Lie Groups and Lie Algebras for Physicists

Harold Steinacker

Lecture Notes¹, spring 2015
University of Vienna

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

Email: harold.steinacker@univie.ac.at

¹These notes are un-finished and undoubtedly contain many mistakes. They are not intended for publication, but may serve as a hopefully useful resource, keeping in mind these limitations.
6 A first look at representation theory

6.1 Definitions .................................................. 36
6.2 The representation theory of $su(2)$ revisited ................. 39
6.3 The adjoint representation .................................. 41
6.4 The Baker-Campbell-Hausdorff Formula ..................... 44
6.5 Constructing and reducing representations .................. 45

7 $SU(3)$ and Quarks ........................................... 48

7.1 More about Quarks ........................................... 52

8 The structure of simple Lie algebras ............................. 53

8.1 Cartan subalgebras, Roots and the Cartan-Weyl basis .......... 53
  8.1.1 Example: $su(3)$ ........................................ 56
  8.1.2 Commutation relations for a Cartan-Weyl basis .......... 58
8.2 Useful concepts for representations: Roots and Weights ........ 58
  8.2.1 $su(2)$ strings and the Weyl group ................... 60
8.3 More on the structure of the adjoint representation .......... 62
8.4 Simple roots and the Cartan matrix .......................... 62
8.5 The Chevalley-Serre form of the Lie algebra ................. 66
8.6 Dynkin Diagrams ............................................ 67

9 The classical Lie algebras ...................................... 68

10 The exceptional Lie algebras ................................... 72

11 Representation theory II ....................................... 73

  11.1 Fundamental weights, Dynkin labels and highest weight representations . 73
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

\[ \Phi : \mathcal{M} \rightarrow \mathcal{M} \]  

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 $SU(5), SO(8), E_6, \ldots$

  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

\[
\mu : \ G \times G \rightarrow G, \quad (g_1, g_2) \mapsto g_1 \cdot g_2
\]

with the following properties:

1. Associativity: for all \( g_1, g_2 \in G \),

\[
g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3.
\]

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

\[
g \cdot e = e \cdot g = g.
\]

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

\[
g \cdot g^{-1} = g^{-1} \cdot g = e.
\]

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 \( SU(n), SO(n), GL(n), \ldots \).

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 \( \phi : G \rightarrow H \) 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, \( \phi \) is bijective, then \( \phi \) 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 \( \phi : G \rightarrow H \) 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

\[
G \times \mathcal{M} \rightarrow \mathcal{M},
\]

\[(g, \psi) \mapsto g \triangleleft \psi \quad (6)\]

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

\[
\pi : G \rightarrow \text{Map}(\mathcal{M}, \mathcal{M}) \quad (7)
\]

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

Usually one only needs linear transformations, i.e. maps \( \pi : G \rightarrow GL(V) \) 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

\[
\pi : G \rightarrow GL(V)
\]

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 \( G \times V \rightarrow V \) 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

$$SO(3) \times \mathbb{R}^3 \rightarrow \mathbb{R}^3,$$

$$(g, \vec{x}) \mapsto g \cdot \vec{x} \quad (8)$$

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 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 $2 \times 2$ complex matrices which are hermitean and have trace zero,

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

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 $\phi_U$ of $V$ to itself by the formula

$$
\phi : \quad SU(2) \times V \rightarrow V \\
(U, x) \rightarrow \phi_U = UxU^{-1}
$$

Moreover, given $U \in SU(2)$, 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 $\phi_U$ 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 $\pm 1$. Since $SU(2)$ is connected (Exercise), and the map $U \rightarrow \phi_U$ is continuous, $\phi_U$ actually maps into $SO(3)$. Thus

$$
\phi : \quad SU(2) \rightarrow SO(3) \\
U \rightarrow \phi_U
$$

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 $\phi_U$ 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

$$
R(\phi \hat{e}_x) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & \sin(\phi) \\ 0 & -\sin(\phi) & \cos(\phi) \end{pmatrix}, \quad R(\phi \hat{e}_y) := \begin{pmatrix} \cos(\phi) & 0 & -\sin(\phi) \\ 0 & 1 & 0 \\ \sin(\phi) & 0 & \cos(\phi) \end{pmatrix}, \quad \text{und}
$$

$$
R(\phi \hat{e}_z) := \begin{pmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix}.
$$
One can show (exercise!) that rotations around the axis $\vec{\phi}$ by the angle $|\vec{\phi}|$ take the form

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

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

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

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

“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

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

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). 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

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

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

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

One can easily verify that the spin $\frac{1}{2}$ representation of a rotation around $2\pi$ is equal to $-\mathbb{1}$, and rotations around $4\pi$ 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

$$\pi : \quad SU(2) \rightarrow GL(V)$$

$$U = e^{i\vec{\phi} \cdot \vec{J}^{(1/2)}} \mapsto e^{i\vec{\phi} \cdot \vec{J}} \quad (13)$$
This is a group homomorphism! For example, the group homomorphism (11) can be written as

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

However, not every representation of $so(3)$ induces a representation of $SO(3)$: this is prevented by global subtleties (related to rotations around $2\pi$). 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 $\left( \begin{array}{c} p \\ n \end{array} \right)$, 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:

$$I_i \left( \begin{array}{c} p \\ n \end{array} \right) = \frac{1}{2}\sigma_i \left( \begin{array}{c} p \\ n \end{array} \right)$$

etc. That is, a proton is represented by

$$|p\rangle = \left( \begin{array}{c} 1 \\ 0 \end{array} \right) \in \mathbb{C}^2,$$

and a neutron by

$$|n\rangle = \left( \begin{array}{c} 0 \\ 1 \end{array} \right) \in \mathbb{C}^2.$$

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

\[ |I = 1, I_3 = 1\rangle = |p\rangle |p\rangle, \]
\[ |I = 1, I_3 = 0\rangle = \frac{1}{\sqrt{2}} (|p\rangle |n\rangle + |n\rangle |p\rangle), \]
\[ |I = 1, I_3 = -1\rangle = |n\rangle |n\rangle, \]
\[ |I = 0, I_3 = 0\rangle = \frac{1}{\sqrt{2}} (|p\rangle |n\rangle - |n\rangle |p\rangle) \]

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

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 \(\left( \begin{array}{c} u \\ d \end{array} \right)\). Later, a third quark flavor ("strange quarks") was discovered, leading to the extension of the \(SU(2)\) isospin to \(SU(3)\), the famous "eight-fold way" of Gell-Mann et al. 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

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

and the following (real) vector space

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

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

$$\det(x) = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \quad (16)$$

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

$$\phi_U : X \rightarrow X$$

$$x \rightarrow \phi_U(x) := UxU^\dagger \quad (17)$$

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 $\phi_U$ 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

$$\phi : SL(2, \mathbb{C}) \rightarrow O(3, 1)$$

$$U \rightarrow \phi_U$$

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

$\phi(\mathbb{1}) = \mathbb{1}$ and $SL(2, \mathbb{C})$ connected implies that $\phi : SL(2, \mathbb{C})$ is in the component of $SO(3, 1)$ connected to the identity, hence $\phi : SL(2, \mathbb{C}) \rightarrow L^+_\uparrow$, 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 not one-to-one, since for any \( U \in SL(2, \mathbb{C}) \), \( \phi_U = \phi_{-U} \).

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. One should therefore also allow objects in relativistic physics which transform under \( SL(2, \mathbb{C}) \) and not necessarily under \( SO(3, 1) \). The basic such objects are columns

\[
\psi = \begin{pmatrix} \psi^+ \\ \psi^- \end{pmatrix} \in \mathbb{C}^2
\]

with the obvious action

\[
SL(2, \mathbb{C}) \times \mathbb{C}^2 \to \mathbb{C}^2,
(M, \psi) \to M \cdot \psi
\]

These are spinors (Weyl spinors; a Dirac spinor consists of 2 Weyl spinors), which are the “smallest” objects (apart from the trivial ones) which are consistent with special relativity. They describe e.g. neutrinos (if the Lorentz group (resp. \( SL(2, \mathbb{C}) \)) is augmented to the Poincare group, and the spinors can depend on spacetime).

Finally the Poincare group is 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

\[
x^\mu \to \Lambda^\mu_\nu x^\nu + a^\mu.
\] (19)

Accordingly, the group law is given by

\[
(\Lambda, a) \cdot (\Lambda', a') = (\Lambda \Lambda', \Lambda a' + a)
\] (20)

It plays a fundamental role in quantum field theory, but since its 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.
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 	extit{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, $\bigcup U_i = M$ and

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

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

(Picture)

Notation:

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

Definition: 	extit{smooth maps} are

$$C^\infty(M) = \{ f : M \to \mathbb{R}, \ C^\infty \}$$

$$C^\infty(M, N) = \{ \Phi : M \to N, \ C^\infty \}$$

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 	extit{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.

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

which means that

$$X[\lambda f + \mu g] = \lambda X[f] + \mu X[g], \quad f, g \in C^\infty(M)$$

$$X[fg] = f(p)X[g] + g(p)X[f]$$

In particular, it follows immediately that

$$X[c] = 0$$

13
for any constant function \( c \).

**Example:**

Let \( \varphi = \begin{pmatrix} x^1 & \vdots & x^m \end{pmatrix} \) be a coordinate system containing \( p \). Then

\[
X_i := \frac{\partial}{\partial x^i} |_p : f \to \mathbb{R} \quad \text{partial derivative at } p
\]

i.e.

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

**Theorem 7**

\[
T_p(M) = \langle \frac{\partial}{\partial x^i} \rangle_{\mathbb{R}} \cong \mathbb{R}^m \quad \text{is a } m\text{-dimensional vector space}
\]

i.e. a general tangent vector has the form

\[
X_p = \sum a_i \frac{\partial}{\partial x^i} |_p, \quad a_i \in \mathbb{R}
\]

The point is that they are *first-order differential operators*, not higher-order ones. This is encoded in the coordinate-independent concept of a derivation.

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

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

(this holds locally, not necessarily globally) and is parametrized by \( m \) “component functions” \( a_i : M \to \mathbb{R} \). They depend of course on the coordinate system, and transform in the usual way under change of coordinates (chain rule):

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

(Exercise).

**The differential of a map or “tangential map”:** let

\[
\Phi : M \to N
\]
be a smooth map. Then one defines

\[ d\Phi : T_p(M) \to T_{\Phi(p)}(N), \]
\[ X \to (d\Phi(X))[f] := X[f \circ \Phi] \]

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

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

\[ \gamma : \mathbb{R} \to M, \quad \gamma(0) = p \in M. \]

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 \( \gamma \) at \( p \) is obtained by

\[ X_p = d\gamma(V_0) \]
i.e. for \( f \in C(\mathbb{R}^n) \) we have

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

which is indeed the “directional derivative” along \( \gamma \). Hence

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

and the components are just the components of \( \frac{d\gamma}{dt} \).

Examples:

1. If \((x^i)\) are coordinates on \( M \) and \((y^i)\) coordinates on \( N \), then

\[ d\Phi(\frac{\partial}{\partial x^i}|_m) = \sum \frac{\partial(y^i \circ \Phi)}{\partial x^i} \frac{\partial}{\partial y^i}|_{\Phi(m)} \] (21)

2. Consider the map \( x^i : M \to \mathbb{R}_t \). Then

\[ dx^i(\frac{\partial}{\partial x^k}|_m) = \sum \frac{\partial x^i}{\partial x^k} \frac{d}{dt} = \delta^i_k \frac{d}{dt} \]

this means that one can identify \( dx^i \) with the dual to \( \frac{\partial}{\partial x^i} \).

3. if \( f : M \to \mathbb{R} \), then

\[ df = \frac{\partial f}{\partial x^i} dx^i \]
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 $X, Y \in T(M)$. Then one can define a new vector field $[X, Y]$ by defining

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

One then easily shows

**Theorem 9**

- $[X, X] = 0$, hence $[X, Y] = -[Y, X]$
- $[X, [Y, Z]] + [Y, [X, Z]] + [Z, [X, Y]] = 0.$
- $[fX, gY] = fg[X,Y] + fX[g]Y - gY[f]X$.
- Given a map $\phi : M \to N$, let $\tilde{X}, \tilde{Y}$ be vector fields on $N$ such that $d\phi(X) = \tilde{X}$ (i.e. $X$ and $\tilde{X}$ are “$\phi$-related”; note that $\phi(M)$ may be a submanifold of $N$). Then $[X, Y]$ is $\phi$-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

\[
[X,Y] = X^i(x) \frac{\partial}{\partial x^i} Y^j(x) \frac{\partial}{\partial x^j} - Y^j(x) \frac{\partial}{\partial x^j} X^i(x) \frac{\partial}{\partial x^i}
\]

\[
= X^i(x) \frac{\partial Y^j(x)}{\partial x^i} - Y^j(x) \frac{\partial X^i(x)}{\partial x^j} \tag{22}
\]

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

\[
\mu : G \times G \rightarrow G \quad (g_1, g_2) \mapsto g_1 \cdot g_2
\]

and

\[
\nu : G \rightarrow G \quad g \mapsto g^{-1}
\]

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

\[
L_g : G \rightarrow G \quad \quad g' \mapsto g \cdot g'
\]

(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.

In particular: a representation of $G$ is a Lie-group homomorphism

\[
\pi : G \rightarrow GL(n, \mathbb{R}) \quad \text{“real” representation}
\]

\[
\pi : G \rightarrow GL(n, \mathbb{C}) \quad \text{“complex” representation}
\]
\[ \pi : G \rightarrow U(n) \quad \text{“unitary” representation} \]

Important problem (both physics and math): find (all) representations of \( G \). This can be solved, and will be explained in this course.

Examples:

- \((\mathbb{R}^n, +)\). The left-translation become here \( L_x(y) = x + y \), i.e. indeed translations of \( y \) by \( x \).

- \( C^* = (\mathbb{C} - \{0\}, \cdot) \)

- the complex numbers with unit modulus \( U(1) = S^1 \) and multiplication

- matrix groups:
  \[
  GL(n, \mathbb{R}) := \{ A \in \text{Mat}(n, \mathbb{R}); \det(A) \neq 0 \}
  \]
  similarly \( GL(n, \mathbb{C}) \), and
  \[
  SL(n, \mathbb{R}) := \{ A \in \text{Mat}(n, \mathbb{R}); \det(A) = 1 \},
  \]
  \[
  O(n) := \{ A \in \text{Mat}(n, \mathbb{R}); \quad A A^T = \mathbb{1} \},
  \]
  \[
  SO(n) := \{ A \in \text{Mat}(n, \mathbb{R}); \quad A A^T = \mathbb{1}, \det(A) = 1 \},
  \]
  \[
  U(n) := \{ A \in \text{Mat}(n, \mathbb{C}); \quad AA^\dagger = \mathbb{1} \},
  \]
  \[
  SU(n) := \{ A \in \text{Mat}(n, \mathbb{C}); \quad AA^\dagger = \mathbb{1}, \det(A) = 1 \},
  \]
  \[
  SP(n, \mathbb{R}) := \{ A \in \text{Mat}(n, \mathbb{C}); \quad A^T J A = J, \quad J = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} \},
  \]

- Lorentz group
  \[
  O(3, 1) = \{ A \in \text{Mat}(n, \mathbb{R}); \quad A \eta A^T = \eta \}, \quad \eta = (1, -1, -1, -1),
  \]
  etc.

- Poincare group (=Lorentz plus translations),

(there are 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)

\[
[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}
\]
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

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

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

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

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:

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

where the Lie algebra is again defined by the commutator.

**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

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

Essentially, one can show that all (finite-dimensional) Lie algebras are subalgebras of $gl(n)$ (Varadarajan, see B.Hall).
A representation of a Lie algebra \( \mathfrak{g} \) is a Lie-algebra homomorphism
\[
\pi : \mathfrak{g} \rightarrow gl(n, \mathbb{R}) \quad \text{“real” representation}
\]
\[
\pi : \mathfrak{g} \rightarrow gl(n, \mathbb{C}) \quad \text{“complex” representation}
\]

One also distinguishes between real and complex Lie algebras, which simply means that the underlying vector space is over \( \mathbb{R} \) resp. \( \mathbb{C} \).

**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,
\[
c_{ij}^k + c_{jk}^i = 0 \quad \text{(antisymmetry)}
\]
\[
\sum_m (c_{ij}^m c_{mk}^l + c_{jk}^m c_{mi}^l + c_{ki}^m c_{mj}^l) = 0 \quad \text{(Jacobi identity)}
\]

### 4.4 The Lie algebra of a Lie group

Let \( G \) be a Lie group. Recall left translations on \( G \), defined by \( L_g : G \rightarrow G, \ g' \mapsto g \cdot g' \).

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, \( dL_g(X_{g'}) = X_{g' \cdot g} \).

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

We have \( L_{\vec{a}}(\vec{x}) = \vec{x} + \vec{a} \). Then (Exercise) \( dL_{\vec{a}}(f^i(x) \frac{\partial}{\partial x^i}) = f^i(x) \frac{\partial L_{\vec{a}} f^j}{\partial x^j} \frac{\partial}{\partial x^i} x + \vec{a} = f^i(x) \frac{\partial}{\partial x^i} x + \vec{a} \), hence \( dL_{\vec{a}}(X) = X \) for all \( \vec{a} \in \mathbb{R}^n \) implies \( f^i(\vec{x} + \vec{a}) = f^i(\vec{x} + \vec{a}) \) \quad \forall \vec{a} \), hence \( f^i = \text{const} \) and
\[
\mathfrak{g} = \{ f^i \frac{\partial}{\partial x^i} \} \cong \mathbb{R}^n.
\]
Observe:

- \( g \cong T_e G \cong \mathbb{R}^n \)
  clear: given \( X_e \), define \( X_g := dL_g(X_e) \)
  can show (easy): is left-invariant V.F.
  (Proof: \( dL_{g'}(X_g) = dL_{g'}(dL_g(X_e)) = d(L_{g'}L_g)(X_e) = dL_{g'g}(X_e) = X_{g'g} \))
- \( g \) is a Lie algebra: for \( X, Y \in g \); define \([X, Y] \ldots \) 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 \( g \) and \( h \), respectively. Then:

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

**Proof:**

1. let \( X, Y \in g \). 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 \( \phi \) 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 \(\hat{X}\) on \(\phi(G) \subset H\). Using Theorem 9, we now get

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

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].

\[
\text{qed}
\]

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, \(\mathfrak{su}(2) \cong \mathfrak{so}(3)\). Moreover \(SU(2)\) is simply connected, hence the statement 2) applies: As soon as we know that \(\mathfrak{su}(2) \cong \mathfrak{so}(3)\) (by simply checking it, see later!) it follows that there exists a group homomorphism \(\phi\) 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. In particular, the Lie algebras of \(\tilde{G}\) and \(G\) coincide by the above theorem, \(\mathfrak{g} = \mathfrak{g}\), and \(\dim(G) = \dim(\tilde{G})\). For example, \(SU(2)\) is the universal cover of \(SO(3)\). Globally, the map \(\Phi\) 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 \(U\)).

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” to use \(SU(2)\) rather than \(SO(3)\), etc.

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 \(G\) 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 $e = I$ is given by the “matrix element coordinates”,

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

where $g = (g^{ij})$. 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

$$X^A_e = A_{ij} \frac{\partial}{\partial x^{ij}}|_e, \quad A_{ij} \in \mathbb{R}$$

(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

$$\text{Lie}(GL(n)) = gl(n)$$

as Lie algebras, not just as vector spaces (the latter is clear). (i.e. commutator for $gl(n)$).

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

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

Then using (21), we have

$$dL_g(\frac{\partial}{\partial x^{ij}}|_e) = \frac{\partial L_gx}{\partial x^{kl}}|_e \frac{\partial}{\partial x^{ij}}|_e = \frac{\partial (g^{km}x^{mj})}{\partial x^{ij}}|_e \frac{\partial}{\partial x^{kl}}|_e = \delta^{ij}g^{kl} \frac{\partial}{\partial x^{kl}}|_e = g^{ki} \frac{\partial}{\partial g^{kj}}|_e$$

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

$$X^A_g = dL_g(X^A_e) = g^{ki}A_{ij} \frac{\partial}{\partial g^{kj}}$$
Now we can calculate the commutator: Let $X^A, X^B$ be left-invariant vector fields as above. Then noting that the $g^{ij} = x^{ij}(g)$ are the coordinate functions on $GL(n)$, i.e. $\frac{\partial}{\partial g^{ij}} |_{g^l} g^{kl} = \delta^k \delta^l$, and using (22) we have

$$[X^A, X^B] = g^{ki} A_{ij} \frac{\partial}{\partial g^{kj}} g^{l'j'} B_{l'j'} - g^{ki} B_{ij} \frac{\partial}{\partial g^{kj}} g^{l'j'} A_{l'j'} \frac{\partial}{\partial g^{kj'}}$$

$$= g^{ki} A_{ij} B_{j'j} \frac{\partial}{\partial g^{kj}} - g^{ki} B_{ij} A_{j'j} \frac{\partial}{\partial g^{kj'}}$$

$$= g^{ki} (A_{ij} B_{j'j} - B_{ij} A_{j'j}) \frac{\partial}{\partial g^{kj}} = g^{ki} [A, B]_{ij} \frac{\partial}{\partial g^{kj'}} = X^{[A, B]} \quad (23)$$

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

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

which we considered before. Similarly one obtains

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

### 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

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

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 continuous 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 $(e^X)^{-1} = e^{-X}$. In particular, $e^X \in GL(n)$.

3. $e^{(\alpha + \beta)X} = e^{\alpha X} e^{\beta X}$ for all real or complex numbers $\alpha, \beta$.

4. If $[X, Y] = 0$, then $e^{X+Y} = e^X e^Y = e^Y e^X$.

5. If $C$ is invertible, then $e^{CX C^{-1}} = C e^X C^{-1}$.

The proof is elementary (Analysis lecture).

It is not true in general that $e^{X+Y} = e^X e^Y$, although by 4) it is true if $X$ and $Y$ commute. This is a crucial point which one should never forget. (There exists a formula (the Baker-Campbell-Hausdorff formula) 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

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

Let's see if this is true: using $J^2_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, we get

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

as desired.

Remark: One good way to calculate the exponential of a Matrix is to diagonalize it if possible: if $X = UDU^{-1}$, then $e^X = U(e^D)U^{-1} = U\text{diag}(e^{d_i})U^{-1}$ 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

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

**Proof:**

**Case 1:** $A$ 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. q.e.d

Also, check it in the explicit example above.
5.2 One-parameter subgroups and the exponential map

Let’s consider again rotations: we wrote finite rotations in the form

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

(26)

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

Fix the axis $\vec{v}$. Then 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

$$R((\phi + \psi) \vec{v}) = R(\phi \vec{v}) R(\psi \vec{v}).$$

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

$$\left. \frac{d}{d\phi} \right|_{\phi=0} R(\phi \vec{v}) \vec{v} = i\vec{v} \cdot \vec{J}$$

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

Let’s 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

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

Hence “infinitesimal rotations” are given by

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

This really IS an ‘infinitesimal rotation” for “infinitesimal” $\phi$, hence the claim is justified. Note that an infinitesimal rotation has the form $(1 + \varepsilon J)$ for $J \in so(3)$.

More generally, each $X \in Lie(G)$ determines uniquely a 1-parameter subgroup in $G$. This provides a nice, useful connection between the Lie algebra and the Lie group, which generalizes to any Lie group:

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 $\mathbb{I}$ . Hence to every $X \in gl(n)$ we can associate a curve $\gamma(t) = e^{tX}$, which satisfies

$$\gamma(t + s) = \gamma(t) \gamma(s) \quad \forall t, s \in \mathbb{R}$$
and
\[ \frac{d}{dt} \bigg|_{t=0} \gamma(t) = X \] 
by the properties of exp. This means that we have a Lie group homomorphism
\[ \gamma : \mathbb{R} \to GL(n) \]
In fact, by theorem 10 there exists a unique Lie group homomorphism \( \gamma : \mathbb{R} \to G \) such that \( d\gamma \big[ \frac{d}{dt} \big]_{t=0} = X_e \in \mathfrak{g} \) for any Lie group \( G \). 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
\[ d \exp_X \bigg( \frac{d}{dt} \bigg) = X. \]
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
\[ \log A = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(A - I)^m}{m} \] 
(28)
is defined and continuous on the set of all \( n \times n \) complex matrices \( A \) with \( \|A - I\| < 1 \), and \( \log A \) is real if \( A \) is real.

29
For all $A$ with $\|A - I\| < 1$, 
\[ e^{\log A} = A. \]

For all $X$ with $\|X\| < \log 2$, $\|e^X - 1\| < 1$ 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

\[
\begin{align*}
G & \xrightarrow{\phi} H \\
\exp & \downarrow \\
\mathfrak{g} & \xrightarrow{d\phi} \mathfrak{h}
\end{align*}
\]

(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 $X \in \text{Lie}(G)$ and “infinitesimal” $\epsilon$.

The elements $X \in \text{Lie}(G)$ 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
\[ \det(e^A) = e^{\text{tr}A} \]

This follows from the following diagram

\[
\begin{align*}
\text{Gl}(V) & \xrightarrow{\det} \mathbb{R}_+ \\
\exp & \downarrow \\
\text{gl}(V) & \xrightarrow{\text{tr}} \mathbb{R}
\end{align*}
\]

noting that $d(\det)|_I = \text{tr}$, 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 $1$, so that $g = e^A$ for some $A \in gl(n)$ (by theorem 14). Then $e^{AT} = g^T = g^{-1} = e^{-A}$. 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

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

$SL(n)$ and $sl(n)$:

Recall that

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

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

$$\exp(sl(n)) = SL(n) \subset GL(n).$$
therefore \( sl(n) = \text{Lie}(SL(n)) \).

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

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

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 \( \mathbb{I} \), 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 \( u(n) = \text{Lie}(U(n)) \).

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

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

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

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

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

32
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

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

Their Lie algebra is

$$[X_i, X_j] = -\epsilon_{ijk} X_k.$$  \hspace{1cm} (31)

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

In physics, one often allows complex coefficients, defining

$$J_k := -i X_k$$

which are hermitian $J_k^\dagger = J_k$ and satisfy the “rotation algebra”

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

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 $g \cong \mathbb{R}^n$ by $g_\mathbb{C} \cong \mathbb{C}^n$, 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

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

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

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

Again, one defines the complexified generators

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

which satisfy

$$[J_i, J_j] = i\epsilon_{ijk} J_k, \quad [X_i, X_j] = -\epsilon_{ijk} X_k.$$  \hspace{1cm} (32)

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 \textit{SO}(3, 1) \text{ and } so(3, 1)

The Lorentz group $\textit{SO}(3, 1)$ is defined by

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

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

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

The structure constants are:

$$[K_x, K_y] = -i J_z, \text{ etc.,}$$

$$[J_x, K_x] = 0, \text{ etc.,}$$

$$[J_x, K_y] = i K_z, \text{ etc.,}$$

$$[J_x, J_y] = i J_z, \text{ etc.}$$

(32)
It is interesting to note that if we allow complex coefficients in the Lie algebra, then

\[
\vec{A} := \frac{1}{2}(\vec{J} + i\vec{K}),
\]

\[
\vec{B} := \frac{1}{2}(\vec{J} - i\vec{K}),
\]

(33)

commute:

\[
[A_x, A_y] = iA_z, \text{ etc},
\]

\[
[B_x, B_y] = iB_z, \text{ etc},
\]

\[
[A_i, B_j] = 0, \forall i, j
\]

(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

\[
\Lambda = e^{ix_j J_j + iy_j K_j} = e^{\frac{i}{2}((x+iy_j)B_i + (x-iy_j)A_i)}
\]

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.

\[
\psi^a = \psi = \begin{pmatrix} \psi^+ \\ \psi^- \end{pmatrix} \in \mathbb{C}^2
\]

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

\[
SL(2, \mathbb{C}) \times \mathbb{C}^2 \to \mathbb{C}^2,
\]

\[
(M, \psi) \to M \cdot \psi
\]

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

\[
M = e^{\frac{i}{2}((x-iy_j)\sigma_i)}
\]

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

\[
\tilde{\psi}^\tilde{a} = \tilde{\psi} = \begin{pmatrix} \tilde{\psi}^+ \\ \tilde{\psi}^- \end{pmatrix} \in \mathbb{C}^2
\]

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

\[
SL(2, \mathbb{C}) \times \mathbb{C}^2 \to \mathbb{C}^2,
\]

\[
(M, \tilde{\psi}) \to M^* \cdot \tilde{\psi}
\]
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 such Weyl spinors.

**Finite boosts:**

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

\[
K^2_x = -\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix},
\]

we get

\[
e^{iK_x} = \sum_{m=0}^{\infty} \frac{(iK_x)^m}{m!} = 1 - K^2_x \sum_{n=1}^{\infty} \frac{(\epsilon)^{2n}}{(2n)!} + iK_x \sum_{n=0}^{\infty} \frac{(\epsilon)^{2n+1}}{(2n+1)!}
\]

\[
= \mathbb{1} - K^2_x (\cosh(\epsilon) - 1) + iK_x \sinh(\epsilon) = \begin{pmatrix}
\cosh(\epsilon) & \sinh(\epsilon) & 0 & 0 \\
\sinh(\epsilon) & \cosh(\epsilon) & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]

\[
= \begin{pmatrix}
\gamma & \beta \gamma & 0 & 0 \\
\beta \gamma & \gamma & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]

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
\[
G \times \mathcal{H} \to \mathcal{H},
\]
\[
(g, \psi) \mapsto g \triangleright \psi
\]  
(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,
\[
\pi : G \to GL(\mathcal{H})
\]
(37)

should satisfy
\[
\pi(g_1) \pi(g_2) = \pi(g_1 g_2), \quad \pi(e) = \mathbb{1}, \quad \pi(g^{-1}) = \pi(g)^{-1}.
\]
(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 algebra $[J_i, J_j] = \epsilon_{ijk} J_k$ 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 representation of $G$ and we write $\pi(g) = e^{ia,\pi(J_i)} \in G$ where $J_i \in \mathfrak{g}$, then $\pi$ is unitary iff $\pi(J_i)$ is hermitian. Hence unitary representations of $G$ correspond to representations of $\mathfrak{g}$ with hermitian (or anti-hermitian...) operators.

**Definition 17** Let $\pi$ 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 $su(2)$ revisited

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 

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

satisfy 

$$ [J_0, J_\pm] = \pm 2J_\pm, \quad [J_+, J_-] = J_0 

(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 $v_\lambda$ of $J_0$ with 

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

using the above CR, we have 

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

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 $v_\Lambda$ with 

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

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

$$ v_{\Lambda - 2n} := (J_-)^n v_\Lambda, $$

(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 $J_0$, and we have 

$$ v_{m-2} = J_- v_m. $$

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

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

We want to find $N$, and understand the structure in detail.
Now consider
\[
J_+ v_{\Lambda-2n} = J_+ J_- v_{\Lambda-2n+2} = (J_- J_+ + J_0) v_{\Lambda-2n+2} = (r_{n-1} + \Lambda - 2n + 2) v_{\Lambda-2n+2}
\]
(42)
where we introduced \( r_n \) by
\[
J_+ v_{\Lambda-2n} = r_n v_{\Lambda-2n+2}.
\]
\((J_+ v_{\Lambda-2n} \) is indeed proportional to \( v_{\Lambda-2n+2} \); to see this use (41) inductively). Because
\( J_+ v_{\Lambda} = 0 \), we have \( r_0 = 0 \). Therefore we find the recursion relation
\[
r_n = (r_{n-1} + \Lambda - 2n + 2) \quad \text{for} \quad r_n,
\]
which is easy to solve (exercise):
\[
r_n = n(\Lambda - n + 1).
\]
Since by definition \( J_- v_{\Lambda-2N} = 0 \), we have
\[
0 = J_+ J_- v_{\Lambda-2N} = (J_- J_+ + J_0) v_{\Lambda-2N} = (r_N + \Lambda - 2N) v_{\Lambda-2N}.
\]
Substituting \( r_n = n(\Lambda - n + 1) \), this yields the quadratic equation \( N^2 + (1 - \Lambda)N - \Lambda = 0 \),
which has the solutions \( N = -1 \) and \( N = \Lambda \). The first is nonsense, hence
\[
N = \Lambda
\]
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
\[
dim V = 2j + 1.
\]
Notice that we have covered 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
\[
V_{\Lambda} := \{ v_{\Lambda-2n}, \quad n = 1, 2, ..., N, \quad \Lambda = N \},
\]
i.e. they are characterized by their dimension.
Irreps are also characterized by the value of the Casimir operator
\[
\hat{J}^2 = J_1 J_1 + J_2 J_2 + J_3 J_3
\]
which satisfies
\[
[\hat{J}^2, J_i] = 0.
\]
Therefore it takes the same value on any vector in the irrep. \( V_\Lambda \). Using

\[ \hat{J}^2 = \frac{1}{4} J_0(J_0 + 2) + J_- J_+ \]

it is easy to evaluate it on the highest weight vector \( v_\Lambda \), which gives

\[ \hat{J}^2 = \frac{1}{4} \Lambda(\Lambda + 2) = j(j + 1) \]

in the irrep \( V_\Lambda \).

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

\[ J^\dagger_a = J_a. \]

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,

\[ |\Lambda - 2n\rangle := c_n v_{\Lambda - 2n} = c_n (J_-)^n v_\Lambda \] (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

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

defined by the formula

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

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

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

\[ \text{ad}_{J_i}(x) = [J_j, X] \]
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

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

hence the matrix which represents $ad_{X_i}$ is

$$(ad_{X_i})^k_l = c^k_{il}.$$  

Hence the structure constants always define the adjoint representation.

**Group version**

The adjoint representation should correspond 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

$$Ad_g : \mathfrak{g} \rightarrow \mathfrak{g},$$

$$X \rightarrow gXg^{-1}$$  

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

$$Ad : G \rightarrow GL(\mathfrak{g}),$$

$$g \rightarrow [X \rightarrow gXg^{-1}]$$

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

$$ad = d(Ad)$$

(proof:

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

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

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

)

By the diagram 29, this implies

$$Ad(e^X) = e^{adX}$$  

(48)
\[ e^X Y e^{-X} = (e^{adX})(Y) = (1 + ad_X + \frac{1}{2!} ad_X^2 + \ldots) Y = Y + [X,Y] + \frac{1}{2!} [X,[X,Y]] + \ldots \] (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
\[ e^{i \vec{\phi} \cdot \vec{J}} X e^{-i \vec{\phi} \cdot \vec{J}} = R(\vec{\phi}) X R(-\vec{\phi}) = e^{i \vec{\phi} [J_j, \cdot]} X \]

Now \([J_j, X]\) is an infinitesimal rotation of \(X \in so(3)\), hence \(e^{i \vec{\phi} [J_j, \cdot]} X\) 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 \(\mathfrak{g}\):
\[ (X,Y) := \kappa(X,Y) := \text{Tr}_{ad}(ad_X ad_Y) = \text{Tr}_{ad}([X,[Y,\cdot]]) \] (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 \(ad\):
\[
(X, ad_Y(Z)) = (X, [Y,Z]) = \text{Tr}(ad_X ad_Y[Z]) = \text{Tr}(ad_X [ad_Y, ad_Z]) \\
= \text{Tr}([[ad_X, ad_Y]ad_Z]) = \text{Tr}(ad_X [ad_Y, ad_Z]) = ([X,Y],Z) \\
= -(ad_Y(X),Z) \] (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. 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), \(\kappa(\ldots)\) is the unique such invariant form on \(\mathfrak{g}\), and it is non-degenerate. We will see that \(\kappa(\ldots)\) 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:
\[ \text{Tr}_{\mathcal{V}}(\pi(X)\pi(Y)) = \alpha \kappa(X,Y) \]

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 $R_g$ (exists and is unique, cp. $d\mu = dx$ on $\mathbb{R}^n$). This is the Haar measure on $G$.

6.4 The Baker-Campbell-Hausdorff Formula

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

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

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

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

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,

$$\log \left( e^X e^Y \right) = X + \int_0^1 g(e^{\text{ad}X} e^{\text{ad}tY})(Y) \, dt. \quad (53)$$

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 $(u, v)$ be the (any!) 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 $R_g$ (exists and is unique, cp. $d\mu = dx^n$ on $\mathbb{R}^n$).

Then define a new inner product $\langle u, v \rangle$ by
\[ \langle u, v \rangle := \int_G d\mu_g(\pi(g)u, \pi(g)v). \] (54)

This is well-defined because $G$ is compact, and positive definite and non-degenerate because $(,)$ is.

By construction, it is invariant under the action of $G$:
\[ \langle \pi(g)u, \pi(g)v \rangle = \langle u, v \rangle. \] (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 it becomes unitary), and therefore they’re completely reducible by Lemma 18. Specializing to the adjoint representation, it follows by invariance that this 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

Antisymmetry of the structure constants

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 \( c_{ijk} = -c_{kij} \) 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
\[
\pi : G \to GL(V \otimes W) = GL(V) \otimes GL(W) \quad g \mapsto \pi_V(g) \otimes \pi_W(g) \quad (56)
\]
Going over to the Lie algebra, this becomes by differentiating (set \( g = e^{it(x_a J_a)} \))
\[
\pi : \mathfrak{g} \to gl(V \otimes W) = gl(V) \otimes gl(W) \quad \mathfrak{g} \mapsto \pi_V(\mathfrak{g}) \otimes \mathbb{I} + \mathbb{I} \otimes \pi_W(\mathfrak{g}) \quad (57)
\]
It is easy to check directly that this is a representation of \( \mathfrak{g} \).

For example, adding angular momenta in Quantum Mechanics: \( J_i = L_i + S_i \), etc.

The Lie algebras we will consider have the property that all representations are completely reducible (“semisimple”, compact!). Then the tensor product of representations \( V \otimes W \) must decompose into the direct sum of irreps,
\[ V \otimes W \cong \bigoplus_j n_j V_j \]
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 \([J_i,] \) on \( A \otimes B \). For example, consider the Casimir for \( su(2) \),
\[ \vec{J}^2 := J_1 J_1 + J_2 J_2 + J_3 J_3 \in Mat(3) \]
where \( J_i \in so(3) \) transform like a vector under \( so(3) \). Clearly \( \vec{J}^2 \) transforms like a scalar under \([J_i,] \), 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. Later)

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.

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

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

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

is proportional to $J_i$:

$$K_i = \alpha J_i$$

for some constant $\alpha$.

**Proof:**

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

$$\pi : su(2) \times Mat(2j + 1, \mathbb{C}) \rightarrow Mat(2j + 1, \mathbb{C}),$$

$$(J_i, M) \rightarrow i[\tilde{J}_i, M]$$  \hspace{1cm} (59)

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

$$\pi : SU(2) \times Mat(2j + 1, \mathbb{C}) \rightarrow Mat(2j + 1, \mathbb{C}),$$

$$(U, M) \rightarrow \pi(U)M\pi(U)^{-1}$$  \hspace{1cm} (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 $(\mathbb{C}^{2j+1})^* \rightarrow -(\mathbb{C}^{2j+1})^*$ is a rep., equivalent to $(\mathbb{C}^{2j+1}) \rightarrow J_i(\mathbb{C}^{2j+1})$)

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

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

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 $\alpha$. Hence any vector operator in an irreducible representation is proportional to $\tilde{J}_i$. \text{qed}

Spherical harmonics: Polynomials, $\Lambda^2$, etc.
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

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

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

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

and the corresponding complex combinations

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

e.g.

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

e etc. To make things more transparent, also introduce

$$
V_3 := [V_+, V_-], \quad U_3 := [U_+, U_-]
$$
which are both linear combinations of $\lambda_3$ and $\lambda_8$. 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
\[
H_3 := \lambda_3, \quad Y := \lambda_8
\]
are diagonal and commute:
\[
(H_3, H_3) = (Y, Y) = 2, \quad (H_3, Y) = 0
\]
(\text{use } (X, Y) \propto \text{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:
\[
H_3|m, y\rangle = m|m, y\rangle, \quad Y|m, y\rangle = y|m, y\rangle
\]
(Isospin and Hypercharge). In particular, in the “defining” 3-dimensional representation we have 3 states
\[
|1, \frac{1}{\sqrt{3}}\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |-1, \frac{1}{\sqrt{3}}\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |0, -\frac{2}{\sqrt{3}}\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}
\]
Moreover since they’re orthogonal, we draw them as orthogonal axis in a 2-dimensional “weight space” (=space of possible eigenvalues):

These label the Quarks $u, d, s$.

They form an equilateral triangle!

Now note 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 $su(2)$:

\[
T_+|m, y\rangle = |m + 2, y\rangle,
U_+|m, y\rangle = |m - 1, y + \sqrt{3}\rangle,
V_+|m, y\rangle = |m + 1, y + \sqrt{3}\rangle,
\]

This will be so in any representation, because of the commutation relations

\[
[H_3, T_\pm] = \pm 2T_\pm, \quad [Y, T_\pm] = 0
\]
\[
[H_3, U_\pm] = \mp U_\pm, \quad [Y, U_\pm] = \pm \sqrt{3}U_\pm
\]
\[
[H_3, V_\pm] = \pm V_\pm, \quad [Y, V_\pm] = \pm \sqrt{3}V_\pm
\]

(65)

etc.

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

\[V \equiv (8) \cong \mathbb{C}^8 = \langle \lambda_1, \ldots, \lambda_8 \rangle\]

and $H$ is represented by $ad_H$, and $Y$ by $ad_Y$. Again they commute (rep!!), hence they can be diagonalized, and we should be able to label the states in (8) by their eigenvalues (=weights!),

\[\lambda_i = \{ |m_i, y_i\rangle, i = 1, 2, \ldots, 8 \}.\]

In fact this is easy from the above CR: Note that the rising- and lowering operators are eigenvectors of $ad_H$ and $ad_Y$, hence

\[
T_\pm \propto |\pm 2, 0\rangle, \quad U_\pm \propto |\mp 1, \pm \sqrt{3}\rangle, \quad V_\pm \propto |\pm 1, \pm \sqrt{3}\rangle
\]

(67)

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

\[H \propto |0, 0\rangle_1, \quad Y \propto |0, 0\rangle_2\]

(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 (composed of u,d,s Quarks)

They form an 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. Useful!

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

\[
\alpha_{T_+} = (2, 0), \quad \alpha_{T_-} = (-2, 0),
\]
\[
\alpha_{U_+} = (-1, \sqrt{3}), \quad \alpha_{U_-} = (1, -\sqrt{3}),
\]
\[
\alpha_{V_+} = (1, \sqrt{3}), \quad \alpha_{V_-} = (-1, -\sqrt{3})
\]

(69)
This will hold for any representation, therefore we give them a name: these 6 α’s are the \textit{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. For example, other important reps are $(3)$:

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

There is also another octet $(8)$ of baryons:

and a heavier decouplet $(10)$ of baryons:
This is also the way to proceed in general:

- choose a maximal set of commuting elements \( \{ H_i \} \) in \( g \) (“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 wavefunctions of the form

\[
\begin{pmatrix}
ungluon 
desu 
s
\end{pmatrix}
\]

They 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 \( (q\bar{q}) \) of a Quark and an Antiquark, hence they are in \( (3) \otimes (\overline{3}) \). We’ll see that

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

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

The point is that the Hamiltonian for strong interactions are invariant under the \( SU(3) \) group acting on these quarks. Actually not quite - there are “soft breaking” terms in the full Hamiltonian of the standar model (due to weak interactions, CKM etc) which do not:

\[
H_{\text{strong}} = H_{\text{su(3)-invar}} + H'
\]

where \( H' \) commutes with hypercharge \( Y \) and isospin, but not with the other generators of \( su(3) \). This then leads to a mass spectrum as above.
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 \( su(2)_{\mathbb{C}} = \langle X_1, X_2, X_3 \rangle_{\mathbb{C}} \supseteq X^+ = X_1 + iX_2 \). The original generators remain linearly independent over \( \mathbb{C} \).

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

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

In particular, in simple Lie algebras there are no generators which commute with the rest of \( \mathfrak{g} \). Also, abelian Lie algebras are excluded by definition...

8.1 Cartan subalgebras, Roots and the Cartan-Weyl basis

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

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

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!).

Actually: if we consider compact Lie groups, then all representations (in particular \( ad! \)) can be made unitary (Weyls unitary trick), hence the generators are hermitian, and can therefore be diagonalized. In fact, all \( \mathfrak{g} \) coming from compact Lie groups with no abelian ideal \( U(1) \) are semisimple.

A \emph{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 \( \{H_i\}_{i=1,...,r} \), and \( [H_i,H_j] = 0 \). Using the Jacobi identity, it follows that \( \text{ad}_{H_i} \) 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 \( \text{ad}_{H_i} \), i.e.

\[
\text{ad}_{H_i}(y_\alpha) = \alpha_i y_\alpha 
\]

(70)

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

\[
\text{ad}_{H}(y_\alpha) = \alpha(H)y_\alpha 
\]

(71)

where

\[
\alpha(H) := \sum c_i \alpha_i. 
\]

That is, \( \alpha \) 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 \( \alpha \) does not depend on the normalization of the eigenvector \( y_\alpha \), they are unique.

Because eigenvectors \( y \) corresponding to different eigenvalues resp. roots \( \alpha \) are linearly independent, it follows that

\[
\mathfrak{g} = \oplus_\alpha \mathfrak{g}_\alpha = \mathfrak{g}_0 \oplus (\oplus_{\alpha \neq 0} \mathfrak{g}_\alpha) 
\]

(72)

where \( \mathfrak{g}_\alpha = \{y \in \mathfrak{g}; \text{ad}_{H}(y) = \alpha(H)y\} \). 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

\[
\Phi = \{\alpha \neq 0\} \subset \mathfrak{g}_0^* 
\]

This is called the root space decomposition of \( \mathfrak{g} \) (relative to \( \mathfrak{g}_0 \)).

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 \( \alpha \). Moreover, the only multiples of \( \alpha \) which are also roots are \(-\alpha\).

This means that there is a basis of \( \mathfrak{g} \) which apart from the basis \( \{H_i\} \) of \( \mathfrak{g}_0 \) consists of elements \( X_\alpha \), one for each root \( \alpha \in \Phi \), which satisfy

\[
\begin{align*}
[H_i,H_j] &= 0 \\
[H_i,X_\alpha] &= \alpha_i X_\alpha \\
& \quad i = 1,2,...,r.
\end{align*}
\]

(73)
where $\alpha_i = \alpha(H_i)$. The $r$-dimensional vector $(\alpha_i)$ resp. the element $\alpha \in g_0^*$ are the roots of $g$. Such a basis is called a Cartan-Weyl basis of $g$. The generators $X_\alpha$ 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 $[J_0, J_\pm] = \pm 2J_\pm$, the roots spaces are

$$g_+ = \mathbb{C} J_+, \quad g_- = \mathbb{C} J_-, \quad g_0 = \mathbb{C} H$$

Similarly for $su(3)$, the Cartan subalgebra is generated by $H_3, Y$ (resp. $\lambda_3, \lambda_8$), 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

$$\langle X, Y \rangle := Tr(X^\dagger Y) \quad (74)$$

and can hence be diagonalized with real eigenvalues $\alpha_H$.)

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

$$g_{0,\mathbb{R}} := \langle H_i \rangle_{\mathbb{R}} \quad (75)$$

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 $g_0$ to $\mathbb{R}$, i.e.

$$\alpha \in g_{0,\mathbb{R}}^* \quad \{ \alpha \in g_{0,\mathbb{R}}^* \}$$

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

Now consider the Killing from. We use the basis $\{X_\alpha, H_i\}$ of $g$. Then for any $H \in g_0$, the matrix representing $ad_H$ is diagonal,

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

(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 $g_{0,\mathbb{R}}$ 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

$g_{0,\mathbb{R}}$ is a Euclidean space

We’ll omit the $\mathbb{R}$ in $g_{0,\mathbb{R}}$.

Because it is nondegenerate, the Killing form defines an isomorphism

$$\alpha \mapsto H_\alpha$$

between $g_0^*$ and $g_0$, which is the space of roots, by

$$\alpha(X) = \kappa(H_\alpha, X)$$  \hspace{1cm} (76)

And, one also obtains an inner product on $g_0^*$ by

$$\langle \alpha, \beta \rangle := \kappa(H_\alpha, H_\beta)$$  \hspace{1cm} (77)

which is again Euclidean. Therefore

*the space $g_0^*$ 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 $g_0^*$.

### 8.1.1 Example: $su(3)$.

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

$$g_0 = \langle H_3, Y \rangle$$  \hspace{1cm} (78)

which are orthogonal: using

\begin{align*}
(H_3, Y) &= \text{Tr}_{C^3}(H_3 Y) = 0, \\
(H_3, H_3) &= \text{Tr}_{C^3}(\lambda_3 \lambda_3) = 2 = (Y, Y)
\end{align*}  \hspace{1cm} (79)
The eigenvectors of $ad_{H_3}, ad_{T_3}$ are

\[
[H_3, T_\pm] = \pm 2T_\pm = \alpha_{T_\pm}(H_3)T_\pm, \quad [Y, T_\pm] = 0 = \alpha_{T_\pm}(Y)Y
\]
\[
[H_3, U_\pm] = \mp U_\pm = \alpha_{U_\pm}(H_3)U_\pm, \quad [Y, U_\pm] = \pm \sqrt{3}U_\pm = \alpha_{U_\pm}(Y)U_\pm
\]
\[
[H_3, V_\pm] = \pm V_\pm = \alpha_{V_\pm}(H_3)V_\pm, \quad [Y, V_\pm] = \pm \sqrt{3}V_\pm = \alpha_{V_\pm}(Y)V_\pm
\]

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

\[
H_{T_\pm} = \pm H_3,
\]
\[
H_{U_\pm} = \pm \left(-\frac{1}{2}H_3 + \frac{\sqrt{3}}{2}Y\right),
\]
\[
H_{V_\pm} = \pm \left(\frac{1}{2}H_3 + \frac{\sqrt{3}}{2}Y\right)
\]

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

\[
\langle \alpha_{T_+}, \alpha_{T_+} \rangle = (H_3, H_3) = 2 = \langle \alpha_{U_+}, \alpha_{U_+} \rangle = \langle \alpha_{V_+}, \alpha_{V_+} \rangle,
\]
\[
\langle \alpha_{T_+}, \alpha_{U_+} \rangle = (H_3, -\frac{1}{2}H_3 + \frac{\sqrt{3}}{2}Y) = -1 = -\langle \alpha_{T_+}, \alpha_{V_+} \rangle,
\]
\[
\langle \alpha_{U_+}, \alpha_{V_+} \rangle = (-\frac{1}{2}H_3 + \frac{\sqrt{3}}{2}Y, \frac{1}{2}H_3 + \frac{\sqrt{3}}{2}Y) = -\frac{1}{2} + \frac{3}{2} = 1
\]

etc. Since $\cos(60) = 1/2$, it follows that all the angles are 60°. 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 $H \in g_0$, consider

$$ [H, [X_\alpha, X_\beta]] = -[X_\alpha, [X_\beta, H]] - [X_\beta, [H, X_\alpha]] $$

$$ = \beta(H)[X_\alpha, X_\beta] + \alpha(H)[X_\alpha, X_\beta] $$

$$ = (\beta(H) + \alpha(H))[X_\alpha, X_\beta] \quad (83) $$

This means that either $[X_\alpha, X_\beta] = 0$, or it is a root vector corresponding to the root $\alpha + \beta$. We will prove in an exercise that if $\alpha + \beta$ is a root, then in fact

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

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

$$ [X_\alpha, X_{-\alpha}] \in g_0 $$

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

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

since $\alpha(H) = (H, H_\alpha)$, for all $H \in g_0$. Therefore

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

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

$$ [H_i, H_j] = 0 $$

$$ [H_i, X_\alpha] = \alpha(H_i)X_\alpha $$

$$ [X_\alpha, X_\beta] = e_{\alpha\beta} X_{\alpha + \beta}, \quad \alpha + \beta \in \Phi $$

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

$$ [X_\alpha, X_{\beta}] = 0, \quad \alpha + \beta \notin \Phi, \alpha + \beta \neq 0 \quad (85) $$

8.2 Useful concepts for representations: Roots and Weights

Now consider a representation $V$ of $g$. We choose a basis of $V$ which are common eigenvectors of the Cartan subalgebra $H_i$: $H_i |\lambda_i; j\rangle = \lambda_i |\lambda_i; j\rangle$
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 $\lambda$ on $g$, just like the roots: if we define $\lambda(H) := \sum c_i \lambda_i$, then

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

(omitting $j$). Note that both roots $\alpha$ and weights $\lambda$ are in $g^*$. This makes sense: recalling that $\text{ad}_H(X\alpha) = \alpha(H)X\alpha$, 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 $\lambda$ is called the weight space of $V$ with weight $\lambda$. An element $v_\lambda \in V_\lambda$ is called a weight vector with weight $\lambda$.

The crucial point is now the following:

$$HX\alpha|\lambda\rangle = (X\alpha H + \alpha(H)X\alpha)|\lambda\rangle = \lambda(H)X\alpha|\lambda\rangle$$

(86)

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

the root vectors $X\alpha$ relate weight vectors with weights differing by $\alpha$. 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).

$$v_{\lambda_1} = |1, \frac{1}{\sqrt{3}}\rangle = \left( \begin{array}{c} 1 \\ 0 \\ 0 \end{array} \right), \quad v_{\lambda_2} = |-1, \frac{1}{\sqrt{3}}\rangle = \left( \begin{array}{c} 0 \\ 1 \\ 0 \end{array} \right), \quad v_{\lambda_3} = |0, -\frac{1}{\sqrt{3}}\rangle = \left( \begin{array}{c} 0 \\ 0 \\ 1 \end{array} \right),$$

(87)

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

$$\lambda_1 = \frac{2}{3} \alpha_{T_+} + \frac{1}{3} \alpha_{U_+},$$

$$\lambda_2 = -\frac{1}{3} \alpha_{T_+} + \frac{1}{3} \alpha_{U_+},$$

$$\lambda_3 = -\frac{1}{3} \alpha_{T_+} - \frac{2}{3} \alpha_{U_+},$$

They form an equilateral triangle.
8.2.1 \textit{su}(2) strings and the Weyl group

The first important thing to note is that $X_\alpha, X_{-\alpha}, H_\alpha$ defines a \textit{su}(2) subalgebra:

\begin{align*}
[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{align*}

since $\alpha(H_\alpha) = (H_\alpha, H_\alpha) = \langle \alpha, \alpha \rangle$. Hence for any pair $\alpha, -\alpha$ of roots there is a corresponding \textit{su}(2). We have seen this explicitly for \textit{su}(3). These various \textit{su}(2)'s are linked in a nontrivial way.

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

\begin{equation}
[X_\alpha, X_{-\alpha}] = H_\alpha
\end{equation}

The \textit{su}(2) algebra generated by $X_\alpha, X_{-\alpha}, H_\alpha$ 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 know how to determine $q$ from $\mu$.

Repeating our analysis of the \textit{su}(2) representations, we define

\begin{equation}
v_j := X_{-\alpha}^j v_0
\end{equation}

where $v_0$ is the weight vector in that string with weight $\mu$. Then $q$ is determined by

\begin{equation}
X_{-\alpha} v_q = 0
\end{equation}

since there is no lower weight.

Using

\begin{equation}
\mu(H_\alpha) = (H_\mu, H_\alpha) = \langle \mu, \alpha \rangle
\end{equation}

we get

\begin{align*}
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{align*}

The solution with $r_0 = 0$ is

\begin{equation}
\begin{array}{c}
r_k = k \langle \mu, \alpha \rangle - \frac{1}{2} k(k - 1)\langle \alpha, \alpha \rangle
\end{array}
\end{equation}

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

\begin{equation}
q = \frac{2 \langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle}
\end{equation}

60
In practice, we often have a weight \( \lambda \) 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

\[
m + p = \frac{2(\lambda + p \alpha, \alpha)}{\langle \alpha, \alpha \rangle}
\]

it follows

\[
m - p = \frac{2(\lambda, \alpha)}{\langle \alpha, \alpha \rangle}
\]

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

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

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

Indeed, \( p = 0 \) and \( m = 1 \).

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

(Picture)

This works as follows:

\[
S_\alpha : \lambda \mapsto \lambda - \frac{2(\lambda, \alpha)}{\langle \alpha, \alpha \rangle} \alpha = \lambda - (m - p)\alpha
\]

because \( \frac{2(\lambda, \alpha)}{\langle \alpha, \alpha \rangle} \alpha \) is just the projection of \( \lambda \) on \( \alpha \). This reflection is called a Weyl reflection. It works for any \( su(2) \) string in any representation: Consider the action of such a \( su(2)_\alpha \) on any irrep \( V \) of \( g \). It follows that \( V \) decomposes into the direct sum of irreps of this \( su(2)_\alpha \). By the above argument, ALL THESE IRREPS (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 \( S_\alpha \).

(Example: \( \text{ad} \) of \( su(3) \).

The various Weyl reflections (in weight space) can be combined, and generate a group called the Weyl group of \( g \). 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...later!
8.3 More on the structure of the adjoint representation

Consider again the $su(2)_\alpha$ subgroups, and the $\alpha$-strings they generate in the adjoint representation. We claim that

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

which has far-reaching consequences.

Assume to the contrary that there is a string which contains 5 weights $\beta - 2\alpha, \beta - \alpha, \beta, \beta + \alpha, \beta + 2\alpha$. We have stated before (without proof...) that if $\alpha$ is a root, then $k\alpha$ can be a root only for $k = \pm 1$. Hence $2\alpha = (\beta + 2\alpha) - \beta$ is not a root, nor is $2(\beta + \alpha) = (\beta + 2\alpha) + \beta$. Thus $\beta + 2\alpha$ is in a $\beta$-string of roots with only one element. Now eqn. (91) implies that

$$\langle \beta + 2\alpha, \beta \rangle = 0.$$ 

Similarly

$$\langle \beta - 2\alpha, \beta \rangle = 0.$$ 

But then $\langle \beta, \beta \rangle = 0$ which is impossible (the Killing form is nondegenerate!). Geometrically, $\beta$ is perpendicular to both $(\beta + 2\alpha)$ and $(\beta - 2\alpha)$ which is possible only if $\beta = 0$.

There is more to learn from (91). Using it twice, we get

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

where $m$ and $n$ are integers given by appropriate values $m' - p'$. The lhs is $\cos^2(\theta)$ where $\theta$ is the angle between the vectors $\alpha$ and $\beta$. Using the Cauchy-Schwarz inequality, it follows that $\frac{1}{4}mn \leq 1$. Hence

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

This means that the possible angles between roots are multiples of $30^\circ, 45^\circ$. 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 \( \{\alpha_1, \ldots, \alpha_r\} \) be an ordered set of roots \((r = \text{rank! not all!})\) such that every element \( \lambda \in \mathfrak{g}_0^* \) can be written uniquely as \( \lambda = \sum c_i \alpha_i \). We then call \( \lambda \) \textit{positive} \((\lambda > 0)\) if \( c_1 > 0 \), or \( c_1 = 0, c_2 > 0, \) or ... etc. ... If the first non-zero \( c_i \) is negative, we call \( \lambda \) negative. (Note: we only consider real linear combinations!!) This is equivalent to choosing some Hyperplane, which divides the space in positive and negative vectors.

We also write \( \lambda > \mu \) if \( \lambda - \mu > 0 \).

In particular, we can now say which roots are positive and which are negative. A \textit{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 \( \alpha_{V+} \) 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 \( \alpha_{V+} \), 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,

\[ \langle \alpha, \beta \rangle \leq 0 \quad \text{if} \quad \alpha, \beta \in \Phi_s \quad (94) \]

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

\[ \sum_{\alpha_i \in \Phi_s} a_i \alpha_i = \sum_{\alpha_i \in \Phi_s} b_i \alpha_i \quad (95) \]

where all \( a_i \geq 0 \) and \( b_i \geq 0 \). (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 \( \alpha_i \) 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 (can also show that they’re 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 \( g \), i.e. \( r = \text{dim}(g_0) = \text{dim}(g_0^*) \). Then the Cartan matrix is the \( r \times r \) matrix

\[
A_{ij} = 2 \frac{\langle \alpha_i, \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle}
\]

where \( \alpha_i \) 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, ±1, ±2, ±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 \( \alpha_i \) and \( \alpha_j \) are proportional. This cannot happen for \( i \neq j \) since the simple roots are linearly independent. Hence

\[
A_{ij} A_{ji} < 4.
\]

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 \( \alpha_U = \alpha_1 + \alpha_2 \), the simple roots are also \( \alpha_1, \alpha_2 \). We already computed the
relevant scalar products:
\[
\begin{align*}
\langle \alpha_1, \alpha_1 \rangle &= 2, \\
\langle \alpha_1, \alpha_2 \rangle &= -1, \\
\langle \alpha_2, \alpha_2 \rangle &= 2,
\end{align*}
\]
(98)

Therefore
\[
A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}
\] (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 \( \beta \). Hence the simple roots have level one. Assume that we found all roots at the \( n \)-th level. Then for each root \( \beta \) 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 \( X_\alpha \) for each root \( \alpha \), 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 get
\[
m - p = 2 \frac{\langle \beta, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = \sum_j 2k_j \frac{\langle \alpha_j, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = \sum_j k_j A_{ji}
\]
(100)

In particular, \( \beta + \alpha_i \) is a root if \( p = m - \sum_j k_j A_{ji} > 0 \). Also, the \( \alpha_i \)-string through \( \alpha_j \) has length \( -A_{ji} + 1 \).

Hence we should keep track of the quantities \( (n_i) = 2 \frac{\langle \beta, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = (\sum_j k_j A_{ji}) \), which are the Dynkin indices of \( \beta \) (see below) for each root. Then (92) simply says that in an \( \alpha_i \) string, the \( i \)-th Dynkin index \( n_i \) takes values in some symmetric set \( \{-m, -m + 2, ..., +m\} \).

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 \( \alpha_i \) string through \( \alpha_j \) takes values in \( \{A_{ji}, A_{ji} + 2, ..., -A_{ji}\} \)).

Consider \( su(3) \). We start by writing down the rows of the Cartan matrix, which represent the simple roots \( \alpha_1, \alpha_2 \):
Start with $\alpha_1$, which has Dynkin indices $A_1$. We ask whether the addition of $\alpha_2$ produces a root at level 2. (recall that $2\alpha_1$ is not a root). Since the second entry in the box representing $\alpha_1$ is negative, the corresponding value of $p$ in (100) is positive. The same conclusion is reached if beginning with $\alpha_2$. Is there a root at level 3? Looking back in the $\alpha_1$ 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 $\alpha_2$. 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 form $-n, \ldots, +n$.

Other example: exceptional algebra $G_2$. The corresponding Cartan matrix is (see later) Hence $G_2$ has 6 positive roots.

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:

\[
\begin{align*}
X_i^+ &:= c_i X_{\alpha_i}, \\
X_i^- &:= c_i X_{-\alpha_i}, \\
H_i &:= \frac{2}{\langle \alpha_i, \alpha_i \rangle} H_{\alpha_i}
\end{align*}
\]

where

\[
c_i = \sqrt{\frac{2}{(X_{\alpha_i}, X_{-\alpha_i}) \langle \alpha_i, \alpha_i \rangle}}
\]

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

\[
\begin{align*}
[H_i, H_j] &= 0 \\
[H_i, X^\pm_j] &= \pm A_{ji} X^\pm_j \\
[X^+_i, X^-_j] &= \delta_{i,j} H_i,
\end{align*}
\]

(103)

The last relation follows because \(\alpha_i - \alpha_j\) is not a root. Finally, since the \(\alpha_i\)-string through \(\alpha_j\) has length \(1 - A_{ji}\) (see below (100)) and \(X^+_j \in g\) (considered as ad-rep) corresponds to the root \(\alpha_j\), it follows that

\[
(ad_{X^+_j})^{1-A_{ji}} X^\pm_j = 0
\]

(104)

These commutation relations are called the Chevalley-Serre commutation relations. One can now follow the above algorithm to find the roots \(\alpha\) and construct the corresponding \(X_\alpha\) at the same time. For example,

\[
X_{\alpha_1 + \alpha_2} = [X^+_1, X^+_2],
\]

e tc. One can show that all root vectors \(X_\alpha\) can be obtained in this way via commutators of the \(X^\pm_i\). This means that it is enough to work with the simple roots and the corresponding \(X^\pm_i\), 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 \(\alpha_i\), 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)).

More interesting: \(SU(3)\):

\[G_2\]: The dark dot in \(G_2\) corresponds to the second root \(\alpha_2\), from the Cartan matrix.
We will later determine all possible Dynkin diagrams, hence all possible simple Lie algebras. To anticipate the close relation between the Cartan matrix and Dynkin diagrams, consider e.g. the following Dynkin Diagram

\[
\begin{bmatrix}
2 & -3 \\
-1 & 2
\end{bmatrix}
\]

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 than the second, it must be \( A_{23} = -2 \) and \( A_{32} = -1 \). Thus

\[
A = \begin{pmatrix}
2 & -1 & 0 \\
-1 & 2 & -2 \\
0 & -1 & 2
\end{pmatrix}
\]

(105)

9 The classical Lie algebras

Let us apply these considerations to the classical Lie algebras \( su(n), so(n) \) and \( sp(n) \). Start with
Recall that $su(n) = \text{Lie}(SU(n)) = \{A \in \text{Mat}(n, \mathbb{C}); A^\dagger = -A = 0, Tr(A) = 0\}$. 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 $su(n)_\mathbb{C}$, see later.)

A basis of $A_{n-1} = su(n)_\mathbb{C}$ is given by all

$$e_{ab} = (e_{ab})_{ij} = \delta_{ia}\delta_{jb}, \quad a \neq b$$

(explain! 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

$$[e_{ab}, e_{cd}] = e_{ad}\delta_{bc} - e_{cb}\delta_{ad} \quad (106)$$

and in particular

$$[H, e_{ab}] = (c_a - c_b)e_{ab} \quad (107)$$

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:

$$\alpha_1(\sum c_i e_{ii}) = c_1 - c_2,$$

$$\alpha_2(\sum c_i e_{ii}) = c_2 - c_3,$$

$$\ldots$$

$$\alpha_{n-1}(\sum c_i e_{ii}) = c_{n-1} - c_n,$$

and declare these positive roots with ordering $\alpha_1 > \ldots > \alpha_{n-1}$. 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

$$\left(\sum c_i e_{ii}, \sum c'_j e_{jj}\right) = Tr(ad\sum c_i e_{ii} ad\sum c'_j e_{jj})$$

$$= \sum_{a,b} (c_a - c_b)(c'_a - c'_b)$$

$$= 2n \sum_a c_a c'_a \quad (108)$$

69
In particular, \((e_{ii}, e_{jj}) = 2n\delta_{i,j}\). This gives the duality between \(\alpha_i\) and \(H_{\alpha_i}\):

\[
(H_{\alpha_i}, \sum_j e_j e_{jj}) = \alpha_i(\sum_j c_j e_{jj}) = \alpha_i - c_{i+1}
\]

Therefore

\[
H_{\alpha_i} = \frac{1}{2n}(e_{ii} - e_{i+1,i+1})
\]

and

\[
\langle \alpha_i, \alpha_j \rangle = \frac{1}{2n}(2\delta_{ij} - \delta_{i,j+1} - \delta_{i+1,j})
\]

Therefore the Cartan Matrix is

\[
A_n = \begin{pmatrix}
2 & -1 & 0 \\
-1 & 2 & -1 \\
0 & -1 & 2 \\
2 & -1 & 0 \\
-1 & 2 & -1 \\
0 & -1 & 2 \\
\end{pmatrix}
\]

and the Dynkin Diagram is

\(sp(2n)\):

The symplectic group \(Sp(2n)\) is defined as

\[
Sp(2n) = \{A \in Mat(2n); A^TJA = J\}
\]

where

\[
J = \begin{pmatrix}
0 & \mathbb{I} \\
-\mathbb{I} & 0 \\
\end{pmatrix}
\]

The corresponding Lie algebra is obtained from \(\exp(X) \approx 1 + X + o(X^2)\), which implies

\[
C_n := sp(2n) = \{X \in Mat(2n); X^T = JXJ\}.
\]

(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

\[
C_n = \begin{pmatrix}
2 & -1 & 0 \\
-1 & 2 & -1 \\
0 & -1 & 2 \\
2 & -1 & 0 \\
-1 & 2 & -1 \\
0 & -2 & 2 \\
\end{pmatrix}
\]
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:

\( D_n := \text{so}(2n) \):

This has rank \( n \), with Cartan matrix

\[
D_n : \begin{pmatrix}
2 & -1 & 0 \\
-1 & 2 & -1 \\
0 & -1 & 2
\end{pmatrix}
\]

and Dynkin diagram

\( B_n := \text{so}(2n + 1) \):

This has also rank \( n \), with Cartan matrix

\[
B_n : \begin{pmatrix}
2 & -1 & 0 \\
-1 & 2 & -1 \\
0 & -1 & 2
\end{pmatrix}
\]

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 disconnected points, and the Cartan matrix is $2I_{2\times 2}$. 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:

This can be proved by considering sets of vectors $\gamma_i \in g^\ast_0$ (candidates for simple roots) and defining the associated matrix

$$M_{ij} = 2\frac{\langle \gamma_i, \gamma_j \rangle}{\langle \gamma_j, \gamma_j \rangle}$$

(119)

and an associated diagram (analogous to the Dynkin diagram) where the $i$th and the $j$th points are joined by $M_{ij}M_{ji}$ lines. The set $\{\gamma_i\}$ is called allowable (zulässig) if all the following conditions hold:

- the $\gamma_i$ are linearly independent, i.e. $det(M) \neq 0$
- $M_{ij} \leq 0$ if $i \neq j$
- $M_{ij}M_{ji} \in \{1, 2, 3\}$

(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 $|\lambda; j\rangle$ of $V$ which are common eigenvectors of the Cartan subalgebra $\mathfrak{g}_0$: $H_i |\lambda; j\rangle = \lambda_i |\lambda; j\rangle$ or in general

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

(omitting $j$). Moreover, the root vectors $X_\alpha$ relate weight vectors with weights differing by $\alpha$,

$$X_\alpha |\lambda\rangle \cong |\lambda + \alpha\rangle.$$

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

$$m - p = \frac{2\langle \lambda, \alpha \rangle}{\langle \alpha, \alpha \rangle}. \quad (120)$$
in particular there are no “holes”. For each weight $\lambda$ of $V$, one defines the *Dynkin labels* (coefficients)

$$\lambda^i := \frac{2\langle \lambda, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} \in \mathbb{Z}, \quad i = 1, 2, \ldots, r$$  \hspace{1cm} (121)

where $\alpha_i$ 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 $\mu$. Since $\mu$ 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:

$$\mu^i := \frac{2\langle \mu, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} \in \mathbb{Z}_{\geq 0}$$  \hspace{1cm} (122)

are positive integers in $\mathbb{Z}_{\geq 0}$. Conversely, one can show (later: using Verma modules) that for each weight $\mu$ with positive Dynkin labels $\mu^i \geq 0$ there is a corresponding (unique!) irrep $V = V_\mu$ with highest weight $\mu$, and furthermore that the corresponding weight space with weight $\mu$ 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 $\mu$, which has non-negative Dynkin labels $\mu^i \geq 0$. It is called highest-weight module with highest weight $\mu$.\*

This can be made more transparent by introducing the *fundamental weights* $\Lambda_{(i)} \in \mathfrak{g}_0^*$, $i = 1, 2, \ldots, r$ by the requirement

$$\frac{2\langle \Lambda_{(i)}, \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle} := \delta_{i,j}$$  \hspace{1cm} (123)

They are just another basis of weight space $\mathfrak{g}_0^*$, then we can expand any weight in terms of this new basis,

$$\lambda = \sum_i \lambda^i \Lambda_{(i)}$$  \hspace{1cm} (124)

Multiplying this with $\frac{2\langle \lambda, \alpha_j \rangle}{\langle \alpha_j, \alpha_j \rangle}$, we see that the $\lambda^i$ are just the Dynkin labels of $\lambda$:

$$\frac{2\langle \lambda, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = \frac{2\langle \lambda \Lambda_{(j)}, \alpha_i \rangle}{\langle \alpha_i, \alpha_i \rangle} = \lambda^i.$$  \hspace{1cm} (125)

Since all Dynkin labels of weights in $V$ are integers, it follows that

*all weights of (finite-dim) an irrep $V$ are integral linear combinations of the fundamental weights.*
In other words, all weights live in the **weight lattice**

\[ L_w := \{ z_i \Lambda(i) \mid z_i \in \mathbb{Z} \} \tag{126} \]

In particular, this applies to the highest weight of \( V \).

**Example: \( su(2) \)**

Let \( \alpha \) 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 \( (2j)\Lambda(1), 2j\Lambda(1) - \alpha, \ldots, (-2j)\Lambda(1) \).

**Example: \( su(3) \)**

(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 \( V_\mu \), we can now calculate all the weights similar as we calculated the roots (=weights of the adjoint rep). Given a weight \( \lambda \) 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 \( (\lambda^i)_{i=1,\ldots,r} \), we know the value of \( p \) in (120). Hence if

\[ m_j = p_j + \lambda^j > 0, \tag{127} \]

then \( \lambda - \alpha_j \) is a weight in \( V \). We record the Dynkin labels of \( \lambda - \alpha_j \) by subtracting \( (A_{ji})_i \) from \( \lambda^i \) (because

\[ \frac{2(\lambda - \alpha_j, \alpha_i)}{\langle \alpha_i, \alpha_i \rangle} = \frac{2(\lambda, \alpha_i)}{\langle \alpha_i, \alpha_i \rangle} - \frac{2(\alpha_j, \alpha_i)}{\langle \alpha_i, \alpha_i \rangle} \]

) Hence the Dynkin labels \( \lambda_j \) of a weight \( \lambda \) are just the \( m_j - p_j \) of the corresponding \( \alpha_j \) string, and go from \( (n_i, n_i - 2, \ldots, -n_i + 2, -n_i) \).

**11.1.1 Examples**

\((1,0)\) of \( A_2 = su(3) \)

Let us determine the weights of the irrep with highest weight \( \mu = \Lambda(1) = (1,0) \):

(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 $\mu = (1, 1)$:

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.

an irrep of $G_2$

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 $\eta$
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.

\( SO(10) \)

As a more complicated example, consider the rep of \( SO(10) (=D_5) \) with highest weight \((1, 0, 0, 0, 0)\):

constructing more representations : tensor products

\subsection*{11.2 Tensor products \( \Pi^2 \)}

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

\[
\begin{align*}
\pi : \mathfrak{g} & \rightarrow gl(V_1 \otimes V_2) = gl(V_1) \otimes gl(V_2) \\
\mathfrak{g} & \mapsto \pi_{V_1}(g) \otimes \mathbb{I} + \mathbb{I} \otimes \pi_{V_2}(g)
\end{align*}
\]

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

\[ V_{\mu_1} \otimes V_{\mu_2} = \bigoplus_{\mu_3} N_{\mu_1 \mu_2}^{\mu_3} V_{\mu_3} \]  

(129)

where \( N_{\mu_1 \mu_2}^{\mu_3} \in \mathbb{N} \) are the multiplicities (Littlewood-Richards coefficients), and \( V_{\mu_1} \) 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 \( \lambda_1 \) resp. \( \lambda_2 \), 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 \( \lambda_1, \lambda_2 \) in \( V_1, V_2 \), 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 \( \mu_1 \) resp. \( \mu_2 \), then the highest weight in \( V_1 \otimes V_2 \) is \( \mu_1 + \mu_2 \), and

\[ V_{\mu_1} \otimes V_{\mu_2} = V_{\mu_1 + \mu_2} \oplus \left( \bigoplus_{\mu_3 < \mu_1 + \mu_2} N_{\mu_1 \mu_2}^{\mu_3} V_{\mu_3} \right) \]  

(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 \( V_{\mu_1 + \mu_2}^\perp \subset V_{\mu_1} \otimes V_{\mu_2} \), etc. However this is practical only for \( su(2) \), and there are more powerful methods.

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 \( \pi_{j_1} (H) \otimes \mathbb{1} \) and \( \mathbb{1} \otimes \pi_{j_2} (H) \).

Because \( V_{j_1} \otimes V_{j_2} = \bigoplus_{j_3} N_{j_1 j_2}^{j_3} V_{j_3} \), there is another basis given by the eigenvectors of \( \pi_{j_1} (H) \otimes \mathbb{1} + \mathbb{1} \otimes \pi_{j_2} (H) \) 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

\[ |j_1, j_2; j_3, m_3\rangle = \sum_{m_1 m_2} |j_1, j_2; m_1 m_2\rangle \langle j_1, j_2; m_1 m_2 |j_3, m_3\rangle \]  

(131)
The coefficients
\[ \langle j_1, j_2; m_1, m_2 | j_3, m_3 \rangle \equiv \langle j_1, j_2; m_1, m_2 | j_1, j_2; j_3, m_3 \rangle \] (132)
are called \textit{Clebsch-Gordan Coefficients}. They determine the unitary transformation between the 2 ONB’s of \( V_{j_1} \otimes V_{j_2} = \bigoplus_{j_3} N_{j_1,j_2}^j V_{j_3} \), 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
\[
N_{j_1,j_2}^{j_3} = 1 \quad |j_1 - j_2| \leq j_3 \leq j_1 + j_2,
N_{j_1,j_2}^{j_3} = 0 \quad \text{otherwise}.
\]
Sometimes one uses the notation
\[ C_{m_1,m_2,m_3}^{j_1,j_2,j_3} = \langle j_1, j_2; m_1, m_2 | j, m \rangle \] (133)
or the so-called \textit{Wigner 3j-symbols}
\[ \langle j_1, j_2; m_1, m_2 | j, m \rangle = \frac{(-1)^{j_1-j_2+m}}{2j+1} \binom{j_1 \ j_2 \ j_3}{m_1 \ m_2 \ -m} \] (134)
which are invariant under cyclic permutations of the 3 columns.

\underline{Example: su(3)}

Hence to understand the multiplet structure of mesons etc, we must consider e.g. \((3) \otimes (3)\).

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 \((\bar{3}) = (0, 1) = V_{\lambda(2)}\). Therefore the bound states of a Quark and an Antiquark are described by the tensor product of the constituent Hilbert spaces, \((3) \otimes (\bar{3})\) (in obvious notation...). I.e.
\[ |\text{meson}; \lambda \rangle = \sum C_{\lambda_1,\lambda_2,\lambda} |\text{Quark}; \lambda_1 \rangle \otimes |\text{Antiquark}; \lambda_2 \rangle \in V_{\lambda(1)} \otimes V_{\lambda(2)}.\]

And similarly for baryons. 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, \((3) \otimes (\bar{3})\) contains \(V_{\lambda(1)+\lambda(2)} = (1, 1) = (8)\), which is the 8-dim. adjoint rep. It follows that
\[ (3) \otimes (\bar{3}) = (8) \oplus (1).\]

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 \(\eta'\) meson).
This is easy to understand graphically:

(draw weight graphs)

Often this method allows to find the complete decomposition, by counting dimensions as well.

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 \( g \), therefore they are either irreducible or themselves reducible (both is possible).

If \( \mu \) 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 \( 2\mu \). 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 \( (3) \). The dimensions add up to 9, hence this is it:

\[
(3) \otimes (3) = (6) \oplus (3)
\]

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 \( V_{n\Lambda(1)} \subset V^{\otimes_{S^n}} \), and we will see that in fact for \( su(n) \) one has

\[
V_{n\Lambda(1)} = V^{\otimes_{s^n}}_{\Lambda(1)} \tag{136}
\]

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

$$H = \frac{\vec{P}^2}{2m} + \frac{1}{2}m\omega^2 \vec{x}^2 = \hbar \omega (a^+_k a_k + \frac{3}{2})$$

where

$$a_k = \sqrt{\frac{m\omega}{2\hbar}}x_k + \frac{i}{\sqrt{2m\hbar\omega}}p_k, \quad a^+_k = \sqrt{\frac{m\omega}{2\hbar}}x_k - \frac{i}{\sqrt{2m\hbar\omega}}p_k, \quad k = 1, 2, 3. \quad (138)$$

The $a^+_k$ resp. $a_k$ are rising resp. lowering operators, satisfying

$$[a_k, a^+_l] = \delta_{kl},$$

$$[a^+_k a_k, a^+_l] = a^+_l \delta_{kl},$$

$$[a^+_k a_k, a_l] = -a_l \delta_{kl}.$$

Let $|0\rangle$ be the ground state, which satisfies

$$a_k |0\rangle = 0 \quad (139)$$

Then the eigenstates of the Hamiltonian, i.e. the energy eigenstates, are

$$\mathcal{H}_n := a^+_{k_1} \ldots a^+_{k_n} |0\rangle \quad (140)$$

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

$$Q_a := a^+_k (T_a) a_l$$

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

$$[T_a, T_b] = i f_{abc} T_c \quad (142)$$
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:

$$[Q_a, Q_b] = if_{abc}Q_c,$$  \hspace{1cm} (143)

hence they define a representation of $su(3)$ on the Hilbert space of the 3d harmon. osc. Moreover,

$$[Q_a, H] = 0,$$  \hspace{1cm} (144)

thus we can fix an energy eigenvalue $E_n = h\omega(n + \frac{3}{2})$, 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

$$Q_a|0\rangle = 0 \hspace{1cm} (145)$$

The precise action of $su(3)$ on $\mathcal{H}_n$ can be found noting that

$$[Q_a, a^+_k] = a^+_l (T_a)_{lk} \hspace{1cm} (146)$$

hence $a^+_k$ transforms like a $(3)$. Therefore $a^+_k ... a^+_n|0\rangle$ 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 \( g \). Let \( \text{mult}_V(\lambda) \) be the multiplicity of the weight space \( \lambda \) in \( V \), i.e. the number of lin. independent elements in \( V \) with weight \( \lambda \). Then

\[
\chi_V := \sum_\lambda \text{mult}_V(\lambda) \, e^\lambda
\]

This is to be understood as a generating function. The \( e^\lambda \) are “formal”, and linearly independent for each \( \lambda \). One can understand \( e^\lambda \) as function on the weight space, by

\[
\mu \mapsto e^\lambda(\mu) := e^{\langle \lambda, \mu \rangle}
\]

Knowing \( \chi_V \) is equivalent to knowing all states and all multiplicities of the representation \( V \), hence in particular its dimension, etc. \( \chi_V \) is a VERY useful way to encode this information. In terms of a basis \( |i\rangle \) with weights \( \lambda_i \), of the representation \( V \), we can write

\[
\chi_V(\mu) = \sum_\lambda \text{mult}_V(\lambda) \, e^{\langle \lambda, \mu \rangle} = \sum_i \langle i| \mu \rangle e^{\langle \lambda_i, \mu \rangle} = \sum_i \langle i| e^{H_\mu} |i\rangle = Tr_V(e^{H_\mu})
\]

One obvious and useful property is the following:

\[
\chi_{V \otimes W} = \chi(V) \chi(W)
\]

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 \( \alpha \) 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

\[
\chi_{V_j} = \sum_{m=-j}^{j} e^{m\alpha} = \frac{e^{(j+\frac{1}{2})\alpha} - e^{-(j+\frac{1}{2})\alpha}}{e^{\frac{1}{2}\alpha} - e^{-\frac{1}{2}\alpha}}\]

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\[ S_\alpha : \lambda \mapsto \lambda - \frac{2(\lambda, \alpha)}{\langle \alpha, \alpha \rangle} \alpha \]which reflects the states (weights) of the \( su(2)_\alpha \) strings, and therefore it also preserved the multiplicities: \( \text{mult}_V(S_\alpha(\lambda)) = \text{mult}_V(\lambda) \), therefore\[ \text{mult}_V(\omega(\lambda)) = \text{mult}_V(\lambda) \]
for any \( \omega \in \mathcal{W} \). Extending this action of the Weyl group to \( \omega(e^\lambda) := e^{\omega(\lambda)} \), it follows that\[ \chi_V(\omega(\mu)) = \chi_V(\mu) \]
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 \( V_\lambda \) is the highest weight irrep with h.w. \( \lambda \):

\[
\chi_\lambda(\mu) = \frac{\sum_{\omega \in \mathcal{W}} \text{sign}(\omega) e^{\langle \omega(\lambda + \rho), \mu \rangle}}{\sum_{\omega \in \mathcal{W}} \text{sign}(\omega) e^{\langle \rho, \mu \rangle}}
\]

Here \( \text{sign}(\omega) = \text{det}(\omega) = \pm 1 \) is the signum of \( \omega \in \mathcal{W} \) (i.e. even/odd number of reflections), and \( \rho := \frac{1}{2} \sum_{\alpha \in \Phi_{>0}} \alpha \) 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 \( g = su(2) \), 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 \( \alpha \) is the (only) positive root of \( su(2) \).

For \( su(n) \), this can be written in an even more compact form: recall the identification \( H_\lambda = \lambda \), and the explicit realization of the Cartan generators as diagonal matrices

\[
H_\lambda = \text{diag}(h^\lambda_1, \ldots, h^\lambda_n) \cong \text{diag}(\lambda_i)
\]
with \( \sum \lambda_i = 0 \) (these are not Dynkin indices!), e.g. \( H_{\alpha_1} \propto \text{diag}(1, -1, 0, \ldots, 0) \) etc. Then the Killing form was essentially \( \langle \lambda, \mu \rangle = Tr(H_\lambda H_\mu) = \sum \lambda_i \mu_i \). Furthermore, one can show that the Weyl group for \( SU(n) \) is just \( S^n \), the permutation group of \( n \) elements. It acts on \( \lambda \) resp. \( H_\lambda \) as \( \text{diag}(h_1, \ldots, h_n) \to \text{diag}(h'_1, \ldots, h'_n) \) by permuting the elements. Furthermore, the highest weight are in the fundamental Weyl chamber which is now characterized by \( h_1 \geq \ldots \geq h_n \).
Then replacing \((\lambda_j)_{i=1,\ldots,n} \leftrightarrow \lambda\), we get
\[
\sum_{\omega \in W} \text{sign}(\omega)e^{(\omega(\lambda+\rho),\mu)} = \sum_{\omega \in S_n} (-1)^\omega e^{\sum_i \omega(\lambda+\rho)\mu_i} = \sum_{\omega \in S_n} (-1)^\omega e^{(\lambda+\rho)\omega(1)\mu_1 \ldots e^{(\lambda+\rho)\omega(n)\mu_n}} = \det(\{e^{(\lambda+\rho)\mu_i}\}_{ij})
\] (156)
Furthermore, one can show that \(\rho_i = (m,m-1,\ldots,-m)\), so that we can rewrite (154) as
\[
\chi(\lambda) = \frac{\det(\{e^{(\lambda+n-i)\mu_j}\})}{\det(\{e^{(n-i)\mu_j}\})}
\] (157)
Note further that
\[
\det(\{e^{(n-i)\mu_j}\}) = \prod_{i<j} (e^{\mu_i} - e^{\mu_j}) = \Delta(e^{\mu_i})
\] (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 \(\lambda + n - i = (3,1,-1)\), and we have
\[
\chi(8)(\mu) = \frac{\det(\{e^{(3,1,-1),(\mu_1,\mu_2,\mu_3)j}\})}{\Delta(e^{\mu_i})} = e^{-(\mu_1+\mu_2+\mu_3)} \det(\{e^{(4,2,0),(\mu_1,\mu_2,\mu_3)j}\})
\]
\[
= e^{-(\mu_1+\mu_2+\mu_3)} \frac{\Delta(e^{2\mu_i})}{\Delta(e^{\mu_i})} = \prod_{i<j} \frac{(e^{2\mu_i} - e^{2\mu_j})}{(e^{\mu_i} - e^{\mu_j})}
\]
\[
= \prod_{i<j} (e^{\mu_i} + e^{\mu_j}) = (e^{\mu_1} + e^{\mu_2})(e^{\mu_1} + e^{\mu_3})(e^{\mu_2} + e^{\mu_3})
\]
\[
= e^{2\mu_1+\mu_2} + e^{2\mu_2+\mu_3} + e^{\mu_1+2\mu_2} + e^{2\mu_1+\mu_3} + e^{\mu_1+2\mu_3} + e^{\mu_2+2\mu_3} + 2e^{\mu_1+\mu_2+\mu_3}
\]
\[
= e^{\mu_1-\mu_3} + e^{-\mu_1+\mu_2} + e^{\mu_2-\mu_3} + e^{\mu_1-\mu_2} + e^{-\mu_2+\mu_3} + e^{-\mu_1+\mu_3} + 2e^0
\]
\] (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
\[
V_j \otimes V_k = N^j_{jk} V_l
\]
where \(V_j\) is the spin \(j\) irrep of \(su(2)\).

Recall that the weights of \(V_j\) are
\[
V_k = \langle v_\lambda \rangle, \quad \lambda \in \{-k\alpha, \ldots, (k-1)\alpha, k\alpha\}
\]

85
where $\alpha$ is the root of $su(2)$. Therefore

$$\chi_{V_j \otimes V_k} = \chi_{V_j} \left( \sum_{m=-k}^{k} e^{ma} \right) = \left( \frac{e^{(j+\frac{1}{2})\alpha} - e^{-(j+\frac{1}{2})\alpha}}{e^{\frac{1}{2}\alpha} - e^{-\frac{1}{2}\alpha}} \right) \left( \sum_{m=-k}^{k} e^{ma} \right)$$

reversing the summation index in the second term. Now assume that $k \leq j$. Then the arguments of the exponentials are $\{(j+k + \frac{1}{2}), \ldots, (j-k + \frac{1}{2})\}$ for the first term, and $\{-(j+k + \frac{1}{2}), \ldots, -(j-k + \frac{1}{2})\}$ for the first term. Therefore the rhs can be rewritten as

$$\chi_{V_j \otimes V_k} = \sum_{l=j-k}^{j+k} \frac{e^{(l+\frac{1}{2})\alpha} - e^{-(l+\frac{1}{2})\alpha}}{e^{\frac{1}{2}\alpha} - e^{-\frac{1}{2}\alpha}} = \sum_{l=j-k}^{j+k} \chi_{V_l}$$

This implies that

$$V_j \otimes V_k = \bigoplus_{l=j-k}^{j+k} V_l.$$

(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 weights 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 $\omega \in W$ preserves the Killing form:

$$\langle S_{\alpha\mu}, S_{\alpha\nu} \rangle = \langle \mu - 2\frac{\langle \mu, \alpha \rangle}{\langle \alpha, \alpha \rangle} \alpha, \nu - 2\frac{\langle \nu, \alpha \rangle}{\langle \alpha, \alpha \rangle} \alpha \rangle = \langle \mu, \nu \rangle$$

(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 $\omega \in W$ maps any Weyl chamber to another one. There is one which is a special Weyl chamber, the fundamental Weyl chamber which is

$$P_+ = \{ \lambda = \sum c_i \Lambda(i), \quad c_i \geq 0 \} \quad (162)$$

Note that the walls of $P_+$ are the hyperplanes $H_{\alpha_i}$, since $c_i = 0 \Leftrightarrow \langle \lambda, \alpha_i \rangle = 0$. 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 $\mu$ in some Weyl chamber, and the orbit of $W$ acting on $\mu$, 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 $\mu^*$ 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 $\mu^*$). It follows that all Weyl chambers are congruent, since they are mapped by some $\omega \in W$ to $P_+$. One can show that $W$ 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 $\alpha_i$, 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 $\alpha_i$ itself. (To see 2., write

$$\beta = \sum k_i \alpha_i$$

Then

$$S_1(\beta) = \sum k_i \alpha_i - 2\alpha_1 \sum \frac{k_i \langle \alpha_i, \alpha_1 \rangle}{\langle \alpha_1, \alpha_1 \rangle} = \sum_{i>1} k_i \alpha_i + \alpha_1(something) \quad (163)$$

ans similar for the other $S_i$. Therefore if $\beta \neq \alpha_1$, then $\beta$ 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

$$\langle \rho - \alpha_i, \alpha_i \rangle = \langle \rho, -\alpha_i \rangle,$$

$$2\langle \rho, \alpha_i \rangle = \langle \alpha_i, \alpha_i \rangle.$$
which implies that

\[ \rho = \sum_i \Lambda(i) \]  

(164)

Consider now the function

\[ Q(\mu) := \prod_{\alpha > 0} \left( e^{\frac{1}{2} (\alpha, \mu)} - e^{-\frac{1}{2} (\alpha, \mu)} \right) \]  

(165)

We want to see how this transforms under the Weyl group. Consider

\[ Q(S_i \mu) = \prod_{\alpha > 0} \left( e^{\frac{1}{2} (\alpha, S_i \mu)} - e^{-\frac{1}{2} (\alpha, S_i \mu)} \right) = \prod_{\alpha > 0} \left( e^{\frac{1}{2} (S_i \alpha, \mu)} - e^{-\frac{1}{2} (S_i \alpha, \mu)} \right) \]  

(166)

Now \( S_i \) interchanges all the positive roots except itself, whose sign it changes. It follows that

\[ Q(S_i \mu) = -Q(\mu) \]  

(167)

Since \( W \) is generated by the \( S_i \) and \( \text{sign}(S_i) = -1 \), it follows immediately that

\[ Q(\omega(\mu)) = \text{sign}(\omega) Q(\mu) \]  

(168)

for any \( \omega \in W \). Hence \( Q(\mu) \) is a totally antisymmetric function under the Weyl group. There is another totally antisymmetric function under the Weyl group, given by

\[ \tilde{Q}(\mu) := \sum_{\omega \in W} \text{sign}(\omega) e^{\langle \omega(\rho), \mu \rangle} \]  

(169)

and we claim that they coincide:

\[ Q(\mu) = \prod_{\alpha > 0} \left( e^{\frac{1}{2} (\alpha, \mu)} - e^{-\frac{1}{2} (\alpha, \mu)} \right) = \sum_{\omega \in W} \text{sign}(\omega) e^{\langle \omega(\rho), \mu \rangle} \]  

(170)

Since the rhs is the denominator of (154)), this is the so-called \textit{denominator identity}, which we need to derive \textit{Weyl’s 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 \( \beta \) is a sum of distinct positive roots. Since both \( Q \) and \( \tilde{Q} \) are antisymmetric and since \( W \) 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. $\lambda$. The dimension of $V$ is given by $Tr_V(1) = Tr_V(e^0) = \chi_V(0)$. However, this gives $0/0$, and we must use a suitable limit. We choose $\mu = t\rho$ and let $t \to 0$. This gives using (154)

$$
\chi_\lambda(t\rho) = \frac{\sum_{\omega \in W} \text{sign}(\omega) e^{\langle (\omega t)(\lambda + \rho), t\rho \rangle}}{\sum_{\omega \in W} \text{sign}(\omega) e^{\langle (\omega t)\rho, t\rho \rangle}} = \frac{Q(t(\lambda + \rho))}{Q(t\rho)}
$$

Now we can take the limit $t \to 0$ and find

$$
\dim(V_\lambda) = \prod_{\alpha > 0} \frac{\langle (\alpha, \lambda + \rho) \rangle}{\langle (\alpha, \rho) \rangle} \sum_{k_\alpha}^n
$$

This is a very useful formula. To evaluate it, we write each positive root in terms of the simple roots:

$$
\alpha = \sum_i k_\alpha^i \alpha_i
$$

Suppose

$$
\lambda = \sum n_i \Lambda_{(i)}
$$

for $n_i \geq 0$. Then

$$
\dim(V_\lambda) = \prod_{\alpha > 0} \frac{\sum k_\alpha^i (n_i + 1) \langle \alpha_i, \alpha_i \rangle}{\sum k_\alpha^i \langle \alpha_i, \alpha_i \rangle}
$$

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

$$
\dim(V_\lambda) = \prod_{\alpha > 0} \frac{\sum k_\alpha^i (n_i + 1)}{\sum k_\alpha^i}, \quad \mathfrak{g} \in \{A_n, D_n, E_n\}.
$$

Consider some examples. For $su(2)$, there is only one positive root $\alpha$, and the spin $j$ rep has highest weight $n = 2j$. Hence we get $\dim(V_{\text{spin}j}) = (2j + 1)$, 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

$$
\dim(V_\lambda) = \left( \frac{n_1 + 1}{1} \right) \left( \frac{n_2 + 1}{1} \right) \left( \frac{n_1 + n_2 + 2}{2} \right)
$$

For example,

$$
\begin{align*}
\dim(V_{(1,0)}) & = 3 = \dim(V_{(0,1)}), \\
\dim(V_{(1,1)}) & = 8, \\
\dim(V_{(2,0)}) & = 6 = \dim(V_{(0,2)}), \\
\dim(V_{(2,1)}) & = 15 = \dim(V_{(1,2)}), \\
\dim(V_{(3,0)}) & = 10 = \dim(V_{(0,3)}),
\end{align*}
$$

89
Similarly, one easily finds the following formula for the dimension of any $V_{(m_1,...,m_n)}$ of $A_n = su(n+1)$:

$$
\dim V_{(m_1,...,m_n)} = \frac{m_1 + 1}{1} \frac{m_2 + 1}{2} \cdots \frac{m_n + 1}{n} \frac{m_1 + m_2 + 2 m_2 + m_3 + 2}{1} \frac{m_{n-1} + m_n + 2}{2} \cdots \frac{m_1 + m_2 + \cdots + m_n + n}{n}
$$

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 $H_N := a_1^+ a_2^+ a_3^+ \langle 0 \rangle$ with 3 different (species of) creation- and anihilation operators we encountered with the 3-dim. harmonic oscillator. The latter is of course $H_n := V_{\Lambda(1)}^{\otimes_s n}$, 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

$$
\dim V_{(m,0,...,0)} = \frac{(m+1)(m+2)...(m+n)}{12...n} = \binom{m+n}{n} = \dim V_{\Lambda(1)}^{\otimes_s m} \tag{176}
$$

This implies that $V_{n\Lambda(1)} = V_{\Lambda(1)}^{\otimes_s n}$ as claimed earlier.

Let's 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 $\alpha_1, \alpha_2$ with the latter being smaller, The positive roots are $\alpha - 1, \alpha_2, \alpha_1 + \alpha_2, \alpha_1 + 2\alpha_2, \alpha_1 + 3\alpha_2, 2\alpha_1 + 3\alpha_2$. Then one finds (...

$$
\dim(V_{(0,1)}) = 7, \quad \dim(V_{(1,0)}) = 14
$$

(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 Weyl’s character formula (as used above for $su(2)$) to arbitrary simple $\mathfrak{g}$.

The starting point is again (150),

$$
\chi_{V_\lambda \otimes V_\mu} = \chi(V_\lambda) \cdot \chi(V_\mu) = \sum N^\nu_{\lambda\mu} \chi(V_\nu)
$$

(177)

We use again Weyl’s character formula for $\chi(V_\lambda)$ and $\chi(V_\nu)$. Noting that the denominator of Weyl’s formula is independent of the representation, we get

$$
\sum_{\omega \in \mathcal{W}} \text{sign}(\omega) e^{\omega(\lambda + \rho)} \cdot \chi(V_\mu) = \sum_{\omega \in \mathcal{W}} \text{sign}(\omega) \sum N^\nu_{\lambda\mu} e^{\omega(\nu + \rho)}
$$

Plugging in

$$
\chi(V_\mu) = \sum_{\kappa} \text{mult}_\mu(\kappa) e^\kappa
$$

(recall that $\text{mult}_\mu(\kappa)$ is the multiplicity of the weight $\kappa$ in $V_\mu$), we get

$$
\sum_{\omega \in \mathcal{W}} \text{sign}(\omega) \sum_{\kappa} \text{mult}_\mu(\kappa) e^{\omega(\lambda + \rho)} e^\kappa = \sum_{\omega \in \mathcal{W}} \text{sign}(\omega) \sum N^\nu_{\lambda\mu} e^{\omega(\nu + \rho)}
$$

Now recall that $\text{mult}_\mu(\kappa) = \text{mult}_\mu(\omega(\kappa))$, hence we can replace $\sum_{\kappa}$ by $\sum_{\omega(\kappa)}$, and get

$$
\sum_{\omega \in \mathcal{W}} \text{sign}(\omega) \sum_{\omega(\kappa) \in V_\mu} \text{mult}_\mu(\kappa) e^{\omega(\lambda + \rho + \kappa + \rho)} = \sum_{\omega \in \mathcal{W}} \text{sign}(\omega) \sum N^\nu_{\lambda\mu} e^{\omega(\nu + \rho)}.
$$

(178)

Now both sides are formal linear combinations of exponentials $e^\eta$ of some weights $\eta$. These $e^\eta$ are linearly independent for different weights $\eta$, and we can compare their coefficients on both sides. Furthermore, recall that for any given $\eta$, 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 $e^\eta$ on both sides with $\eta \in P_+$ in the fundamental Weyl chamber.

Now consider the set of weights

$$
\eta := \lambda + \kappa + \rho \quad \kappa \in V_\mu.
$$

Let us assume first that all of these weights $\eta$ are in the interior of the fundamental Weyl chamber (not on the boundary), this often happens for small $\mu$. 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:

$$
\text{mult}_\mu(\kappa) e^{\omega(\lambda + \kappa + \rho)} = N^\nu_{\lambda\mu} e^{\omega(\nu + \rho)}
$$
hence $\text{mult}_\mu(\nu - \lambda) = N_{\lambda\mu}^\nu$, or equivalently

\[ N_{\lambda\mu}^{\lambda+\nu} = \text{mult}_\mu(\nu) \quad \text{if } \lambda + \kappa + \rho \in P_+ \quad \forall \kappa \in V_\mu. \]  

(179)

Hence we know $N_{\lambda\mu}^{\lambda+\nu}$ once we know $\text{mult}_\mu(\nu)$ (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

$$ V_{(1,1)} \otimes V_{(1,0)} = V_{(2,1)} \oplus V_{(0,2)} \oplus V_{(1,1)}. $$

(180)

or

$$ (8) \otimes (3) = (15) \oplus (6) \oplus (3) $$

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 $\omega \in W$ such that $\omega(\lambda + \kappa + \rho) \in P_+$. Then this gives a contribution

$$ \text{sign}(\omega) \text{e}^{\omega(\lambda+\kappa+\rho)} \text{mult}_\mu(\kappa) $$

(181)

from the lhs, which if summed up must match some $N_{\lambda\mu}^\nu \text{e}^{\nu+\rho}$ on the rhs. Therefore the weights are related by

$$ \omega(\lambda + \kappa + \rho) = \nu + \rho $$

or

$$ \kappa = \omega^{-1}(\nu + \rho) - \rho - \lambda. $$

Therefore

$$ N_{\lambda\mu}^\nu = \sum'_{\omega \in W} \text{sign}(\omega) \text{mult}_\mu(\omega^{-1}(\nu + \rho) - \rho - \lambda) $$

(182)

where $\sum'$ denotes the sum over those $\omega \in W$ such that $\omega^{-1}(\nu + \rho) - \rho - \lambda$ is a weight of $V_\mu$ (or set $\text{mult}_\mu = 0$ 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

$$ \eta := \lambda + \kappa + \rho \quad \kappa \in V_\mu. $$

Assume that $\eta$ is not on the boundary of any Weyl chamber. Then there is a unique $\omega \in W$ such that $\omega(\eta) \in P^\circ_\mu$ is inside the fundamental Weyl chamber. In this case we have a contribution of

$$\text{sign}(\omega)\text{mult}_\mu(\kappa)$$

to the Littlewood-Richards coefficient $N^{\nu}_{\lambda\mu}$ with $\nu = \omega(\eta) - \rho$. All these contributions have to be summed for all $\kappa \in V_\mu$. This computes all the $N^{\nu}_{\lambda\mu}$ at once (If $\lambda + \kappa + \rho$ is on the boundary of some Weyl chamber, this $\kappa$ must be ignored since there is no $\omega \in W$ which maps it in the interior of the fundamental one).

This is best done by drawing the weights of $V_\mu$ on top of $\lambda + \rho$, which gives the relevant $\kappa$ and their $\omega$ with possible sign.

Example 1: $(8) \otimes (8)$

Claim:

$$V_{(1,1)} \otimes V_{(1,1)} = V_{(2,2)} \oplus V_{(3,0)} \oplus V_{(0,3)} \oplus 2V_{(1,1)} \oplus V_{(0,0)}$$

(183)

Example 2: $(8) \otimes (6)$

Claim:

$$V_{(1,1)} \otimes V_{(2,0)} = V_{(3,1)} \oplus V_{(1,2)} \oplus V_{(0,1)} \oplus V_{(2,0)}$$

(184)
References


