3 := \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

hence any $u \in so(3)$ can be written uniquely as $u = u_k X_k$, and any element of $SO(3)$ can be written as

$$e^u = e^{u_k X_k} \qquad \in SO(3).$$

Their Lie algebra is

$$[X_i, X_j] = -\epsilon_{ijk} X_k. \tag{31}$$

It is easy to calculate the exponentials explicitly, reproducing finite rotation matrices.

In physics, one often allows complex coefficients, defining

$$J_k := -iX_k$$

which are hermitian $J_i^\dagger = J_i$ and satisfy the "rotation algebra"

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

as known from Quantum Mechanics. Technically speaking one *complexifies the Lie algebra*: Given any "real" Lie algebra such as $so(3) = \langle X_1, X_2, X_3 \rangle_{\mathbb{R}}$ with some basis X_i, one simply allows linear combinations over $\mathbb{C}$, i.e. replaces $\mathfrak{g} \cong \mathbb{R}^n$ by $\mathfrak{g}_{\mathbb{C}} \cong \mathbb{C}^n$, extending the commutation relations linearly over $\mathbb{C}$: $so(3)_{\mathbb{C}} = \langle X_1, X_2, X_3 \rangle_{\mathbb{C}}$. From now on we work with Lie algebras over $\mathbb{C}$, which is very useful and much easier than $\mathbb{R}$. Then finite rotations are given by

$$e^u = e^{iu_k J_k} = R(\vec{u}) \qquad \in SO(3)$$

Similarly, consider $SU(2)$. According to the above, its Lie algebra is

$$su(2) := \{A \in Mat(n, \mathbb{C});\ A^\dagger = -A = 0, Tr(A) = 0\}.$$

A convenient basis of $su(2)$ is given by (i) times the Pauli matrices, $X_i = \frac{i}{2}\sigma_i$ for

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

hence any $u \in su(2)$ can be written uniquely as $u = u_j(i\sigma_j)$. Then

$$e^u = e^{u_j(i\sigma_j)} \qquad \in SU(2).$$

Again, one defines the complexified generators

$$J_k = \frac{1}{2}\,\sigma_k,$$

which satisfy

$$\boxed{[J_i, J_j] = i\epsilon_{ijk}J_k, \qquad [X_i, X_j] = -\epsilon_{ijk}X_k.}$$

therefore

$$so(3) \cong su(2).$$

This is the "algebraic" reason why $SO(3)$ and $SU(2)$ are "locally isomorphic", and according to Theorem 10 it implies that there is a Group-homomorphism

$$\Phi : SU(2) \mapsto SO(3)$$

We have seen this explicitly in the beginning.

5.4 $SO(3,1)$ and $so(3,1)$

The Lorentz group $SO(3,1)$ is defined by

$$M^i_{i'} M^j_{j'} \eta^{i'j'} = \eta^{ij},$$

where

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

i.e.

$$M\eta M^T = \eta$$

or

$$M\eta = \eta M^{-1T}$$

and $\det M = 1$. The set of these M is certainly a Lie group. Considering "infinitesimal group elements" or

$$M = e^{iL}$$

this leads to

$$L\eta = -\eta L^T$$

A (complexified) basis is given by

$$K_x = -i \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad K_y = -i \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad K_z = -i \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$

("boost generators"), and the usual "space-like" generators of rotations

$$J_x = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad J_y = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad J_z = -i \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

The structure constants are:

$$\begin{aligned} [K_x, K_y] &= -iJ_z, etc, \\ [J_x, K_x] &= 0, etc, \\ [J_x, K_y] &= iK_z, etc, \\ [J_x, J_y] &= iJ_z, etc, \end{aligned} \tag{32}$$

It is interesting to note that if we allow complex coefficients in the Lie algebra, then

$$\begin{aligned} \vec{A} &:= \frac{1}{2}(\vec{J} + i\vec{K}), \\ \vec{B} &:= \frac{1}{2}(\vec{J} - i\vec{K}), \end{aligned} \tag{33}$$

commute:

$$\begin{aligned} [A_x, A_y] &= iA_z, etc, \\ [B_x, B_y] &= iB_z, etc, \\ [A_i, B_j] &= 0, \forall i, j \end{aligned} \tag{34}$$

Hence formally, $so(3,1)_{\mathbb{C}} \cong su(2)_{\mathbb{C}} \oplus su(2)_{\mathbb{C}}$. However, this is only for the complexified Lie algebra! "Real" elements of $SO(3,1)$ have the form

$$\Lambda = e^{i(x_j J_j + y_j K_j)} = e^{\frac{i}{2}((x+iy)_i B_i + (x-iy)_i A_i)}$$

with real x_i, y_i. In terms of the generators A and B, the coefficients are not real any more! Nevertheless, this is very useful to find representations . In particular, there are 2 inequivalent 2-dimensional representations :

a) $A_i = \frac{1}{2}\sigma_i, B_i = 0$: "undotted spinors", corresponding to the 2-dim. rep.

$$\psi^\alpha = \psi = \begin{pmatrix} \psi^+ \\ \psi^- \end{pmatrix} \in \mathbb{C}^2$$

of $SL(2,\mathbb{C})$. with the obvious action

$$\begin{aligned} SL(2,\mathbb{C}) \times \mathbb{C}^2 &\rightarrow \mathbb{C}^2, \\ (M,\psi) &\rightarrow M \cdot \psi \end{aligned}$$

This is NOT a rep. of $SO(3,1)$!! The exponential map takes the form

$$M = e^{\frac{i}{4}(x-iy)_i \sigma_i}$$

b) $A_i = 0, B_i = \frac{1}{2}\sigma_i$: "dotted spinors", corresponding to the 2-dim. rep.

$$\tilde{\psi}^{\tilde{\alpha}} = \tilde{\psi} = \begin{pmatrix} \tilde{\psi}^+ \\ \tilde{\psi}^- \end{pmatrix} \in \mathbb{C}^2$$

of $SL(2,\mathbb{C})$. with the action

$$\begin{aligned} SL(2,\mathbb{C}) \times \mathbb{C}^2 &\rightarrow \mathbb{C}^2, \\ (M,\tilde{\psi}) &\rightarrow M^* \cdot \tilde{\psi} \end{aligned}$$

This is also NOT a rep. of $SO(3,1)$, and it is in no sense equivalent the the above one.

These are Weyl spinors. In the standard model, all leptons are described by (or built up by) such Weyl spinors.

<u>Finite boosts:</u>

We already know finite rotations. Finite boosts can be calculated similarly, e.g.

$$K_x^2 = -\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

we get

$$\begin{aligned} e^{i\epsilon K_x} &= \sum_{m=0}^{\infty} \frac{(i\epsilon K_x)^m}{m!} = 1 - K_x^2 \sum_{n=1}^{\infty} \frac{(\epsilon)^{2n}}{(2n)!} + iK_x \sum_{n=0}^{\infty} \frac{(\epsilon)^{2n+1}}{(2n+1)!} \\ &= \mathbb{1} - K_x^2(\cosh(\epsilon) - 1) + iK_x \sinh(\epsilon) = \begin{pmatrix} \cosh(\epsilon) & \sinh(\epsilon) & 0 & 0 \\ \sinh(\epsilon) & \cosh(\epsilon) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{aligned} \tag{35}$$

as desired, where $\beta = \tanh(\epsilon)$ and $\gamma = \cosh(\epsilon) = \frac{1}{\sqrt{1-\beta^2}}$.

Observe that $\mathfrak{so}(3,1)_{\mathbb{C}} = \mathfrak{so}(4)_{\mathbb{C}}$. More generally, to understand the structure of the (finite-dimensional) representations, we can (and will) restrict ourselves to Lie algebras corresponding to compact Lie groups.

6 A first look at representation theory

6.1 Definitions

The main application of groups in physics is to exploit *symmetries* of physical systems. A symmetry is given by a group (e.g. rotations, permutations, reflections, ...), which can "act" on a physical system and puts it in another but "equivalent" state. This is particularly simple in Quantum Mechanics: The states of the system form a Hilbert space $\mathcal{H}$, which is a vector space. A symmetry of the system therefore amounts to an action of a group G (rotations, say) on $\mathcal{H}$. Hence we have a map

$$\begin{aligned} G \times \mathcal{H} &\to \mathcal{H}, \\ (g, \psi) &\mapsto g \triangleright \psi \end{aligned} \tag{36}$$

which of course should respect the group law, $(g_1 g_2) \triangleright \psi = (g_1) \triangleright (g_2 \triangleright \psi)$ and $e \triangleright \psi = \psi$. Due to the superposition principle, it should be linear in the second argument. Equivalently,

$$\pi : G \to GL(\mathcal{H}) \tag{37}$$

should satisfy

$$\pi(g_1)\pi(g_2) = \pi(g_1 g_2), \qquad \pi(e) = 1\!\!1\ , \qquad \pi(g^{-1}) = \pi(g)^{-1}. \tag{38}$$

This is precisely the definition of a *representation of G on $\mathcal{H}$*:

Definition 15 *Let G be a group. Then a* ***(real, complex) representation*** *of G is a group homomorphism*

$$\pi : G \to GL(V)$$

where V is a (real, complex) vector space (i.e. $\mathbb{R}^n$ resp. $\mathbb{C}^n$).

A ***unitary representation*** *of G is a group homomorphism*

$$\pi : G \to U(\mathcal{H})$$

into the unitary operators on some Hilbert space $\mathcal{H}$.

We will mainly consider finite-dimensional representations. Understanding the representations is one of the main issues in group theory (and crucial in physics).

Now we can apply theorem 10, and we obtain "by differentiating" from each representation of G a Lie algebra homomorphism $d\pi : \mathfrak{g} \to gl(V)$. This yields the following definition:

Definition 16 *A* ***(finite-dimensional, real, complex) representation*** *of a Lie algebra $\mathfrak{g}$ is a Lie algebra homomorphism*

$$\mathfrak{g} \to gl(V)$$

where V is a (finite-dimensional, real, complex) vector space.

Note that if G is simply-connected, then theorem 10 implies conversely that every representation of the Lie algebra $\mathfrak{g}$ induces a representation of the Lie group G. For example, recall that the spin 1/2 rep of the angular momentum aglebra $[J_i, J_j] = \epsilon_{ijk} J_k$ leads to a representation of $SU(2)$, but not of $SO(3)$. This means that we can basically restrict ourselves to studying representations of Lie algebras.

Furthermore, note that if $\pi : G \to U(\mathcal{H})$ is a unitary representationof G and we write $\pi(g) = e^{i a_i \pi(J_i)} \in G$ where $J_i \in \mathfrak{g}$, then π is unitary if and only if $\pi(J_i)$ is hermitian. Hence unitary representations of G correspond to representations of $\mathfrak{g}$ with hermitian (or anti-hermitian...) operators.

Definition 17 *Let π be a representation of a group G, acting on a space V. A subspace W of V is called* ***invariant*** *if $\pi(g)w \in W$ for all $w \in W$ and all $g \in G$. A representation with no non-trivial invariant subspaces (apart from $W = \{0\}$ and $W = V$) is called* ***irreducible****. A representation which can be written as the direct sum of irreps $V = V_1 \oplus V_2, \pi = \pi_1 \oplus \pi_2$ (or more) is called* ***completely reducible****.*

Note that

Lemma 18 (finite-dimensional) Unitary representations are always completely reducible.

proof: Assume that the unitary representation $\mathcal{H}$ is not irreducible, and let $W \subset \mathcal{H}$ be an invariant subspace. Then $W^\perp$ is also invariant (since $\langle w, \pi(g)v\rangle = \langle \pi(g)^\dagger w, v\rangle = 0$), and

$$\mathcal{H} = W \oplus W^\perp.$$

Repeat if $W^\perp$ is not irreducible. qed

For example, the *basic representation* of $SO(3)$ is the one in which $SO(3)$ acts in the usual way on $\mathbb{R}^3$. More generally, If G is a subgroup of $GL(n;\mathbb{R})$ or $GL(n;\mathbb{C})$, it acts naturally on $\mathbb{R}^n$ resp. $\mathbb{C}^n$. There are many different and non-equivalent representations, though. A less trivial example is the action of $SO(3)$ on "fields", i.e. functions on $\mathbb{R}^3$ or S^2, via $g \triangleright f(x) = f(g^{-1}x)$. This is an infinite-dimensional representation, which however can be decomposed further. This leads to the spherical harmonics, which are precisely the (finite-dimensional) irreps of $SO(3)$.

(Exercise: work this out. Consider polynomial functions (on S^2), decompose by degree, express in spherical coordinates ...).

6.2 The representation theory of $su(2)$

Recall the rotation algebra (39) $su(2)_\mathbb{C} \cong so(3)_\mathbb{C}$,

Rising-and lowering operators

A convenient basis of $su(2)_\mathbb{C}$ is given by the rising-and lowering operators, which in terms of

$$J_\pm := J_1 \pm iJ_2, \qquad J_0 := 2J_3$$

satisfy

$$[J_0, J_\pm] = \pm 2J_\pm, \qquad [J_+, J_-] = J_0 \tag{39}$$

This is very useful if one studies representations, and we want to determine all finite-dimensional irreps $V = \mathbb{C}^n$ of this Lie algebra.

Because V is finite-dimensional (and we work over $\mathbb{C}$!), there is surely an eigenvector v_λ of J_0 with

$$J_0\, v_\lambda = \lambda\, v_\lambda.$$

using the above CR, we have

$$J_0(J_\pm\, v_\lambda) = J_\pm J_0\, v_\lambda \pm 2J_\pm\, v_\lambda = (\lambda \pm 2)\, (J_\pm\, v_\lambda)$$

Hence $J_\pm v_\lambda$ is again an eigenvector of J_0, with eigenvalue $(\lambda \pm 2)$. This is why $J_\pm$ are called rising-and lowering operators. We can continue like this acting with J_+. Each time we get an eigenvector of J_0 whose eigenvalues is increased by 2. From linear algebra we know that these are all linearly independent, so at some some point we must have a v_Λ with

$$J_+\, v_\Lambda = 0 \tag{40}$$

This $v_\Lambda \neq 0 \in V$ is called the "highest weight vector" of V. Now consider

$$v_{\Lambda-2n} := (J_-)^n \, v_\Lambda, \qquad \text{i.e.} \quad v_{m-2} = J_- v_m \tag{41}$$

(hence $v_{m-2} = J_- v_m$) which for the same reason have eigenvalues

$$J_0 \, v_{\Lambda-2n} = (\Lambda - 2n) \, v_{\Lambda-2n}.$$

One says that $(\Lambda - 2n)$ is the weight of $v_{\Lambda-2n}$, i.e. the eigenvalue of J_0. Now we claim that $J_+ \, v_{\Lambda-2n}$ is proportional to $v_{\Lambda-2n+2}$, i.e.

$$J_+ \, v_{\Lambda-2n} = r_n \, v_{\Lambda-2n+2}.$$

for some $r_n \in \mathbb{R}$. To see this, consider

$$J_+ \, v_{\Lambda-2n} = J_+ J_- \, v_{\Lambda-2n+2} = (J_- J_+ + J_0) \, v_{\Lambda-2n+2} \ .$$

The claim now follows by induction on n, starting with $r_0 = 0$. We can thus proceed as

$$J_+ \, v_{\Lambda-2n} = (J_- J_+ + J_0) \, v_{\Lambda-2n+2} = (r_{n-1} + \Lambda - 2n + 2) \, v_{\Lambda-2n+2} \tag{42}$$

Hence we obtain the recursion relation $r_n = (r_{n-1} + \Lambda - 2n + 2)$ for r_n, which is easy to solve (exercise):

$$r_n = n(\Lambda - n + 1).$$

Because V is finite-dimensional, there is a maximal integer N such that $v_{\Lambda-2N} \neq 0$ but

$$J_- \, v_{\Lambda-2N} = 0.$$

This implies

$$0 = J_+ J_- \, v_{\Lambda-2N} = (J_- J_+ + J_0) \, v_{\Lambda-2N} = (r_N + \Lambda - 2N) \, v_{\Lambda-2N}.$$

Substituting $r_n = n(\Lambda - n + 1)$, this yields the equation $N^2 + (1 - \Lambda)N - \Lambda = 0$ which we can solve for Λ. This gives

$$N = \Lambda$$

which is a non-negative integer. The dimension of V is then

$$dim V = N + 1 = \Lambda + 1$$

In physics, one usually defines the *spin* as

$$j = \frac{1}{2}\Lambda.$$

Then

$$dim V = 2j + 1.$$

We have thereby classified *all possible* finite-dimensional representations of $su(2)$. This means that up to a choice of basis, all irreps are equivalent to some "highest weight" irrep

$$V_\Lambda := \{v_{\Lambda-2n}, n = 0, 1, 2, ..., N, \qquad \Lambda = N\},$$

i.e. they are characterized by their dimension.

Irreps are also characterized by the value of the Casimir operator

$$\vec{J}^2 = J_1 J_1 + J_2 J_2 + J_3 J_3$$

which satisfies

$$[\vec{J}^2, J_i] = 0.$$

Therefore it takes the same value on any vector in the irrep. V_Λ. Using

$$\vec{J}^2 = \frac{1}{4} J_0(J_0 + 2) + J_- J_+$$

it is easy to evaluate it on the highest weight vector v_Λ, which gives

$$\vec{J}^2 = \frac{1}{4}\Lambda(\Lambda + 2) = j(j + 1)$$

in the irrep V_Λ.

In physics, one is usually interested in unitary representations of the group $SU(2)$. This means that $e^{ix_a J_a}$ is unitary, hence

$$J_a^\dagger = J_a.$$

Hence this is equivalent to a representation of $su(2)$ with hermitian generators J_a. One can easily show that all the above representations are actually unitary in this sense, if one defines an inner product on $\mathbb{C}^n$ such that states with different weight are orthogonal: for suitable normalization,

$$|\Lambda - 2n\rangle := c_n\, v_{\Lambda-2n} = c_n\, (J_-)^n\, v_\Lambda \tag{43}$$

satisfy

$$\langle \Lambda - 2n, \Lambda - 2m\rangle = \delta_{n,m}.$$

It is easy to see that

$$J_+|2(m-1)\rangle = \sqrt{\frac{1}{2}(\Lambda + 2m)(\Lambda - 2m + 2)}|2m\rangle.$$

6.3 The adjoint representation

For every Lie algebra $\mathfrak{g}$, there is a natural representation on itself considered as a vector space. One defines

$$\text{ad} : \mathfrak{g} \to gl(\mathfrak{g})$$

defined by the formula

$$\mathsf{ad}_X(Y) = [X, Y].$$

It is easy to see (check !!!! Jacobi) that ad is a Lie algebra homomorphism, and is therefore a representation of $\mathfrak{g}$, called the **adjoint representation**.

For example, consider $\mathsf{SO}(3)$ with generators J_i. Then $so(3)$ acts on $X \in so(3)$ as

$$ad_{J_i}(x) = [J_j, X]$$

This is an infinitesimal rotation of $X \in so(3)$. In fact it is 3-dimensional, which is the dimension of the vector (3) rep. By uniqueness, it follows that the basic representation and the adjoint representation are equivalent.

In a basis with $[X_i, X_k] = c_{ij}^k X_k$, this is

$$ad_{X_i}(X_j) = c_{ij}^k X_k \tag{44}$$

hence the matrix which represents ad_{X_i} is

$$(ad_{X_i})_l^k = c_{il}^k.$$

Hence the structure constants always define the adjoint representation.

<u>Group version</u>

The adjoint representation of $\mathfrak{g}$ on $\mathfrak{g}$ should lift to a representation of G on $\mathfrak{g}$. This works as follows:

Let G be a Lie group with Lie algebra $\mathfrak{g}$. For each $g \in G$, consider the map

$$\begin{aligned} Ad_g : \mathfrak{g} &\to \mathfrak{g}, \\ X &\to gXg^{-1} \end{aligned} \tag{45}$$

The rhs is in fact $\in \mathfrak{g}$. One way to see this is the following (for matrix groups): $e^{gXg^{-1}} = ge^X g^{-1} \in G$, therefore (exp is locally invertible!) $gXg^{-1} \in \mathfrak{g}$. We can view Ad as map

$$\begin{aligned} Ad : G &\to GL(\mathfrak{g}), \\ g &\to [X \to gXg^{-1}] \end{aligned} \tag{46}$$

which is clearly a representation (group homom.). Then we have

$$ad = d(Ad) \tag{47}$$

(proof:

$$d(Ad)_X = \frac{d}{dt}|_0 \; Ad(e^{tX}) \in gl(\mathfrak{g}),$$

which if applied to $Y \in \mathfrak{g}$ gives

$$d(Ad)_X(Y) = \frac{d}{dt}|_0 \; Ad(e^{tX})(Y) = \frac{d}{dt}|_0 \; e^{tX} Y e^{-tX} = [X, Y].$$

)

By the diagram 29, this implies

$$Ad(e^X) = e^{ad_X} \tag{48}$$

i.e.

$$e^X Y e^{-X} = (e^{ad_X})(Y) = (1 + ad_X + \frac{1}{2!} ad_X^2 + ...)Y = Y + [X, Y] + \frac{1}{2!}[X, [X, Y]] + ...) \tag{49}$$

(of course this can also be verified directly). Equation (48) is correct for any Lie group resp. algebra.

For example, for $so(3)$ this means that

$$e^{i\vec{\phi}\cdot\vec{J}} X e^{-i\vec{\phi}\cdot\vec{J}} = R(\vec{\phi}) X R(-\vec{\phi}) = e^{i\vec{\phi}_j [J_j, .]} X$$

Now $[J_j, X]$ is an infinitesimal rotation of $X \in so(3)$, hence $e^{i\vec{\phi}_j [J_j, .]} X$ is the corresponding finite rotation of X. This means that a vector $X \in so(3)$ can be rotated by either rotating it directly (rhs) or conjugating it with the corresponding rotation matrices. We know this from Quantum mechanics: The rotation of angular momentum operators can be achieved by conjugation with rotation operators, i.e. by rotating the states in the Hilbert space.

The Killing form Define the following bilinear inner product on $\mathfrak{g}$:

$$(X, Y) := \kappa(X, Y) := Tr_{ad}(ad_X ad_Y) = Tr_{ad}([X, [Y, .]]) \tag{50}$$

This makes sense because ad_X is a map from $\mathfrak{g}$ to $\mathfrak{g}$. It is easy to show that this is a symmetric bilinear form which is invariant under ad:

$$\begin{aligned}
(X, ad_Y(Z)) &= (X, [Y, Z]) = Tr(ad_X ad_{[Y,Z]}) = Tr(ad_X [ad_Y, ad_Z]) \\
&= Tr([ad_X, ad_Y] ad_Z) = Tr(ad_{[X,Y]} ad_Z) = ([X, Y], Z) \\
&= -(ad_Y(X), Z)
\end{aligned} \tag{51}$$

This means that ad_Y acts as an anti-symmetric matrix on $\mathfrak{g}$, and therefore $Ad(e^Y) = e^{ad_Y}$ acts as an orthogonal matrix. Note that (51) is the infinitesimal version of

$$(X, Ad(e^Y)(Z)) = (Ad(e^{-Y})(X), Y),$$

i.e. $Ad(e^Y)$ is an orthogonal matrix w.r.t. the Killing form, i.e. the Killing from is invariant under G.

One can show that for semi-simple Lie algebras $\mathfrak{g}$ (these are by definition the direct sum of simple Lie algebras; simple Lie algebras are those which contain no nontrivial ideal and are not abelian), $\kappa(.,.)$ is the unique such invariant form on $\mathfrak{g}$, and it is non-degenerate. We will see that $\kappa(.,.)$ is positive definite for compact Lie groups, so that $\mathfrak{g}$ is a *Euclidean space* $\mathbb{R}^n$. Because it is unique, one can calculate it up to proportionality in any representation:

$$Tr_V(\pi(X)\pi(Y)) = \alpha\kappa(X,Y)$$

because this is also invariant (same proof as above).

Since $\mathfrak{g} = T_e(G)$, we can transport this Killing-metric to any $T_g(G)$ via dL_g, like the left-invariant vector fields. In this way, G becomes a Riemannian Manifold (i.e. with metric). Since the Killing metric is invariant under Ad, it follows that this metric is invariant under both left- and right translations. This also shows that there is a measure $d\mu_g$ on G which is invariant under L_g and R_g. This is the *Haar measure* on G, which exists and is unique on any (reasonable) Lie group. For example, $d\mu = d^n x$ on $\mathbb{R}^n$.

6.4 The Baker-Campbell-Hausdorff Formula

For X and Y sufficiently small elements of $\mathfrak{g}$, the following formula holds:

$$\exp(X)\exp(Y) = \exp(X + Y + \tfrac{1}{2}[X,Y] + \tfrac{1}{12}[X,[X,Y]] - \tfrac{1}{12}[Y,[X,Y]] + \cdots). \quad (52)$$

where .... stands for further terms which are always given by Lie brackets of ..., i.e. *terms in the Lie algebra*. This is a "pedestrian's" version of theorem 10, because we only need to know the Commutators, i.e. the Lie algebra of $\mathfrak{g}$ in order to calculate products of the group. This formula therefore tells us how to pass from the Lie algebra to the Lie group. In particular, if we have a Lie algebra homomorphism, we will also get a Lie group homomorphism just like in theorem 10

There is a "closed" BCH formula: Consider the function

$$g(z) = \frac{\log z}{1 - \frac{1}{z}}.$$

which is well-defined and analytic in the disk $\{|z-1|<1\}$, and thus for z in this set, $g(z)$ can be expressed as

$$g(z)=\sum_{m=0}^{\infty}a_m(z-1)^m.$$

This series has radius of convergence one. Then for any operator A on V with $\|A-I\|<1$, we can define

$$g(A)=\sum_{m=0}^{\infty}a_m(A-1)^m.$$

We are now ready to state the integral form of the Baker-Campbell-Hausdorff formula.

Theorem 19 (Baker-Campbell-Hausdorff) *For all $n\times n$ complex matrices X and Y with $\|X\|$ and $\|Y\|$ sufficiently small,*

$$\log\left(e^Xe^Y\right)=X+\int_0^1 g(e^{\mathrm{ad}X}e^{t\mathrm{ad}Y})(Y)\,dt. \tag{53}$$

Proof: omitted

6.5 Constructing and reducing representations

Weyls unitarity trick Assume that G is a compact Lie group, and

$$\pi:G\to GL(\mathcal{H})$$

is any finite-dimensional representation of G on some Hilbert space $\mathcal{H}$, not necessarily unitary. Let (u,v) denote the inner product on $\mathcal{H}$.

One can then obtain a *unitary* representation from it as follows: Let $d\mu_g$ be the Haar measure on G, i.e. the unique measure on G which is invariant under L_g and R_g. This exists and is unique as explained above.

Then define a *new* inner product $\langle u,v\rangle$ by

$$\langle u,v\rangle:=\int_G d\mu_g(\pi(g)u,\pi(g)v). \tag{54}$$

This is well-defined because G is compact, and positiv definite and non-degenerate because $(,)$ is. Most importantly, it is invariant under the action of G:

$$\langle\pi(g)u,\pi(g)v\rangle=\langle u,v\rangle. \tag{55}$$

This means that all $\pi(g)$ are unitary operators w.r.t. this new inner product. In particular, it follows that all finite-dimensional representations of compact G are unitarizable (i.e. there is a suitable inner product such that is becomes unitary), and therefore they're *completely reducible* by Lemma 18. This also means that the corresponding Lie algebras $Lie(G)$ are "semisimple".

Applying this result to the adjoint representation (starting with any positive-definite real inner product instead of a sesquilinear form), it follows by invariance that the thus-obtained invariant inner product is the Killing form. Therefore

The Killing form is non-degenerate and Euclidean for compact G.

This is not true for non-compact groups: e.g. the 4-dimensional representation of $SO(3,1)$ is not unitarizable. In fact the opposite is true: only infinite-dimensional representations of $SO(3,1)$ are unitary! This is why one has to go to Field theory in order to have a relativistic Quantum theory, where Lorentz transformations should preserve probability and therefore be unitary.

We can now show

<u>Antisymmetry of the structure constants</u>

Let c_{ijk} be the structure constants in an ON basis (defined by the Killing from!), $[X_i, X_j] = c_{ijk} X_k$ for any representation of $\mathfrak{g}$.

Then

$$c_{ijk} = Tr([X_i, X_j]X_k) = Tr(X_j[X_k, X_i]) = c_{kij}$$

hence c_{ijk} is cyclic. Together with $c_{ijk} = -c_{jik}$ it follows that c_{ijk} is totally antisymmetric.

Tensor products: If V and W are 2 representations of the Lie group G, then so is $V \otimes W$ by

$$\begin{aligned} \pi : G &\rightarrow GL(V \otimes W) \\ g &\mapsto \pi_V(g) \otimes \pi_W(g) \end{aligned} \tag{56}$$

Passing to the Lie algebra by differentiating, this becomes (set $g = e^{it(x_a J_a)}$)

$$\begin{aligned} \pi : \mathfrak{g} &\rightarrow gl(V \otimes W) = gl(V) \otimes gl(W) \\ \mathfrak{g} &\mapsto \pi_V(\mathfrak{g}) \otimes \mathbb{1} + \mathbb{1} \otimes \pi_W(\mathfrak{g}) \end{aligned} \tag{57}$$

It is easy to check directly that this is a representation of $\mathfrak{g}$.

For example, this is the meaning of "adding angular momenta" in Quantum Mechanics, via $J_i = L_i + S_i$.

For semi-simple Lie algebras, the tensor product of representations $V \otimes W$ always decomposes into the direct sum of irreps,

$$V \otimes W \cong \oplus_j n_j V_j$$

where V_j denote all possible irreps, and n_i are the "multiplicities".

Actually any products of representations transform like tensor products. e.g.,

$$[J_i, AB] = [J_i, A]B + A[J_i, B]$$

looks just like the action of $[J, .]$ on $A \otimes B$. For example, consider the Casimir for $su(2)$,

$$\vec{J}^2 := J_1 J_1 + J_2 J_2 + J_3 J_3 \qquad \in Mat(3)$$

where $J_i \in so(3)$ transform like a vector under $so(3)$. Clearly $\vec{J}^2$ transforms like a scalar under $[J, .]$, i.e. trivially. According to the above, this means that

$$[\vec{J}^2, J_i] = 0$$

Of course this can be checked explicitly.

In general, Casimirs are expressions in the generators which commute with all generators.

As another application we can quickly derive the

Wigner-Eckart theorem: Let $\tilde{J}_i$ be the spin j irrep of $su(2)$, i.e.

$$\tilde{J}_i \in Mat(2j + 1, \mathbb{C}) \tag{58}$$

The Wigner-Eckart theorem states that every "vector operator" $K_i \in Mat(2j + 1, \mathbb{C})$, i.e.

$$[J_i, K_j] = i\varepsilon_{ijk} K_k$$

is proportional to J_i:

$$K_i = \alpha J_i$$

for some constant α.

<u>Proof:</u>

Consider the following action of $su(2)$ on $Mat(2j + 1, \mathbb{C})$:

$$\begin{aligned} \pi : su(2) \times Mat(2j + 1, \mathbb{C}) &\to Mat(2j + 1, \mathbb{C}), \\ (J_i, M) &\mapsto i[\tilde{J}_i, M] \end{aligned} \tag{59}$$

(cp. the adjoint!). (This corresponds to the rep of $SU(2)$

$$\begin{aligned} \pi : SU(2) \times Mat(2j+1,\mathbb{C}) &\to Mat(2j+1,\mathbb{C}), \\ (U,M) &\mapsto \pi(U)M\pi(U)^{-1} \end{aligned} \tag{60}$$

)

Under this action,

$$Mat(2j+1,\mathbb{C}) \cong \mathbb{C}^{2j+1} \otimes (\mathbb{C}^{2j+1})^* = (0) + (1) + ... + (2j)$$

denoting the irreps by their spin.

(note that $(\mathbb{C}^{2j+1})^* \mapsto -(\mathbb{C}^{2j+1})^* J_i$ is a rep., equivalent to $(\mathbb{C}^{2j+1}) \mapsto J_i(\mathbb{C}^{2j+1})$!)

Now assume that $X_i \in Mat(2j+1,\mathbb{C})$ are "vector operators", i.e.

$$[J_i, K_j] = i\varepsilon_{ijk} K_k$$

This means that X_i transforms like a spin 1 rep. under this action. But there is only one (1) in the above decomposition, given by J_i. Hence

$$K_i = \alpha J_i$$

for some constant α. Hence any vector operator in an irreducible representation is proportional to $\tilde{J}_i$. qed

Spherical harmonics: Polynomials, Λ^2, etc.

7 $SU(3)$ and Quarks

Let us choose a basis of $i\, su(3) = \{M;\ M^\dagger = M, Tr(M) = 0\}$: The standard choice is given by the Gell-Mann matrices

$$\begin{aligned}
\lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
\lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \\
\lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \\
\lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.
\end{aligned} \tag{61}$$

These are the analogs of the Pauli matrices for $su(2)$. They are normalized such that

$$(\lambda_a, \lambda_b) := Tr(\lambda_a \lambda_b) = 2\delta_{ab}.$$

The first 3 are just the Pauli matrices with an extra column, the other also have some similarity. Hence we define

$$\begin{aligned} T_x &= \frac{1}{2}\lambda_1, \quad T_y = \frac{1}{2}\lambda_2, \quad T_z = \frac{1}{2}\lambda_3, \\ V_x &= \frac{1}{2}\lambda_4, \quad V_y = \frac{1}{2}\lambda_5, \\ U_x &= \frac{1}{2}\lambda_6, \quad U_y = \frac{1}{2}\lambda_7, \end{aligned} \tag{62}$$

and the corresponding complex combinations

$$T_\pm = T_x \pm iT_y, \quad V_\pm = V_x \pm iV_y, \quad U_\pm = U_x \pm iU_y,$$

e.g.

$$U_+ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, V_+ = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

etc. To make things more transparent, also introduce

$$V_3 := [V_+, V_-], \qquad U_3 := [U_+, U_-]$$

which are both linear combinations of λ_3 and λ_8. Then it is easy to check that $U_i,\ V_i,\ T_i$ form 3 different representations of $su(2)$, called $su(2)_T,\ su(2)_U,\ su(2)_V$.

Now note that

$$H_3 := \lambda_3, \qquad Y := \lambda_8$$

are diagonal and orthogonal, and commute with each other:

$$(H_3, H_3) = (Y, Y) = 2, \quad (H_3, Y) = 0$$

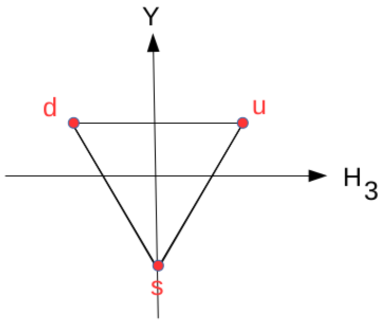
(use $(X, Y) \propto Tr(\pi(X)\pi(Y))$!). Also, there is no other element in $su(3)$ which commutes with them. Hence H_3 and Y form a "maximal set of commuting observables", and one can diagonalize them in any representation (recall that any rep. is unitary here). We will therefore label the eigenstates with the eigenvalues of these observables:

$$H_3|m, y\rangle = m|m, y\rangle, \qquad Y|m, y\rangle = y|m, y\rangle \tag{63}$$

(denoted isospin and hypercharge). In particular, in the above "defining" 3-dimensional representation we have 3 common eigenstates

$$|1, \frac{1}{\sqrt{3}}\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |-1, \frac{1}{\sqrt{3}}\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |0, -\frac{2}{\sqrt{3}}\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{64}$$

Since H_3 and Y orthogonal, we draw their eigenvalues as orthogonal axis in a 2-dimensional "weight space" (=space of possible eigenvalues):



Then the 3 eigenstates form an equilateral triangle.

In particle physics, these are the quarks u, d, s, called "up", "down" and "strange"!

We observe how the $T_\pm, V_\pm, U_\pm$ act on this representation: they are rising- and lowering operators on the 3 sides of the triangle, which each form a spin $\frac{1}{2}$ representation of $su(2)$:

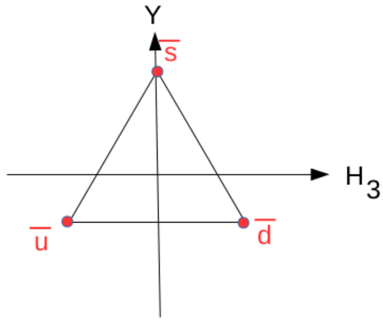
$$\begin{aligned} T_+|m,y\rangle &= |m+2,y\rangle, \\ U_+|m,y\rangle &= |m-1,y+\sqrt{3}\rangle, \\ V_+|m,y\rangle &= |m+1,y+\sqrt{3}\rangle, \end{aligned} \tag{65}$$

This holds in any representation, because of the commutation relations

$$\begin{aligned} &[H_3, T_\pm] = \pm 2T_\pm, \quad [Y, T_\pm] = 0 \\ &[H_3, U_\pm] = \mp U_\pm, \quad [Y, U_\pm] = \pm\sqrt{3}U_\pm \\ &[H_3, V_\pm] = \pm V_\pm, \quad [Y, V_\pm] = \pm\sqrt{3}V_\pm \end{aligned} \tag{66}$$

There exists another 3-dimensional representation of $\mathfrak{su}(3)$ which is inequivalent to the above (3) representation, which is obtained by complex conjugating (but not transposing) the Gell-mann matrices.

$\underline{(\overline{3})}$:



These label the Antiquarks $\overline{u}, \overline{d}, \overline{s}$.

Now consider another representation, the adjoint representation. Hence the vector space is

$$V \equiv (8) \cong \mathbb{C}^8 = \langle \lambda_1, ..., \lambda_8 \rangle$$

and H is represented by ad_H, and Y by ad_Y. Again they commute (since it is a representation), hence they can be diagonalized, and we label the states in (8) again by their eigenvalues (=weights),

$$(8) = \{ |m_i, y_i\rangle, i = 1, 2, ..., 8\}.$$

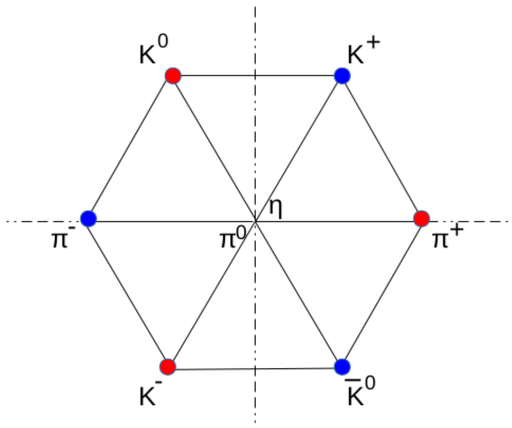
The weights are easily obtained from the commutation relations: Note that the rising- and lowering operators are eigenvectors of ad_H and ad_Y, hence

$$T_\pm \propto | \pm 2, 0\rangle, \qquad U_\pm \propto | \mp 1, \pm\sqrt{3}\rangle, \qquad V_\pm \propto | \pm 1, \pm\sqrt{3}\rangle \tag{67}$$

These are 6 eigenvectors of H_3, Y; the 2 missing ones are H_3 and Y themselves, which have eigenvalues 0:

$$H_3 \propto |0, 0\rangle_1, \qquad Y \propto |0, 0\rangle_2 \tag{68}$$

Hence there is a degeneracy: not all states in the adjoint rep can be labeled uniquely by the "weights" (m, y). Drawing the weight diagram, we get



These label the mesons, which are bound states of one quark and one anti-quark.

They form a regular hexagon. All angles are $\frac{\pi}{3}$. The action of the rising-and lowering operators is clear since they form irreps of $su(2)$ along the lines, and can be read off from the picture. This is very useful!

Notice that the differences of weights of states connected by the rising-and lowering ops are always the same: from (65), they are

$$\begin{aligned}
\alpha_{T_+} &= (2,0), & \alpha_{T_-} &= (-2,0), \\
\alpha_{U_+} &= (-1,\sqrt{3}), & \alpha_{U_-} &= (1,-\sqrt{3}), \\
\alpha_{V_+} &= (1,\sqrt{3}), & \alpha_{V_-} &= (-1,-\sqrt{3})
\end{aligned} \tag{69}$$

This will hold for any representation, therefore we give them a name: these 6 α'S are the *roots* of $su(3)$. Notice that they correspond precisely to the various rising- and lowering operators.

It turns out that all irreps of $su(3)$ have this kind of pattern: there is a "weight lattice", and all weights are linked by the above roots. Other representations include a six–dimensional irrep (6), a decouplet $\underline{(10)}$, and so on.

This is also the way to proceed in general:

- choose a maximal set of commuting elements $\{H_i\}$ in $\mathfrak{g}$ ("Cartan subalgebra")
- consider the adjoint representation, and find the eigenvectors of ad_{H_i}: these are the rising- and lowering operators, and correspond to the roots.

Moreover using the Killing form, we get a Euclidean metric on the weight space, which leads to lattices as above.

7.1 More about Quarks

The reason why this is successful for particle physics is that baryons and mesons are composite objects of Quarks, which are represented by

$$\begin{pmatrix} u \\ d \\ s \end{pmatrix}$$

which transform as (3) under $SU(3)$. The antiquarks

$$\begin{pmatrix} \overline{u} \\ \overline{d} \\ \overline{s} \end{pmatrix}$$

transform as $(\overline{3})$ under $SU(3)$. This generalizes isospin, see before. In fact, mesons are bound states $(q\overline{q})$ of a Quark and an Antiquark, hence they are in $(3) \otimes (\overline{3})$. We will see that

$$(3) \otimes (\overline{3}) = (8) + (1)$$

and the (8) is just the meson representation above. The other multiplets arise similarly. In fact, some of the particles were successfully predicted because they were missing in these diagrams.

The point is that the Hamiltonian for strong interactions is invariant under the $SU(3)$ group acting on these quarks, but there are some terms H' in the full Hamiltonian of the standard model (due to weak interactions) which break this symmetry:

$$H_{full} = H_{su(3)-invar} + H' \ .$$

This breaking leads to different masses for the various baryons in a given multiplet.

There is also an octet $(8) \subset (3) \otimes (3) \otimes (3)$ of baryons describing bound states of 3 quarks which includes the neutron and proton, and a decouplet (10) of baryons which arises also as a bound states or 3 quarks.

8 The structure of simple Lie algebras

The central technique to understand and work with Lie algebras is to identify the rising- and lowering operators. They rise and lower eigenvalues of the Cartan subalgebra.

From now on we work with Lie algebras over $\mathbb{C}$, which is crucial and much easier than $\mathbb{R}$. Given any "real" Lie algebra, one simply considers the "complexified Lie algebra" by just "allowing linear combinations over $\mathbb{C}$. For example, the complexification of $su(2) = \langle X_1, X_2, X_3 \rangle_{\mathbb{R}}$ is $su(2)_{\mathbb{C}} = \langle X_1, X_2, X_3 \rangle_{\mathbb{C}} \ni X^+ = X_1 + iX_2$. The original generators remain linearly independent over $\mathbb{C}$.

We consider only *semisimple* and *simple* Lie algebras. *Semisimple* Lie algebras are direct sums of simple Lie algebras. *Simple* Lie algebras are those which are not abelian and contain no (nontrivial) ideals, i.e. subspace $\mathfrak{h} \subset \mathfrak{g}$ with

$$[\mathfrak{h}, \mathfrak{g}] \subset \mathfrak{h}.$$

We have seen that all Lie algebras arising from compact Lie groups are semi-simple (except for $U(1)$ which is excluded by definition), and this is all we will need here.

8.1 Cartan subalgebras, Roots and the Cartan-Weyl basis

Consider the adjoint representation. Any $x \in \mathfrak{g}$ defines a linear map

$$ad_x : \mathfrak{g} \to \mathfrak{g}, \qquad ad_x(y) = [x, y].$$

Because we consider only (semi)simple Lie algebras, one can show that there exist $x \in \mathfrak{g}$ such that ad_x is diagonalizable (this is not trivial!).

For compact Lie groups, all representations (in particular ad) can be made unitary (by Weyls unitary trick), hence the generators are hermitian, and can therefore be diagonalized. This reflects the fact that all $\mathfrak{g}$ coming from compact Lie groups (with no $U(1)$ factor) are semisimple.

A *Cartan subalgebra* $\mathfrak{g}_0 \subset \mathfrak{g}$ is a maximal abelian subalgebra of $\mathfrak{g}$ whose elements $x \in \mathfrak{g}_0$ are all such that $ad_x : \mathfrak{g} \to \mathfrak{g}$ is diagonalizable. They exist by the above remark.

A (semisimple) Lie algebra $\mathfrak{g}$ can posses many different Cartan subalgebras. However, it turns out that they are related through some automorphisms, hence they are essentially equivalent. One can show that they all have the same dimension r, which is called the "rank" of $\mathfrak{g}$.

Since $\mathfrak{g}_0$ is a vector space (over $\mathbb{C}$, clearly), we can choose some basis $\{H_i\}_{i=1,\dots,r}$, and $[H_i, H_j] = 0$. Using the Jacobi identity, it follows that ad_{H_i} commute with each other, and one can diagonalize them simultaneously ("commuting observables"). Therefore $\mathfrak{g}$ is spanned by elements $y_\alpha \in \mathfrak{g}$ which are simultaneous eigenvectors of all ad_{H_i}, i.e.

$$ad_{H_i}(y_\alpha) = \alpha_i y_\alpha \tag{70}$$

where $\alpha = (\alpha_1, ..., \alpha_r)$ are the respective eigenvalues. In fact, such a given y_α is an eigenvector of any $H = \sum c_i H_i \in \mathfrak{g}_0$ in the Cartan subalgebra, with eigenvalue

$$ad_H(y_\alpha) = \alpha(H) y_\alpha \tag{71}$$

where

$$\alpha(H) := \sum c_i \alpha_i.$$

That is, α is a linear function from $\mathfrak{g}_0$ to $\mathbb{C}$, i.e.

$$\alpha \in \mathfrak{g}_0^*$$

(dual space). Such a non-zero $\alpha \in \mathfrak{g}_0^*$ corresponding to some common eigenvector $y_\alpha \in \mathfrak{g}$ is called a *root* of $\mathfrak{g}$. Note that a root α does not depend on the normalization of the eigenvector y_α, they are *unique.*

Because eigenvectors y corresponding to different eigenvalues resp. roots α are linearly independent, it follows that

$$\mathfrak{g} = \oplus_\alpha \mathfrak{g}_\alpha = \mathfrak{g}_0 \oplus (\oplus_{\alpha \neq 0} \mathfrak{g}_\alpha) \tag{72}$$

where $\mathfrak{g}_\alpha = \{y \in \mathfrak{g}; \quad ad_H(y) = \alpha(H) y\}$. This is called the *root space decomposition* of $\mathfrak{g}$ (relative to $\mathfrak{g}_0$). Clearly there are only finitely many roots! Usually one does not consider 0 to be a root, and separates the Cartan subalgebra $\mathfrak{g}_0$ (Notice that $\mathfrak{g}_0$ really IS the Cartan subalgebra!). The set of roots is denoted by

$$\boxed{\Phi = \{\alpha \neq 0\} \subset \mathfrak{g}_0^*}$$

One can show that

the "root spaces" $\mathfrak{g}_\alpha$ are one-dimensional, i.e. there is only one rising-resp. lowering operator (up to scalar mult.) for each root α. Moreover, the only multiples of α which are also roots are $-\alpha$.

(This will become more clear in section 8.1.2). This means that there is a basis of $\mathfrak{g}$ which apart from the basis $\{H_i\}$ of $\mathfrak{g}_0$ consists of elements X_α, one for each root $\alpha \in \Phi$, which satisfy

$$\boxed{\begin{aligned} [H_i, H_j] &= 0 \\ [H_i, X_\alpha] &= \alpha_i X_\alpha \qquad i = 1, 2, ..., r. \end{aligned}} \tag{73}$$

where $\alpha_i = \alpha(H_i)$. The r-dimensional vector (α_i) resp. the element $\alpha \in \mathfrak{g}_0^*$ are the roots of $\mathfrak{g}$. Such a basis is called a *Cartan-Weyl basis* of $\mathfrak{g}$. The generators X_α are called *ladder-operators*, because they rise resp. lower the eigenvalues of the H_i in a representation, just like for $su(2)$.

For example, consider $su(2)$ in the basis $J_0, J_\pm$. Here the Cartan subalgebra is given by $\mathbb{C}\, J_0$, and since $[J_0, J_\pm] = \pm 2 J_\pm$, the roots spaces are

$$\mathfrak{g}_+ = \mathbb{C} J_+, \qquad \mathfrak{g}_- = \mathbb{C} J_-, \quad \mathfrak{g}_0 = \mathbb{C} H$$

Similarly for $su(3)$, the Cartan subalgebra is generated by H_3, Y (resp. λ_3, λ_8), and the rising- and lowering operators are $T_\pm, U_\pm, V_\pm$.

One can also show that

Lemma: one can choose the basis H_i of the Cartan subalgebra such that all $\alpha_i = \alpha(H_i)$ are real (in fact, integers!) for each i.

(this is easy to see: if we choose all H_i to be hermitian, then all eigenvalues of ad_{H_i} are real, because then ad_{H_i} is hermitian w.r.t. the inner product

$$\langle X, Y \rangle := Tr(X^\dagger Y) \tag{74}$$

and can hence be diagonalized with real eigenvalues α_H.)

This is very crucial: it allows to define the "real Cartan subalgebra"

$$\mathfrak{g}_{0,\mathbb{R}} := \langle H_i \rangle_\mathbb{R} \tag{75}$$

as the space of *real* linear combinations of the H_i. For compact Lie algebras, this is what one starts with before the complexification, since the H_i are then (anti)hermitian.

The roots are therefore linear function from $\mathfrak{g}_0$ to $\mathbb{R}$, i.e.

$$\boxed{\alpha \in \mathfrak{g}_{0,\mathbb{R}}^*},$$

This will be understood from now on, and we omit the $\mathbb{R}$.

Now consider the Killing from. We use the basis $\{X_{\alpha_i}, H_i\}$ of $\mathfrak{g}$. Then for any $H \in \mathfrak{g}_0$, the matrix representing ad_H is diagonal,

$$ad_H = diag\left(\alpha_1(H), ..., \alpha_n(H), 0, ..., 0\right).$$

(since $ad_H(H_i) = 0$). Therefore the Killing form is

$$\kappa(H_1, H_2) = \sum_{\alpha \in \Phi} \alpha(H_1)\alpha(H_2)$$

Because the $\alpha(H_1)$ are real as stated above, it follows that

Lemma: The Killing form on the "real" Cartan subalgebra $\mathfrak{g}_{0,\mathbb{R}}$ is real and positive definite.

(again, this is clear for compact groups...)

Note that this is not trivial: for example,

$$\kappa(iH_1, iH_1) = -\kappa(H_1, H_1)$$

The point is that the roots are uniquely defined, and the "real" Cartan subalgebra is defined as the "real" dual of $\langle\alpha\rangle_{\mathbb{R}}$. Therefore

$$\boxed{\mathfrak{g}_{0,\mathbb{R}} \quad \text{is a Euclidean space}}$$

We'll omit the $\mathbb{R}$ in $\mathfrak{g}_{0,\mathbb{R}}$.

Because it is nondegenerate, the Killing form defines an isomorphism

$$\alpha \mapsto H_\alpha$$

between $\mathfrak{g}_0^*$ and $\mathfrak{g}_0$, which is the space of roots, by

$$\alpha(X) = \kappa(H_\alpha, X) \tag{76}$$

And, one also obtains an inner product on $\mathfrak{g}_0^*$ by

$$\langle\alpha, \beta\rangle := \kappa(H_\alpha, H_\beta) \tag{77}$$

which is again Euclidean. Therefore

the space $\mathfrak{g}_0^*$ spanned by the roots is an Euclidean space with inner product $\langle\alpha, \beta\rangle$

Furthermore, one can show that the roots span all of $\mathfrak{g}_0^*$.

8.1.1 Example: $su(3)$.

To illustrate this, consider $su(3)$. The Cartan subalgebra is generated by

$$\mathfrak{g}_0 = \langle\{H_3, Y\}\rangle_{\mathbb{R}} \tag{78}$$

which are orthogonal: using

$$\begin{aligned} (H_3, Y) &= Tr_{\mathbb{C}^3}(H_3 Y) = 0, \\ (H_3, H_3) &= Tr_{\mathbb{C}^3}(\lambda_3\lambda_3) = 2 = (Y, Y) \end{aligned} \tag{79}$$

The eigenvectors of ad_{H_3}, ad_{T_3} are

$$\begin{aligned} &[H_3, T_\pm] = \pm 2T_\pm = \alpha_{T_\pm}(H_3)T_\pm, \quad [Y, T_\pm] = 0 = \alpha_{T_\pm}(Y)Y \\ &[H_3, U_\pm] = \mp U_\pm = \alpha_{U_\pm}(H_3)U_\pm, \quad [Y, U_\pm] = \pm\sqrt{3}U_\pm = \alpha_{U_\pm}(Y)U_\pm \\ &[H_3, V_\pm] = \pm V_\pm = \alpha_{V_\pm}(H_3)V_\pm, \quad [Y, V_\pm] = \pm\sqrt{3}V_\pm = \alpha_{V_\pm}(Y)V_\pm \end{aligned} \tag{80}$$

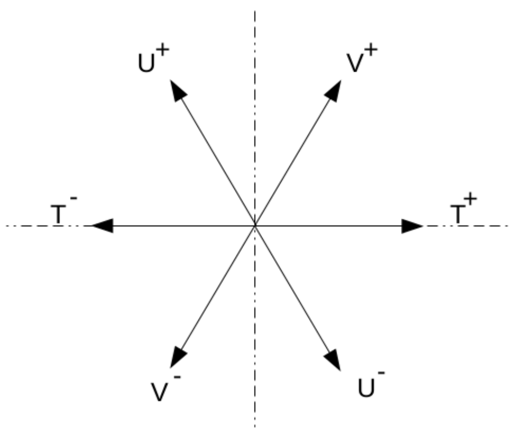
which defines the roots $\alpha_{T_\pm}, \alpha_{U_\pm}, \alpha_{V_\pm}$. One then obtains

$$\begin{aligned} H_{T_\pm} &= \pm H_3, \\ H_{U_\pm} &= \pm\left(-\frac{1}{2}H_3 + \frac{\sqrt{3}}{2}Y\right), \\ H_{V_\pm} &= \pm\left(\frac{1}{2}H_3 + \frac{\sqrt{3}}{2}Y\right) \end{aligned} \tag{81}$$

which satisfy $\alpha_{T_\pm}(H_3) = (H_{T_\pm}, H_3)$ etc. (check) Therefore we can calculate the inner products:

$$\begin{aligned} \langle\alpha_{T_+}, \alpha_{T_+}\rangle &= (H_3, H_3) = 2 = \langle\alpha_{U_+}, \alpha_{U_+}\rangle = \langle\alpha_{V_+}, \alpha_{V_+}\rangle, \\ \langle\alpha_{T_+}, \alpha_{U_+}\rangle &= (H_3, -\frac{1}{2}H_3 + \frac{\sqrt{3}}{2}Y) = -1 = -\langle\alpha_{T_+}, \alpha_{V_+}\rangle, \\ \langle\alpha_{U_+}, \alpha_{V_+}\rangle &= (-\frac{1}{2}H_3 + \frac{\sqrt{3}}{2}Y, \frac{1}{2}H_3 + \frac{\sqrt{3}}{2}Y) = -\frac{1}{2} + \frac{3}{2} = 1 \end{aligned} \tag{82}$$

etc. Since $cos(60) = 1/2$, it follows that all the angles are 60^o. Hence the 6 roots form a regular hexagon:



(also include Y and H_3).

8.1.2 Commutation relations for a Cartan-Weyl basis

We proceed with the general analysis. For $H \in \mathfrak{g}_0$, consider

$$\begin{aligned}
[H,[X_\alpha,X_\beta]] &= -[X_\alpha,[X_\beta,H]]-[X_\beta,[H,X_\alpha]] \\
&= \beta(H)[X_\alpha,X_\beta]+\alpha(H)[X_\alpha,X_\beta] \\
&= (\beta(H)+\alpha(H))[X_\alpha,X_\beta] \qquad (83)
\end{aligned}$$

This means that either $[X_\alpha, X_\beta] = 0$, or it is a root vector corresponding to the root $\alpha+\beta$. It is not hard to show (exercise) that if $\alpha+\beta$ is a root, then in fact

$$\boxed{0 \neq [X_\alpha, X_\beta] = e_{\alpha\beta}\, X_{\alpha+\beta}}$$

If $\beta = -\alpha$, it follows that $[H, [X_\alpha, X_{-\alpha}]] = 0$ hence

$$[X_\alpha, X_{-\alpha}] \in \mathfrak{g}_0$$

is in the Cartan subalgebra. Moreover it is nonzero, since

$$(H,[X_\alpha,X_{-\alpha}]) = ([H,X_\alpha],X_{-\alpha}) = \alpha(H)(X_\alpha,X_{-\alpha}) = (H,(X_\alpha,X_{-\alpha})H_\alpha) \qquad (84)$$

since $\alpha(H) = (H, H_\alpha)$ (76), for all $H \in \mathfrak{g}_0$. Therefore[3]

$$\boxed{[X_\alpha, X_{-\alpha}] = (X_\alpha, X_{-\alpha})H_\alpha}$$

Collecting these results, we constructed a “Cartan-Weyl” basis $\{H_i, X_\alpha\}$ consisting of Cartan generators and root vectors (one for each root α), with commutation relations

$$\begin{aligned}
[H_i,H_j] &= 0 \\
[H_i,X_\alpha] &= \alpha(H_i)X_\alpha \\
[X_\alpha,X_\beta] &= e_{\alpha\beta}\,X_{\alpha+\beta}, && \alpha+\beta\in\Phi \\
[X_\alpha,X_{-\alpha}] &= (X_\alpha,X_{-\alpha})\,H_\alpha, \\
[X_\alpha,X_\beta] &= 0, && \alpha+\beta\notin\Phi, \alpha+\beta\neq 0 \qquad (85)
\end{aligned}$$

8.2 Useful concepts for representations: Roots and Weights

Now consider a representation V of $\mathfrak{g}$. We choose a basis of V which are common eigenvectors of the Cartan subalgebra H_i:

$$H_i|\lambda_i; j\rangle = \lambda_i|\lambda_i; j\rangle$$

[3]From these considerations it follows easily that the root spaces $\mathfrak{g}_\alpha$ are one-dimensional, since otherwise one could extend the Cartan subalgebra.

where the extra index j denotes possible other indices (for degeneracies). Of course it follows $(\sum c_i H_i)|\lambda_i; j\rangle = (\sum_i c_i\lambda_i)|\lambda_i; j\rangle$, hence the common eigenstates really define a linear functional λ on $\mathfrak{g}$, just like the roots: if we define $\lambda(H) := \sum_i c_i\lambda_i$, then

$$H|\lambda\rangle = \lambda(H)|\lambda\rangle$$

(omitting j). These $\lambda \in \mathfrak{g}_0^*$ are called *weights* of the representation V. Note that both roots α and weights λ are in $\mathfrak{g}_0^*$. This makes sense: recalling that $ad_H(X_\alpha) = \alpha(H)X_\alpha$, we see that

the roots are just the weights of the adjoint representation

The space

$$V_\lambda := \{|\lambda\rangle \in V; H|\lambda\rangle = \lambda(H)|\lambda\rangle\}$$

of eigenvectors in V with given weight λ is called the *weight space* of V with weight λ. An element $v_\lambda \in V_\lambda$ is called a *weight vector* with weight λ.

The crucial point is now the following:

$$\begin{aligned} HX_\alpha|\lambda\rangle &= (X_\alpha H + \alpha(H)X_\alpha)|\lambda\rangle \\ &= (\lambda(H) + \alpha(H))X_\alpha|\lambda\rangle \end{aligned} \tag{86}$$

This means that $X_\alpha|\lambda\rangle \cong |\lambda + \alpha\rangle$ has weight $\lambda + \alpha$. Similarly $X_{-\alpha}$ lowers the weights by α, hence

the root vectors X_α relate weight vectors with weights differing by α. They are therefore called raising and lowering operators.

Example: For the 3-dim. representation of $su(3)$ defined by the Gell-Mann matrices, there are 3 different weights, see before. We have already calculated the eigenvalues in (64).

$$v_{\lambda_1} = |1, \frac{1}{\sqrt{3}}\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_{\lambda_2} = |-1, \frac{1}{\sqrt{3}}\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_{\lambda_3} = |0, -\frac{2}{\sqrt{3}}\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{87}$$

One can easily check using (69) that their weights are

$$\begin{aligned} \lambda_1 &= \frac{2}{3}\alpha_{T_+} + \frac{1}{3}\alpha_{U_+}, \\ \lambda_2 &= -\frac{1}{3}\alpha_{T_+} + \frac{1}{3}\alpha_{U_+}, \\ \lambda_3 &= -\frac{1}{3}\alpha_{T_+} - \frac{2}{3}\alpha_{U_+}, \end{aligned}$$

They form an equilateral triangle.

8.2.1 $su(2)$ strings and the Weyl group

The first important thing to note is that for any root α, the generators $X_\alpha, X_{-\alpha}, H_a$ define a $su(2)$ subalgebra:

$$\begin{aligned}
[H_\alpha, X_\alpha] &= \alpha(H_\alpha) X_\alpha = \langle \alpha, \alpha \rangle X_\alpha \\
[H_\alpha, X_{-\alpha}] &= -\langle \alpha, \alpha \rangle X_{-\alpha} \\
[X_\alpha, X_{-\alpha}] &= (X_\alpha, X_{-\alpha})\ H_\alpha,
\end{aligned}$$

since $\alpha(H_\alpha) = (H_\alpha, H_\alpha) = \langle \alpha, \alpha \rangle$. Hence for any pair $\alpha, -\alpha$ of roots there is a corresponding $su(2)_\alpha$. We have seen this explicitly for $su(3)$. These different $su(2)$ are entangled in a nontrivial way.

Choose some root α, and normalize $X_{\pm\alpha}$ such that $(X_\alpha, X_{-\alpha}) = 1$. Then

$$[X_\alpha, X_{-\alpha}] = H_\alpha \tag{88}$$

This $su(2)$ algebra generated by $X_\alpha, X_{-\alpha}, H_a$ acts on V, which therefore decomposes into a direct sum of irreps. Pick one such irrep. Its weights consist of "weight strings" $\mu, \mu - \alpha, ..., \mu - q\alpha$ for some integer q. We want to determine q from μ. Repeating our analysis of the $su(2)$ representations , we define

$$v_j := X_{-\alpha}^j v_0 \tag{89}$$

where v_0 is the (highest) weight vector in that string with weight μ. Then q is determined by

$$X_{-\alpha} v_q = 0$$

since there is no lower weight.

Using

$$\mu(H_\alpha) = (H_\mu, H_\alpha) = \langle \mu, \alpha \rangle$$

we get

$$\begin{aligned}
X_\alpha v_k =: r_k v_{k-1} &= X_\alpha X_{-\alpha} v_{k-1} \\
&= (X_{-\alpha} X_\alpha + H_\alpha) v_{k-1} \\
&= r_{k-1} v_{k-1} + (\mu(H_\alpha) - (k-1)\alpha(H_\alpha)) v_{k-1} \\
&= (r_{k-1} v_{k-1} + \langle \mu, \alpha \rangle - (k-1)\langle \alpha, \alpha \rangle) v_{k-1}
\end{aligned}$$

The solution or this recursion relation with $r_0 = 0$ is

$$r_k = k\langle \mu, \alpha \rangle - \frac{1}{2} k(k-1) \langle \alpha, \alpha \rangle$$

Since $r_{q+1} = 0$, we get

$$q = \frac{2\langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle} \tag{90}$$

In practice, we often have a weight λ which may or may not be the highest weight in this weight string $\lambda + p\alpha, ..., \lambda, ..., \lambda - m\alpha$. Since the above formula says

$$q = m + p = \frac{2\langle \lambda + p\alpha, \alpha \rangle}{\langle \alpha, \alpha \rangle}$$

it follows

$$\boxed{m - p = \tfrac{2\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle} \quad \in \{-q, -q+2, ..., q-2, q\}} \tag{91}$$

This is a very useful formula: it tells us where λ may lie in a $su(2)$ string.

For example, consider the 3-dim. rep. of $su(3)$. Consider the α_{T_+} string (horizontal) through $v_{\lambda_1} = |1, \frac{1}{\sqrt{3}}\rangle$. We get

$$m - p = \frac{2\langle (\frac{2}{3}\alpha_{T_+} + \frac{1}{3}\alpha_{U_+}), \alpha_{T_+} \rangle}{\langle \alpha_{T_+}, \alpha_{T_+} \rangle} = 1.$$

Indeed, $p = 0$ and $m = 1$.

Therefore there is a reflection symmetry in such a string, reflecting along α through the hyperplane $H_\alpha = \{x;\ \langle x, \alpha \rangle = 0\}$:

(Picture)

This reflection acts on weights as follows:

$$\boxed{S_\alpha : \lambda \mapsto \lambda - \tfrac{2\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle}\alpha = \lambda - (m-p)\alpha} \tag{92}$$

because $\frac{\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle}$ is the projection of λ on α. This reflection is called a *Weyl reflection*. This construction applies to any root α in any representation.

Now consider the action of such a $su(2)_\alpha$ on any irrep V of $\mathfrak{g}$. It follows that V decomposes into the direct sum of irreps of this $su(2)_\alpha$. By the above argument, all these irreps (or strings) of $su(2)_\alpha$ have the Weyl reflection symmetry described by the same formula (92). This means that all these $su(2)_\alpha$ strings are symmetric i.e. preserved under S_α.

(Example: ad of $su(3)$.)

The various Weyl reflections (in weight space) can be combined, and generate a group called the *Weyl group* of $\mathfrak{g}$. The Weyl group maps weights into weights, for any representation.

(Exercise: determine the Weyl group for $su(3)$, and $su(2)$).

The Weyl group is very useful to understand the structure of the irreps, and to calculate their characters. This will be exploited further below.

8.3 More on the structure of the adjoint representation

Consider again the $su(2)_\alpha$ subgroups, and the α-strings they generate in the adjoint representation. Using (91) twice for roots $\alpha \neq \beta$, we get

$$\cos^2(\theta) := \frac{\langle\alpha,\beta\rangle\langle\alpha,\beta\rangle}{\langle\alpha,\alpha\rangle\langle\beta,\beta\rangle} = \frac{1}{4}mn \tag{93}$$

for roots $\alpha \neq \beta$, where m and n are integers given by appropriate values $m'-p'$. The lhs is $\cos^2(\theta)$ where θ is the angle between the vectors α and β. Using the Cauchy-Schwarz inequality, it follows that $\frac{1}{4}mn \leq 1$. Moreover it follows that $|m| \leq 3$ and $|n| \leq 3$, since $\alpha \neq \beta$ and using the fact that if α is a root, then $k\alpha$ can be a root only for $k = \pm 1$. It follows that

In the adjoint representation, a string can have no more than 4 weights in it.

which has far-reaching consequences. Furthermore,

$\cos(\theta)$ *can only take the values* $0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$, *and* $|m|, |n| \leq 3$.

This means that the possible angles between roots are multiples of $30^o, 45^o$. This will severely restrict the possible Lie algebras: it will imply that every (simple) Lie algebra is one of the classical matrix algebras $su(n), so(n), sp(n)$ or one of 5 “exceptional” Lie algebras first enumerated by Killing.

8.4 Simple roots and the Cartan matrix

The next step is to define an ordering among the roots. This is somewhat arbitrary, but nevertheless very useful.

Let $\{\alpha_1, ..., \alpha_r\} \subset \Phi$ be an ordered set of linearly independent roots (note that r is the rank, hence this is only a subset of all roots) such that every element $\lambda \in \mathfrak{g}_0^*$ can be written uniquely as $\lambda = \sum c_i\alpha_i$. We then call λ *positive* (denoted by $\lambda > 0$) if the first non-vanishing c_i is positive, i.e. $c_1 > 0$, or $c_1 = 0, c_2 > 0$, etc. If the first non-zero c_i is negative, we call λ negative. This is a refined version of choosing some hyperplane in $\mathfrak{g}_{0,\mathbb{R}}^*$ which divides the space in positive and negative vectors.

The set of positive λ is closed under addition, and we can write $\lambda > \mu$ if $\lambda - \mu > 0$. In particular, we can now say which roots are positive and which are negative.

A *simple root* is a positive root which cannot be written as the sum of two positive roots.

For example, consider $su(3)$. The roots are $\alpha_{T_+}, \alpha_{U_+}, \alpha_{V_+}$ and their negatives. We note that $\alpha_{V_+} = \alpha_{T_+} + \alpha_{U_+}$. If we choose $(\alpha_{T_+}, \alpha_{U_+})$ as basis, then these are simple roots while α_{V_+} is not. But we can also choose $(\alpha_{T_+}, \alpha_{V_+})$ as basis. Then $\alpha_{U_+} = -\alpha_{T_+} + \alpha_{V_+}$, and the positive roots are $\alpha_{T_+}, -\alpha_{U_+}, \alpha_{V_+}$ and the simple roots are $-\alpha_{U_+}$ and α_{V_+}, with $-\alpha_{U_+} > \alpha_{V_+}$.

Denote the set of simple roots with $\Phi_s \subset \Phi$. One very important property of the simple roots is that the difference of two simple roots is not a root at all:

$$\alpha, \beta \in \Phi_s \Rightarrow \alpha - \beta \notin \Phi$$

Assume to the contrary that $\alpha - \beta \in \Phi$. Then either $\alpha - \beta$ is positive or $\beta - \alpha$ is positive. Thus either $\alpha = (\alpha - \beta) + \beta$ or $\beta = (\beta - \alpha) + \alpha$ can be written as the sum of two positive roots, in contradiction with the definition.

Furthermore,

$$\langle \alpha, \beta \rangle \leq 0 \quad if \quad \alpha, \beta \in \Phi_s \tag{94}$$

This follows from (91) applied to the adjoint representation, because $\beta =: \lambda$ is a root, but $\beta - \alpha$ is not (recall that the roots are the weights in the adjoint). Thus $m = 0$ in (91), so $m - p \leq 0$.

This implies that the simple roots are linearly independent. If they were linearly dependent, we could write

$$\sum_{\alpha_i \in \Phi_s} a_i \alpha_i = \sum_{\alpha_i \in \Phi_s} b_i \alpha_i \tag{95}$$

where $a_i \geq 0$ and $b_i \geq 0$ and $a_i b_i = 0$. (There cannot be a relation $\sum_{\alpha_i \in \Phi_s} c_i \alpha_i = 0$ with all $c_i \geq 0$, since the simple roots α_i are all positive). Now multiplying both sides with $\sum_{\alpha_i \in \Phi_s} a_i \alpha_i$ we get

$$\langle \sum_i a_i \alpha_i, \sum_i a_i \alpha_i \rangle = \langle \sum_i b_i \alpha_i, \sum_i a_i \alpha_i \rangle$$

The lhs is positive because it is a square, but the rhs is negative by (94). Hence both sides must be zero, establishing the linear independence of the simple roots. Hence we can take as a basis of roots (as before) the simple roots. (one can also show that they are complete).

The crucial property is now the following:

every positive root can be written as a positive (integer!) sum of simple roots

This is certainly true for the positive roots which happen to be simple. Now consider the smallest positive root for which this is not true. Since it is not simple, it can be written as sum of two positive roots. But these are smaller than their sum, hence each can by hypothesis be written as positive sum of simple roots. QED.

Furthermore, note that every root is either positive or negative, and every negative root can be written as -(positive sum of simple roots).

From the simple roots we can form the *Cartan matrix*, which summarizes all properties of the simple Lie algebra to which it corresponds. Let r be the rank of $\mathfrak{g}$, i.e. $r = dim(\mathfrak{g}_0) = dim(\mathfrak{g}_0^*)$. Then the Cartan matrix is the $r \times r$ matrix

$$A_{ij} = 2\frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle} \tag{96}$$

where α_i are the simple roots.

The diagonal elements are all equal to 2. The matrix is not necessarily symmetric, but if $A_{ij} \neq 0$ then $A_{ji} \neq 0$. In fact, we have shown before that the only possible values for the off-diagonal elements are $0, \pm 1, \pm 2, \pm 3$ (because the length of any $su(2)$-string is at most 4). Actually they can only be $0, -1, -2, -3$ using (94).

There is more: we have

$$\langle \alpha_i, \alpha_j \rangle^2 \leq \langle \alpha_i, \alpha_i \rangle \langle \alpha_j, \alpha_j \rangle$$

where the inequality is strict unless α_i and α_j are proportional. This cannot happen for $i \neq j$ since the simple roots are linearly independent. Hence

$$A_{ij} A_{ji} < 4. \tag{97}$$

i.e.. if $A_{ij} = -2$ or -3, then $A_{ji} = -1$.

Consider $su(3)$ as an example. We take the basis

$$(\alpha_1, \alpha_2) = (\alpha_{T_+}, \alpha_{U_+})$$

Since then $\alpha_{V_+} = \alpha_1 + \alpha_2$, the simple roots are also α_1, α_2. We already computed the relevant scalar products:

$$\begin{aligned} \langle \alpha_1, \alpha_1 \rangle &= 2, \\ \langle \alpha_1, \alpha_2 \rangle &= -1, \\ \langle \alpha_2, \alpha_2 \rangle &= 2, \end{aligned} \tag{98}$$

Therefore

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \tag{99}$$

The Cartan matrix, together with (91), allows to determine all the roots of a given simple Lie algebra. It is enough to determine the positive roots $\beta = \sum k_i \alpha_i$ which have $k_i \geq 0$. One calls $n := (\sum k_i)$ the *level* of the root β. Hence the simple roots have level one. Assume that we found all roots at the n^{th} level. Then for each root β with level n, we must determine whether $\beta + \alpha_i$ is a root or not.

Since all the roots at level n are known, it is known how far back the root strings $\beta, \beta - \alpha_i, ..., \beta - m\alpha_i$ extends (recall that there is only one X_α for each root α, and that all roots are either a positive or a negative sum of the simple roots). From this, we compute how far the string extends: $\beta, \beta + \alpha_i, ..., \beta + p\alpha_i$. From (92) we define the *Dynkin indices* of β as

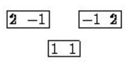
$$n_i(\beta) := 2\frac{\langle \beta, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = \sum_j 2k_j \frac{\langle \alpha_j, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = \sum_j k_j A_{ji} \quad = m - p \tag{100}$$

which take values in $\{-q, ..., q\}$ in the α_i string through β. In particular, $\beta + \alpha_i$ is a root if $p = m - \sum_j k_j A_{ji} > 0$. Also, the α_i-string through α_j has length $-A_{ji} + 1$.

Hence we should keep track of the Dynkin indices (n_i) of the roots β. Then (92) states that in an α_i string, the $i - th$ Dynkin index n_i takes values in some symmetric set $\{-q, -q + 2, ..., +q\}$. This is easy to keep track of, by simply adding the j'th row of the Cartan Matrix whenever the j'th simple root is added to a root. (In particular, the $i - th$ Dynkin index in the α_i string through α_j takes the values $\{A_{ji}, A_{ji} + 2, ..., -A_{ji}\}$).

Consider $su(3)$. We start by writing down the rows of the Cartan matrix, which represent the simple roots α_1, α_2:

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$



Start with α_1, which has Dynkin indices A_{1i}. We ask whether the addition of α_2 produces a root at level 2. (recall that $2\alpha_1$ is not a root). Since the second entry in the box representing α_1 is negative, the corresponding value of p in (100) is positive. The same conclution is reached if beginning with α_2. Is there a root at level 3? Looking back in the α_1 direction, we have $m = 1$. Since the first entry in the box for $\alpha_1 + \alpha_2$ is 1, we have $p = 0$, so this is the end. Similarly for α_2. Hence there are no more positive roots.

To summarize, we can determine the $m - p$ in (100) in this graphical representation, by noting that each string must have the symmwtric form $\{-q, ..., +q\}$.

Other example: exceptional algebra G_2. The corresponding Cartan matrix is

$$\begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$$

[2 −3]   [−1 2]

[1 −1]

[0 1]

[−1 3]

[1 0]

Hence G_2 has 6 positive roots. The set of roots looks as follows:

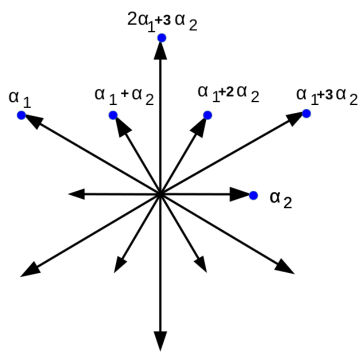


Figure 2: roots of G_2.

The Cartan matrix also contains all the information about the commutation relations:

8.5 The Chevalley-Serre form of the Lie algebra

Lets use the following normalization:

$$\begin{aligned} X_i^+ &:= c_i X_{\alpha_i}, \\ X_i^- &:= c_i X_{-\alpha_i}, \\ H_i &:= \frac{2}{\langle \alpha_i, \alpha_i \rangle} H_{\alpha_i} \end{aligned} \tag{101}$$

where

$$c_i = \sqrt{\frac{2}{(X_{\alpha_i}, X_{-\alpha_i}) \langle \alpha_i, \alpha_i \rangle}} \tag{102}$$

Then the commutation relations (85) of the Cartan-Weyl basis imply e.g. $[H_{\alpha_i}, X_j^+] = \alpha_j(H_{\alpha_i})X_j^+ = \langle\alpha_j, \alpha_i\rangle X_j^+$, hence

$$[H_i, X_j^+] = \frac{2}{\langle\alpha_i, \alpha_i\rangle}\,[H_{\alpha_i}, X_j^+] = \frac{2\langle\alpha_j, \alpha_i\rangle}{\langle\alpha_i, \alpha_i\rangle} X_j^+,$$

etc, hence

$$\begin{aligned}
[H_i, H_j] &= 0 \\
[H_i, X_j^\pm] &= \pm A_{ji} X_j^\pm \\
[X_i^+, X_j^-] &= \delta_{i,j}\, H_i,
\end{aligned} \tag{103}$$

The last relation follows because $\alpha_i - \alpha_j$ is not a root. Finally, since the α_i-string through α_j has length $1 - A_{ji}$ (see below (100)) and $X_j^+ \in \mathfrak{g}$ (considered as ad-rep) corresponds to the root α_j, it follows that

$$(ad_{X_i^\pm})^{1-A_{ji}} X_j^\pm = 0 \tag{104}$$

These commutation relations are called the *Chevalley-Serre* commutation relations. One can now follow the above algorithm to find the roots α and construct the corresponding X_α at the same time. For example,

$$X_{\alpha_1+\alpha_2} = [X_{\alpha_1}^+, X_{\alpha_2}^+],$$

etc. One can show that all root vectors X_α can be obtained in this way via commutators of the $X_i^\pm$. This means that it is enough to work with the simple roots and the corresponding $X_i^\pm$, which is a big simplification.

Therefore the Cartan matrix contains all the information necessary to reconstruct the full Lie algebra. Its content can be summarized in a very elegant and useful way using the so-called *Dynkin diagram*:

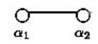
8.6 Dynkin Diagrams

The Dynkin diagram of a (semi-simple) Lie algebra is constructed as follows: For every simple root α_i, make a dot. We will see that the length of the simple roots $\langle\alpha_i, \alpha_i\rangle$ can take at most 2 different values. Hence one makes a dark dot for the short roots, and an "empty" dot for the long ones. Now one connects the i^{th} and the j^{th} dot with a number of straight lines equal to $A_{ij}A_{ji}$, which can be either 1,2 or 3 (indicating the angle between the simple roots). It turns out that for simple algebras one obtains a connected graph, while for direct sums one obtains disconnected graphs (since the roots are orthogonal).

The Dynkin diagram for $su(2)$ is just a point (the Cartan matrix being (2)).

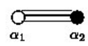
The Cartan matrix and the Dynkin diagram for $\mathfrak{su}(3)$ is

$$A_{ij} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$



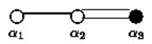
The Cartan matrix and the Dynkin diagram for the exceptional Lie algebra G_2 is

$$G_2 \qquad \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$$



The dark dot in G_2 corresponds to the second root α_2, which follows from the asymmetric Cartan matrix.

To illustrate the relation between the Cartan matrix and Dynkin diagrams, consider the Dynkin Diagram for B_3.



It follows that $A_{31} = A_{13} = 0$. Since one line connects the first and the second point, we must have $A_{21} = A_{12} = -1$. The second and third points are related by 2 lines, hence $A_{23}A_{32} = 2$. Since the third root is smaller that the second, it must be $A_{23} = -2$ and $A_{32} = -1$. Thus

$$A_{ij} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -2 \\ 0 & -1 & 2 \end{pmatrix} \tag{105}$$

We will determine all possible Dynkin diagrams below, hence all possible simple Lie algebras.

9 The classical Lie algebras

Let us apply these considerations to the classical Lie algebras $su(n), so(n)$ and $sp(n)$. Start with

$\underline{su(n)}$:

Recall that $su(n) = Lie(SU(n)) = \{A \in Mat(n,\mathbb{C});\ A^\dagger = -A = 0, Tr(A) = 0\}$. Recall that we always work with the complexified Lie algebra $su(n)_\mathbb{C}$, which is denoted now by

$$A_{n-1} = su(n)_\mathbb{C}$$

where $n-1$ is the rank as we'll see. (Incidentally, A_{n-1} coincides with $sl(n)_\mathbb{C}$. Hence these are the same if complexified, i.e. the structure of their Lie algebras is the same. Then $su(n)$ and $sl(n,\mathbb{R})$ are different real sectors of A_{n-1}.)

A basis of $A_{n-1} = su(n)_\mathbb{C}$ is given by all

$$e_{ab} = (e_{ab})_{ij} = \delta_{ia}\delta_{jb}, \qquad a \neq b$$

(due to complexification!) and of $n-1$ elements

$$H = \sum c_i e_{ii}, \quad \sum c_i = 0.$$

Alltogether these are n^2-1 independent generators (over $\mathbb{C}$), as it should be (this is the same as for the real Lie group $su(n)$). We easily check the commutation relations

$$[e_{ab}, e_{cd}] = e_{ad}\delta_{bc} - e_{cb}\delta_{ad} \tag{106}$$

and in particular

$$[H, e_{ab}] = (c_a - c_b)e_{ab} \tag{107}$$

Thus e_{ab} is a root vector corresponding to the root $\alpha(H = \sum c_i e_{ii}) = (c_a - c_b)$.

Let us choose a basis for the root space:

$$\begin{aligned}
\alpha_1(\sum c_i e_{ii}) &= c_1 - c_2,\\
\alpha_2(\sum c_i e_{ii}) &= c_2 - c_3,\\
&\dots\\
\alpha_{n-1}(\sum c_i e_{ii}) &= c_{n-1} - c_n,
\end{aligned}$$

and declare these positive roots with ordering $\alpha_1 > ... > \alpha_{n-1}$. It is easy to see that these are just the simple roots.

Next we need the Killing-form applied to the Cartan subalgebra. It is

$$
\begin{aligned}
(\sum c_i e_{ii}, \sum c'_j e_{jj}) &= Tr(ad_{\sum c_i e_{ii}} ad_{\sum c'_j e_{jj}}) \\
&= \sum_{a,b} (c_a - c_b)(c'_a - c'_b) \\
&= 2n \sum_a c_a c'_a
\end{aligned}
\tag{108}
$$

In particular, $(e_{ii}, e_{jj}) = 2n\delta_{i,j}$. This gives the duality between α_i and H_{α_i}:

$$
(H_{\alpha_i}, \sum c_j e_{jj}) = \alpha_i(\sum c_j e_{jj}) = c_i - c_{i+1} \tag{109}
$$

Therefore

$$
H_{\alpha_i} = \frac{1}{2n}(e_{ii} - e_{i+1\,i+1}) \tag{110}
$$

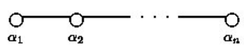
and

$$
\langle \alpha_i, \alpha_j \rangle = \frac{1}{2n}(2\delta_{ij} - \delta_{i,j+1} - \delta_{i+1,j}) \tag{111}
$$

Therefore the Cartan Matrix is

$$
A_n : \quad \begin{pmatrix}
2 & -1 & 0 & & & \\
-1 & 2 & -1 & & & \\
0 & -1 & 2 & & & \\
 & & & 2 & -1 & 0 \\
 & & & -1 & 2 & -1 \\
 & & & 0 & -1 & 2
\end{pmatrix} \tag{112}
$$

and the Dynkin Diagram is



<u>$sp(2n)$</u>:

The symplectic group $Sp(2n)$ is defined as

$$
Sp(2n) = \{A \in Mat(2n);\; A^T J A = J\} \tag{113}
$$

where

$$
J = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} \tag{114}
$$

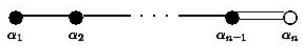
The corresponding Lie algebra is obtained from $\exp(X) \approx 1 + X + o(X^2)$, which implies

$$
C_n := sp(2n) = \{X \in Mat(2n)\;\; X^T = JXJ\}. \tag{115}
$$

(complexified). Again one chooses a suitable basis of $sp(2n)$, a Cartan subalgebra, calculates the Killing form etc. It turns out that the rank is n, hence there are n simple roots $\alpha_1, ..., \alpha_n$ One finds the Cartan matrix

$$C_n : \begin{pmatrix} 2 & -1 & 0 & & & \\ -1 & 2 & -1 & & & \\ 0 & -1 & 2 & & & \\ & & & 2 & -1 & 0 \\ & & & -1 & 2 & -1 \\ & & & 0 & -2 & 2 \end{pmatrix} \qquad (116)$$

note that the difference is only in the last line. The corresponding Dynkin diagram is



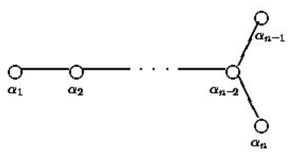
For the orthogonal groups, one must distinguish between even and odd dimensions:

$D_n := so(2n)$:

This has rank n, with Cartan matrix

$$D_n : \begin{pmatrix} 2 & -1 & 0 & & & & \\ -1 & 2 & -1 & & & & \\ 0 & -1 & 2 & & & & \\ & & & 2 & -1 & 0 & 0 \\ & & & -1 & 2 & -1 & -1 \\ & & & 0 & -1 & 2 & 0 \\ & & & 0 & -1 & 0 & 2 \end{pmatrix} \qquad (117)$$

and Dynkin diagram

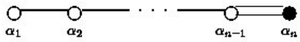


$B_n := so(2n+1)$:

This has also rank n, with Cartan matrix

$$B_n: \begin{pmatrix} 2 & -1 & 0 & & & \\ -1 & 2 & -1 & & & \\ 0 & -1 & 2 & & & \\ & & & 2 & -1 & 0 \\ & & & -1 & 2 & -2 \\ & & & 0 & -1 & 2 \end{pmatrix} \tag{118}$$

and Dynkin diagram



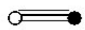
Note the similarity between B_n and C_n, which differ only by interchanging the last off-diagonal elements. This amounts to reversing the shading of the dots.

Note also that the Dynkin diagram for $D_2 = so(4)$ consists of 2 disconencted points, and the Cartan matrix is $2\mathbb{1}_{2\times 2}$. This corresponds (and in fact shows!) to the fact that $so(4) \cong su(2) \times su(2)$. Similarly, the Dynkin diagram for $so(6)$ is the same as the one for $su(4)$, which again implies that $so(6) \cong su(4)$. This indicates their usefulness!

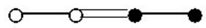
10 The exceptional Lie algebras

Surprisingly, there are only five other simple Lie algebras besides the series A_n, B_n, C_n and D_n. This amounts to a complete classification of the simple Lie algebras. These additional ones are the following:

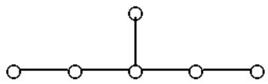
G_2



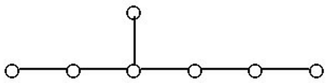
F_4



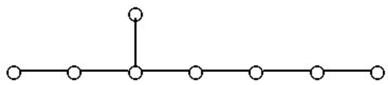
E_6



E_7



E_8



This can be proved by considering sets of vectors $\gamma_i \in \mathfrak{g}_0^*$ (candidates for simple roots) and defining the associated matrix

$$M_{ij} = 2\frac{\langle\gamma_i, \gamma_j\rangle}{\langle\gamma_j, \gamma_j\rangle} \tag{119}$$

and an associated diagram (analogous to the Dynkin diagram) where the ith and the jth points are joined by $M_{ij}M_{ji}$ lines. The set $\{\gamma_i\}$ is called *allowable (zulässig)* if all the following conditions hold:

- the γ_i are linearly independent, i.e. $det(M) \neq 0$
- $M_{ij} \leq 0$ if $i \neq j$
- $M_{ij}M_{ji} \in \{0, 1, 2, 3\}$

(recall that this is satisfied for simple roots). One can now show that the only allowable diagrams have the above form, by systematically analyzing these conditions in terms of the diagrams. The details can be found e.g. in the book [Cahn].

11 Representation theory II

11.1 Fundamental weights, Dynkin labels and highest weight representations

Recall that in a representation V of $\mathfrak{g}$, the generators become operators acting on V, and in particular one can choose a basis $|\lambda_i; j\rangle$ of V which are common eigenvectors of the Cartan subalgebra $\mathfrak{g}_0$: $H_i|\lambda_i; j\rangle = \lambda_i|\lambda_i; j\rangle$ or in general

$$H|\lambda\rangle = \lambda(H)|\lambda\rangle$$

(omitting j). Moreover, the root vectors X_α relate weight vectors with weights differing by α,

$$X_\alpha|\lambda\rangle \cong |\lambda + \alpha\rangle.$$

This can be used to find all the weights of a given (finite-dimensional) irreducible representation (irrep) V.

First, we can again consider V as representation of any of the $su(2)_\alpha$ subalgebras. Then by (91), the weights again come in strings $\lambda + p\alpha, ..., \lambda, ..., \lambda - m\alpha$ with

$$m - p = \frac{2\langle\lambda, \alpha\rangle}{\langle\alpha, \alpha\rangle}. \tag{120}$$

in particular there are no "holes". For each weight λ of V, one defines the *Dynkin labels (coefficients)*

$$\lambda^i := \frac{2\langle\lambda, \alpha_i\rangle}{\langle\alpha_i, \alpha_i\rangle} \in \mathbb{Z}, \qquad i = 1, 2, ..., r \tag{121}$$

where α_i are the simple roots. These are convenient "coordinates" of the weights $\lambda \in \mathfrak{g}_0^*$. For finite-dimensional representations , they are always integers.

Furthermore, recall that we introduced an ordering relation among the elements of $\mathfrak{g}_0^*$, i.e. among the weights. Therefore among the weight of V, there is a maximal one: call it μ. Since μ is the maximal (or "highest") weight, $\mu + \alpha_i$ is not a weight of V. Therefore by (120), its Dynkin labels are non-negative integers:

$$\mu^i := \frac{2\langle\mu, \alpha_i\rangle}{\langle\alpha_i, \alpha_i\rangle} \in \mathbb{Z}_{\geq 0} \tag{122}$$

are positive integers. Conversely, one can show that for each weight μ with positive Dynkin labels $\mu^i \geq 0$ there is a corresponding (unique!) irrep $V = V_\mu$ with highest weight μ, and furthermore that the corresponding weight space with weight μ in V is one-dimensional, i.e. there is only one vector $|\mu\rangle \in V$. To summarize, *Any (finite-dim) irrep $V = V_\mu$ is characterized by its highest weight μ, which has non-negative Dynkin labels $\mu^i \geq 0$. It is called highest-weight module with highest weight μ.*

This can be made more transparent by introducing the *fundamental weights* $\Lambda_{(i)} \in \mathfrak{g}_0^*, \quad i = 1, 2, ..., r$ by the requirement

$$\frac{2\langle \Lambda_{(i)}, \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle} := \delta_{i,j} \tag{123}$$

They are just another basis of weight space $\mathfrak{g}_0^*$. Then we can expand any weight in terms of this new basis,

$$\lambda = \sum_i \lambda^i \Lambda_{(i)} \tag{124}$$

Multiplying this with $\frac{2\langle ., \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle}$, we see that the λ^i are just the Dynkin labels of λ:

$$\frac{2\langle \lambda, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = \frac{2\langle \lambda^j \Lambda_{(j)}, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = \lambda^i. \tag{125}$$

Since all Dynkin labels of weights in V are integers, it follows that

all weights of a (finite-dimensional) irrep V are integral linear combinations of the fundamental weights.

In other words, all weights live in the *weight lattice*

$$L_w := \{z_i \Lambda_{(i)} \;\; z_i \in \mathbb{Z}\} \tag{126}$$

In particular, this applies to the highest weight of V.

<u>example: $su(2)$</u>

Let α be the (only) positive root of $su(2)$. Then $\Lambda_{(1)} = \frac{1}{2}\alpha$, and the spin $j \in \frac{1}{2}\mathbb{Z}$ irrep is the highest weight rep with h.w. $(2j)\Lambda_{(1)}$. Indeed, its weights are then $(2j)\Lambda_{(1)}, 2j\Lambda_{(1)} - \alpha, ..., (-2j)\Lambda_{(1)}$.

<u>example: $su(3)$</u>

(draw picture).

The set of allowed highest weights (=positive integer Dynkin labels) hence forms a cone of a lattice in weight space (draw picture).

For each given V_μ, we can now calculate all the weights similar as we calculated the roots (=weights of the adjoint rep). Given a weight λ of V, we need to determine whether

$\lambda - \alpha_j$ is also a weight. Starting with the highest weight and keeping track of their Dynkin labels $(\lambda^i)_{i=1,\dots,r}$, we know the value of p in (120). Hence if

$$m_j = p_j + \lambda^j > 0, \qquad (127)$$

then $\lambda - \alpha_j$ is a weight in V. We obtain the Dynkin labels of $\lambda - \alpha_j$ by subtracting $(A_{ji})_i$ from λ^i, because

$$\frac{2\langle \lambda - \alpha_j, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = \frac{2\langle \lambda, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} - \frac{2\langle \alpha_j, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle}.$$

Hence the Dynkin labels λ_j of a weight λ are just the $m_j - p_j$ of the corresponding α_j string, and go from $(n_i, n_i - 2, \dots, -n_i + 2, -n_i)$.

11.1.1 Examples

$(1, 0)$ of $A_2 = su(3)$

Let us determine the weights of the irrep with highest weight $\mu = \Lambda_{(1)} = (1, 0)$:

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$\boxed{1 \;\; 0}$

$\boxed{-1 \;\; 1}$

$\boxed{0 \;\; -1}$

(make also drawing of weights...)

This is the 3-dimensional representation corresponding to the Quanks u, d, s.

What about the irrep with highest weight $\mu = \Lambda_{(2)} = (0, 1)$? These are the Anti-Quarks:

(make also drawing of weights...)

Actually we didn't really show that these reps are 3-dim, we only showed that the only possible weights are as above. In general it happens that there are more than one weight vector for a given weight. This "multiplicity" must be determined by other means, e.g. using Weyls character formula (later).

Next, consider the irrep with highest weight $\mu = (1, 1)$:

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$\boxed{1\ \ 1}$

$\boxed{-1\ \ 2}$ $\quad$ $\boxed{2\ \ -1}$

$\boxed{0\ \ 0}$

$\boxed{1\ \ -2}$ $\quad$ $\boxed{-2\ \ 1}$

$\boxed{-1\ \ -1}$

This is clearly the adjoint representation, which is 8-dimensional. The weight with Dynkin labels $(0,0)$ occurs twice, and corresponds to the 2 Cartan generators Y and T_3 which have weight 0.

All the irreps of $su(3)$ can be determined in this way, e.g. the decouplet of baryons etc etc. In particular, this gives all possible “families” of particles consisting of these 3 Quarks.

<u>an irrep of G_2</u>

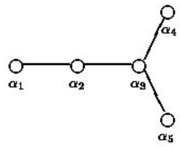
Next, consider the rep of G_2 with highest weight $(0,1)$: recall the Cartan matrix is

$$\begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$$

and the weights are (...). This is a 7-dimensional representation, which looks like the adjoint of $SU(3)$ except that the weight 0 has only multiplicity 1. Before the η meson was discovered, it was thought that G_2 might be a useful symmetry of the strong interaction, with this 7-dimensional representation describing the $\pi^+, \pi^0, \pi^-, K^+, K^0$ and $\bar{K^0}, K^-$ mesons. But it didn't work out.

<u>$SO(10)$</u>

As a more complicated example, consider the rep of $SO(10)$ (=D_5) with highest weight $(1,0,0,0,0)$:



We have then the schemes:

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

$\boxed{1\ 0\ 0\ 0\ 0}$

$\boxed{-1\ 1\ 0\ 0\ 0}$

$\boxed{0\ -1\ 1\ 0\ 0}$

$\boxed{0\ 0\ -1\ 1\ 1}$

$\boxed{0\ 0\ 0\ -1\ 1}$ $\boxed{0\ 0\ 0\ 1\ -1}$

$\boxed{0\ 0\ 1\ -1\ -1}$

$\boxed{0\ 1\ -1\ 0\ 0}$

$\boxed{1\ -1\ 0\ 0\ 0}$

$\boxed{-1\ 0\ 0\ 0\ 0}$

constructing more representations : tensor products

11.2 Tensor products II

Recall that if V_1 and V_2 are 2 representations of the Lie algebra $\mathfrak{g}$, then so is $V_1 \otimes V_2$ by

$$\begin{aligned} \pi : \mathfrak{g} &\rightarrow gl(V_1 \otimes V_2) = gl(V_1) \otimes gl(V_2) \\ \mathfrak{g} &\mapsto \pi_{V_1}(\mathfrak{g}) \otimes \mathbb{1} + \mathbb{1} \otimes \pi_{V_2}(\mathfrak{g}) \end{aligned} \quad (128)$$

For example, adding angular momenta in Quantum Mechanics: $J_i = L_i + S_i$, etc. Also, the Hilbert space of systems of several particles (e.g. baryons, mesons consisting of Quarks) is the tensor product of the individual Hilbert spaces.

Since all (finite-dim. ...) reps are completely reducible, we have

$$V_{\mu_1} \otimes V_{\mu_2} = \bigoplus_{\mu_3} N^{\mu_3}_{\mu_1 \mu_2} V_{\mu_3} \quad (129)$$

where $N^{\mu_3}_{\mu_1 \mu_2} \in \mathbb{N}$ are the *multiplicities* (Littlewood-Richards coefficients), and V_{μ_1} etc are the highest weight irreps. For $su(2)$, they're 0 or 1. One of the goals is to determine this decomposition explicitly.

Now consider the weights of the elements: if $v_1 \in V_1$ and $v_2 \in V_2$ are weight vectors with weights λ_1 resp. λ_2, then $v_1 \otimes v_2 \in V_1 \otimes V_2$ has weight $\lambda_1 + \lambda_2$. Hence all weights in $V_1 \otimes V_2$ have the form $\lambda_1 + \lambda_2$ for weights $\lambda_{1,2}$ in $V_{1,2}$, and

$$V_1 \otimes V_2 = \sum_{v_\lambda \in V_1, v'_\lambda \in V_2} v_\lambda \otimes v'_\lambda.$$

Therefore if the highest weights of V_1 resp V_2 are μ_1 resp. μ_2, then the highest weight in $V_1 \otimes V_2$ is $\mu_1 + \mu_2$, and

$$V_{\mu_1} \otimes V_{\mu_2} = V_{\mu_1+\mu_2} \oplus \Big(\bigoplus_{\mu_3 < \mu_1+\mu_2} N'^{\mu_3}_{\mu_1\mu_2}\, V_{\mu_3} \Big) \tag{130}$$

In principle, one can now proceed by finding the space $V_{\mu_1+\mu_2}$ starting with its highest weight vector, then finding the highest weight vector and its irrep in the orthogonal complement $V^{\perp}_{\mu_1+\mu_2} \subset V_{\mu_1} \otimes V_{\mu_2}$, etc. However this is practical only for $su(2)$ or small representations of other groups, and there are more powerful methods. The most powerful method is by using the characters together with (150). This is basically a formalization of the above method, and combined with Weyls character formula (later) this gives a systematic allgoithm known as Racah - Speiser algoithm, cf. [**?**]. However for small reps, one can proceed in a pedestrian way by matching the weights and their multiplicities, as above.

11.2.1 Clebsch-Gordon Coefficients

We consider only $su(2)$ here for simplicity, but everything generalizes to other Lie algebras.

Consider again the tensor product $V_{j_1} \otimes V_{j_2}$ of 2 irreps with spin j_1 resp. j_2. One basis is given by the vectors

$$|j_1 j_2; m_1 m_2\rangle = |j_1; m_1\rangle |j_2; m_2\rangle.$$

which are eigenvectors of $\pi_{j_1}(H) \otimes 1\!\!1$ and $1\!\!1 \otimes \pi_{j_2}(H)$.

Because $V_{j_1} \otimes V_{j_2} = \bigoplus_{j_3} N^{j_3}_{j_1 j_2}\, V_{j_3}$, there is another basis given by the eigenvectors of $\pi_{j_1}(H) \otimes 1\!\!1 + 1\!\!1 \otimes \pi_{j_2}(H)$ in some component $V_{j_3} \subset V_{j_1} \otimes V_{j_2}$, which we denote by $|j_1, j_2; j_3, m_3\rangle$. Therefore we have

$$|j_1, j_2; j_3, m_3\rangle = \sum_{m_1 m_2} |j_1, j_2; m_1, m_2\rangle \langle j_1, j_2; m_1 m_2 | j_3, m_3\rangle \tag{131}$$

The coefficients

$$\langle j_1, j_2; m_1, m_2 | j_3, m_3\rangle \equiv \langle j_1, j_2; m_1, m_2 | j_1, j_2; j_3, m_3\rangle \tag{132}$$

are called *Clebsch-Gordan Coefficients*. They determine the unitary transformation between the 2 ONB's of $V_{j_1} \otimes V_{j_2} = \bigoplus_{j_3} N_{j_1 j_2}^{j_3} V_{j_3}$, hence their phases are just a convention. They can be calculated by finding the space $V_{j_1+j_2}$ starting with the highest weight sector of $V_{j_1+j_2}$, then finding the highest weight submodule in the orthogonal complement $V_{j_1+j_2}^{\perp} \subset V_{j_1} \otimes V_{j_2}$, etc. See QMI. One finds the multiplicities

$$\begin{aligned} N_{j_1 j_2}^{j_3} &= 1 \quad |j_1 - j_2| \leq j_3 \leq j_1 + j_2, \\ N_{j_1 j_2}^{j_3} &= 0 \quad \text{otherwise.} \end{aligned}$$

Sometimes one uses the notation

$$C_{m_1,m_2,m_3}^{j_1,j_2,j_3} = \langle j_1, j_2; m_1, m_2 | j, m \rangle \tag{133}$$

or the so-called *Wigner 3j-symbols*

$$\langle j_1, j_2; m_1, m_2 | j, m \rangle = \frac{(-1)^{j_1 - j_2 + m}}{2j+1} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m \end{pmatrix} \tag{134}$$

which are invariant under cyclic permutations of the 3 columns.

Example: $su(3)$

To understand the multiplet structure of mesons etc, we must consider e.g. $(3) \otimes (\overline{3})$.

The 3-dimensional representation of the Quarks is $(3) = (1, 0) = V_{\Lambda_{(1)}}$, and the representation of the Antiquarks is the other 3-dimensional representation$(\overline{3}) = (0, 1) = V_{\Lambda_{(2)}}$. Therefore the bound states of a Quark and an Antiquark are described by the tensor product of the constituent Hilbert spaces, $(3) \otimes (\overline{3})$ (in obvious notation...). I.e.

$$|meson; \lambda\rangle = \sum C_{\lambda_1,\lambda_2,\lambda} |Quark; \lambda_1\rangle \otimes |Antiquark; \lambda_2\rangle \qquad \in V_{\Lambda_{(1)}} \otimes V_{\Lambda_{(2)}}.$$

This describes the mesons. Similarly, baryons are obtained from triple tensor products $(3) \otimes (3) \otimes (3)$. To understand the multiplet structure of mesons etc, we therefore must find the decomposition of that tensor product into irreps.

According to the above results, $(3) \otimes (\overline{3})$ contains $V_{\Lambda_{(1)}+\Lambda_{(2)}} = (1, 1) = (8)$, which is the 8-dim. adjoint rep. It follows that

$$(3) \otimes (\overline{3}) = (8) \oplus (1).$$

This is the reason why there are octets of mesons, which are bound states of 1 Quark and 1 Antiquark! (the (1) indeed exists, and is the η' meson).

This is easy to understand graphically:

(draw weight graphs)

Often this method allows to find the complete decomposition, by counting dimensions as well. These can be computed by Weyls dimension formula.

11.2.2 (Anti)symmetric tensor products

Now consider the tensor product of two identical representations , $V \otimes V$. This can be decomposed into two parts,

$$V \otimes V = (V \otimes_S V) \oplus (V \otimes_{AS} V) \tag{135}$$

the symmetric tensor product with basis $v_i \otimes v_j + v_j \otimes v_i$, and the antisymmetric part with basis $v_i \otimes v_j - v_j \otimes v_i$ (cp. identical particles in Quantum Mechanics!!). Notice that these subspaces are invariant (preserved) under the action of $\mathfrak{g}$, therefore they are either irreducible or themselves reducible (both is possible).

If μ is the highest weight of V, then clearly $v_\mu \otimes v_\mu \in V \otimes_S V$ is in the symmetric part, therefore the highest weight of $V \otimes_S V$ is 2μ. The anti-symmetric part does not contain this weight, rather its highest weight is the sum of the highest and the next-to-highest weights of V.

For example, consider $(3) \otimes (3)$ of $su(3)$. The symmetric tensor product contains the weight $(2,0)$, which is the highest weight of the (6) of $su(3)$. The next-to-highest weight in $(1,0)$ is $(-1,1)$, therefore the antisymmetric part contains the highest weight representation with h.w. $(0,1)$, which is $(\overline{3})$. The dimensions add up to 9, hence this is it:

$$(3) \otimes (3) = (6) \oplus (\overline{3})$$

In general there are more than 2 components! (There are no such particles, because they cannot be color singlets).

This procedure can be generalized to $V^{\otimes n}$, e.g. considering the n-fold totally symmetric or totally antisymmetric tensor product (identical particles!). The totally symmetric part always contains the h.w. rep. with highest weight $n\mu$. For example, we surely have $V_{n\Lambda_{(1)}} \subset V_{\Lambda_{(1)}}^{\otimes_S n}$, and we will see that in fact for $su(n)$ one has

$$V_{n\Lambda_{(1)}} = V_{\Lambda_{(1)}}^{\otimes_S n} \tag{136}$$

In fact there are also components with "mixed" symmetry; this can be studied systematically and leads to the method of *Young Diagrams*, which give the complete decomposition of $V^{\otimes n}$ for $su(n)$. However this is not complete for other groups, and the method doesn't work for general $V \otimes W$. We will describe a generaly method which always works.

Remark: One can show that for $su(n)$, one can produce ANY rep as part of $V_{\Lambda_{(1)}}^{\otimes n}$. Therefore $V_{\Lambda_{(1)}}$ is called the *fundamental* representation.

11.3 An application of $SU(3)$: the three-dimensional harmonic oscillator

cp. [Georgi, chapter 14]

The Hamiltonian for the three-dimensional harmonic oscillator is

$$H = \frac{\vec{P}^2}{2m} + \frac{1}{2}m\omega^2\vec{x}^2 = \hbar\omega(a_k^+ a_k + \frac{3}{2}) \tag{137}$$

where

$$a_k = \sqrt{\frac{m\omega}{2\hbar}}x_k + \frac{i}{\sqrt{2m\hbar\omega}}p_k, \qquad a_k^+ = \sqrt{\frac{m\omega}{2\hbar}}x_k - \frac{i}{\sqrt{2m\hbar\omega}}p_k, \quad k = 1,2,3. \tag{138}$$

The a_k^+ resp. a_k are rising-resp. lowering operators, satisfying

$$\begin{aligned} [a_k, a_l^+] &= \delta_{kl}, \\ [a_k^+ a_k, a_l^+] &= a_l^+ \, \delta_{kl}, \\ [a_k^+ a_k, a_l] &= -a_l \, \delta_{kl} \end{aligned}$$

Let $|0\rangle$ be the ground state, which satisfies

$$a_k|0\rangle = 0 \tag{139}$$

Then the eigenstates of the Hamiltonian, i.e. the energy eigenstates, are

$$\mathcal{H}_n := a_{k_1}^+ ... a_{k_n}^+ |0\rangle \tag{140}$$

with energy $\hbar\omega(n+\frac{3}{2})$. However, except for the ground state these are degenerate energy levels, because any choice of n generators out of the $\{a_1^+, a_2^+, a_3^+\}$ gives the same energy. This degeneracy is - as usual! - due to a symmetry, which is $su(3)$ here:

Consider the operators

$$Q_a := a_k^+ (T_a) a_l \tag{141}$$

where $T_a = \frac{1}{2}\lambda_a$ are the Gell-Mann matrices of $su(3)$. Because they are a basis of $su(3)$, they satisfy

$$[T_a, T_b] = i f_{abc} T_c \tag{142}$$

where f_{abc} are the structure constants of $su(3)$ (could be calculated explicitly, but we don't need this). One can easily check that the Q_a satisfy the same commutation relations:

$$[Q_a, Q_b] = i f_{abc} Q_c, \tag{143}$$

hence they define a representation of $su(3)$ on the Hilbert space of the 3d harmon. osc. Moreover,

$$[Q_a, H] = 0, \tag{144}$$

thus we can fix an energy eigenvalue $E_n = \hbar\omega(n+\frac{3}{2})$, and the Q_a act on this subspace of the Hilbert space. This means that the energy eigenstates are representations of $su(3)$, and will decompose into the direct sum of irreps. In fact, they become a single irrep ("multiplet"), see below.

For example, the ground state is a singlet since

$$Q_a|0\rangle = 0 \tag{145}$$

The precise action of $su(3)$ on $\mathcal{H}_n$ can be found noting that

$$[Q_a, a_k^+] = a_l^+(T_a)_{lk} \tag{146}$$

hence a_k^+ transforms like a (3). Therefore $a_{k_1}^+ ... a_{k_n}^+|0\rangle$ transforms like $(3) \otimes ... \otimes (3)$, and in fact $\mathcal{H}_n \subset (3)^{\otimes sn}$ since the a_k^+ commute. But since $(3)^{\otimes sn} = V_{n\Lambda_{(1)}} = (n, 0)$, it follows that

$$\mathcal{H}_n = V_{n\Lambda_{(1)}} = (n, 0).$$

Therefore the degenerate energy levels are precisely the $(n, 0)$ irrep of $su(3)$!!

Of course the harmonic oscillator also has a $SO(3)$ symmetry. But this is weaker, since $SO(3) \subset SU(3)$, and the irrep $V_{n\Lambda_{(1)}}$ decomposes into several irreps of $so(3)$.

This generalizes immediately to n dimensions, i.e. a system with n creation- and anihilation operators naturally has a $su(n)$ symmetry.

11.4 The character of a representation and Weyls character formula

We encountered so far (at least) 2 main open problems in the context of representations:

1. find the weights including the multiplicities of an irrep
2. find the decomposition of the tensor product of 2 irreps into irreps: for example, to find the bound states of quarks.

There is a powerful theorem which provides the answer (or the basis for it...) for both these problems: Weyls character formula. It also gives a very nice formula for the dimension of highest weight irreps.

The crucial concept is the *character* of a representation. Consider a representation V of a (simple, ...) Lie algebra $\mathfrak{g}$. Let $mult_V(\lambda)$ be the multiplicity of the weight space λ in

V, i.e. the number of lin. independent elements in V with weight λ. Then

$$\chi_V := \sum_\lambda mult_V(\lambda)\ e^\lambda \tag{147}$$

This is to be understood as a *generating function*. The e^λ are "formal", and linearly independent for each λ. One can understand e^λ as function on the weight space, by

$$\mu \mapsto e^\lambda(\mu) := e^{\langle \lambda,\mu\rangle} \tag{148}$$

Knowing χ_V is equivalent to knowing all states and all multiplicities of the representation V, hence in particular its dimension, etc. χ_V is a VERY useful way to encode this information. In terms of a basis $|i\rangle$ with weights λ_i, of the representation V, we can write

$$\chi_V(\mu) = \sum_\lambda mult_V(\lambda)\ e^{\langle\lambda,\mu\rangle} = \sum_i \langle i|i\rangle e^{\langle\lambda_i,\mu\rangle} = \sum_i \langle i|e^{H_\mu}|i\rangle = Tr_V(e^{H_\mu}) \tag{149}$$

One obvious and useful property is the following:

$$\chi_{V\otimes W} = \chi(V)\chi(W) \tag{150}$$

using the usual properties of the exponential $e^{\lambda+\mu} = e^\lambda e^\mu$. Thus the weights of $V \otimes W$ have the form $\lambda + \lambda'$, and can be obtained graphically by superimposing the weights of V on the ones of W, or vice versa.

(draw image)

This is the basis of a method to decompose tensor products using the "method of characters", see later.

For example, consider the spin j representation of $su(2)$. If α is the (only) positive root of $su(2)$, then the weights of the spin $j \in \frac{1}{2}\mathbb{Z}$ irrep are $j\alpha, (j-1)\alpha, ..., -j\alpha$, and its character is

$$\chi_{V_j} = \sum_{m=-j}^{j} e^{m\alpha} = \frac{e^{(j+\frac{1}{2})\alpha} - e^{-(j+\frac{1}{2})\alpha}}{e^{\frac{1}{2}\alpha} - e^{-\frac{1}{2}\alpha}} \tag{151}$$

The last form is Weyls character formula for $su(2)$, which generalizes to other Lie algebras as we will see. The point is that by expanding as sum of exponentials, one gets all the multiplicities.

Another important property is that the character is invariant under the Weyl group. Recall that the Weyl group is generated by the reflections

$$S_\alpha : \lambda \mapsto \lambda - \frac{2\langle\lambda,\alpha\rangle}{\langle\alpha,\alpha\rangle}\alpha \tag{152}$$

which reflects the states (weights) of the $su(2)_\alpha$ strings, and therefore it also preserved the multiplicities: $mult_V(S_\alpha(\lambda)) = mult_V(\lambda)$, therefore

$$mult_V(\omega(\lambda)) = mult_V(\lambda)$$

for any $\omega \in \mathcal{W}$. Extending this action of the Weyl group to $\omega(e^\lambda) := e^{\omega(\lambda)}$, it follows that

$$\chi_V(\omega(\mu)) = \chi_V(\mu) \tag{153}$$

for any $\omega \in \mathcal{W}$. Hence the character is invariant under $\mathcal{W}$. Combining this with several other tools (see later), this property leads to the Weyl character formula for $\chi_\lambda := \chi_{V_\lambda}$ where V_λ is the highest weight irrep with h.w. λ:

$$\boxed{\chi_\lambda(\mu) = \frac{\sum_{\omega\in\mathcal{W}} sign(\omega) e^{\langle\omega(\lambda+\rho),\mu\rangle}}{\sum_{\omega\in\mathcal{W}} sign(\omega) e^{\langle\omega(\rho),\mu\rangle}}} \tag{154}$$

Here $sign(\omega) = det(\omega) = \pm 1$ is the signum of $\omega \in \mathcal{W}$ (i.e. even/odd number of reflections), and

$$\rho := \frac{1}{2}\sum_{\alpha\in\Phi_{>0}} \alpha = \sum_i \Lambda_{(i)} \tag{155}$$

is the *Weyl vector*. (154) is one of the great formulas in Mathematics. Its definition implies that the rhs can always be divided out.

For example, consider the spin j irrep of $\mathfrak{g} = su(2)$, which has highest weight $(2j)\Lambda_{(1)} = j\alpha$. The Weyl group contains only one reflection, and (154) reduces to (151) since $\rho = \frac{1}{2}\alpha$ where α is the (only) positive root of $su(2)$.

For $su(n)$, this can be written in an even more compact form: recall the identification $H_\lambda \leftrightarrow \lambda$, and the explicit realization of the Cartan generators as diagonal matrices

$$H_\lambda = diag(h_1^\lambda, .., h_n^\lambda) \qquad \cong diag(\lambda_i)$$

with $\sum \lambda_i = 0$ (these are *not* Dynkin indices!), e.g. $H_{\alpha_1} \propto diag(1, -1, 0, ..., 0)$ etc. Then the Killing form was essentially $(\lambda, \mu) = Tr(H_\lambda H_\mu) = \sum \lambda_i \mu_i$. Furthermore, one can show that the Weyl group for $SU(n)$ is just S^n, the permutation group of n elements. It acts on λ resp. H_λ as $diag(h_1, .., h_n) \to diag(h'_1, .., h'_n)$ by permuting the elements. Furthermore, the highest weight are in the fundamental Weyl chamber which is now characterized by $h_1 \geq ... \geq h_n$.

Then replacing $(\lambda_i)_{i=1,...,n} \leftrightarrow \lambda$, we get

$$\begin{aligned}\sum_{\omega\in\mathcal{W}} sign(\omega) e^{\langle\omega(\lambda+\rho),\mu\rangle} &= \sum_{\omega\in S_n} (-1)^\omega e^{\sum_i \omega(\lambda+\rho)_i \mu_i} = \sum_{\omega\in S_n} (-1)^\omega e^{(\lambda+\rho)_{\omega(1)}\mu_1} ... e^{(\lambda+\rho)_{\omega(n)}\mu_n} \\ &= \det((e^{(\lambda+\rho)_i \mu_j})_{ij})\end{aligned} \tag{156}$$

Furthermore, one can show that $\rho_i = (m, m-1, ..., -m)$, so that we can rewrite (154) as

$$\boxed{\chi_\lambda(\mu) = \frac{\det(e^{(\lambda_i+n-i)\mu_j})}{\det(e^{(n-i)\mu_j})}} \tag{157}$$

Note further that

$$\det(e^{(n-i)\mu_j}) = \prod_{i<j}(e^{\mu_i} - e^{\mu_j}) = \Delta(e^{\mu_i}) \tag{158}$$

is the Vandermonde-determinant.

Example: character of (8) of $su(3)$.

The Dynkin indices of (8) are $\lambda = (1,1) = \rho$, hence $\lambda_i = (1,0,-1) = \rho_i$. Therefore $\lambda_i + n - i = (3,1,-1)$, and we have

$$\begin{aligned}
\chi_{(8)}(\mu) &= \frac{\det(e^{(3,1,-1)_i(\mu_1,\mu_2,\mu_3)_j})}{\Delta(e^{\mu_i})} = e^{-(\mu_1+\mu_2+\mu_3)}\frac{\det(e^{(4,2,0)_i(\mu_1,\mu_2,\mu_3)_j})}{\Delta(e^{\mu_i})} \\
&= e^{-(\mu_1+\mu_2+\mu_3)}\frac{\Delta(e^{2\mu_i})}{\Delta(e^{\mu_i})} = \prod_{i<j}\frac{(e^{2\mu_i} - e^{2\mu_j})}{(e^{\mu_i} - e^{\mu_j})} \\
&= \prod_{i<j}(e^{\mu_i} + e^{\mu_j}) = (e^{\mu_1} + e^{\mu_2})(e^{\mu_1} + e^{\mu_3})(e^{\mu_2} + e^{\mu_3}) \\
&= e^{2\mu_1+\mu_2} + e^{2\mu_2+\mu_3} + e^{\mu_1+2\mu_2} + e^{2\mu_1+\mu_3} + e^{\mu_1+2\mu_3} + e^{\mu_2+2\mu_3} + 2e^{\mu_1+\mu_2+\mu_3} \\
&= e^{\mu_1-\mu_3} + e^{-\mu_1+\mu_2} + e^{\mu_2-\mu_3} + e^{\mu_1-\mu_2} + e^{-\mu_2+\mu_3} + e^{-\mu_1+\mu_3} + 2e^0 \\
&= \sum_\alpha e^{(\alpha,\mu)} + 2e^0
\end{aligned} \tag{159}$$

since $\mu_1 + \mu_2 + \mu_3 = 0$. We see that it is 8-dimensional, with the exponents

Application: tensor product for $su(2)$ As a first application, we derive the decomposition of the tensor product

$$V_j \otimes V_k = N_{jk}^l V_l$$

where V_j is the spin j irrep of $su(2)$.

Recall that the weights of V_j are

$$V_k = \langle v_\lambda \rangle, \quad \lambda \in \{-k\alpha, ..., (k-1)\alpha, k\alpha\}$$

where α is the root of $su(2)$. Therefore

$$\begin{aligned}
\chi_{V_j \otimes V_k} &= \chi_{V_j} \left(\sum_{m=-k}^{k} e^{m\alpha} \right) = \left(\frac{e^{(j+\frac{1}{2})\alpha} - e^{-(j+\frac{1}{2})\alpha}}{e^{\frac{1}{2}\alpha} - e^{-\frac{1}{2}\alpha}} \right) \left(\sum_{m=-k}^{k} e^{m\alpha} \right) \\
&= \left(\frac{1}{e^{\frac{1}{2}\alpha} - e^{-\frac{1}{2}\alpha}} \right) \left(\sum_{m=-k}^{k} e^{(j+\frac{1}{2})\alpha} e^{m\alpha} - e^{-(j+\frac{1}{2})\alpha} e^{m\alpha} \right) \\
&= \left(\frac{1}{e^{\frac{1}{2}\alpha} - e^{-\frac{1}{2}\alpha}} \right) \left(\sum_{m=-k}^{k} e^{(j+\frac{1}{2})\alpha + m\alpha} - e^{-(j+\frac{1}{2})\alpha - m\alpha} \right).
\end{aligned}$$

reversing the summation index in the second term. Now assume that $k \leq j$. Then the arguments of the exponentials are $\{(j+k+\frac{1}{2}), ..., (j-k+\frac{1}{2})\}$ for the first term, and $\{-(j+k+\frac{1}{2}), ..., -(j-k+\frac{1}{2})\}$ for the first term. Therefore the rhs can be rewritten as

$$\chi_{V_j \otimes V_k} = \sum_{l=j-k}^{j+k} \frac{e^{(l+\frac{1}{2})\alpha} - e^{-(l+\frac{1}{2})\alpha}}{e^{\frac{1}{2}\alpha} - e^{-\frac{1}{2}\alpha}} = \sum_{l=j-k}^{j+k} \chi_{V_l} \tag{160}$$

This implies that

$$V_j \otimes V_k = \oplus_{l=j-k}^{j+k} V_l.$$

(the reason is that the characters of different irreps are always linearly independent. This is easy to understand, since the "leading exponential" for different highest weigths is different. (In a sense, the characters even form a ONB ...)).

This shows the power of the characters. The same idea works for the other simple Lie algebras, but we need some better understanding of the geometry involved.

Note also that if we assumed $k > j$, then there would be some cancellations before we can identify the irreducible characters. This will happen also below.

11.4.1 Some properties of the Weyl group, Weyl chambers.

We need some more technical background. First, any element $\omega \in \mathcal{W}$ preserves the Killing form:

$$\langle S_\alpha \mu, S_\alpha \nu \rangle = \langle \mu - \frac{2\langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle} \alpha, \nu - \frac{2\langle \nu, \alpha \rangle}{\langle \alpha, \alpha \rangle} \alpha \rangle = \langle \mu, \nu \rangle \tag{161}$$

(this is quite clear, since any reflection preserves the inner product on an Euclidean space).

There is a nice graphical way to understand the Weyl group. Consider again the hyperplanes $H_\alpha = \{x;\ \langle x, \alpha \rangle = 0\}$ introduced before. They divide weight space into cones, the so-called *Weyl chambers*. Clearly any $\omega \in \mathcal{W}$ maps any Weyl chamber to another one. There is one which is a special Weyl chamber, the *fundamental Weyl chamber* which is

$$P_+ = \{\lambda = \sum c_i \Lambda_{(i)}, \quad c_i \geq 0\} \tag{162}$$

Note that the walls of P_+ are the hyperplanes H_{α_i}, since $c_i = 0 \Leftrightarrow (\lambda, \alpha_i) = 0$. Notice also that the highest weights of irreps are precisely the weights in the lattice $L_w = \{z_i \Lambda_{(i)}\}$ which are in P_+!

(Picture for $su(3)$.)

Now consider for any given weight μ in some Weyl chamber, and the orbit of $\mathcal{W}$ acting on μ, i.e. the set of all weights $\omega(\mu)$. It is easy to see that there is one among them which lies in this fundamental Weyl chamber (consider the weight μ^* the orbit which is maximal. If some Dynkin coefficient $2\langle\mu^*, \alpha_i\rangle/\langle\alpha_i, \alpha_i\rangle < 0$, then $S_i\mu^* = \mu^* - 2\alpha_i\langle\mu^*, \alpha_i\rangle/\langle\alpha_i, \alpha_i\rangle$ is an even higher weight, in contradiction to maximality of μ^*). It follows that all Weyl chambers are congruent, since they are mapped by some $\omega \in \mathcal{W}$ to P_+. One can show that $\mathcal{W}$ acts freely and transitively on the Weyl chambers.

Moreover, since the Weyl group maps weights into weights, it also maps the roots into roots. For any fixed simple root α_i, consider $S_i := S_{\alpha_i}$; the following holds:

1. $S_i(\alpha_i) = -\alpha_i$ (obviously)
2. $S_i(\beta)$ is a positive root if $\beta \neq \alpha_i$ is a positive root

That is, S_i interchanges all the positive roots except for α_i itself. (To see 2., write

$$\beta = \sum k_i \alpha_i$$

Then

$$S_1(\beta) = \sum k_i\alpha_i - 2\alpha_1 \sum k_i \frac{\langle\alpha_i, \alpha_1\rangle}{\langle\alpha_1, \alpha_1\rangle} = \sum_{i>1} k_i\alpha_i + \alpha_1(\textit{something}) \tag{163}$$

ans similar for the other S_i. Therefore if $\beta \neq \alpha_1$, then β has SOME positive coefficients if expanded into the simple roots, which implies that all coefficients are positive so that $S_{\alpha_1}(\beta)$ is a positive root.)

It follows that

$$S_{\alpha_i}\rho = \rho - \alpha_i.$$

Therefore

$$\begin{aligned} \langle\rho - \alpha_i, \alpha_i\rangle &= \langle\rho, -\alpha_i\rangle, \\ 2\langle\rho, \alpha_i\rangle &= \langle\alpha_i, \alpha_i\rangle \end{aligned}$$

which implies that

$$\rho = \sum_i \Lambda_{(i)} \tag{164}$$

Consider now the function

$$Q(\mu) := \prod_{\alpha>0} (e^{\frac{1}{2}\langle\alpha,\mu\rangle} - e^{-\frac{1}{2}\langle\alpha,\mu\rangle}) \tag{165}$$

We want to see how this transforms under the Weyl group. Consider

$$Q(S_i\mu) = \prod_{\alpha>0} (e^{\frac{1}{2}\langle\alpha,S_i\mu\rangle} - e^{-\frac{1}{2}\langle\alpha,S_i\mu\rangle}) = \prod_{\alpha>0} (e^{\frac{1}{2}\langle S_i\alpha,\mu\rangle} - e^{-\frac{1}{2}\langle S_i\alpha,\mu\rangle}) \tag{166}$$

Now S_i interchanges all the positive roots except itself, whose sign it changes. It follows that

$$Q(S_i\mu) = -Q(\mu) \tag{167}$$

Since $\mathcal{W}$ is generated by the S_i and $sign(S_i) = -1$, it follows immediately that

$$Q(\omega(\mu)) = sign(\omega)Q(\mu) \tag{168}$$

for any $\omega \in \mathcal{W}$. Hence $Q(\mu)$ is a totally antisymmetric function under the Weyl group. There is another totally antisymmetric function under the Weyl group, given by

$$\tilde{Q}(\mu) := \sum_{\omega\in\mathcal{W}} sign(\omega) e^{\langle\omega(\rho),\mu\rangle} \tag{169}$$

and we claim that they coincide:

$$Q(\mu) = \prod_{\alpha>0} (e^{\frac{1}{2}\langle\alpha,\mu\rangle} - e^{-\frac{1}{2}\langle\alpha,\mu\rangle}) = \sum_{\omega\in\mathcal{W}} sign(\omega) e^{\langle\omega(\rho),\mu\rangle} \tag{170}$$

Since the rhs is the denominator of (154)), this is the so-called *denominator identity*, which we need to derive *Weyls dimension formula*.

To see (170), we expand the lhs into a sum of the form

$$Q(\mu) = \sum_\beta c_\beta e^{\langle\rho-\beta,\mu\rangle}$$

where β is a sum of distinct positive roots. Since both Q and $\tilde{Q}$ are antisymmetric and since $\mathcal{W}$ freely permutes the Weyl chambers, it is enough to show that the terms in (170) where $\rho - \beta$ resp. $\omega(\rho)$ lie in P_+ coincide (can show: terms on the boundary cancel). Using (164) $\rho = \sum_i \Lambda_{(i)}$, the only possible term is $\beta = 0$. Comparing the coefficients, (170) follows.

11.5 Weyls dimension formula

Consider the highest weight irrep $V = V_\lambda$ with h.w. λ. The dimension of V is given by $Tr_V(1) = Tr_V(e^0) = \chi_V(0)$. However, this gives $\frac{0}{0}$, and we must use a suitable limit. We choose $\mu = t\rho$ and let $t \to 0$. This gives using (154)

$$\begin{aligned} \chi_\lambda(t\rho) &= \frac{\sum_{\omega\in\mathcal{W}} sign(\omega) e^{\langle\omega(\lambda+\rho),t\rho\rangle}}{\sum_{\omega\in\mathcal{W}} sign(\omega) e^{\langle\omega(\rho),t\rho\rangle}} = \frac{Q(t(\lambda+\rho))}{Q(t\rho)} \\ &= e^{-\langle\rho,t\lambda\rangle} \prod_{\alpha>0} \frac{e^{\langle\alpha,t(\lambda+\rho)\rangle}-1}{e^{\langle\alpha,t\rho\rangle}-1} \end{aligned} \qquad (171)$$

Now we can take the limit $t \to 0$ and find

$$\boxed{dim(V_\lambda) = \prod_{\alpha>0} \frac{\langle\alpha,\lambda+\rho\rangle}{\langle\alpha,\rho\rangle}} \qquad (172)$$

This is a very useful formula. To evaluate it, we write each positive root in terms of the simple roots:

$$\alpha = \sum_i k^i_\alpha \alpha_i$$

Suppose

$$\lambda = \sum n_i \Lambda_{(i)}$$

for $n_i \geq 0$. Then

$$dim(V_\lambda) = \prod_{\alpha>0} \frac{\sum k^i_\alpha (n_i+1)\langle\alpha_i,\alpha_i\rangle}{\sum k^i_\alpha \langle\alpha_i,\alpha_i\rangle} \qquad (173)$$

using $\langle\alpha_i, \Lambda_{(j)}\rangle = \frac{1}{2}\delta_{ij}\langle\alpha_i,\alpha_i\rangle$. For A_n, D_n, E_n all the simple roots have the same size, so that

$$dim(V_\lambda) = \prod_{\alpha>0} \frac{\sum k^i_\alpha (n_i+1)}{\sum k^i_\alpha}, \qquad \mathfrak{g} \in \{A_n, D_n, E_n\}. \qquad (174)$$

Consider some examples. For $su(2)$, there is only one positive root α, and the spin j rep has highest weight $n = 2j$. Hence we get $dim(V_{spinj}) = (2j+1)$, which is correct.

Consider now $su(3) = A_2$. The positive roots are $\alpha_1, \alpha_2, \alpha_1+\alpha_2$. Then the dimension of the highest weight representation with highest weight $\lambda = n_1\Lambda_{(1)} + n_2\Lambda_{(2)}$ has dimension

$$dim(V_\lambda) = \left(\frac{n_1+1}{1}\right)\left(\frac{n_2+1}{1}\right)\left(\frac{n_1+n_2+2}{2}\right) \qquad (175)$$

For example,

$$\begin{aligned} dim(V_{(1,0)}) &= 3 = dim(V_{(0,1)}), \\ dim(V_{(1,1)}) &= 8, \\ dim(V_{(2,0)}) &= 6 = dim(V_{(0,2)}), \\ dim(V_{(2,1)}) &= 15 = dim(V_{(1,2)}), \\ dim(V_{(3,0)}) &= 10 = dim(V_{(0,3)}), \end{aligned}$$

etc.

Similarly, one easily finds the following formula for the dimension of any $V_{(m_1,...,m_n)}$ of $A_n = su(n+1)$:

$$
\begin{aligned}
dimV_{(m_1,...,m_n)} = \quad & \frac{m_1+1}{1}\frac{m_2+1}{1}...\frac{m_n+1}{1} \\
& \frac{m_1+m_2+2}{2}\frac{m_2+m_3+2}{2}...\frac{m_{n-1}+m_n+2}{2} \\
& ... \\
& \frac{m_1+m_2+...+m_n+n}{n}
\end{aligned}
$$

Comment: here one sees the relation with Young Diagrams. The irrep with Dynkin indices $(m_1, ..., m_n)$ corresponds to a Young tableau with m_k columns of k boxes (since $(1, 0...)$ is the fundamental, and $(0, 1, 0, ...)$ is the 2-fold antisymmetric product of the fundamental, etc).

In particular, for $su(3)$ we have $dim(V_{(n,0)}) = \frac{1}{2}(n+1)(n+2)$, which is the dimension of the Fock space $\mathcal{H}_N := a^+_{k_1}...a^+_{k_n}|0\rangle$ with 3 different (species of) creation- and anihilation operators we encountered with the 3-dim. harmonic oscillator. The latter is of course $\mathcal{H}_n :\cong V_{\Lambda_{(1)}}^{\otimes_S n}$, and is therefore an irrep of $su(3)$. Hence the energy levels of the 3-dim. harmonic oscillators are precisely $(n, 0)$ irreps of $su(3)$.

Similarly one can show that the m-particle (bosonic) Fock space with $n+1$ different (species of) creation- and anihilation operators (which is $\cong V_{\Lambda_{(1)}}^{\otimes_S m}$ for $su(n+1)$) has the same dimension as $V_{(m,0,..,0)}$ of $su(n+1)$, namely

$$dimV_{(m,0,..,0)} = \frac{(m+1)...(m+n)}{12...n} = \begin{pmatrix} m+n \\ n \end{pmatrix} = dimV_{\Lambda_{(1)}}^{\otimes_S m} \tag{176}$$

This implies that $V_{n\Lambda_{(1)}} = V_{\Lambda_{(1)}}^{\otimes_S n}$ as claimed earlier.

Lets consider a more complicated application, for G_2. We have seen that $dim(G_2) = 14$, hence it has 6 positive roots. If the simple roots are denoted by α_1, α_2 with the latter being smaller, The positive roots are $\alpha - 1, \alpha_2, \alpha_1 + a_2, \alpha_1 + 2a_2, \alpha_1 + 3a_2, 2\alpha_1 + 3a_2$. Then one finds (...)

$$dim(V_{(0,1)}) = 7, \qquad dim(V_{(1,0)}) = 14$$

(cp. earlier). The latter is in fact the adjoint representation.

11.6 *Decomposing tensor products: the Racah-Speiser algorithm

Now we generalize the method of decomposing tensor products using Weyls character formula (as used above for $su(2)$) to arbitrary simple $\mathfrak{g}$.

The starting point is again (150),

$$\chi_{V_\lambda \otimes V_\mu} = \chi(V_\lambda) \cdot \chi(V_\mu) = \sum N_{\lambda\mu}^{\nu} \chi(V_\nu) \tag{177}$$

We use again Weyls character formula for $\chi(V_\lambda)$ and $\chi(V_\nu)$. Noting that the denominator of Weyls formula is independent of the representation, we get

$$\sum_{\omega \in \mathcal{W}} sign(\omega) e^{\omega(\lambda+\rho)} \cdot \chi(V_\mu) = \sum_{\omega \in \mathcal{W}} sign(\omega) \sum N_{\lambda\mu}^{\nu} e^{\omega(\nu+\rho)}$$

Plugging in

$$\chi(V_\mu) = \sum_{\kappa} mult_\mu(\kappa)\, e^{\kappa}$$

(recall that $mult_\mu(\kappa)$ is the multiplicity of the weight κ in V_μ), we get

$$\sum_{\omega \in \mathcal{W}} sign(\omega) \sum_{\kappa} mult_\mu(\kappa)\, e^{\omega(\lambda+\rho)} e^{\kappa} = \sum_{\omega \in \mathcal{W}} sign(\omega) \sum N_{\lambda\mu}^{\nu} e^{\omega(\nu+\rho)}$$

Now recall that $mult_\mu(\kappa) = mult_\mu(\omega(\kappa))$, hence we can replace $\sum_\kappa$ by $\sum_{\omega(\kappa)}$, and get

$$\sum_{\omega \in \mathcal{W}} sign(\omega) \sum_{\kappa \in V_\mu} mult_\mu(\kappa)\, e^{\omega(\lambda+\kappa+\rho)} = \sum_{\omega \in \mathcal{W}} sign(\omega) \sum N_{\lambda\mu}^{\nu} e^{\omega(\nu+\rho)}. \tag{178}$$

Now both sides are formal linear combinations of exponentials e^η of some weights η. These e^η are linearly independent for different weights η, and we can compare their coefficients on both sides. Furthermore, recall that for any given η, there is precisely one $\omega \in \mathcal{W}$ such that $\omega(\eta)$ is in the fundamental Weyl chamber P_+ (the orbit $\omega(\eta)$ hits all the different Weyl chambers precisely once). Therefore it is sufficient to compare the terms e^η on both sides with $\eta \in P_+$ in the fundamental Weyl chamber.

Now consider the set of weights

$$\eta := \lambda + \kappa + \rho \qquad \kappa \in V_\mu.$$

Let us assume first that all of these weights η are in the interior of the fundamental Weyl chamber (not on the boundary), this often happens for small μ. Then only the terms with $\omega = id \in \mathcal{W}$ on the lhs give exponentials with weights in the fundamental Weyl chamber, and we can compare their coefficients:

$$mult_\mu(\kappa) e^{\omega(\lambda+\kappa+\rho)} = N_{\lambda\mu}^{\nu} e^{\omega(\nu+\rho)}$$

hence $mult_\mu(\nu - \lambda) = N^\nu_{\lambda\mu}$, or equivalently

$$\boxed{N^{\lambda+\nu}_{\lambda\mu} = mult_\mu(\nu) \qquad \text{if } \lambda + \kappa + \rho \in P^o_+ \ \ \forall \kappa \in V_\mu.} \tag{179}$$

Hence we know $N^{\lambda+\nu}_{\lambda\mu}$ once we know $mult_\mu(\nu)$ (e.g. from Weyls character formula).

This is very easy to understand graphically: for $su(2)$, this is just... the previous result,

(Picture)

Consider another example: for $su(3)$, consider $(8)\otimes(3)$. The weights of (3) are $(1,0)$, $(-1,1)$ and $(0,-1)$. Since

$$\rho = (1,1)$$

and the highest weight of (8) is $(1,1)$, we see that the above condition is satisfied. Therefore

$$V_{(1,1)} \otimes V_{(1,0)} = V_{(2,1)} \oplus V_{(0,2)} \oplus V_{(1,0)}. \tag{180}$$

or

$$(8) \otimes (3) = (15) \oplus (\overline{6}) \oplus (3)$$

which seems to match.

In general, it is not true that all the $\lambda + \kappa + \rho$ for $\kappa \in V_\mu$ are in the fundamental Weyl chamber. Pick one such $\kappa \in V_\mu$. Then there exists precisely one $\omega \in \mathcal{W}$ such that $\omega(\lambda + \kappa + \rho) \in P^o_+$. Then this gives a contribution

$$sign(\omega) e^{\omega(\lambda+\kappa+\rho)} mult_\mu(\kappa) \tag{181}$$

from the lhs, which if summed up must match some $N^\nu_{\lambda\mu} e^{\nu+\rho}$ on the rhs. Therefore the weights are related by

$$\omega(\lambda + \kappa + \rho) = \nu + \rho$$

or

$$\kappa = \omega^{-1}(\nu + \rho) - \rho - \lambda.$$

Therefore

$$\boxed{N^\nu_{\lambda\mu} = \textstyle\sum'_{\omega\in\mathcal{W}} sign(\omega) mult_\mu(\omega^{-1}(\nu+\rho) - \rho - \lambda)} \tag{182}$$

where $\sum'$ denotes the sum over those $\omega \in \mathcal{W}$ such that $\omega^{-1}(\nu + \rho) - \rho - \lambda$ is a weight of V_μ (or set $mult_\mu = 0$ otherwise). This is the general formula.

In practice, to calculate $V_\lambda \otimes V_\mu$ it is easiest to go back to (178). One considers all $\kappa \in V_\mu$, and forms all corresponding

$$\eta := \lambda + \kappa + \rho \qquad \kappa \in V_\mu.$$

Assume that η is not on the boundary of any Weyl chamber. Then there is a unique $\omega \in \mathcal{W}$ such that $\omega(\eta) \in P_+^o$ is inside the fundamental Weyl chamber. In this case we have a contribution of

$$sign(\omega) mult_\mu(\kappa)$$

to the Littlewood-Richards coefficient $N_{\lambda\mu}^\nu$ with $\nu = \omega(\eta) - \rho$. All these contributions have to be summed for all $\kappa \in V_\mu$. This computes all the $N_{\lambda\mu}^\nu$ at once (If $\lambda + \kappa + \rho$ is on the boundary of some Weyl chamber, this κ must be ignored since there is no $\omega \in \mathcal{W}$ which maps it in the interior of the fundamental one).

This is best done by drawing the weights of V_μ on top of $\lambda + \rho$, which gives the relevant κ and their ω with possible sign.

Example 1: $(8) \otimes (8)$

Claim:

$$V_{(1,1)} \otimes V_{(1,1)} = V_{(2,2)} \oplus V_{(3,0)} \oplus V_{(0,3)} \oplus 2V_{(1,1)} \oplus V_{(0,0)} \tag{183}$$

Example 2: $(8) \otimes (6)$

Claim:

$$V_{(1,1)} \otimes V_{(2,0)} = V_{(3,1)} \oplus V_{(1,2)} \oplus V_{(0,1)} \oplus V_{(2,0)} \tag{184}$$

References

[1] B. Cahn, "Semisimple Lie Algebras and their Representations", Frontiers in Physics, Addison-Wesley publishing company, 1984; http://phyweb.lbl.gov/~rncahn/www/liealgebras/texall.pdf

[2] J. Fuchs, C. Schweigert, "Symmetries, Lie Algebras and Representations" Cambridge University Press 1997

[3] W. Fulton J. Harris, "Representation theory" Springer

[4] H. Georgi, "Lie Algebras in Particle Physics". Frontiers in Physics, Addison-Wesley publishing company, 1982

[5] F. Warner, "Foundations of differentiable manifolds and Lie groups". Scott, Foresman and Company, 1971

[6] C. Isham, "Modern Differential Geometry for Physicists"

[7] J. Humphreys, "Introduction to Lie Algebras and Representation Theory". Springer, Graduate Texts in Mathematics 9, 1972

[8] J.F. Cornwell, "Group Theory in Physics I,II,(III?)". Academic Press 1984

[9] H.F. Jones, Groups, Representations, and Physics Bristol