Notes On Group Theory

George Thompson

Third Draft: Still Very Preliminary!

CONTENTS

CHAPTER 1

INTRODUCTION

These lectures start with three basic examples of Lie algebras. These examples are $U(1)$, $SU(2)$ and $SU(3)$. While we only deal with the first few chapters, most of the statements are made in the generality that is required for dealing with general matrix groups. The reason for insisting on these three examples is, on the one hand, that they are the most important for the particle physics and field theory courses and, on the other hand, as the simplest of the classical Lie groups, they give the diploma student the most 'painless' entry into this subject.

Further on in these the notes we will be concerned with the general theory of Lie algebras. This is necessarily more abstract. However, the explicit understanding of $SU(2)$ obtained in the first part is essential for understanding Lie groups and algebras in general as it is used essentially in the general theory. My advice here is simple: understand $SU(2)$. We will infact apply our knowledge of the representation theory of $SU(2)$ to

Throughout the basic idea to get is that there are groups and algebras and that they act on things (including themselves). The only way you will understand this and really feel it is if you do all the exercises, and there are many of those!

In the second part we will deal with general matrix groups and their Lie algebras. We will begin a study of the structure of general Lie algebras and then come back to those associated to the general matrix groups that we have already studied.

The main tool in this part the Killing-Cartan form. This form is, by definition, non-degenerate on semi-simple Lie algebras so we also introduce the concepts of semi-simple and simple Lie algebras.

There are only two conscious sources for these notes. The first is a book by Cahn which is for physicists and the second is a set of lecture notes by Ziller for mathematicians. Where I use these I have tried to achieve a balance both in level as well as in notation.

CHAPTER 2

BASICS OF GROUPS

In this chapter we introduce the concept of a group and then we give some examples. Not only do we see groups as abstract objects we also come across matrix realizations of them, that is of representations of them, though a proper definition of a representation will have to wait till a later chapter.

2.1 Definitions

Definition 2.1. A group G is a set of elements (this can be finite or infinite, countable or not)

$$G = \{g_1, g_2, \dots\}$$

with a group 'product' rule satisfying the following 4 conditions

1. **Closure:** $\forall\, g,\ g' \in G$ the product is in G, $g \circ g' \in G$.

2. **Associativity:** $\forall\, g,\, g',\, g'' \in G$, $(g \circ g') \circ g'' = g \circ (g' \circ g'')$.

3. **Identity Element:** There exists a unique element of G called the identity element, denoted by e, and such that $\forall\, g \in G$, $e \circ g = g \circ e = g$.

4. **Inverse Element:** For every element $g \in G\ \exists\, g^{-1} \in G$ s.t. $g \circ g^{-1} = g^{-1} \circ g = e$.

Definition 2.2. If $\forall g,\ g' \in G,\ g \circ g' = g' \circ g$ the group is called an Abelian group, otherwise it is called a non-Abelian group

That is the formal definition. The product $\circ$ need not be multiplication but any rule that satisfies all of the conditions. To make the definition palatable here are some examples.

Examples of Finite Groups:

1. $\mathbb{Z}_n \equiv \{1, \omega, \ldots, \omega^{n-1}\}$ where $\omega = \exp(2\pi i/n)$ (an n-th root of unity). The product rule $\circ = .$ where $.$ stands for usual multiplication of complex numbers. This is an Abelian group of n elements.

2. The permutation or symmetry group S_n of n objects is defined as follows. Consider n objects $(x_1, \ldots, x_n)$ in this order the elements of S_n are all the possible permutations of these objects. S_n has $n!$ elements. Lets see how this works for small n. If $n = 1$, there is only one possible ordering (x_1) and that is it, hence $S_1 = e$. If $n = 2$, the two possible orderings are (x_1, x_2) and (x_2, x_1). What are the elements of S_2? They are

$$e(x_1, x_2) = (x_1, x_2), \quad g_{(12)}(x_1, x_2) = (x_2, x_1),$$

so that $S_2 = \{e, g_{(12)}\}$. Note that the inverse of $g_{(12)}$ is $g_{(12)}$ itself since

$$g_{(12)}\left(g_{(12)}(x_1, x_2)\right) = g_{(12)}(x_2, x_1) = (x_1, x_2) = e(x_1, x_2).$$

This implies that $S_2 = \mathbb{Z}_2$. Abstractly $\mathbb{Z}_2$ is made up of 2 elements $\{e, \omega\}$ with the product rule $\omega \circ \omega = e$ (the others are obvious), while S_2 has 2 elements $\{e, g_{(12)}\}$ and product $g_{(12)} \circ g_{(12)} = e$. The groups are the same if we simply change the name of $g_{(12)}$ to ω. We see that what appear, by definition, to be different groups at the outset can turn out to be the same group.

Now let $n = 3$. The group $S_3 = \{e, g_{(12)}, g_{(13)}, g_{(23)}, g_{(123)}, g_{(132)}\}$. The notation (123) means that the element in the first position goes to the second. The element that was in the second goes to the third position and the element that was in the third goes to the first (cyclic),

$$g_{(123)}(x, y, z) = (z, x, y), \quad g_{(132)}(x, y, z) = (y, z, x).$$

Here we learn that there is a big difference between the space a group acts on, 3 objects, and the number of elements of the group which, in this case, is 6 elements

Note the different way in which we defined $\mathbb{Z}_n$ and S_n. In the former we gave the elements directly, as the n-th roots of unity, in the latter we gave the elements by their action on another set of objects, called a 'representation'.

Exercise 2.1. Show, for S_3 that $g_{(132)} = g_{(23)} \circ g_{(12)}$ and $g_{(123)} = g_{(12)} \circ g_{(23)}$.

Exercise 2.2. Is S_3 an Abelian group?

Example of an Infinite Discrete Group:

The example is the group of integers $\mathbb{Z} \equiv \{\text{all integers}\}$. This time the product $\circ$ is addition $+$. The identity element $e := 0$ and the inverse of $n \in \mathbb{Z}$ is $-n$. There is a second operation defined on $\mathbb{Z}$ which is usual multiplication, however, $\mathbb{Z}$ is **not** a group with respect to this product since the inverse of n is $1/n$ but $1/n \notin \mathbb{Z}$.

Examples of Continuous Groups:

These groups have a continuous number of elements. I will give two examples here. The course will be about, more general examples of, groups of this type.

1. The real line $\mathbb{R}$ is a group under addition, '$\circ = +'$. Since every number is an element we have an (uncountable) infinity of elements. This group has $\mathbb{Z}$ as a subgroup. $\mathbb{R}$ is an Abelian group since $a + b = b + a\ \forall a,\, b \in \mathbb{R}$.

2. The one dimensional unitary group, $U(1) = \{\exp(i\theta) \mid 0 \le \theta < 2\pi\}$. The product rule is multiplication in $\mathbb{C}$, if $g(\theta) = \exp(i\theta)$ and $g(\theta') = \exp(i\theta')$ then $g(\theta) \circ g(\theta') \equiv g(\theta).g(\theta') == \exp i\theta. \exp i\theta' = \exp(i(\theta + \theta')) = g(\theta + \theta')$.

2.2 A Little on Representations

We noted above that that the group $\mathbb{Z}_2$ is correctly expressed as the group of two elements $(e,\, \omega)$ with the multiplication rules

$$e \circ e = e, \quad e \circ \omega = \omega \circ e = \omega, \quad \omega \circ \omega = e \tag{2.2.1}$$

One may exhibit these rules, for a fiinite group, in the form of a table:

$$\begin{array}{c||c|c} * & e & \omega \\ \hline\hline e & e & \omega \\ \hline \omega & \omega & e \end{array} \tag{2.2.2}$$

with the rule (i-th element in the left column) $\circ$ (j-th element of the top row) = (i-j-th entry in the table). In any case, from these rules we see that e is the identity element and ω is its own inverse. In the first example of finite groups we did not use this abstract notation rather we defined the group $\mathbb{Z}_2 \equiv \{1,\, -1\}$ with multiplication $\circ$ being multiplication in the real numbers. What is the relationship between the two?

To answer the question we complicate things a little. Consider the 2- dimensional real vector space $\mathbb{R}^2$. On this space 2×2 matrices act. We let

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } \omega = -e \tag{2.2.3}$$

then these matices satisfy the equations (2.2.1) when $\circ$ is understood to be matrix multiplication. We say that these matrices are a **representation** of the group $\mathbb{Z}_2$. There is nothing unique about this, for example, still in dimension 2 we note that

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \omega = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{2.2.4}$$

also satisfy (2.2.1) and part of our problem is to work out when two representations are in a sense equivalent. In the examples just noted equivalence would mean that they are related by a similarity transformation (since this is just a re-arrangement of the basis of the vector space and the matrices are symmetric). One can easily see that they are not equivalent since taking the trace is an operation which is **invariant** under similarity transformations and the trace of ω in (2.2.3) is -2 while that in (2.2.4) is 0.

The 'representation' (2.2.4) is actually S_2. To see this think of the 2-tuple (x_1, x_2) as a vector then the matrices e and ω act as

$$e.\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{and} \quad \omega.\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \tag{2.2.5}$$

which is the way we saw that S_2 acts.

You should try to come up with representation of $\mathbb{Z}_2$ as 3×3 matrices, just to be sure you have the idea straight.

So the answer to our question is that we have the abstract group given by some definition and then we have the possibility to represent the group with matrices (including 1×1 matrices or numbers). You can also find other ways to represent a group including as complex matrices or even as differential operators.

The important point is that groups also **act** on things. Certainly, they act on themselves (by the product rule) but we see that they can, when we think of representing them in terms of matrices, act on vectors in a vector space (and so their representations act on the vector space).

We will not be concerned any more with finite groups or with the representations of finite groups. However, the notion of the definition of the group, what groups act on, its representations, their equivalnces and invariants will be recurring themes in these notes.

CHAPTER 3

1-DIMENSIONAL LIE GROUPS

The 1 in the title refers to one real dimension. The groups that we will consider here are continuous groups and some are compact while others are not.

3.1 The Groups $U(1)$ and $SO(2, \mathbb{R})$

The groups $\mathbb{R}$ and $U(1)$ are one-dimensional continuous groups. They are one-dimensional because they depend on one parameter and they are continuous since that parameter is. The easiest way to convince yourself of this is to draw the groups. $\mathbb{R}$ is easy to draw, its just the straight line. What about $U(1)$? Well the easiest place to draw it is in the complex plane $\mathbb{C}$, with complex coordinate z. Now any complex number can be written in polar form as

$$z = r\, e^{i\theta}, \quad 0 \leq r < \infty, \quad 0 \leq \theta < 2\pi.$$

The elements of $U(1)$ are all those points in $\mathbb{C}$ with $r = 1$. This is the definition of a circle.

Let us introduce another group called $SO(2, \mathbb{R})$. There is an intrinsic way to define it which also has a nice geometrical interpretation (which is where it comes from). $SO(2, \mathbb{R})$ is defined to be all of those 2×2 matrices with real coefficients, M, which satisfy the following conditions

$$\begin{aligned} M^T.M &= \mathbb{I}_2, \\ \mathrm{Det}\, M &= 1. \end{aligned} \tag{3.1.1}$$

We read from the notation $SO(2, \mathbb{R})$ that the $\mathbb{R}$ means real coefficients, the 2 means a 2×2 matrix, that O stands for orthogonal which is the first condition in (3.1.1) while S stands for special which means the determinant is unity and is the second condition in (3.1.1). This is a baby version of the rotation group that you are probably familiar with. The geometry is the following. Consider a non-zero vector, $\mathbf{v}$, in $\mathbb{R}^2$ then the group of rotations is just that: any rotation that is possible on the vector $\mathbf{v}$ about the origin. When one rotates v through an angle θ one gets a new vector, call it $\mathbf{v}(\theta)$.

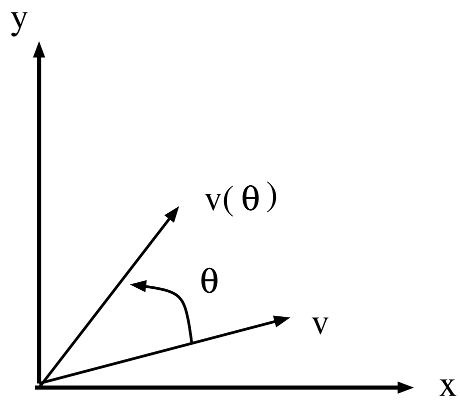


Figure 3.1: The action of the representation of group element $g(\theta)$ on the vector $\mathbf{v}$ to obtain $\mathbf{v}(\theta)$.

There is always a matrix $M(\theta)$ such that $\mathbf{v}(\theta) = M(\theta).\mathbf{v}$. Since a rotation does not change the length of $\mathbf{v}$ we must have $\mathbf{v}^T(\theta).\mathbf{v}(\theta) = \mathbf{v}^T.\mathbf{v}$, so that $M^T.M = \mathbb{I}_2$ (as we demand the formula works for all $\mathbf{v}$). At this point we can only conclude that $\mathrm{Det}\, M = \pm 1$. If we stop here we still get a group, it is called $O(2, \mathbb{R})$, since we drop the 'special' condition S.

Exercise 3.1. Show that $M \in O(2, \mathbb{R})$ with $\mathrm{Det}\, M = -1$ corresponds to a rotation plus a reflection.

The rotations in $\mathbb{R}^2$ are parameterized by a circle and so we can expect that $U(1) \simeq SO(2, \mathbb{R})$. Lets establish this. We have described $SO(2, \mathbb{R})$ as it acts on a vector in $\mathbb{R}^2$ and likewise an element $g(\theta)$ of $U(1)$ acts on $z \in \mathbb{C}$ by multiplication $g(\theta)z$. Lets see how $U(1)$ acts on real objects. Write $z = x + iy$, so that

$$g(\theta).z = e^{i\theta}.(x + iy) = (\cos\theta x - \sin\theta y) + i(\sin\theta x + \cos\theta y) \tag{3.1.2}$$

which we can write as a matrix equation

$$g(\theta).\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix},$$

that is

$$g(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{3.1.3}$$

It is an exercise to show that all $M \in SO(2, \mathbb{R})$ can be written as the matrix in (3.1.3), which shows the equivalence of the two groups $SO(2, \mathbb{R})$ and $U(1)$.

Exercise 3.2. Prove that any element of $SO(2, \mathbb{R})$ takes the form given in (3.1.3).

It is important to emphasize something here. We said that a rotation does not change the length of a vector. Another way of saying that is that the length of a vector is **invariant** under rotations. But you should now ask: which length? Well if $\mathbf{v} = (x, y)$ the length that is invariant under $O(2, \mathbb{R})$ is the Euclidean length, namely

$$|\mathbf{v}|^2 = v^i \delta_{ij} v^j = x^2 + y^2 \tag{3.1.4}$$

and the Kronecker delta, δ_{ij}, is the **Euclidean metric**. From this we get another way of expressing definition of the group $O(2, \mathbb{R})$.

Definition 3.1. $O(2, \mathbb{R})$ is the group of all linear transformations on $\mathbb{R}^2$ which preserves the Euclidean metric.

Before leaving these groups we note that we have been describing them by their action on something. A quick look at (3.1) shows us that this is the case as we are describing $SO(2, \mathbb{R})$ as the group which rotates the vector $\mathbf{v}$. Notice that if we go through all possible angles we would be drawing a circle centred at the origin and whose radius is the length of $\mathbf{v}$. This circle is the (image of the) group itself. Only if the length of $\mathbf{v}$ is zero, so that $\mathbf{v}$ is the zero vector, do we not get a circle since the origin does not move under a rotation (and we say that it is fixed under the group action).

3.2 The Groups $SO(1, 1, \mathbb{R})$ and $\mathbb{R}_+$

In special relativity the invariant distance is proper distance. To make life simple lets think that space-time is two dimensional, so that we have one time coordinate x^0 and one space coordinate x^1. The metric we now use is **not** the Euclidean metric δ but rather the Minkowski metric η. The position vector is denoted x^μ and its length is defined to be

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

You should really pay attention to the fact that space-time is still $\mathbb{R}^2$ it is just that we have put a different metric on it. This means that we have decided to measure length in a different way to the more usual Euclidean one. But the space always remains the same.

Write the metric as

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

Obviously $O(2, \mathbb{R})$ does not preserve this metric. We define $O(1, 1, \mathbb{R})$ to be those linear transformations which preserve the Minkowski metric, or put another way, which preserve the length of a vector. Let $M \in O(1,1,\mathbb{R})$ with

$$M = \begin{pmatrix} p & q \\ r & s \end{pmatrix},$$

and it must satisfy

$$M^T.\eta.\,M = \eta, \tag{3.2.1}$$

or

$$p^2 - r^2 = 1, \quad pq - rs = 0, \quad q^2 - s^2 = 1.$$

Also we get for free that $\det M = \pm 1$. If we restrict our attention to $SO(1,1)$, so that $\det M = 1$, the solution to the equations is

$$M = \pm \begin{pmatrix} \cosh\lambda & \sinh\lambda \\ \sinh\lambda & \cosh\lambda \end{pmatrix} \tag{3.2.2}$$

for $\lambda \in \mathbb{R}$.

Exercise 3.3. Show that (3.2.2) is the general matrix for an element of $SO(1,1,\mathbb{R})$.

Hence we see that the difference between $SO(2,\mathbb{R})$ and $SO(1,1,\mathbb{R})$ is that trigonometric functions pass to hyperbolic functions. In the process the group goes from being **compact** (which means that every entry in the matrix is bounded as both cos and sine are) to a **non-compact** group (as cosh and sinh can be as big as we like).

Now we need another notion namely

Definition 3.2. A subgroup, H, of a group G, written $H \subseteq G$, is a subset of the set G such that the elements of H form a group with respect to the product rule in G.

Example of a Subgroup: Let $\mathbb{R}$ be the group of real numbers under addition, then $\mathbb{Z}$, the group of integers under addition, is a subgroup.

A Non-Example of a Subgroup: The complex plane (some mathematicians call it a complex line), $\mathbb{C}$, is a group under addition. The elements of $\mathbb{Z}_n$ are contained in $\mathbb{C}$, but it is **not** a subgroup of $\mathbb{C}$ because the product rule for $\mathbb{Z}_n$ is multiplication and not addition.

The Lorentz group has an important subgroup. This is called the **proper Lorentz group**, $SO(1,1,\mathbb{R})_+$, and corresponds to those matrices which also satisfy the condition

$$M_0{}^0 \geq 1.$$

Such transformations do not change the sign of the time coordinate x^0. This means the matrices we are interested in are those with the plus sign in (3.2.2) as $\cosh \geq 1$.

In the case of $SO(2,\mathbb{R})$ it was convenient to take complex combinations to pass to $U(1)$. This works because $e^{i\theta} = \cos\theta + i\sin\theta$ (and there is an explicit factor of i). For

$SO(1,1,\mathbb{R})_+$ we should take real linear combinations of x^0 and x^1, since in this case $e^\lambda = \cosh\lambda + \sinh\lambda$. With this in mind we define **light-cone** coordinates

$$x^\pm = x^0 \pm x^1, \tag{3.2.3}$$

and these transform as

$$x^\pm \to e^{\pm\lambda}\, x^\pm. \tag{3.2.4}$$

Exercise 3.4. Prove that the light cone coordinates transform as in (3.2.4).

This shows us that $x^\pm$ transform by scaling. Note that e^λ defines a group namely the group under multiplication of all positive reals which is denoted by $\mathbb{R}_+$. So just as $SO(2,\mathbb{R}) \simeq U(1)$, we have shown that $SO(1,1,\mathbb{R})_+ \simeq \mathbb{R}_+$.

We started from the fact that we wanted to preserve the Minkowski metric, which in light cone coordinates reads

$$|x|^2 = x^+\, x^-$$

and it is obvious that the scalings (3.2.4) preserve the metric.

Exercise 3.5. Is $\mathbb{R}_-$, the set of all negative reals, a group under multiplication?

3.3 Things to do and Keep in Mind

An important message to keep from this chapter is that a group is both a **set** as well a **product rule**. To equate $U(1)$ and $SO(2,\mathbb{R})$ we showed that as **sets** they are the same, namely both are the circle S^1 and that the product rule of one goes over, naturally, to the product rule of the other. Similar things can be said of $SO(1,1,\mathbb{R})_+$ and its equality with $\mathbb{R}_+$.

We have considered a number of groups in this chapter and in the previous one. It is very good practice to check that they really are groups. There is only one way to do that. You do it by checking the 4 basic properties that go into the definition of a group.

CHAPTER 4

INFINITESIMAL TRANSFORMATIONS, SYMMETRIES AND CONSERVED CHARGES

This chapter is designed to convince the reader that the mathematics that we have introduced thus far is important for understanding physics. In particular, since groups act on spaces, we use this action to discover conserved quantities.

4.1 Close to the Identity

We will now look at the continuous group $U(1)$ near the identity element. For $U(1)$ the identity element is 1 (or $\mathbb{I}_2$ if you think of it at $SO(2,\mathbb{R})$). Let $g(\epsilon) \in U(1)$ and write

$$g(\epsilon) = \exp(i\epsilon) = 1 + \epsilon\, T + O(\epsilon^2) \tag{4.1.1}$$

where ϵ is an infinitesimal parameter (meaning as small as you like) and $O(\epsilon^2)$ means of order ϵ^2. T is called the generator of an infinitesimal $U(1)$ transformation. Indeed from our description of $U(1)$ we have that if $\theta = \epsilon$ then the generator T is $i = \sqrt{-1}$. Notice that (4.1.1) is essentially a Taylor series expansion about the identity,

$$g(\theta) = g(0) + \theta \left.\frac{dg}{d\theta}\right|_{\theta=0} + O(\theta^2). \tag{4.1.2}$$

The number of generators is equal to the number of parameters in the group. Since $U(1)$ is 1 dimensional we only have one generator.

If we were working with $SO(2,\mathbb{R})$ the generator would come from expanding the matrix in (3.1.3) out to first order

$$g(\theta) = \mathbb{I}_2 + \theta \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + O(\theta^2) \tag{4.1.3}$$

In this case the generator is given by

$$T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{4.1.4}$$

and we notice that regardless of whether we think of T as this matrix or as i it has the property that $T^2 = -\mathbb{I}$.

Lets turn our attention to $\mathbb{R}$. Before we showed that any group element in $U(1)$ could be written in matrix form as $g(\theta)$ where θ parameterizes the group (it is a circle). If we want to do the same for $\mathbb{R}$ we would write $g(\alpha) = \alpha$ where α parameterizes $\mathbb{R}$. Suppose we want to 'expand' any group element in $\mathbb{R}$ around the identity element, namely 0, we would find

$$\begin{aligned} g(\alpha) &= g(0) + \alpha \left.\frac{dg}{d\alpha}\right|_{\alpha=0} \\ &= \alpha T \end{aligned}$$

without any corrections and where the generator T is given by

$$T = \left.\frac{dg}{d\alpha}\right|_{\alpha=0} = 1. \tag{4.1.5}$$

The generators of groups are a central object of study in Lie groups and Lie algebras. They are at the heart of a mathematical understanding of the structure of Lie groups and algebras. However, what is good mathematics is usually also good physics and to give a feel of this we turn to the relationship between symmetries and conserved quantities.

Both in classical physics as well as in quantum physics the role of conserved charges are very important. Charge conservation and momentum conservation are two of the building blocks of our understanding of basic physics. So important are these that there is a fundamental theorem in physics which associates symmetries with conserved charges. This is, one of, Emmy Noether's many theorems and in the physics world it is just referred to as Noether's theorem[1].

Theorem 4.1. To every global continuous symmetry of the Lagrangian there is an associated conserved charge.

Proof:
The Lagrangian (density), $\mathcal{L}(q_i, \dot{q}_i)$ is a function of positions, q_i and velocities $\dot{q}_i$. A symmetry transformation, f, corresponds to a map of the coordinates and an induced

[1]The statement that follows is slightly stronger than necessary, however, I want to keep things as simple as possible. In the Relativistic quantum mechanics course you will see this again in the Hamiltonian formulation. In your field theory course you will see it for a third time stated in terms of the action.

map on the velocities,

$$f : q_i \to f_i(q; s^a), \quad f : \dot{q}_i \to \dot{f}_i(q, s^a) = \sum_j \frac{\partial f_i}{\partial q_j}(q_i; s^a)\, \dot{q}_j. \tag{4.1.6}$$

Since f is supposed to come from a global continuous group it will depend on some continuous parameters, all of which have been denoted by s^a. Also, think of $s^a = 0$ as corresponding to the identity element in the group, so to the identity map, $f(q; 0) = q$. To say that the Lagrangian is invariant is to say that

$$\mathcal{L}(q, \dot{q}) = \mathcal{L}(f_i(q; s^a), \dot{f}_i(q, s^a)). \tag{4.1.7}$$

Since the left hand side of (4.1.7) has no parameter dependence we can differentiate with respect to s to get zero, so

$$\begin{aligned} 0 &= \frac{d}{ds}\mathcal{L}(f_i(q; s^a), \dot{f}_i(q; s^a))\bigg|_{s=0} \\ &= \sum_j \frac{df_j}{ds^a}\bigg|_{s=0} \frac{\partial \mathcal{L}}{\partial q_j}(q_i, \dot{q}_i) + \sum_i \frac{d}{ds^a}\frac{df_j}{dt}\bigg|_{s=0} \frac{\partial \mathcal{L}}{\partial \dot{q}_j}(q_i, \dot{q}_i) \\ &= \frac{d}{dt}\left(\sum_j \frac{df_j}{ds^a}\bigg|_{s=0} \frac{\partial \mathcal{L}}{\partial \dot{q}_j}(q_i, \dot{q}_i)\right). \end{aligned}$$

To pass from the second to the third line one makes use of the equation of motion

$$\frac{\partial \mathcal{L}}{\partial q_j}(q_i, \dot{q}_i) = \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j}(q_i, \dot{q}_i).$$

From this we deduce that the conserved charges Q_a, are

$$Q_a = \left(\sum_j \frac{df_j}{ds^a}\bigg|_{s=0} \frac{\partial \mathcal{L}}{\partial \dot{q}_j}(q_i, \dot{q}_i)\right). \tag{4.1.8}$$

□

From our discussion above we see that the objects

$$\frac{df}{ds^a}\bigg|_{s=0}$$

are the generators of the group **as they act on the coordinates**. So they play a very important role in the study of symmetries and conserved charges. We will return to these in later lectures.

Examples:

1. Consider the two dimensional Lagrangian

$$L(\dot{x}, \dot{y}) = \frac{m}{2}\left(\dot{x}^2 + \dot{y}^2\right), \tag{4.1.9}$$

and the $SO(2)$ transformation $g(\theta)$ in (3.1.3). From (4.1.2) we have

$$\left.\frac{dg}{d\theta}(\theta)\right|_{s=0} = T.$$

The conserved charge (4.1.8) in this case is

$$Q = m(x\,\dot{y} - y\,\dot{x}) = x\,p_y - y\,p_x,$$

where the momenta are

$$p_j = \frac{\partial \mathcal{L}}{\partial \dot{q}_j}(q_i, \dot{q}_i).$$

So this shows us that the conserved charge is the 2-dimensional angular momentum.

2. Consider the same Lagrangian as before (4.1.9). But now we take the group of translations in x and y. This group is $\mathbb{R} \times \mathbb{R}$. That is 2 copies of the group $\mathbb{R}$. It acts by addition as before. If $(a, b) \in \mathbb{R} \times \mathbb{R}$ and $(a', b') \in \mathbb{R} \times \mathbb{R}$ the product rule is $(a, b) \circ (a', b') = (a + a', b + b')$. The action on the coordinates (x, y) is the same,

$$(a, b).(x, y) = (x + a, y + b)$$

There will be two generators T_1 and T_2 both of which are 1 as follows from (4.1.5). If we now determine the conserved charges we find

$$Q_1 = p_x, \quad Q_2 = p_y, \tag{4.1.10}$$

so that translation invariance gives rise to momentum conservation!

The example above shows that we need a definition for products of groups. Here it is

Definition 4.1. A **Direct Product** $G_1 \times G_2$ of two groups, G_1 and G_2, is defined as

$$G_1 \times G_2 = \{(g_1, g_2)|\ g_1 \in G_1,\ g_2 \in G_2\}$$

with the product rule

$$(g_1, g_2) \circ (g_1', g_2') = (g_1 \circ_1 g_1', g_2 \circ_2 g_2')$$

where $\circ_1$ and $\circ_2$ are the product rules on G_1 and G_2 respectively.

Exercise 4.1. Write the Lagrangian (4.1.9) as

$$L = \frac{m}{2}\,\dot{\bar{z}}\dot{z}$$

where $z = x + iy$, $\overline{z} = x - iy$, and redo the above symmetry arguments in the examples not in the language of $SO(2, \mathbb{R})$ and $\mathbb{R} \times \mathbb{R}$ but rather in that of $U(1)$ and $\mathbb{C}$ (you will have to define the additive group structure).

CHAPTER 5

THE GROUPS $SU(2)$ AND $SO(3, \mathbb{R})$

The two groups in the title are perhaps the most widely used in physics. In this chapter we will define these groups and understand what they are as manifolds (smooth spaces), so as sets. The product rule will also become evident thus giving us groups as advertised.

5.1 $SU(2)$, its Parameterization and its Lie Algebra

Definition 5.1. $SU(2)$ is the group of 2×2 matrices with complex coefficients which satisfy

$$U\,.\,U^{\dagger} = \mathbb{I}_2, \quad \mathrm{Det}\,U = 1,$$

$\forall\, U \in SU(2)$.

You can understand the name of the group by reading it from right to left. 2 tells us that it is a 2×2 matrix group, U tells us that it is unitary and S tells us that it is special, that is the determinant is one.

Exercise 5.1. Show that such matrices do form a group.

We have the definition so that we can determine the group elements, or how the group is parameterized. Write

$$U = \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix} \tag{5.1.1}$$

where the z_i are complex numbers. From the definition we deduce that

$$\begin{aligned} z_1\overline{z}_1 + z_2\overline{z}_2 &= 1 \\ z_1\overline{z}_3 + z_2\overline{z}_4 &= 0 \\ z_3\overline{z}_3 + z_4\overline{z}_4 &= 1 \end{aligned} \tag{5.1.2}$$

from the unitarity condition and

$$z_1 z_4 - z_2 z_3 \quad = \quad 1 \tag{5.1.3}$$

from the fact that the determinant is one. The space of all z_i which satisfy the above equations is the space of all $SU(2)$ elements. In order to find out what that space is we have to solve these equations.

Multiply (5.1.3) by $\overline{z}_4$ to get

$$z_1 z_4 \overline{z}_4 - z_2 \overline{z}_4 z_3 = \overline{z}_4$$

and use the last equation in (5.1.2) to arrive at

$$z_1 - z_3(z_1 \overline{z}_3 + z_2 \overline{z}_4) = \overline{z}_4$$

but now the second equation of (5.1.2) allows us to deduce that

$$z_1 = \overline{z}_4.$$

A similar exercise allows us to deduce that $z_3 = -\overline{z}_2$ (you can do this by multiplying (5.1.3) with $\overline{z}_2$ and make use of the first of (5.1.2) then of the second). Notice that the only equation we now need to satisfy is (5.1.3) but written as

$$z_1 \overline{z}_1 + z_2 \overline{z}_2 \quad = \quad 1. \tag{5.1.4}$$

We conclude that any $U \in SU(2)$ is a matrix of the form

$$U = \begin{pmatrix} z_1 & z_2 \\ -\overline{z}_2 & \overline{z}_1 \end{pmatrix} \tag{5.1.5}$$

such that (5.1.4) holds.

Now we can do some geometry. Write the complex numbers in terms of real numbers

$$z_1 = x_1 + i x_2, \quad z_2 = x_3 + i x_4,$$

then the determinant is one condition (5.1.4) is

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1. \tag{5.1.6}$$

This is fantastic, it is the equation for a unit radius 3-dimensional sphere. It is analogous to the equations for the 1-dimensional sphere, S^1, which we normally call a circle, $x^2 + y^2 = 1$ and for the 2-dimensional sphere S^2, $x^2 + y^2 + z^2 = 1$. Likewise, we denote the 3-sphere by S^3. Quite generally a unit radius n-dimensional sphere is defined in $n+1$-dimensional Euclidean space, $\mathbb{R}^{n+1}$, by the equation

$$\sum_{i=1}^{n+1} x_i^2 = 1. \tag{5.1.7}$$

(To see if you have the ideas straight what is the 0-dimensional sphere S^0?)

In any case what we have found is that given a point on the 3-sphere, that is given a point (x_1, x_2, x_3, x_4) satisfying (5.1.6), we are really given a group element (5.1.5) in $SU(2)$. So as a space $SU(2)$ is S^3 which means that parameterizing $SU(2)$ amounts to parameterizing S^3.

We have seen that as a space $U(1)$ is S^1 and now we have that $SU(2)$ as a space is S^3. Are other spheres groups also? Unfortunately, the answer is no, S^2, for example, is not a group. Other spheres do play an important role, not as groups but as 'cosets' as we will see later on in this course.

Exercise 5.2. Show that $SU(2)$ is a non-Abelian group.

Exercise 5.3. Every point on S^3 (5.1.6) corresponds to an element of $SU(2)$ by assigning it to a matrix (5.1.5). Like wise every such matrix gives us a point in S^3. Let $z = (z_1,\, z_2) \in \mathbb{C}^2$ with $|z|^2 = 1$ and $w = (w_1,\, w_2) \in \mathbb{C}^2$ with $|w|^2 = 1$ be points on S^3 what is the point $z \circ w \in S^3$?

5.2 Infinitesimal Generators of $SU(2)$

To find the infinitesimal generators of $SU(2)$ we can proceed in a number of ways. Lets start with a formal derivation. We want to write

$$U = \mathbb{I}_2 + \epsilon M + O(\epsilon^2),$$

where ϵ is a real infinitesimal parameter. Now,

$$U^\dagger = \mathbb{I}_2 + \epsilon M^\dagger + O(\epsilon^2)$$

and $U.U^\dagger = \mathbb{I}_2$ implies that

$$M^\dagger = -M$$

that is the infinitesimal generators are anti-hermitian. The requirement that the determinant is one gives rise to the equation

$$\mathrm{Tr}\,(M) = 0.$$

To prove this we need a digression on linear algebra and the relationship between determinants and traces. As you will find the resulting formula useful in other courses as well the derivation will be quite general. We assume, in the following, that A is an $n \times n$ matrix which has no zero eigenvalues (otherwise the determinant is zero). Denote the eigenvalues of A by λ_i.

The determinant of A is[1]

$$\begin{aligned}
\mathrm{Det}\, A &= \prod_{i=1}^{n} \lambda_i \\
&= \prod_{i=1}^{n} \exp\left(\ln \lambda_i\right) \\
&= \exp\left(\sum_{i=1}^{n} \ln \lambda_i\right) \\
&= \exp\left(\mathrm{Tr} \ln A\right). \qquad (5.2.1)
\end{aligned}$$

We should really explain what we mean by the log of a matrix, but here we take it to be that matrix so that $\mathrm{Tr}\left(\ln A\right) = \sum_{i=1}^{n} \ln \lambda_i$.

Now we let $A = U = \mathbb{I}_2 + \epsilon M + \dots$ in (5.2.1). We expand

$$\ln U = \ln\left(\mathbb{I}_2 + \epsilon M + \dots\right) = \epsilon M + O(\epsilon^2),$$

from which one finds

$$\mathrm{Det}\, U = \exp \epsilon\ \mathrm{Tr}\, M + \dots = 1 + \epsilon\ \mathrm{Tr}\, M + O(\epsilon^2),$$

and which allows us to deduce that M must be traceless as the determinant of U should be 1.

Thus far we see that the generators of $SU(2)$ are any matrices M which are anti-hermitian and trace-less.

Exercise 5.4. Show that any 2×2, anti-hermitian and traceless matrix M can be written as

$$M = iC^a\, \sigma_a,$$

where the σ_a are the three hermitian Pauli matrices

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

and the C^a are real numbers. (We sometimes use the Einstein summation convention.)

From the exercise we learn that near the identity any group element U can be written as

$$U(\epsilon^a) = \mathbb{I}_2 + i\epsilon^a\, \sigma_a + \dots \quad . \qquad (5.2.2)$$

where the coefficients ϵ^a are real numbers. The σ_a are the hermitian generators of the group $SU(2)$. (if you want the anti-hermitian generators use $i\sigma_a$ instead.)

[1] The manipulations that follow actually make sense for matrices since non-degenerate matrices are dense in the space of matrices.

5.2.1 The Tangent Space

We now would like to interpret what we did in the previous section geometrically. To do so we start with an easier example; it is easier in the sense that one can visualize it. Consider once more the group $U(1)$, depicted in Figure 5.1, whose defining equation is

$$x^2 + y^2 = \mathbf{r}^2 = 1 \tag{5.2.3}$$

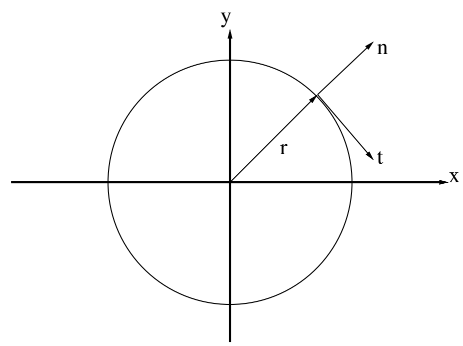


Figure 5.1: The group $U(1)$ embedded in $\mathbb{R}^2$ showing that the radial vector $\mathbf{r}$ lies in the direction of the normal vector $\mathbf{n}$ and both are orthogonal to the tangent vector $\mathbf{t}$.

We can see, directly from the figure, that the radial vector and the normal vector lie in the same direction. The radial vector from the origin to $(x_0,\, y_0)$ is given as

$$\mathbf{r} = (x_0,\, y_0), \quad \text{where } \ x_0^2 + y_0^2 = 1 \tag{5.2.4}$$

A tangent vector $\mathbf{t} = (a,\, b)$ to $U(1)$ at the point $(x_0,\, y_0)$ satisfies,

$$\mathbf{r}.\mathbf{t} = 0, \quad \text{or } \ x_0.a + y_0.b = 0 \tag{5.2.5}$$

Another way to get the tangent vector to $(x_0,\, y_0)$ is to ask that a point $(x_0+\delta x_0,\, y_0+\delta y_0)$ infinitesimally near to $(x_0,\, y_0)$ is also on the circle, or

$$(x_0 + \delta x_0)^2 + (y_0 + \delta y_0)^2 = 1, \ \text{ so that } \ x_0.\delta x_0 + y_0.\delta y_0 \simeq 0 \tag{5.2.6}$$

which agrees with the previous equation for the tangent vector $\mathbf{t}$.

An equivalent, but somewhat more intrinsic, way of expressing all of this is to let $f(x, y) = x^2 + y^2$, then tangents are simply defined by

$$\left.\frac{\partial f}{\partial x^i}\right|_{(x_0,\, y_0)} .\mathbf{t}^i = 0. \tag{5.2.7}$$

this agrees with (5.2.5) since $f = \mathbf{r}^2$ and also conveys the idea of differentiation which is implicit in (5.2.6).

Whichever way one likes to think about this the $\mathbf{r}$ that corresponds to the identity element in $U(1)$ is $(1,0)$ and the unit tangent vector there is $(0,1)$. Passing to complex coordinates we see that the unit tangent vector corresponds to $i=\sqrt{-1}$ which is also the generator of $U(1)$. Hence, what we have learnt is that the generator of $U(1)$ spans the tangent space to $U(1)$ at the identity.

Now let's reconsider the derivation of the generators of $SU(2)$ within the context of our explicit parametrization. Firstly we note that the identity element of the group is $\mathbb{I}_2$ which corresponds to the point $(1,0,0,0)$ (call this the north pole of S^3). We want to Taylor series around the identity element which means that we want to consider points on S^3 very close to the north pole, i.e. points $(1+\delta x_1, \delta x_2, \delta x_3, \delta x_4) \in S^3$ where the δx_i are infinitesimal. Now $(1+\delta x_1, \delta x_2, \delta x_3, \delta x_4)$ must satisfy (5.1.6) which implies that

$$2\delta x_1 + O((\delta x)^2) = 0,$$

or $\delta x_1 = 0 + \ldots$, where the ellipses stand for terms quadratic in the other infinitesimals. We are now in good shape to Taylor series

$$\begin{aligned} U &= \mathbb{I}_2 + \delta U + \ldots \\ &= \mathbb{I}_2 + \begin{pmatrix} i\delta x_2 & \delta x_3 + i\delta x_4 \\ -\delta x_3 + i\delta x_4 & -i\delta x_2 \end{pmatrix} + \ldots \\ &= \mathbb{I}_2 + +i\delta x_4\,\sigma_1 + i\delta x_3\,\sigma_2 + i\delta x_2\,\sigma_3 + \ldots \quad . \end{aligned} \tag{5.2.8}$$

We see that the parameters ϵ^a in (5.2.2) are nothing other than infinitesimal parameters of S^3 around the north pole. Consequently the space of generators of $SU(2)$ is the tangent space to $SU(2)$ at the identity. The tangent space is a vector space so we are saying that teh space of generators is a vector space.

5.3 The Lie Algebra of $SU(2)$

There is a great deal of structure is associated to the space of generators of a group. In this part we will consider these structures for $SU(2)$. Firstly, for any group G of matrices we get a vector space. To see this note that if we write two group elements, $g_1, g_2 \in G$ as

$$g_i = \mathbb{I} + \epsilon M_i + O(\epsilon^2)$$

where the M_i are infinitesimal generators. Then, as $g_1.g_2 \in G$, we have

$$g_1.g_2 = \mathbb{I} + \epsilon(M_1 + M_2) + O(\epsilon^2)$$

which tells us that M_1+M_2 is also an infinitesimal generator. So we have a vector space: elements of a vector space can be added and subtracted to give other elements of the vector space[2] and you can multiply them with real (or complex numbers depending on

[2]You will recall that this is the definition of an Abelian group under addition.

the situation) to give an element of the vector space. Call the vector space of generators V. The dimension of V is equal to the dimension of the group G (at least this is immediate for $SU(2)$ from our study above). In fact just as for any vector space we would like to find a basis for V. In the case of $SU(2)$, as we saw in the exercise above, the Pauli matrices are the basis 'vectors' of V and there are 3 of them because we need 3 parameters to parameterize the 3-sphere.

But we do not only have a vector space. There is an extra structure called a bracket, $[,]$, or commutation relation on V. This means that the commutator of two generators is again a generator. For the basis of generators of $SU(2)$, we have

$$\left[\frac{\sigma_a}{2}, \frac{\sigma_b}{2}\right] = i\epsilon_{abc}\frac{\sigma_c}{2} \tag{5.3.1}$$

where ϵ_{abc} is a totally antisymmetric tensor such that $\epsilon_{123} = 1$.

Exercise 5.5. Show that if M_i, $i = 1$, 2, are generators of $SU(2)$ then $[M_1, M_2]$ is also a generator of $SU(2)$.

Definition 5.2. The vector space V of generators of $SU(2)$ together with the commutation relation is called the **Lie Algebra** of $SU(2)$.

5.4 Counting Equations

Lets see how many parameters are need to parameterize $SU(2)$ **without** solving the equations. Consider an arbitrary complex matrix U. Such matrices have 4 complex entries so are parameterized by 8 real variables. The determinant of a complex matrix U is some complex number, call it w. Now suppose that U just satisfies the unitary condition $U.U^\dagger = 1$. The determinant of the unitarity condition tells us that

$$\mathrm{Det}\, U \,.\, \mathrm{Det}\, U^\dagger = 1,$$

but since $\mathrm{Det}\, U^\dagger = (\mathrm{Det}\, U)^\dagger$ this tells us that $w\overline{w} = 1$. Let $w = re^{i\theta}$, then we have that the unitarity condition implies that $r = 1$ but really tells us nothing about θ. On the other hand, if we want U to be an element of $SU(2)$ the determinant condition is really an equation for θ, that is $\theta = 0$. So apart from the unitarity condition the determinant condition is only one equation. But how many equations are living in the unitarity condition? If we take the hermitian conjugate of the equation $U.U^\dagger = 1$ we see that it does not change. This means that equations are basically real. Well we have 4 entries in the matrix equation so this means 4 real equations which together with the equation from the determinant condition makes 5 real equations. We have 5 equations and 8 unknowns, so that leaves 3 unknowns. This is the dimension of $SU(2)$.

5.5 The Group $SO(3)$

Definition 5.3. $SO(3)$ is the group of real 3×3 orthogonal matrices with unit determinant, $M \in SO(3)$,

$$M \,.\, M^T = \mathbb{I}_3, \quad \mathrm{Det}\, M = 1.$$

Geometrically these matrices are the possible rotations in $\mathbb{R}^3$. Just as for $SU(2)$ we would like to find all matrices which satisfy these conditions. One way to do that is to use the geometry of the situation. But first lets count equations.

Consider M to be a real 3×3 matrix, so it has 9 real variables. Now take a closer look at the the orthogonality condition. From that we see that the determinant of M must satisfy

$$\mathrm{Det}\, M = \pm 1.$$

If you have a matrix, M which satisfies this with $+1$ then you automatically get a matrix which has the negative determinant, namely $-M$. So the number of equations does not change with either sign, so we do not have to count the determinant equation. How many equations do we have from the orthogonality equation? Well that equation is invariant under transposition which means it is symmetric. A symmetric 3×3 matrix has 6 independent components, so we really only have 6 equations. The sum is: 6 equations with 9 unknowns which leaves us with 3 unknowns. The dimension of $SO(3)$ is therefore 3, which is the same as the dimension of $SU(2)$. But this does not mean that $SO(3)$ and $SU(2)$ are the same group.

5.6 Parameterization of $SO(3)$

We will show that any element of $SU(2)$ gives us an element of $SO(3)$. The construction introduced here is useful for general matrix groups.

Definition 5.4. The adjoint action, $\mathrm{Ad} : SU(2) \times su(2) \to su(2)$, of the group $SU(2)$ on its Lie algebra is, for $g \in SU(2)$,

$$\mathrm{Ad}_g(\sigma_a) \equiv g^{-1} \sigma_a \, g.$$

This means that if you are given a group element $g \in SU(2)$ then what you have is a mapping of the Lie algebra to itself. Recall that the generators of $SU(2)$ are traceless and hermitian. Now we check that this is true of $g^{-1}\sigma_a \, g$,

$$\left(g^{-1}\sigma_a \, g\right)^\dagger = g^{-1}\sigma_a \, g$$

since $g^\dagger = g^{-1}$ by the unitarity condition, so $g^{-1}\sigma_a \, g$ is hermitian. Also we have

$$\mathrm{Tr}\left(g^{-1}\sigma_a \, g\right) = \mathrm{Tr}\,\sigma_a = 0,$$

so it is traceless. As promised $g^{-1}\sigma_a\, g$ is a generator of $SU(2)$. As it is a generator of $SU(2)$ we can expand it in a basis of Pauli matrices,

$$g^{-1}\sigma_a\, g = \sum_{b=1}^{3} M_{ab}(g)\, \sigma_b \tag{5.6.1}$$

where the $M_{ab}(g)$ are real and they depend on the group element g. The important point is that

Theorem 5.1. The coefficients $M_{ab}(g)$ thought of as matrices are elements of $SO(3)$ and all elements of $SO(3)$ can be obtained in this way.

proof:

We first prove that $M_{ab}(g) \in SO(3)$. On the one hand we have that

$$g^{-1}\sigma_a\, g.g^{-1}\sigma_b\, g = g^{-1}\sigma_a\sigma_b g = \delta_{ab}\mathbb{I}_2 + i\sum_c \epsilon_{abc} g^{-1}\sigma_c\, g = \delta_{ab}\mathbb{I}_2 + i\sum_{c,d}\epsilon_{abc}M_{cd}\sigma_d$$

while on the other we have

$$g^{-1}\sigma_a\, g.g^{-1}\sigma_b\, g = \sum_{cd} M_{ac}M_{bd}\sigma_c\sigma_d = \sum_{cd} M_{ac}M_{bd}\left(\mathbb{I}_2\delta_{cd} + i\sum_e \epsilon_{cde}\sigma_e\right).$$

Comparing the two equations we find that

$$\sum_c M_{ac}M_{bc} = \delta_{ab}, \quad \sum_d \epsilon_{abd}M_{dc} = \sum_{d,e}\epsilon_{edc}M_{ae}M_{bd} \tag{5.6.2}$$

which are really just the conditions that the matrix be in $SO(3)$. (Using the first equation, the second can be written as $\epsilon_{abc} = \sum_{d,e,f}\epsilon_{def}M_{ad}\,M_{eb}\,M_{cf}$ which is the same as $\mathrm{Det}\, M = 1$.)

That every $SO(3)$ element is obtained in this way is a an exercise. $\square$

Exercise 5.6. Show that every $M \in SO(3)$ can be written as $M_{ab}(g)$ for some $g \in SU(2)$ (Don't do this one yet).

We have not shown that $SU(2)$ and $SO(3)$ are the same. Instead we have shown that every element of $SO(3)$ corresponds to **two** elements of $SU(2)$. The groups elements $g,\ -g \in SU(2)$ give the same element $M(g) \in SO(3)$. One says that $SU(2)$ is a **double cover** of $SO(3)$. Since $SU(2) \equiv S^3$, we find that $SO(3) \equiv S^3/\mathbb{Z}_2 \equiv \mathbb{RP}^3$.

5.7 The Lie Algebra of $SO(3)$

Write the element $M \in SO(3)$ near the identity element as

$$M = \mathbb{I}_3 + \epsilon T + O(\epsilon^2). \tag{5.7.1}$$

Orthogonality tells us that

$$(\mathbb{I}_3 + \epsilon T + \dots) \, . \, (\mathbb{I}_3 + \epsilon T + \dots)^T = \mathbb{I}_3$$

or that T is an anti-symmetric matrix,

$$T_{ab} = -T_{ba}. \tag{5.7.2}$$

Exercise 5.7. Show that the requirement that $\mathrm{Det}\, M = 1$ does not impose any extra conditions on T.

Exercise 5.8. Show that any real antisymmetric 3×3 matrix can be written as a real linear combination of the matrices

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Exercise 5.9. Show that the J_a satisfy the commutation relations

$$[J_a, \, J_b] = -\epsilon_{abc} \, J_c.$$

So we have that the space of generators of $SO(3)$, V is the set of 3×3 real anti-symmetric matrices. The Lie algebra of $SO(3)$ is V together with the commutation relation found in the exercises. The vector space and commutation relations of $SO(3)$ and $SU(2)$ are isomorphic, so the Lie Algebras of the two are isomorphic. The vector spaces are isomorphic as they are both $\mathbb{R}^3$ and the commutation relations are the same if one identifies J_a with $i\sigma_a/2$.

The Lie algebras are the same even though the groups are not since, to find the Lie Algebra we do not look at the whole space of the group but only near the identity. Near the identity elements $SU(2)$ and $SO(3)$ are the same.

5.8 The ad Action on the Lie Algebra

Given what has been said so far it should not come as a surprise that there is a map from the Pauli matrices to the rotation matrices. This map is the Lie algebra version of the adjoint action Ad.

Definition 5.5. The adjoint action, $\mathrm{ad} : su(2) \times su(2) \to su(2)$ of $su(2)$ to $su(2)$ is, for $X \in su(2)$, given by

$$\mathrm{ad}_X(\sigma_a) = [X, \sigma_a].$$

To see how this gives us the J_a we recall that since the commutator gives us a Lie algebra element that

$$\mathrm{ad}_X(\frac{i\sigma_a}{2}) = [X, \frac{i\sigma_a}{2}] = \sum_b \frac{i\sigma_b}{2} \, \Phi^b{}_a(X). \tag{5.8.1}$$

Exercise 5.10. Show that the ${\Phi^b}_a(X)$ are in the Lie algebra of $SO(3)$ and that ${\Phi^b}_a(X) = -\epsilon_{abc}X_c$ where $X = X^a i\sigma_a/2$.

Note: The position of the matrix ${\Phi^b}_a(X)$ in the formula (5.8.1), where it is defined, is dictated by the fact that the sigma matrices form a basis of the vector space and this is matrix multiplication on the basis vectors. If you as how it acts on the coeeficients of, say $Y = Y^a i\sigma_a/2$ we would say that its action is $Y^a \to \Phi^a_b(X)\, Y^b$.

The exercise tells us that if we are given an element, X, of the Lie algebra of $SU(2)$ then we get an element, $\Phi(X)$ of the Lie algebra of $SO(3)$ which has components ${\Phi^b}_a(X)$. When $X = i\sigma_c/2$ call the element that one gets J_c, it has components

$${\Phi^b}_a(\frac{i\sigma_c}{2}) = -\epsilon_{abc} = {(J_c)^b}_a\,.$$

Exercise 5.11. Show that the matrix elements that we obtain in this way for J_c agree with the explicit matrices we found before.

Consider the $SO(3)$ rotation about the 3-rd axis

$$\exp{(\theta J_3)} = \cos{(\theta)} + \; J_3 \, \sin{(\theta)}$$

is periodic with period $\theta \to \theta + 2\pi$. The equivalent element in $SU(2)$ is

$$\exp{(i\theta\sigma_3/2)} = \cos{(\theta/2)} + \; i\sigma_3 \, \sin{(\theta/2)}$$

but the periodicity is now $\theta \to \theta + 4\pi$. The reason is that, as we saw the elements g and $-g$ in $SU(2)$ are distinct but correspond to the same element in $SO(3)$. This means that $\theta = 0$ and $\theta = 2\pi$ are the same in $SO(3)$ (the identity), but in $SU(2)$ correspond to the identity and to minus the identity respectively.

CHAPTER 6

REPRESENTATIONS OF THE LIE ALGEBRA OF $SU(2)$

We start with some general definitions.

6.1 Representations

Definition 6.1. A representation of a Lie Algebra, V, is a linear map, R from V to a set of matrices (say $n \times n$) such that for $X,\ Y \in V$ R preserves the bracket or commutation relations,

$$R([X, Y]) = R(X)R(Y) - R(Y)R(X).$$

Recall that a linear map is one that satisfies

$$R(aX + bY) = aR(X) + bR(Y),$$

for a, $b \in \mathbb{R}$ (or $\mathbb{C}$ depending on the situation).

Definition 6.2. The dimension of a representation is the dimension of the vector space that the matrix representatives act on.

This means that if the elements of the Lie algebra are represented by $n \times n$ matrices then they act on n component column vectors. These vectors are vectors in an n-dimensional vector space. The dimension of a representation is the dimension of that vector space.

Examples:

The Lie algebra of $SU(2)$ as an abstract object is given by

$$[T_a, T_b] = -\epsilon_{abc} T_c.$$

One representation, called the fundamental (or defining) representation, is given by

$$R_{\mathrm{f}}(T_a) = \frac{i\sigma_a}{2}.$$

This is a 2-dimensional representation. Another representation that you have seen is the adjoint representation,

$$R_{\text{ad}}(T_a) = J_a,$$

which is a 3-dimensional representation. There is always a trivial representation,

Definition 6.3. A representation of the Lie algebra is said to be trivial if the matrix representatives are the zero matrix.

6.2 Reducible and Irreducible Representations

For a given representation of dimension n we think of the representation matrix as a map from the **representation space** to itself. The representation space is an n-dimensional vector space, call it W. So we have for a vector $w \in W$, $X \in V$

$$w \xrightarrow{X} R_W(X).w \tag{6.2.1}$$

The representation matrix carries the label W just so that you remember that the representation depends on the vector space.

It can happen that matrices we use to represent the Lie algebra do not really see all of the representation space W.

Exercise 6.1. Show that the following matrices form a 4-dimensional representation of the Lie algebra of $SU(2)$

$$R(T_a) = \begin{pmatrix} i\sigma_a/2 & 0_2 \\ 0_2 & 0_2 \end{pmatrix}.$$

where 0_2 stands for the 2×2 zero matrix.

This exercise shows us the point that if we split $W = \mathbb{R}^4 = \mathbb{R}^2 \oplus \mathbb{R}^2$, so that any $w \in W$ is written as

$$w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

where w_1 is a vector of the first $\mathbb{R}^2$ and w_2 is a vector of the second $\mathbb{R}^2$, then the representation is acting on the first $\mathbb{R}^2$ one way (in this case non-trivially) and on the second $\mathbb{R}^2$ in another (in this case trivially). This leads us to the rather more general

Definition 6.4. A representation is called **reducible** if there exists an m-dimensional subspace, U, of the representation space W, which is invariant under the transformations, meaning that $\forall X \in V$, $R_W(X) : U \to U$.

In such a situation we can choose a basis so that the m vectors that span U take up the first m places of a column vector of W,

$$w = \begin{pmatrix} u_1 \\ \vdots \\ u_m \\ - \\ \vdots \\ \vdots \end{pmatrix} \tag{6.2.2}$$

and the representation matrix will take the form, for $X \in V$,

$$R_W(X) = \begin{pmatrix} R_U(X) & A(X) \\ 0_{(n-m)\times m} & B(X) \end{pmatrix} \tag{6.2.3}$$

where $A(X)$ and $B(X)$ are $m \times (n-m)$ and $(n-m) \times (n-m)$ matrices respectively. Mostly the Lie algebras that we will be dealing with are such that there is a basis where the representation, when it is reducible, can be put in block diagonal form. This means that we can, in those cases, write $W = U + Y$ and

$$R_W(X) = \begin{pmatrix} R_U(X) & 0_{m\times(n-m)} \\ 0_{(n-m)\times m} & R_Y(X) \end{pmatrix} \tag{6.2.4}$$

Exercise 6.2. Show that if the representation matrix has the form shown in (6.2.4) then R_U and R_Y are representations.

Definition 6.5. An **irreducible representation** is a representation for which there are **no proper** invariant subspaces.

Proper means neither the original space W nor the zero vector space.

If a representation is reducible, then we know everything about it by understanding the smaller representations. This leads us to the study of irreducible representations.

Lets write the Lie algebra of $SU(2)$ (or of $SO(3)$) as

$$[T_a, T_b] = i\epsilon_{abc}T_c. \tag{6.2.5}$$

Since ϵ_{abc} are just real numbers, if the T_a are matrices (that is when we work with a representation) they must be hermitian matrices.

Exercise 6.3. Prove the last statement.

The eigenvalues of a hermitian matrix are real and this is the advantage of working with this form of the algebra. Furthermore, we know that we can simultaneously diagonalize a commuting set of hermitian matrices. In this case none of the matrices, T_a, commute with any of the others, so at best we can diagonalize one of them. We will understand to have diagonalized T_3.

In the notation above, a representation for the matrices, amounts to associating to T_a its matrix representative $R(T_a)$. This way of labeling things is not very convenient so we will adopt the following notation. Set $h = T_3$, $e_\pm = T_1 \pm iT_2$, the commutation relations become

$$[h, e_\pm] = \pm e_\pm, \quad [e_+, e_-] = 2h. \tag{6.2.6}$$

Exercise 6.4. Derive these commutation relations from those for T.

We will call h the **Cartan element** and e_+ a **positive root vector** and e_- a **negative root vector**.

Thats the algebra as an abstract object, now when dealing with a representation we denote the matrices with capital letters.

$$h \to H, \quad e_\pm \to E_\pm \tag{6.2.7}$$

6.3 Finite Dimensional Irreducible Representations of Lie $su(2)$

Let us presume that our representation space W is n dimensional, though we do not know what n is yet in order to have an irreducible representation. Since H is hermitian we presume that it is diagonalized. Let $\phi \in W$ be an eigenvector of H with eigenvalue m which we also do not know yet,

$$H\,\phi = m\,\phi. \tag{6.3.1}$$

Now apply the roots on ϕ

$$H\,(E_\pm\,\phi) = ([H, E_\pm] + E_\pm H)\;\phi = (m \pm 1)\;(E_\pm\,\phi)\,, \tag{6.3.2}$$

which tells us that $E_\pm\,\phi$ are also eigenvectors of the Cartan element H with eigenvalues $m \pm 1$. We can continue this process so, for example,

$$H\left(E_+^k\,\phi\right) = (m + k)\;\left(E_+^k\,\phi\right). \tag{6.3.3}$$

But we cannot really continue the process indefinitely. Non-zero eigenvectors of H with different eigenvalues are orthogonal to each other. Since the dimension of the space is n we cannot have more than n linearly independent vectors, so for some value of k we have that $E_+^k\,\phi = \phi' \neq 0$, but $E_+^{k+1}\,\phi = 0$. ϕ' is called a highest weight.

So now that we have shown that there must be a highest weight vector, we start from it. A highest weight is the vector that satisfies

$$H\,\phi' = j\,\phi', \quad E_+\,\phi' = 0, \tag{6.3.4}$$

for some j still to be determined. Now we act with E_-, the same argument as before tells us that there is an l such that

$$E_-^l\,\phi' \neq 0, \quad E_-^{l+1}\,\phi' = 0. \tag{6.3.5}$$

From this we can deduce a relationship between j and l,

$$
\begin{aligned}
E_+ E_-^{l+1}\,\phi' = 0 &= ([E_+, E_-] + E_- E_+)\, E_-^l \phi' \\
&= (2H + E_- E_+)\, E_-^l \phi' \\
&= 2(j-l)\, E_-^l \phi' + E_- E_+ E_-^l \phi' \\
&= (2(j-l) + 2(j-l+1))\; E_-^l \phi' + E_-^2 E_+ E_-^{l-1} \phi' \\
&= \dots \\
&= \sum_{r=0}^{l} 2(j-l+r)\, E_-^l \phi' + E_-^{l+1} E_+ \phi' \\
&= \sum_{r=0}^{l} 2(j-l+r)\, E_-^l \phi'. \qquad (6.3.6)
\end{aligned}
$$

However, as $E_-^l \phi' \neq 0$ we conclude that

$$
\sum_{r=0}^{l} 2(j-l+r) = (l+1)\,(2j-l) = 0, \qquad (6.3.7)
$$

and as l is a nonnegative integer we learn that

$$
j = l/2, \qquad (6.3.8)
$$

so that j comes in $1/2$ integral units and $j \geq 0$. Thus the states that we have are

State	ϕ'	$E_-\phi'$	$\dots$	$\dots$	$E_-^l\phi'$
H eigenvalue	j	$j-1$	$\dots$	$\dots$	$j-l=-j$

(6.3.9)

The question that we must now ask ourselves is: are these all the vectors in the representation? Clearly if we act with H we do not generate new vectors (they are eigenstates) and acting with E_- on any vector in the table gives another vector in the table or zero. So we have to see what E_+ does. Consider the action of E_+ on any one of the states, that is $E_+\, E_-^q\, \phi'$ for $0 \leq q \leq l$. We have,

$$
\begin{aligned}
E_+\, E_-^q\, \phi' &= ([E_+, E_-] + E_- E_+)\, E_-^{q-1}\, \phi' \\
&= (2H + E_- E_+)\, E_-^{q-1}\, \phi' \\
&= (2(j-q+1) + E_- E_+)\, E_-^{q-1}\, \phi' \\
&= q(2j-q+1)\, E_-^{q-1} \phi'. \qquad (6.3.10)
\end{aligned}
$$

Exercise 6.5. Derive the last line of this formula.

This shows us that E_+ maps any vector in the table to another vector in the table or to zero (the highest weight is defined by this property).

So our vector space W, the representation space, is made up of the vectors in the table and no others. There are $2j+1$ vectors so the dimension of the representation is $n = 2j+1$. There is also, for each eigenvalue $j-q$ of H exactly one eigenvector $E_-^q \phi'$.

But is the representation irreducible? If you start from any one of the vectors in the table, then by acting with either E_+ or E_- enough times you get all of the others. So it is not possible to have a proper sub-vector space Y spanned by a subset of the vectors which is invariant under the action of the Lie algebra.

WARNING!

We have made use of the $e_\pm$ to help to build irreducible representations, however, they are themselves **not** generators with our definition. The generators of the Lie algebra of $SU(2)$ are, for us, traceless and hermitian (with real coefficients). However, $e_\pm^\dagger = e_\mp$ so these are definitely not hermitian. Nevertheless, one can always expand a Lie algebra element X as

$$X = \alpha\, h + \beta\, e_+ + \overline{\beta} e_-$$

which is certainly hermitian when $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{C}$. Note that the number of real parameters is 3 as it should be.

6.4 The Quadratic Casimir

A very useful concept in physics in general is that of invariants, things which do not change, for example things which are conserved. There are such objects in Lie algebras as well, that is objects, C, which commute with all the generators,

$$[T_a, C] = 0.$$

For $SU(2)$ the object in question is

$$C_2 = \sum_{a=1}^{3} T_a T_a = h^2 + \frac{1}{2}\left(e_+ e_- + e_- e_+\right) = h^2 + h + e_- e_+.$$

This is known as the quadratic casimir. It is called quadratic since the generators are squared. For other groups there are also higher order casimirs.

Exercise 6.6. Show that the second equality is correct and that C_2 commutes with h, e_+ and e_-.

The usefulness of this is that C_2 is takes the same value on every vector of an irreducible representation.

Exercise 6.7. Prove that for every vector in an irreducible representation of dimension $2j + 1$ that C_2 takes the value $j(j + 1)$. (**Hint:** What value does it take on the highest weight?)

Exercise 6.8. What do we call the $j = 0$ representation?

6.5 Complex Conjugate Representation and an Invariant Inner Product

Suppose that we have found matrices, R, with which we can represent the Lie algebra, then we get another representation for free, it is the complex conjugate representation. We have that if

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

then

$$[-\overline{R}(X), -\overline{R}(Y)] = -\overline{R}([X, Y]) \tag{6.5.1}$$

as well. For the defining 2-dimensional representation of $su(2)$ we have

$$\left[\frac{\sigma_a}{2}, \frac{\sigma_b}{2}\right] = i\epsilon_{abc}\frac{\sigma_c}{2} \implies \left[\frac{\overline{\sigma}_a}{2}, \frac{\overline{\sigma}_b}{2}\right] = -i\epsilon_{abc}\frac{\overline{\sigma}_c}{2}$$

Exercise 6.9. Derive the formula (6.5.1).

Such a representation need not be a new one. It might be the same as a representation that we already know but in a different basis. We say that

Definition 6.6. A representation R is (pseudo-) real if there exists a matrix S such that

$$S\,R(X)\,S^{-1} = -\overline{R}(X).$$

The Pauli representation of $su(2)$ is pseudo-real. This is because

$$\sigma_2\,\sigma_a\,\sigma_2 = -\overline{\sigma}_a,$$

and one writes that $\overline{\mathbf{2}} = \mathbf{2}$ (the 2 is just the dimension of the representation). Indeed all of the representations of $su(2)$ are (pseudo-) real. To see this we note that the representative of the Cartan generator H is diagonal with real entries. But $-\overline{H} = -H$ so the eigenvalues of the Cartan element go over to minus the eigenvalues of the original Cartan element. In the representations that we have found we get all eigenvalues from $-j$ up to j in the Cartan element and so $-H$ is just a reordering of H.

However, while it is true for $SU(2)$ that all representations are (pseudo-) real, it is far from being true in general. The next group that we come across is $SU(3)$ and the fundamental representation and the complex conjugate fundamental representation are quite different, while the adjoint representation is real.

An important point to remember is that if we take the transpose of the complex conjugate representation then the representation matrices go over to (minus) their hermitian conjugates. In equations the last sentence means

$$\left(-\overline{R}(X)\,.\,v\right)^{\mathrm{T}} = v^{\mathrm{T}}.\,-R^{\dagger}(X).$$

Remember generators came from expanding the group element out around the identity so that, for $g \in SU(2)$

$$g = \exp\left(i\lambda^a \frac{\sigma_a}{2}\right) = \mathbb{I}_2 + i\lambda^a \frac{\sigma_a}{2} + \dots\ ,$$

where the λ^a are real. We have given the group element in the 2 dimensional representation. So we should really write

$$R_{\mathrm{f}}(g)_i^{\ j}$$

for this element but I will just call it $g_i^{\ j}$. The group element that acts in the complex conjugate representation is obtains by sending $\sigma_a \to -\overline{\sigma}_a$

$$\overline{g} = \exp\left(-i\lambda^a \frac{\overline{\sigma}_a}{2}\right).$$

Suppose ϕ_i, $i = 1,\ 2$, transforms in the fundamental representation of $su(2)$, then it also transforms in the fundamental representation of $SU(2)$,

$$\phi_i \to g_i^{\ j}\, \phi_j. \tag{6.5.2}$$

Let $\overline{\phi}^{\,i}$ transform in the complex conjugate representation of $su(2)$ then it also transforms in the complex conjugate representation of $SU(2)$ as

$$\overline{\phi}^{\,i} \to \overline{g}^{i}{}_{j}\, \overline{\phi}^{\,j} = \overline{\phi}^{\,j}\, g_j^{\dagger\, i}. \tag{6.5.3}$$

We have adopted the convention that the complex conjugate vector gets its label 'upstairs' so that we have a rationality to the labeling system.

We now have an important conclusion. There is an inner product which is invariant under $SU(2)$ transformations, namely if ϕ transforms as a $\mathbf{2}$ and $\overline{\psi}$ transforms as a $\overline{\mathbf{2}}$ then

$$\overline{\psi}.\phi \equiv \sum_{i=1}^{2} \overline{\psi}^{\,i}\, \phi_i$$

is invariant under $SU(2)$ transformations. This is easily checked, the transformation of the inner product under $g \in SU(2)$ is

$$\overline{\psi}.\phi \to \overline{\psi}\, g^\dagger\, g\, \phi = \overline{\psi}.\phi.$$

(Check this in components.)

CHAPTER 7

TENSOR REPRESENTATIONS AND YOUNG TABLEAUX

A rather normal situation in physics is that we have two objects, both of which transform in some irreducible representations of a group, they are brought together and we would like to know exactly how the combined object behaves under the group action. For example when two spin 1/2 particles (or isospin half or ...) are brought together we would like to know what spins are possible for the composite object. Another way to say this is that we have the product of two representations and we would like to write this as a sum of irreducible representations. Such products form tensors and that is what we study next.

7.1 Tensors

A tensor is a generalization of a matrix. It is an object with many labels (we will take them all to be of the same kind). The rank of a tensor is just the number of labels that it carries. Here is an example of a rank r tensor, $i_j = 1, \ldots, n$, $\forall j$,

$$T_{i_1\, i_2\, \ldots i_r}.$$

Usually we consider that tensors satisfy certain properties. In special relativity, for example, vectors P^μ (tensors of rank 1), the metric tensor $\eta^{\mu\nu}$ (a tensor of rank 2) and so on, depend on their position in space-time and under a Lorentz transformation, $x'^{\,\mu} = \Lambda^\mu_{\ \nu}\, x^\nu$ obey the rule

$$P'^\mu = \Lambda^\mu_{\ \nu}\, P^\nu, \quad \eta^{\mu\nu} = \Lambda^\mu_{\ \rho}\, \Lambda^\nu_{\ \sigma}\, \eta^{\rho\sigma}.$$

This is at the level of the Lorentz group. What happens at the level of the algebra? Well we would say that

$$\Lambda^\mu_{\ \nu} = \delta^\mu_{\ \nu} + \lambda^\mu_{\ \nu} + \ldots$$

for $\lambda^{\mu}_{\ \nu}$ infinitesimal. Then at this level the above equations becomes

$$P'^{\mu} = P^{\mu} + \lambda^{\mu}_{\ \nu} P^{\nu} + \dots, \quad \eta^{\mu\nu} = \eta^{\mu\nu} + \lambda^{\mu}_{\ \rho}\eta^{\rho\nu} + \lambda^{\nu}_{\ \rho}\eta^{\mu\rho} + \dots$$

($\eta^{\mu\nu}$ is invariant but we write the transformation before we use the fact that $\lambda_{\mu\nu} = -\lambda_{\nu\mu}$). We would like to do the same for representations of $su(2)$ where the labels individually run over the set of labels of the fundamental representation.

7.2 Tensors of a Lie Algebra

Suppose that we are given a Lie algebra whose fundamental representation is of dimension r, then we can consider tensors of the form

$$T_{i_1\, i_2 \dots\, i_n} \tag{7.2.1}$$

where $i_j = 1, \dots,\ r\ \forall j$ and the tensors 'transform' as

$$R_{\mathrm{T}}(X) T_{i_1\, i_2 \dots\, i_n} = \sum_{s=1}^{n} R_{\mathrm{f}}(X)_{i_s}^{\ j}\, T_{1_1 \dots i_{s-1}\, j\, i_{s+1} \dots\, i_n}, \tag{7.2.2}$$

where X is in the Lie algebra and R_{f} is the fundamental representation (2-dimensional for $su(2)$). We need to prove that this construction actually gives us a representation. The map is obviously linear as $R_{\mathrm{f}}(X)_i^{\ j}$ is, but that means we need to be sure the bracket is preserved.

Exercise 7.1. Show that the map defined in (7.2.2) preserves the commutation relations.

STOP: Our definition of a representation said we would get a matrix, so where is it? Its there in (7.2.2) but you need a little work to see it. First we note that a tensor can be viewed as a large vector. For example consider the rank 2 tensor T_{ij}, where $i,\ j = 1, \dots,\ n$. For each i we have a vector with components given by j, $(T_i)_j$ so we can write T_{ij} as follows

$$\begin{pmatrix} T_{1j} \\ T_{2j} \\ \vdots \\ T_{nj} \end{pmatrix} = \begin{pmatrix} T_{11} \\ \vdots \\ T_{1n} \\ \vdots \\ T_{n1} \\ \vdots \\ T_{nn} \end{pmatrix}$$

The way to think about this is that the i labels a vector whose components are themselves vectors labeled by j. For a rank 3 tensor, T_{ijk} we apply the same logic so that we first get a column vector whose components are rank 2 tensors, but each rank 2 tensor is itself a vector

$$\begin{pmatrix} T_{1jk} \\ \vdots \\ T_{njk} \end{pmatrix} = \begin{pmatrix} T_{11k} \\ \vdots \\ T_{1nk} \\ \vdots \\ T_{n1k} \\ \vdots \\ T_{nnk} \end{pmatrix} = \begin{pmatrix} T_{111} \\ \vdots \\ T_{11n} \\ \vdots \\ T_{1n1} \\ \vdots \\ T_{1nn} \\ \vdots \\ T_{nnn} \end{pmatrix}$$

This quickly gets out of control! But the point is once you write the tensor as a vector, since the representation map is linear it must be just one big matrix acting on the big vector. Rather than write it out in all its wonderful detail as a vector it is much easier to keep the tensor notation and that is what we will do for most of the lectures.

The way mathematicians say this is the following. If the vector space is $W \equiv \mathbb{R}^n$, then a rank 1 tensor is an element of W a rank 2 tensor is in $W \otimes W$ and a rank p tensor is in $W \otimes \cdots \otimes W$ (with p factors of W). These are called tensor products. What it means is just what we did with writing the vectors above. The dimensions of these spaces are np. (and this is exactly the number of components of the tensor that you start with). So a tensor of W is a vector in a much bigger space, namely a vector in $W \otimes \cdots \otimes W$.

To see how one can construct the matrix representing the action on the tensors, but thought of as vectors, we specialize to a rank 2 tensor in $su(2)$. Write the tensor as before, namely as

$$\begin{pmatrix} T_{1j} \\ T_{2j} \end{pmatrix} = \begin{pmatrix} \underline{T}_1 \\ \underline{T}_2 \end{pmatrix}.$$

In the transformation rule (7.2.2) the first R_{f} acts on the i label, which for our vector means it acts on the positions of the vectors $\underline{T}_1$ and $\underline{T}_2$ but it does not act on their respective components. To fix ideas let $X = \sigma_3$, then

$$R_{\mathrm{f}}(\sigma_3) T_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \begin{pmatrix} \underline{T}_1 \\ \underline{T}_2 \end{pmatrix},$$

but this abuses the notation since we really mean not $1.\underline{T}_1$ but rather $\mathbb{I}_2.\underline{T}_1$ (it should be a 4×4 matrix). The matrix is then, more correctly,

$$\begin{pmatrix} \mathbb{I}_2 & 0 \\ 0 & -\mathbb{I}_2 \end{pmatrix} = \sigma_3 \otimes \mathbb{I}_2$$

where the tensor product notation means you write out the left matrix and then in its components you replace the number 1 with the right hand matrix. With this notation

$$\sigma_2 \otimes \mathbb{I}_2 = \begin{pmatrix} 0 & -i\mathbb{I}_2 \\ i\mathbb{I}_2 & 0 \end{pmatrix}, \quad \mathbb{I}_2 \otimes \sigma_2 = \begin{pmatrix} \sigma_2 & 0 \\ 0 & \sigma_2 \end{pmatrix}.$$

But we have not finished because we also need to consider the action of the second R_{f} on the j label. This is easy to write in the vector notation, and in the tensor product notation it is

$$\begin{pmatrix} R_{\mathrm{f}}(X)_j^{\ k} T_{1k} \\ R_{\mathrm{f}}(X)_j^{\ k} T_{2k} \end{pmatrix} = (\mathbb{I}_2 \otimes R_{\mathrm{f}}(X)) \begin{pmatrix} \underline{T}_1 \\ \underline{T}_2 \end{pmatrix}.$$

Putting all the pieces together we find that the matrix representative for a rank 2 tensor of $su(2)$ is

$$R_T(X) = R_{\mathrm{f}}(X) \otimes \mathbb{I}_2 + \mathbb{I}_2 \otimes R_{\mathrm{f}}(X).$$

Exercise 7.2. Show that a rank p tensor of $su(2)$ transforms as

$$R_T(X) = \sum_{i=1}^{p} \mathbb{I}_2 \otimes \cdots \otimes R_{\mathrm{f}}(X) \otimes \cdots \otimes \mathbb{I}_2$$

where in the sum $R_{\mathrm{f}}(X)$ only appears in the i'th position.

7.3 Symmetry and Reducibility

In general such a representation will be quite reducible and we would like to find a way to pick out the irreducible parts. First lets see where the problem is. Consider a rank 2 tensor, T_{ij}, we can decompose it into its symmetric, $T_{(ij)}$, and anti-symmetric, $T_{[ij]}$, parts

$$T_{ij} = T_{(ij)} + T_{[ij]}, \quad T_{(ij)} = \frac{1}{2}\left(T_{ij} + T_{ji}\right), \quad T_{[ij]} = \frac{1}{2}\left(T_{ij} - T_{ji}\right).$$

Symmetric tensors live in the symmetric part of $W \otimes W$ which is denoted by $\mathrm{S}ym(W \otimes W)$, antisymmetric tensors live in the part denoted by $W \wedge W$ and we have shown that all rank 2 tensors can written as a combination of a symmetric and an anti-symmetric part so we have

$$W \otimes W = \mathrm{S}ym(W \otimes W) + W \wedge W.$$

The 'problem' is that under the representation R_{T} the two parts do not mix, for example

$$R_T T_{ij} = R_{\mathrm{f}}(X)_i^{\ k} T_{kj} + R_{\mathrm{f}}(X)_j^{\ k} T_{ik}$$

implies that

$$R_T\, T_{(ij)} = R_{\mathrm{f}}(X)_i^{\ k} T_{(kj)} + R_{\mathrm{f}}(X)_j^{\ k} T_{(ik)} = R_{\mathrm{f}}(X)_i^{\ k} T_{(kj)} + R_{\mathrm{f}}(X)_j^{\ k} T_{(ki)}$$

the right hand side of which is symmetric in i and j sharing the same symmetry property as the left hand side. This means that if we start with a symmetric tensor we end up with a symmetric tensor. So we have that

$$R_{\mathrm{T}} : \mathrm{S}ym(W \otimes W) \rightarrow \mathrm{S}ym(W \otimes W)$$

and the representation is reducible.

Exercise 7.3. Prove that $R_{\mathrm{T}} : W \wedge W \rightarrow W \wedge W$ as well.

For $su(2)$, if W is the fundamental representation space, then $W \wedge W$ is one dimensional. Another way to pose this statement is to say that the space of antisymmetric 2×2 tensors is one dimensional. Basically we are saying that any anti-symmetric tensor of rank 2 can be written as

$$T_{[ij]} = f\, \epsilon_{ij}.$$

Since this representation for the Lie algebra of $SU(2)$ is one dimensional this must be a $j = 0$ representation, that is a trivial representation. Lets check that this is so. From our tensor product representation rules we can write the tensor as

$$f \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}. \tag{7.3.1}$$

In the tensor product basis above,

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad E_+ = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad E_- = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

Exercise 7.4. Derive these formulae.

Act any of these matrices on the vector (7.3.1) and you get zero. So the matrix representatives on this space for the algebra are zero, i.e. this is the trivial representation.

Exercise 7.5. Give a vector basis for a symmetric rank 2 $su(2)$ tensor where the basis vectors are eigenstates of the Cartan element. Is this an irreducible representation and if so which one?

Exercise 7.6. Prove that for a vectors space of dimension n all totally anti-symmetric tensors of rank n form a one dimensional space and the only such tensor of rank $> n$ is the zero tensor.

Life usually gets a little complicated. Consider the rank 3 tensor, T_{ijk}. We can decompose it into a totally symmetric part a totally anti-symmetric part and what is left over. The part which is left over can then be expressed in terms of tensors of ‘mixed’ symmetry as

$$T_{ijk} = T_{(ijk)} + T_{[ijk]} + T_{ij;k} + T_{ik;j}.$$

One can use the symmetric group to write out everything in detail. The action of the symmetric group is understood to be on the position of the label in the tensor. For example

$$\begin{aligned}
T_{(ijk)} &= \frac{1}{6}\left(1 + g(12) + g(13) + g(23) + g(123) + g(132)\right) T_{ijk}, \\
T_{[ijk]} &= \frac{1}{6}\left(1 - g(12) - g(13) - g(23) + g(123) + g(132)\right) T_{ijk}, \\
T_{ij;k} &= \frac{1}{3}\left(1 - g(13)\right).\left(1 + g(12)\right) T_{ijk}, \\
T_{ik;j} &= \frac{1}{3}\left(1 - g(12)\right).\left(1 + g(13)\right) T_{ijk}.
\end{aligned}$$

The coefficients that appear are known as Clebsch-Gordon coefficients and there is a formalism for determining them, though we will not enter into that. To see that the coefficients are correct just add up the tensors in this list and see if you get T_{ijk} back. Note that the tensors of mixed symmetry are actually anti-symmetric in the first and second labels, with no apparent symmetry between the first and second.

So the general problem we are faced with is to ‘decompose’ arbitrary tensors into all the possible symmetry types. For $su(2)$ life turns out to be simple.

By the exercise we learn that if we have an $su(2)$ tensor rank p of the fundamental representation which is anti-symmetric in any two labels we can simply ignore those labels and consider the tensor to be of rank $p-2$. Again it is simple to see this in equations

$$T_{i_1\ldots i_{r-1}\, i_r\, i_{r+1}\ldots i_{s-1}\, i_s\, i_{s+1}\ldots i_p} = \epsilon_{i_r i_s}\, T_{i_1\ldots i_{r-1}\, i_{r+1}\ldots i_{s-1}\, i_{s+1}\ldots i_p}$$

where

$$T_{i_1\ldots i_{r-1}\, i_{r+1}\ldots i_{s-1}\, i_{s+1}\ldots i_p} = T_{i_1\ldots i_{r-1}\, 1\, i_{r+1}\ldots i_{s-1}\, 2\, i_{s+1}\ldots i_p}.$$

This means that essentially the only symmetry property we need to take care of is total symmetry.

7.4 Young Tableaux

The combinatorics is greatly helped along by what are known as Young tableaux (plural, tableau is singular). The fundamental representation space W is designated by a box $\Box$, then the tensor product space $W \otimes W$ is,

$$\boxed{1} \otimes \boxed{2} = \boxed{1}\boxed{2} \oplus \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline \end{array}$$

The horizontal boxes are understood to be symmetrized while the vertical ones are anti-symmetrized. The numbers 1 and 2 denote the first and second labels of our tensor. The tensor of rank 3 belongs in the space $W \otimes W \otimes W$ and can be decomposed as

$$\boxed{1} \otimes \boxed{2} \otimes \boxed{3} = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline \end{array} \oplus \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline \end{array} \oplus \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array} \oplus \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1} \end{array}$$

The tensors of mixed symmetry are shown as the last two Young tableaux. The rule is that those on horizontal lines are symmetrized while those vertically are anti-symmetrized. Some more rules to remember are that:

1. Rows cannot have gaps between boxes
2. Each row begins flush on the left
3. The i'th row is as long as or longer than the $i+1$'th row.

In general there are more rules but this will do for our purposes here.

Exercise 7.7. Show that for $SU(2)$ one can ignore the vertical part of these tableaux. Or that $2 \times 2 \times 2 = 4 + 0 + 2 + 2$.

Exercise 7.8. Show that if the dimension of W is 3 then for a rank 3 tensor the above decomposition says that $3 \times 3 \times 3 = 10 + 1 + 8 + 8$.

The quadratic Casimir of $SU(2)$ for tensor products is

$$C_{2T} = \sum_{i=1}^{p} \mathbb{I}_2 \otimes \cdots \otimes C_2 \otimes \cdots \otimes \mathbb{I}_2$$

with the sum being over all positions. Now one should

Exercise 7.9. Check that $[C_{2T}\,,\, R_T(X)] = 0 \;\forall\; X \in su(2)$.

7.5 Representations of the Groups $SU(2)$ and $SO(3)$

So far we have focused on the Lie algebra of $SU(2)$ which is the same as the Lie algebra of $SO(3)$ and so their representations are the same. As groups however, the representations are certainly **not** the same. We already know that the 2-dimensional representation of $SU(2)$ is not a representation of $SO(3)$. Lets recall that argument. A J_3 rotation with group element

$$\exp\left(\theta\, J_3\right),$$

is such that the periodicity is $\theta \to \theta + 2\pi$, however, the corresponding 2-dimensional representation element is, thanks to the adjoint map,

$$\exp\left(\theta\, i \frac{\sigma_3}{2}\right), \tag{7.5.1}$$

which has periodicity given by $\theta \to \theta + 4\pi$. Consequently the 2-dimensional representation is not a representation of $SO(3)$ (the group, not the algebra).

What of other representations of the algebra, which of those are representations of $SO(3)$? We can answer this question by looking at tensor products. The analogue of (7.2.2) for groups is

Definition 7.1. The tensor product representation of the fundamental representation of a Lie Group is

$$R_T(g)\, T_{i_1 \dots i_p} = R_{\mathbf{f}}(g)_{i_1}^{\ j_1} \dots R_{\mathbf{f}}(g)_{i_p}^{\ j_p}\, T_{j_1 \dots j_p}, \tag{7.5.2}$$

Notice that this is just like the rule for transformations of tensors in special relativity. It is a good rule since

$$R_T(g_2 \circ g_1) = R_T(g_2)\,.\,R_T(g_1),$$

just as in the defining representation. This is the definition of a representation for a group. The group elements are mapped to matrices and the group product is respected by the map.

Now suppose that we have a rank r tensor transforming under $su(2)$ then it is automatically a representation for $SU(2)$. Lets specialize to the group element (7.5.1). Notice that if the rank of the tensor is odd then R_T is periodic with period 4π, while if the rank is even then R_T is periodic with period 2π. This tells us that those tensor products which are even are indeed representations of $SO(3)$ while those which are odd are not.

The tensor product representations are reducible, but their irreducible parts have the same periodicity properties. The totally symmetric component of a rank r $SU(2)$ tensor has dimension $2.3.\dots.(2+r-1)/r! = r+1$. So if the rank is even we get all the odd dimensional representations and these are then the ones that are representations of $SO(3)$ (as well as of $SU(2)$). When the rank is odd the dimensions are even and these are not representations of $SO(3)$.

7.6 Explicit Actions on Tensor Products

It is useful to see once how the algebra works on a tensor product in all its detail. We will consider rank 2 tensors.

In the case of $SU(2)$ we let e_i be a basis of $V_{\mathbf{f}}$ the two dimensional fundamental representation (the vector space). We take the e_i to form the standard basis, meaning

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{7.6.1}$$

As an eigenstate of $h = \sigma_3/2$ the vector e_1 has eigenvalue 1/2 while e_2 has eigenvalue -1/2, so in Dirac notation we write

$$e_1 = |1/2\rangle, \quad e_2 = |-1/2\rangle \tag{7.6.2}$$

The positive and negative roots act as

$$E_+.|1/2\rangle = 0, \quad E_+.|-1/2\rangle = |1/2\rangle, \quad E_-.|1/2\rangle = |-1/2\rangle, \quad E_-.|-1/2\rangle = 0 \quad (7.6.3)$$

The basis, written as eigenstates of h, for the tensor product space $V_{\mathbf{f}} \otimes V_{\mathbf{f}}$ is

$$\begin{array}{cc} |1/2\rangle \otimes |1/2\rangle & |1/2\rangle \otimes |-1/2\rangle \\ |-1/2\rangle \otimes |1/2\rangle & |-1/2\rangle \otimes |-1/2\rangle \end{array} \quad (7.6.4)$$

Given that we know how Lie $SU(2)$ acts on $V_{\mathbf{f}}$ then on its tensor product $V_{\mathbf{f}} \otimes V_{\mathbf{f}}$ we have

$$e_i \longrightarrow R_{\mathbf{f}}(X).e_i, \quad e_i \otimes e_j \longrightarrow R_T(X).e_i \otimes e_j = R_{\mathbf{f}}(X).e_i \otimes e_j + e_i \otimes R_{\mathbf{f}}(X).e_j$$

which tells us that we add the actions of $E_\pm$ on the two factors of the tensor product. For example

$$\begin{array}{ll} R_T(E_+).|1/2\rangle \otimes |-1/2\rangle = |1/2\rangle \otimes |1/2\rangle, & R_T(E_+).|-1/2\rangle \otimes |1/2\rangle = |1/2\rangle \otimes |1/2\rangle \\ R_T(E_-).|1/2\rangle \otimes |-1/2\rangle = |-1/2\rangle \otimes |-1/2\rangle, & R_T(E_-).|-1/2\rangle \otimes |1/2\rangle = |-1/2\rangle \otimes |-1/2\rangle \\ R_T(h).|1/2\rangle \otimes |-1/2\rangle = 0, & R_T(h).|-1/2\rangle \otimes |1/2\rangle = 0 \end{array}$$

which tells us that the combination

$$\frac{1}{2}\left(|1/2\rangle \otimes |-1/2\rangle - |-1/2\rangle \otimes |1/2\rangle\right) \quad (7.6.5)$$

is invariant under the action of $SU(2)$. It is also evident that this is the anti-symmetric part of the tensor product, which we saw is invariant by other means before. So we are saying that we may decompose the tensor product space as

$$V_{\mathbf{f}} \otimes V_{\mathbf{f}} = V_{\mathbf{3}} \oplus V_{\mathbf{1}} \quad (7.6.6)$$

where $V_{\mathbf{3}}$ is spanned by $|1/2\rangle \otimes |1/2\rangle$, $|-1/2\rangle \otimes |-1/2\rangle$ and $(|1/2\rangle \otimes |-1/2\rangle + |-1/2\rangle \otimes |1/2\rangle)/2$ (and is obviously the symmetric 2 tensor representation). There are no invariant subspaces in $V_{\mathbf{3}}$, since

$$R_T(E_+)\left(a.|1/2\rangle \otimes |1/2\rangle + b.|-1/2\rangle \otimes |-1/2\rangle + c(|1/2\rangle \otimes |-1/2\rangle + |-1/2\rangle \otimes |1/2\rangle)\right) = 0$$

implies that $a = b = c = 0$. Hence the $\mathbf{2} \otimes \mathbf{2}$ representations decomposes into the $\mathbf{3} \oplus \mathbf{1}$ irreducible representations. One notes that the highest weight of the $\mathbf{3}$ is $|1\rangle \otimes |1\rangle$ with $\sigma_3/2$ eigenvalue 1 (i.e $j = 1$).

The highest weight vector of the two dimensional $\mathbf{2}$ (fundamental) representation is the $|1/2\rangle$ while that of the three dimensional representation $\mathbf{3}$ is the $|1\rangle = |1/2\rangle \otimes |1/2\rangle$, where we use the notation that vectors are labelled by their $h = \sigma_3/2$ eigenvalue even for the tensor product. As we saw in Section 6.3 the highest weight states are those that are mapped to the zero vector by E_+. With our present notation the states

$$|n/2\rangle = |1/2\rangle \otimes \cdots \otimes |1/2\rangle \quad (7.6.7)$$

are highest weight vectors. The highest weight for the fundamental is set to be

$$\mu = |1/2\rangle \tag{7.6.8}$$

and, with a slight abuse of notation, one writes

$$|n/2\rangle = n.\mu, \quad \mu = |1/2\rangle \tag{7.6.9}$$

for the highest weight of the $j = n/2$ representation of dimension $n + 1 = 2j + 1$.

CHAPTER 8

$SU(3)$ AND ITS LIE ALGEBRA

This is the next most complicated group. Interestingly enough the theory of the fundamental interactions is, as far as we know, $U(1) \times SU(2) \times SU(3)$. So, once you have this group under control you will have a large portion of particle physics theory at your finger tips.

8.1 The Definition of the Lie Group $SU(3)$

As always we start with a

Definition 8.1. $SU(3)$ is the group of 3×3 matrices with complex coefficients which satisfy

$$U \,.\, U^{\dagger} = \mathbb{I}_3, \quad \operatorname{Det} U = 1,$$

$\forall U \in SU(3)$.

This definition is the same as that for $SU(2)$ except that everywhere the number 2 has been replaced by the number 3. You should check that matrices of this type do form a group under the usual matrix multiplication.

Unlike the case of $SU(2)$ we will not get a nice description of $SU(3)$ as a space (except by its equations). Nevertheless we can work out its dimension by the same discussion that we used in the case of $SU(2)$. Lets repeat that here except in a slightly more general context. Let U be a $n \times n$ unitary matrix with unit determinant. Then U will have n^2 complex entries or $2n^2$ real entries. The unitarity condition is n^2 conditions but (by hermitian conjugation we see that) they are essentially n^2 real conditions. The determinant is unity condition is one real condition (since unitarity in any case implies $\operatorname{Det} U \,.\, \operatorname{Det} U^{\dagger} = 1$) so altogether we have $2n^2$ real unknowns in $n^2 + 1$ real equations. This leaves us with $n^2 - 1$ real undetermined parameters and so the space of $U(n)$ matrices is $n^2 - 1$ dimensional.

Therefore, the dimension of $SU(3)$ as a space is 8. We should get the same dimension by calculating the dimension of the vector space of generators of $SU(3)$.

Though we have not shown it the spaces, or manifolds, that one gets this way are smooth. Smooth, as mentioned before, means that the manifolds have functions on them that can be differentiated as often as we like. Locally, these spaces look just like copies of R^m (where m is the dimension of the space) and all of these local spaces glue together to make a nice smooth manifold.

8.2 $\mathrm{Lie}\, SU(3)$ the Lie Algebra of $SU(3)$

The calculations we need to perform, to determine the Lie algebra of $SU(3)$ are almost identical to those we have already done in the case of $SU(2)$. I will repeat them here but again in a slightly more general context. Let U be a $n \times n$ unitary matrix with unit determinant and expand such a U around the identity element $\mathbb{I}_n$ as

$$U = \mathbb{I}_n + i\epsilon\, M + O(\epsilon^2),$$

with ϵ an infinitesimal parameter. The unitarity condition tells us that

$$U \,.\, U^\dagger = \mathbb{I}_n = \left(\mathbb{I}_n + i\epsilon\, M + O(\epsilon^2)\right) \,.\, (\mathbb{I}_n - i\epsilon\, M^\dagger + O(\epsilon^2))$$

which implies that

$$M = M^\dagger.$$

That the determinant of U is unity implies that

$$\mathrm{Tr}\, M = 0$$

as well. So the generators are Hermitian and traceless. The number of linearly independent such matrices, for $SU(3)$ is 8. One way to prove this is to count equations as before or equivalently we have

Exercise 8.1. Show that the following 8 matrices (called the Gell-Mann matrices) span the space of Hermitian generators of $SU(3)$,

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

The normalization is chosen for convenience so that

$$\mathrm{Tr}\,(\lambda_a\,\lambda_b) = 2\delta_{ab}. \tag{8.2.1}$$

This takes care of the vector space structure of $su(3)$, but for it to be Lie algebra the commutator two elements of $su(3)$ should give you another element of $su(3)$. In fact we have by a direct calculation that

$$\left[\frac{\lambda_a}{2}, \frac{\lambda_b}{2}\right] = i f_{abc} \frac{\lambda_c}{2}$$

where the structure constants f_{abc} are real and totally antisymmetric. The constants f_{abc} are as follows (you should check this for yourself)

$$\begin{aligned} &f_{123} = 1, \quad f_{147} = 1/2, \quad f_{156} = -1/2, \quad f_{257} = 1/2 \\ &f_{345} = 1/2, \quad f_{367} = -1/2, \quad f_{458} = \sqrt{3/2}, \quad f_{678} = \sqrt{3/2}. \end{aligned}$$

To see that the structure constants are totally antisymmetric we use the trace normalization. We have that

$$4i f_{abc} = \mathrm{Tr}\,([\lambda_a, \lambda_b]\;\lambda_c),$$

and now we make use of the cyclicity of the trace to

Exercise 8.2. Show that

$$\mathrm{Tr}\,([\lambda_a, \lambda_b]\;\lambda_c) = \mathrm{Tr}\,([\lambda_b, \lambda_c]\;\lambda_a)$$

so that $f_{abc} = f_{cab}$.

8.3 The Jacobi Identity and the Adjoint Representation

The Jacobi identity is exactly that it is **identically** true for any set of $m \times m$ matrices X, Y and Z (these can be complex). These matrices do not need to belong to a group. The Jacobi identity is

$$[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0 \tag{8.3.1}$$

To prove this just expand out each commutator and you will get 12 terms which graciously cancel pairwise.

Looking ahead, we expect that any Lie algebra commutation relations will take the form

$$[T_a, T_b] = C^c_{ab}\, T_c \tag{8.3.2}$$

for some set of constants C_{abc} and generators T_a. The Jacobi identity puts conditions on what the structure constants C^c_{ab} can be. For example, suppose a friend said to you that there is a Lie algebra with 3 non-zero generators with the commutation relations

$$[Y, Z] = X, \quad [X, Y] = X, \quad [Z, X] = Y.$$

If you plug these into (8.3.1) you would find

$$0 = [X, Z] + [X, X] + [Y, Y] = -Y \neq 0.$$

and you could tell your friend that he or she is wrong!

Now plugging the bracket relation (8.3.2) into the Jacobi identity (with $X = T_a$, $Y = T_b$ and $Z = T_c$) we get

$$\left(C_{ab}^{d}\, C_{dc}^{e} + C_{bc}^{d}\, C_{da}^{e} + C_{ca}^{d}\, C_{db}^{e} \right) T_e = 0.$$

As we will take the Lie algebra to be a vector space spanned by the T_a (as is the case for $SU(3)$ being considered here, $T_a = \lambda_a/2$), we can strip off the generator at the end[1] (but more on this in the second half of the course) to arrive at

$$C_{ab}^{d}\, C_{dc}^{e} + C_{bc}^{d}\, C_{da}^{e} + C_{ca}^{d}\, C_{db}^{e} = 0. \tag{8.3.3}$$

You can check that the structure constants, $f_{abc} = -iC_{bc}^{a}$ that we wrote down before for the Lie algebra of $SU(3)$ do indeed satisfy these equations.

Now comes a great thing: from the Jacobi identity we realize we automatically have another representation of the Lie algebra, the **adjoint** representation by setting

$$(T_a)^c{}_b = C_{ab}^{c}. \tag{8.3.4}$$

Exercise 8.3. Show that (8.3.4) satisfies the algebra (8.3.2) by virtue of the Jacobi identity (8.3.3).

Don't panic, you have seen this before! In the particular case of $SU(2)$, section 5.8, we defined the adjoint action ad_X and we showed that it gave rise to the generators of $SO(3)$ namely the $(J_a)^b_c = \epsilon_{abc}$. Now we have shown that, quite generally, the adjoint action provides us with a representation, the representation space being the Lie algebra itself. (When comparing with section 5.8 watch out for signs.)

Some of this is formalized below.

8.4 The Cartan Sub-Algebra and Roots

If you look at the generators of $SU(3)$ you see that λ_1, λ_2 and λ_3 by themselves form the Lie algebra of $SU(2)$, more explicitly we have, for $i = 1, 2, 3$

$$\lambda_i = \left(\begin{array}{c|c} \sigma_i & \\ \hline & 0 \end{array} \right).$$

Likewise, there is another $SU(2)$ sub-algebra made out of the matrices, λ_6, λ_7 and $\lambda_8/2\sqrt{3} - \lambda_3/2$, which take the form

$$\left(\begin{array}{c|c} 0 & \\ \hline & \sigma_i \end{array} \right).$$

[1] One way to do this, for $SU(3)$, is to multiply by λ_k and use the trace condition (8.2.1).

There are other copies of $su(2)$ in $su(3)$ as well. But we can start with these two. (Note that they do not commute). The other two generators, λ_4 and λ_5 'move' things around and connect the different $su(2)$'s. We will base our study of $su(3)$ on our understanding of $su(2)$.

Just as for $SU(2)$ we pick a maximal set of commuting hermitian elements. In this case there are two such elements,

$$h_1 = E_{11} - E_{22} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad h_2 = E_{22} - E_{33} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{8.4.1}$$

where E_{ij} is the matrix with only one non-zero entry namely at the i'th row on the j'th column where it is unity. These satisfy

$$[h_1, h_2] = 0. \tag{8.4.2}$$

The elements h_1 and h_2 are again called Cartan elements and they form a sub-algebra of $su(3)$ which is known as the Cartan sub-algebra. This sub-algebra is an Abelian sub-algebra as everything in it commutes with everything else in it. Denote the Cartan sub-algebra by $\mathfrak{h}$.

Again just as for $su(2)$ we have positive and negative root vectors, but because we have two copies of $su(2)$ and extra generators there are different types of positive and negative root vectors so we give them slightly different labels. Let the positive root vectors be given by

$$X_{\alpha_1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad X_{\alpha_2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad X_{\alpha_3} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{8.4.3}$$

so that $X_{\alpha_1} = E_{12}$, $X_{\alpha_2} = E_{23}$ and $X_{\alpha_3} = E_{13}$. The negative roots are

$$X_{-\alpha_1} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad X_{-\alpha_2} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad X_{-\alpha_3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{8.4.4}$$

Note that $X_{\alpha_i}^\dagger = X_{-\alpha_i}$ and $X_{-\alpha_1} = E_{21}$, $X_{-\alpha_2} = E_{32}$ and $X_{-\alpha_3} = E_{31}$.

The α_i are called positive roots. We denote the set of all roots by Δ. Δ is referred to as the root system of the Lie algebra. The space of positive roots is denoted by Δ_+.

The positive root vectors are called that because they are eigenvectors of the Cartan elements (under the adjoint action) with positive eigenvalues given by the positive roots. Indeed[2] for all $h \in \mathfrak{h}$ and for all $\alpha_i \in \Delta$ (so regardless of the sign of the root)

$$[h, X_{\alpha_i}] = \alpha_i(h)\, X_{\alpha_i}.$$

[2]Another way of saying this is that the α_i are the solutions, hence roots, of the characteristic polynomial equation $\mathrm{Det}\,(\alpha\mathbb{I} - \mathrm{ad}\, h) = 0$.

One can work out what this means by multiplying everything out. For example we learn that for $i = 1, 2$

$$\alpha_i(h_i) = 2.$$

Exercise 8.4. Show that

$$\alpha_1(ah_1 + bh_2) = 2a - b, \quad \alpha_2(ah_1 + bh_2) = -a + 2b, \quad \alpha_3(ah_1 + bh_2) = a + b.$$

We can formalise this by a definition

Definition 8.2. For $\mathfrak{g}$ a Lie algebra, $\mathfrak{h}$ a Cartan sub algebra $\mathfrak{h} \subset \mathfrak{g}$ and $\alpha \in \mathfrak{h}^*$ let

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} : [H, X] = \alpha(H)\, X, \quad \forall H \in \mathfrak{h}\}$$

then if $\alpha \neq 0$ and $\mathfrak{g}_\alpha \neq 0$ α is called a root and $\mathfrak{g}_\alpha$ is the root space. A non-zero element of $\mathfrak{g}_\alpha$ is called a root vector for the root α.

Notice that in this definition we now think of the α_i as maps from the Lie algebra to the real numbers (in fact integers) so that they live in the dual space of the Cartan algebra. The root spaces $\mathfrak{g}_\alpha$ for the classical groups (coming up in a chapter or two) are all one dimensional, a fact that we have seen both for $SU(2)$ and for $SU(3)$.

From the exercise we learn that $\alpha_1 + \alpha_2 = \alpha_3$. (This does **not** mean that $X_{\alpha_1} + X_{\alpha_2} = X_{\alpha_1}$.) If we denote the set of all roots by Δ then you should also note that $\alpha_1 + \alpha_3 \notin \Delta$ but certainly $\alpha_1 + (-\alpha_3) = -\alpha_2 \in \Delta$.

So far we have determined the commutators of the Cartan elements with root vectors we still need to work out the commutators of the root vectors with each other. We can easily evaluate the commutator of positive root vectors with their respective negative root vectors,

$$[X_{\alpha_1}, X_{-\alpha_1}] = h_1, \quad [X_{\alpha_2}, X_{-\alpha_2}] = h_2, \quad [X_{\alpha_3}, X_{-\alpha_3}] = h_1 + h_2. \tag{8.4.5}$$

There are other commutators as well but we can put the whole thing into one

Exercise 8.5. Show that

$$[X_\alpha, X_\beta] = N_{\alpha\beta}\, X_{\alpha+\beta} \tag{8.4.6}$$

where $\alpha + \beta \in \Delta$ and

$$[X_\alpha, X_\beta] = 0, \tag{8.4.7}$$

with $\alpha + \beta \neq 0$ and $\alpha + \beta \notin \Delta$. The non-zero coefficients $N_{\alpha\beta}$ are, for $su(3)$,

$$N_{\alpha_1\alpha_2} = 1, \quad N_{\alpha_1(-\alpha_3)} = -1, \quad N_{\alpha_2(-\alpha_3)} = 1, \quad N_{\alpha_3(-\alpha_1)} = 1, \quad N_{\alpha_3(-\alpha_2)} = -1,$$

and similar expressions with all roots the negative of the ones here.

The abstract Lie algebra $su(3)$ is understood to be the lie algebra of the 8 objects h_1, h_2 and $X_{\pm\alpha_i}$ with $i = 1, 2, 3$, satisfying the relations (8.4.2-8.4.7).

8.5 Fundamental Representations of $SU(3)$

You should note the use of the plural in the title of this section. We take the fundamental representation to be that given by the matrices (8.4.1, 8.4.3, 8.4.4) but we denote these by capital letters as before

$$h_1 \to H_1, \quad h_2 \to H_2, \quad X_{\pm\alpha_i} \to E_{\pm\alpha_i}.$$

Since the Hermitian matrices H_1 and H_2 commute we can (and have) simultaneously diagonalized them. The eigenvalues are given (since we have the matrices in diagonal form) and the eigenvectors are

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

These are again called weight vectors and the eigenvalues are weights. One defines the **weight vectors**, M, by

$$H\, e_a = M^a(H)\, e_a, \quad H \in \mathfrak{h}.$$

with no sum over a.

If we write $H = aH_1 + bH_2$ we discover that

Exercise 8.6.

$$M^1 = \frac{2}{3}\alpha_1 + \frac{1}{3}\alpha_2, \quad M^2 = -\frac{1}{3}\alpha_1 + \frac{1}{3}\alpha_2, \quad M^3 = -\frac{1}{3}\alpha_1 - \frac{2}{3}\alpha_2$$

Aside: Note the following interesting relations

$$M^1 + M^2 + M^3 = 0, \quad M^2 = M^1 - \alpha_1, \quad M^3 = M^2 - \alpha_2.$$

The first follows from what we know namely $\operatorname{tr} H = 0 = \sum_a M^a(H)$. The other two will become clearer as we go on.

We can label the eigenvectors by the eigenvalues of the Cartan elements or by the weights. For now lets label them by their eigenvalues with respect to the Cartan elements H_1 and H_2. In Dirac notation we write an eigenstate as $|\lambda_1, \lambda_2\rangle$ for the eigenvalues λ_1 of H_1 and λ_2 of H_2. With this notation we have

$$e_1 = |1, 0\rangle\,, \quad e_2 = |-1, 1\rangle\,, e_3 = |0, -1\rangle\,. \tag{8.5.1}$$

This is the 3-dimensional fundamental representation and is denoted **3**.

Now lets think about the complex conjugate representation. Remember this tells us to map $R(X) \to -\overline{R}(X)$ so we learn that we must send $H_1 \to -H_1$, $H_2 \to -H_2$ and it also tells us to map

$$E_{\pm\alpha_i} \to -E_{\mp\alpha_i}. \tag{8.5.2}$$

The latter relation comes about as follows. The root vectors are not in our Lie algebra (they are not Hermitian), the combinations which are in the Lie algebra are

$$(E_{\alpha_i} + E_{-\alpha_i}) \,, \quad i\,(E_{\alpha_i} - E_{-\alpha_i}) \,. \tag{8.5.3}$$

These map to

$$-\,(E_{\alpha_i} + E_{-\alpha_i}) \,, \quad i\,(E_{\alpha_i} - E_{-\alpha_i}) \tag{8.5.4}$$

respectively. If one performs the change (8.5.2) in (8.5.3) then (8.5.4) is also obtained. In any case we get a 3-dimensional representation again of $su(3)$. Notice that the eigenvectors are now

$$\overline{e}^1 = |-1,0\rangle \,, \quad \overline{e}^2 = |1,-1\rangle \,, \overline{e}^3 = |0,1\rangle \,. \tag{8.5.5}$$

As vectors e^a and $\overline{e}^a$ are the same but as weights they are of course different. This representation is also 3-dimensional but because it comes from complex conjugating it is denoted by $\mathbf{3}^*$ (for now).

8.6 The Eigenvectors of the Adjoint Action

Lets look more closely at the adjoint map (and also somewhat more formally). The commutator of two adjoint maps is the adjoint map of the commutator, this means that

$$\mathrm{ad}_X \,.\, \mathrm{ad}_Y \, Z - \mathrm{ad}_Y \,.\, \mathrm{ad}_X \,.Z = \mathrm{ad}_{[X,Y]} \,.Z, \quad \forall X, Y, Z \in \mathrm{L}ie\ G \tag{8.6.1}$$

To prove this we just use the definition of the adjoint map, namely

$$\mathrm{ad}_X \,.Z = [X, Z],$$

together with the Jacobi identity

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

ad maps the Lie algebra to itself linearly. Since a Lie algebra is a vector space we have that ad is a linear map of the vector space to itself, so that ad can be represented by a matrix which we call R_{ad}. This means that we write

$$\mathrm{ad}_X = R_{\mathrm{ad}}(X).$$

The identity (8.6.1) can now be expressed as

$$[R_{\mathrm{ad}}(X), R_{\mathrm{ad}}(Y)] = R_{\mathrm{ad}}([X,Y]),$$

which tells us that we are dealing with a representation of the Lie algebra, namely the adjoint representation.

This is a neat way of capturing what we know, namely that the structure constants furnish a representation of the Lie algebra. Let $X = X^a T_a$ and $Z = Z^b T_b$ then

$$\mathrm{ad}_X\, Z = X^a\, C^b_{ac}\, Z^c\, T_b$$

so that $Z^b \to R_{\mathrm{ad}}(X)^b{}_c\, Z^c$ with

$$R_{\mathrm{ad}}(X)^b{}_c = X^a\, C^b_{ac}.$$

We use the Gell-Mann matrices as a basis and see how they act on the basis

$$R_{\mathrm{ad}}\left(i\frac{\lambda_a}{2}\right).\frac{i\lambda_b}{2} = -f_{abc}i\frac{\lambda_c}{2}$$

so as a matrix we have

$$\left(R_{\mathrm{ad}}\left(i\frac{\lambda_a}{2}\right)\right)_b^{\ c} = -f_{abc}.$$

So, the adjoint representation acts on the Lie algebra itself. The eigenstates of the Cartan subalgebra are the root vectors. We can draw a diagram, called the root diagram, to exhibit this. As you will see it in your particle physics course as well, lets take a break from maths and do some physics.

8.7 A Few Words from Physics

When one mentions the complex conjugate representation in physics what comes to mind is, equivalently, the transpose of the $\mathbf{3}^*$. So instead of column vectors one considers row vectors instead. Then, as we saw for $su(2)$, the representation action goes over to

$$\bar{e}. - R^\dagger_{\mathbf{f}}(X) = -\bar{e}.R_{\mathbf{f}}(X).$$

This is usually what one means when one refers to the $\overline{\mathbf{3}}$ representation, but as I said before this is equivalent to what we called $\mathbf{3}^*$. The reason one thinks of it this way is because we can make invariants using anti-fields, in particular with quarks.

Also in the use of $SU(3)$ it is more appropriate to consider different combinations of the Cartan elements than the ones we considered. These Abelian elements correspond to conserved (or almost conserved) quantities and it is important to get the identification right. One usually takes the combinations

$$I_3 = \frac{1}{2}h_1, \quad Y = \frac{1}{3}h_1 + \frac{2}{3}h_2 \tag{8.7.1}$$

where I_3 is known as isospin and Y is called hyper-charge. Note that these are orthogonal with respect to trace,

$$\mathrm{Tr}\,(I_3 Y) = 0 \tag{8.7.2}$$

One of the important formulae in particle physics is that for the charge of the fundamental particles in terms of their isospin and hyper-charge, that formula is

$$Q = I_3 + \frac{Y}{2} \tag{8.7.3}$$

Also the notation for positive and negative roots is somewhat different (just the notation not the normalization)

$$I_\pm = X_{\pm\alpha_1}, \quad U_\pm = X_{\pm\alpha_2}, \quad V_\pm = X_{\pm\alpha_3}$$

With these generators the Lie algebra takes the form

$$\begin{array}{lll}
[I_3, I_\pm] = \pm I_\pm & [I_3, U_\pm] = \mp\frac{1}{2}U_\pm & [I_3, V_\pm] = \pm V_\pm \\
[Y, I_\pm] = 0 & [Y, U_\pm] = \pm U_\pm & [Y, V_\pm] = \pm V_\pm \\
[I_+, I_-] = 2I_3 & [U_+, U_-] = \frac{3}{2}Y - I_3 & [V_+, V_-] = \frac{3}{2}Y + I_3 \\
[I_+, V_-] = -U_+ & [I_+, U_+] = V_+ & [U_+, V_-] = I_- \\
[I_-, V_+] = U_- & [I_-, U_-] = -V_- & [U_-, V_+] = -I_+
\end{array} \tag{8.7.4}$$

with the other 7 non-trivial commutators being zero (non trivial means not the commutator of a generator with itself nor the negative of one of those above).

Now we should compare this with what we did for $su(2)$. In the case of $su(2)$ the positive root increased the value of j by 1, while the negative root decreased the value of j by 1. One can draw this as in figure 8.1.

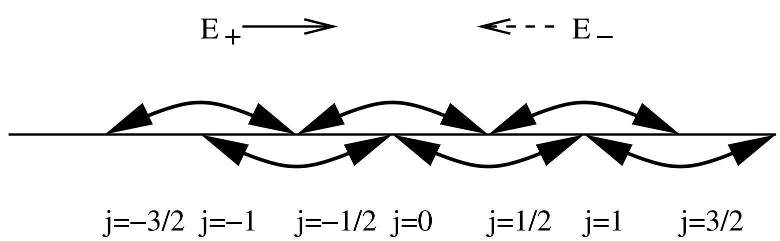


Figure 8.1: The Action of the Roots of SU(2)

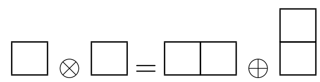
Like wise the positive and negative roots of $su(3)$ act as 'raising' and 'lowering' operators of the eigenvalues of the two Cartan elements. Denote a state by the eigenvalues λ_1 and λ_2 of I_3 and Y respectively. Then one can read off from the commutation relations that

$$\begin{array}{ll}
I_+|\lambda_1, \lambda_2\rangle \to |\lambda_1 + 1, \lambda_2\rangle & I_-|\lambda_1, \lambda_2\rangle \to |\lambda_1 - 1, \lambda_2\rangle \\
U_+|\lambda_1, \lambda_2\rangle \to |\lambda_1 - 1/2, \lambda_2 + 1\rangle & U_-|\lambda_1, \lambda_2\rangle \to |\lambda_1 + 1/2, \lambda_2 - 1\rangle \\
V_+|\lambda_1, \lambda_2\rangle \to |\lambda_1 + 1, \lambda_2 + 1\rangle & V_-|\lambda_1, \lambda_2\rangle \to |\lambda_1 - 1, \lambda_2 - 1\rangle
\end{array}$$

You should check that you believe these formulae.

8.8 Some Representation Theory and Young Tableaux

Given that we know that we have a **3** of $su(3)$ we can take tensor products. You have already shown that



which we can write as

$$\mathbf{3} \otimes \mathbf{3} = \mathbf{6} \oplus \overline{\mathbf{3}}.$$

Certainly the anti-symmetric tensor is 3-dimensional, but how did we know that it must be the $\overline{\mathbf{3}}$ representation and not the **3**? The way to see that we got it right is to first convert the anti-symmetric rank 2 tensor into a vector (so it is now obviously 3-dimensional) and to do that we use the fact that the rank 3 epsilon symbol ϵ^{ijk} is invariant under $SU(3)$. In equations this means that the vector

$$U^i \equiv \epsilon^{ijk} T_{jk}$$

is equivalent to the the anti-symmetric tensor T_{ij}. (We can invert this equation). But how does U^i transform? Well lets just apply the rule that we know for how T_{ij} transforms as well as the fact that ϵ^{ijk} is invariant, at the level of the Lie algebra,

$$U'^i = (R_T(X).U)^i = \epsilon^{ijk} \left(R_{\mathbf{f}}(X)_j^{\ l}\, T_{lk} + R_{\mathbf{f}}(X)_k^{\ l}\, T_{jl} \right) = 2\epsilon^{ijk}\, R_{\mathbf{f}}(X)_j^{\ l}\, T_{lk}.$$

However, to say that ϵ^{ijk} is invariant is to say that

$$\epsilon^{ljk}\, R_{\mathbf{f}}(X)_l^{\ i} + \epsilon^{ilk}\, R_{\mathbf{f}}(X)_l^{\ j} + \epsilon^{ijl}\, R_{\mathbf{f}}(X)_l^{\ k} = 0.$$

This last equation we can check directly. It is antisymmetric in any 2 of the labels (i, j, k) so that it is enough to check it by setting $i = 1$, $j = 2$ and $k = 3$, then it becomes

$$R_{\mathbf{f}}(X)_1^{\ 1} + R(X)_2^{\ 2} + R_{\mathbf{f}}(X)_3^{\ 3} = \mathrm{Tr}\, R_{\mathbf{f}}(X) = 0.$$

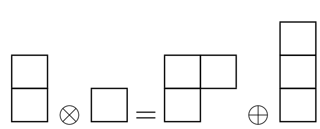
This means that

$$2\epsilon^{ijk} R_{\mathbf{f}}(X)_j^{\ l}\, T_{lk} = -U^l\, R_{\mathbf{f}}(X)_l^{\ i}$$

which is the transformation rule for the complex conjugate representation (after taking the transpose, i.e. after lifting the vector label).

Exercise 8.7. Show that, at the level of the group $SU(3)$ representations , that U^i transforms in the $\overline{\mathbf{3}}$.

We can keep multiplying to get new irreducible representations. Lets multiply the **3** and $\overline{\mathbf{3}}$ together. We obtain



or

$$\overline{\mathbf{3}} \otimes \mathbf{3} = \mathbf{8} \oplus \mathbf{1}.$$

The 8-dimensional representation we know, it is the adjoint representation.

We can write the tensor product of these two representations as

$$\overline{e}^i \, e_j$$

(The product $\overline{e}^i \, e_j$ is the matrix E_{ij} with the only non-zero entry being 1 in the i'th row j'th column) and this transforms as

$$R_T(X).\overline{e}^{\,i} \, e_j = -\overline{e}^{\,l} \, R_{\mathbf{f}}(X)_l^{\;i} \, e_j + \overline{e}^{\,i} \, R_{\mathbf{f}}(X)_j^{\;l} \, e_l.$$

Obviously $\sum_i e_i \, \overline{e}^{\,i} = \mathbb{I}_{3\times 3}$ is invariant and corresponds to the singlet or trivial representation $\mathbf{1}$.

We know that the $\mathbf{8}$ is the adjoint representation, but how can we see it directly? Well any 3×3 matrix can be expanded in the 8 Gell-Mann matrices and the unit matrix (a total of 9 = the number of entries of the matrix) with complex coefficients. Hence

$$e_j \, \overline{e}^i = \sum_{a=1}^{8} A_a \, (\lambda_a)_j^{\;i} + B\delta_j^i.$$

where A_a and B are matrices. To work out what the coefficients are just take the trace (sum over the explicit labels) of both sides to find that

$$B = \frac{1}{3} \sum_i e_i \, \overline{e}^{\,i} = \frac{1}{3} \mathbb{I}_{3\times 3}$$

or multiply by λ_b and take the trace (again sum over the explicit labels) to see that

$$A_a = \overline{e}. \, \frac{\lambda_a}{2} \, .e \, .$$

We know how B transforms-it is the invariant. We can also see how A_a transforms

$$R_T(\frac{\lambda_b}{2}) \, A_a = \overline{e}. \left(-\frac{\lambda_b}{2} \frac{\lambda_a}{2} + \frac{\lambda_a}{2} \frac{\lambda_b}{2} \right) .e = \overline{e}. \left[\frac{\lambda_a}{2}, \frac{\lambda_b}{2} \right] .e = i f_{abc} \, A_c = -i(f_b)_{ac} A_c,$$

that is it transforms in the adjoint representation. Actually if we think about what $\overline{e}^j \, (\lambda_a)_j^{\;i} \, e_i$ means we would conclude that $A_a = \lambda_a/2$. The reason for this is that $E_{ij} \, (\lambda_a)_{ij}$ (no sum) is the matrix E_{ij} but with the coefficient one now changed to $(\lambda_a)_{ij}$ on summing over i and j we just get back the matrix λ_a.

Exercise 8.8. Prove that

$$\begin{aligned} e_j \, \overline{e}^{\,i} \;\; &= \;\; \sum_{k=1}^{3} \left(\overline{e}.E_{\alpha_k}.e \; (E_{-\alpha_k})_j^{\;i} + \overline{e}.E_{-\alpha_k}.e \; (E_{\alpha_k})_j^{\;i} \right) \\ &\qquad + 2\overline{e}.I_3.e \; (I_3)_j^{\;i} + \frac{3}{2}\overline{e}.Y.e \; (Y)_j^{\;i} + \frac{1}{3}\overline{e}.e \, \delta_j^{\;i} \end{aligned}$$

We can get other representations as well in this way. For example the totally symmetric tensor of rank 3 is a **10**.

Unfortunately the rules for the construction of Young tableaux, for general groups, are somewhat more complicated than those for $SU(2)$. These will be explained in a seperate set of notes.

8.9 A Few More Words From Physics

There is (almost) a symmetry group in particle physics called the $SU(3)$ flavour symmetry group. The idea is that there are 3 different quarks, called u, d and s which can be 'confused' with each other. The quarks are themselves Dirac spinors but I will generally ignore the Dirac label. The quarks also transform under another $SU(3)$ group called the colour group, but I will ignore that label as well.

Set

$$u = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad d = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad s = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

and

$$\overline{u} = (1,0,0), \quad \overline{d} = (0,1,0), \quad \overline{s} = (0,0,1).$$

With respect to the commuting set (8.7.1) these vectors can be written as

$$u = |\frac{1}{2}, \frac{1}{3} >, \quad d = |-\frac{1}{2}, \frac{1}{3} >, \quad s = |0, -\frac{2}{3} >,$$

and

$$\overline{u} = |-\frac{1}{2}, -\frac{1}{3} >, \quad \overline{d} = |\frac{1}{2}, -\frac{1}{3} >, \quad \overline{s} = |0, \frac{2}{3} > .$$

Let q_i denote the 3-quarks u, d and s and let $\overline{q}^i$ denote their conjugates $\overline{u}$, $\overline{d}$ and $\overline{s}$ respectively.

Exercise 8.9. Plot the vectors u, d, and s on a two dimensional plane with axes the eigenvalues of I_3 and Y. On the graph show the action of all the generators.

With the above notation the invariant scalar is, up to a normalization factor,

$$\overline{q}^i . q_i = \overline{u}u + \overline{d}d + \overline{s}s = \sqrt{6}\eta'.$$

We determine the matrix $q_j\overline{q}^i$

$$q_j\overline{q}^i = \begin{pmatrix} u\overline{u} & u\overline{d} & u\overline{s} \\ d\overline{u} & d\overline{d} & d\overline{s} \\ s\overline{u} & s\overline{d} & s\overline{s} \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\pi^0 + \frac{1}{\sqrt{6}}\eta^0 & \pi^+ & K^+ \\ \pi^- & -\frac{1}{2}\pi^0 + \frac{1}{\sqrt{6}}\eta^0 & K^0 \\ K^- & \overline{K}^0 & -\frac{2}{\sqrt{6}}\eta^0 \end{pmatrix} + \sqrt{2/3}\eta'\,\delta^i_j$$

Exercise 8.10. Show that

$$\pi^+ = \overline{q}.X_{-\alpha_1}.q, \quad \pi^0 = \overline{q}.I_3.q, \quad \pi^- = \overline{q}.X_{\alpha_1}.q, \quad \eta^0 = \frac{1}{\sqrt{6}}\overline{q}.Y.q, \text{ etc.}$$

Lets draw the representation in a basis of the eigenvalues of the two Cartan elements I_3 and Y.

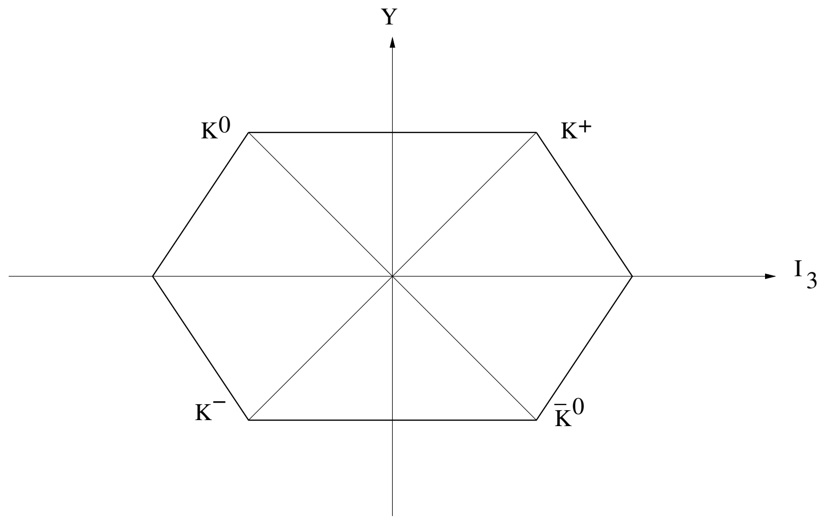


Figure 8.2: The Eightfold Way

We finish with a slightly longer

Exercise 8.11. Show that if one acts by the root vectors on q_j and $\overline{q}^{\,i}$ (in the correct representation) that

$$\begin{array}{rll} R_{\text{ad}}(X_{\alpha_2}): & \overline{q}.X_{\alpha_1}.q \to \overline{q}.X_{\alpha_3}.q & \text{or,} \quad X_{\alpha_2}: \pi^+ \to K^+ \\ R_{\text{ad}}(X_{-\alpha_1}): & \overline{q}.X_{\alpha_1}.q \to 2\overline{q}.I_3.q & \text{or,} \quad X_{-\alpha_1}: \pi^+ \to -\frac{1}{2}\pi^0 + \sqrt{\frac{1}{6}}\,\eta^0 \\ R_{\text{ad}}(X_{-\alpha_3}): & \overline{q}.X_{\alpha_1}.q \to -\overline{q}.X_{-\alpha_2}.q & \text{or,} \quad X_{-\alpha_3}: \pi^+ \to -\overline{K}^0 \end{array}$$

and so on to justify the arrows in Figure 8.2.

8.10 The Tensor Product in Components

There are two tensor products of interest namely $\mathbf{3} \otimes \mathbf{3}$ and $\mathbf{3} \otimes \overline{\mathbf{3}}$. Let e_i be a basis of the fundamental $\mathbf{3}$ and $\overline{e}^i$ be a basis of the complex conjugate fundamental $\overline{\mathbf{3}}$ which we write as before (but drop the reference to quarks)

$$\begin{aligned} &e_1 = |1/2, 1/3\rangle, \;\; e_2 = |-1/2, 1/3\rangle, \;\; e_3 = |0, -2/3\rangle \\ &\overline{e}^1 = |-1/2, -1/3\rangle, \;\; \overline{e}^2 = |1/2, -1/3\rangle, \;\; \overline{e}^3 = |0, 2/3\rangle \end{aligned} \tag{8.10.1}$$

as eigenstates of I_3 and Y. Now the the basis $e_i \otimes \bar{e}^j$ of $\mathbf{3} \otimes \overline{\mathbf{3}}$ is

$$\begin{array}{lll}
|1/2, 1/3\rangle \otimes |-1/2, -1/3\rangle & |1/2, 1/3\rangle \otimes |1/2, -1/3\rangle & |1/2, 1/3\rangle \otimes |0, 2/3\rangle \\
|-1/2, 1/3\rangle \otimes |-1/2, -1/3\rangle & |-1/2, 1/3\rangle \otimes |1/2, -1/3\rangle & |-1/2, 1/3\rangle \otimes |0, 2/3\rangle \\
|0, -2/3\rangle \otimes |-1/2, -1/3\rangle & |0, -2/3\rangle \otimes |1/2, -1/3\rangle & |0, -2/3\rangle \otimes |0, 2/3\rangle
\end{array}$$

Notice that the state

$$|1/2, 1/3\rangle \otimes |-1/2, -1/3\rangle + |-1/2, 1/3\rangle \otimes |1/2, -1/3\rangle + |0, -2/3\rangle \otimes |0, 2/3\rangle$$

is invariant. So once more we get the split of the $\mathbf{3} \otimes \overline{\mathbf{3}}$ as an $\mathbf{8} \oplus \mathbf{1}$.

The basis $e_i \otimes e_j$ for the $\mathbf{3} \otimes \mathbf{3}$ is

$$\begin{array}{lll}
|1/2, 1/3\rangle \otimes |1/2, 1/3\rangle & |1/2, 1/3\rangle \otimes |-1/2, 1/3\rangle & |1/2, 1/3\rangle \otimes |0, -2/3\rangle \\
|-1/2, 1/3\rangle \otimes |1/2, 1/3\rangle & |-1/2, 1/3\rangle \otimes |-1/2, 1/3\rangle & |-1/2, 1/3\rangle \otimes |0, -2/3\rangle \\
|0, -2/3\rangle \otimes |1/2, 1/3\rangle & |0, -2/3\rangle \otimes |-1/2, 1/3\rangle & |0, -2/3\rangle \otimes |0, -2/3\rangle
\end{array}$$

The anti-symmetric symmetric combinations are

$$\begin{array}{l}
\left(|-1/2, 1/3\rangle \otimes |0, -2/3\rangle - |0, -2/3\rangle \otimes |-1/2, 1/3\rangle\right)/2 \\
(|1/2, 1/3\rangle \otimes |0, -2/3\rangle - |0, -2/3\rangle \otimes |1/2, 1/3\rangle)/2 \\
(|1/2, 1/3\rangle \otimes |-1/2, 1/3\rangle - |-1/2, 1/3\rangle \otimes |1/2, 1/3\rangle)/2
\end{array} \tag{8.10.2}$$

Note that as far as the eigenvalues of I_3 and Y are concerned the states that we have here are

$$|-1/2, -1/3\rangle, \quad |1/2, 1/3\rangle \quad |0, -2/3\rangle \tag{8.10.3}$$

respectively, i.e. they for a $\overline{\mathbf{3}}$ as we have already seen (one needs to check that they transform in the right way too -as they do).

The symmetric combinations form the irreducible $\mathbf{6}$ dimensional representation. You should also plot the $\mathbf{6}$ on a graph with axis I_3 and Y as for the the other representations.

8.11 Highest Weights

To define the highest weight of a representation we need to make a choice. The reason for making a choice is that the weight vectors are labeled by two weights not just one as in the case of $su(2)$. In the case of $su(2)$ for the irreducible l dimensional representation we understood that the spin $j = l/2 - 1$ state was the highest weight state (with the largest possible value of j).

In the present situation we choose an ordering. We take the eigenstate with the largest eigenvalue of I_3 to be the highest weight state. So for the two 'fundamental representations' we take the highest weights to be the state $\boldsymbol{\mu}_1 = |1/2, 1/3 >$ in the $\mathbf{3}$, while for the $\overline{\mathbf{3}}$ it is the state $\boldsymbol{\mu}_2 = |1/2, -1/3 >$.

One can show that for **any** irreducible representation of $su(3)$ the highest weight vector is given by

$$n_1\boldsymbol{\mu}_1 + n_2\boldsymbol{\mu}_2$$

for integer n_1 and n_2. This is an abuse of notation since, for example, what we mean by $n_1\boldsymbol{\mu}_1$ is

$$\underbrace{\boldsymbol{\mu}_1 \otimes \cdots \otimes \boldsymbol{\mu}_1}_{n_1 \text{ times}} \tag{8.11.1}$$

the tensor product is multiplicative but the eigenvalues of the Cartan elements are additive, hence the 'mixed' notation. With this notation understood it is clear that in the tensor product

$$\underbrace{\mathbf{3} \otimes \cdots \otimes \mathbf{3}}_{n_1 \text{ times}} \otimes \underbrace{\overline{\mathbf{3}} \otimes \cdots \otimes \overline{\mathbf{3}}}_{n_2 \text{ times}} \tag{8.11.2}$$

the vector

$$n_1\boldsymbol{\mu}_1 + n_2\boldsymbol{\mu}_2 \equiv \underbrace{\boldsymbol{\mu}_1 \otimes \cdots \otimes \boldsymbol{\mu}_1}_{n_1 \text{ times}} \otimes \underbrace{\boldsymbol{\mu}_2 \otimes \cdots \otimes \boldsymbol{\mu}_2}_{n_2 \text{ times}} \tag{8.11.3}$$

is certainly the one with the largest eigenvalue value of I_3.

For example the **8** has highest weight $|1, 0>$ (the π^+ particle) and we can write that as

$$|1, 0 >= \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2 \equiv |1/2, 1/3 > \otimes |1/2, -1/3 >,$$

i.e. $n_1 = n_2 = 1$. The weights $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ are called **fundamental weights** since one can obtain any highest weight from these two and as they are the heighest weights of both fundamental representations. By starting from the highest weight one can obtain all the other vectors in the irreducible representation by acting with the root vectors.

Exercise 8.12. Which irreducible representation of $su(3)$ has highest weight $2\boldsymbol{\mu}_1$?

CHAPTER 9

LIE ALGEBRAS, SUB-ALGEBRAS AND IDEALS

From our study of the Lie algebras of the groups $U(1)$, $SU(2)$ and $SU(3)$ we can see that the following abstraction is the correct definition of a Lie Algebra.

Definition 9.1. A Lie Algebra is a vector space V together with a Commutator $[\,,\,]$ satisfying the following conditions

1. $\forall X, Y \in V$,
$$[X, Y] \in V$$
2. $\forall X, Y \in V$, the bracket is anti-symmetric,
$$[X, Y] = -[Y, X]$$
3. The Jacobi identity is satisfied, that is, $\forall X, Y, Z \in V$
$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$$
4. The bracket is linear, $\forall X, Y, Z \in V$
$$[aX + Yb, Z] = a\,[X, Z] + b\,[Y, Z]$$

We also have that

Definition 9.2. A Subalgebra W of a Lie algebra V is a subvector space of V which is closed under the Lie bracket. We write $W \subseteq V$.

This means that if X, Y are in W (not just in V) then W is a sub Lie algebra if
$$[X, Y] \in W.$$

We already have examples of this. The Cartan subalgebras of $SU(2)$ and $SU(3)$. The $SU(2)$ subalgebra of $SU(3)$.

There is another useful notion having to do with very special sub algebras,

Definition 9.3. An Ideal I of a Lie algebra V is a subalgebra of V such that $\forall X \in I$ and $\forall Y \in V$

$$[X, Y] \in I.$$

A simple example here is the Lie algebra of the group $U(1) \times SU(2)$. The two groups act independently of each other, so they commute. Likewise then the Lie algebras commute. But this means that for the $U(1)$ element h we have that $[h, X] = 0 \in u(1)$ for all $X \in su(2)$. So $u(1)$ is an ideal of $u(1) \oplus su(2)$. However, we could have written $[h, X] = 0 \in su(2)$ so that $su(2)$ is also an ideal of $u(1) \oplus su(2)$.

Though we will not really make direct use of Ideals, the notion of an Ideal plays a role in the concepts of simple and semi-simple Lie algebras which will be a recurrent theme later in these notes.

CHAPTER 10

MATRIX GROUPS AND THEIR LIE ALGEBRAS

As the title of this chapter suggests we look at those groups which can be defined in terms of matrices. This means that they already come together with a given representation, namely the fundamental representation. Since we are dealing with finite dimensional matrices the group multiplication rule will always be matrix multiplication. This means that associativity is automatic and we always have an identity element, namely the identity matrix. For all of the groups listed below you should check closure and the existence of an inverse element in the group. The matrix groups are related to certain geometric properties that are put onto the vector spaces on which they act and partially explains the ineterest in these groups. The geometric properties are also described as we go along.

The Lie algebras of these matrix groups will automatically have brackets that satisfy the anti-symmetry property, the Jacobi identity and linearity (as we are dealing with finite dimensional matrices). However, you should, for all the Lie algebras of the groups listed below, check that the commutator of two Lie algebra elements is a Lie algebra element.

One word about the (inconsistent) notation that I use. Very often we have different groups depending on whether the entries of the matrices are taken to be real or complex. If a statement is independent of the choice I will usually just denote the group by its defining letters and dimension e.g. $G(n)$, however, if the distinction is important I will exhibit reality properties of the entries of the matrices e.g. $G(n,\mathbb{R})$ or $G(n,\mathbb{C})$.

10.1 The General Linear Group Gl$(n,*)$

Definition 10.1. The General Linear Group $\mathrm{Gl}(n,\mathbb{R})$ $(\mathrm{Gl}(n,\mathbb{C}))$, is the group of all $n\times n$ matrices with real (complex) entries and whose determinant is non-zero.

Note that if $g \in \mathrm{Gl}(n)$ then since $\mathrm{Det}\, g \neq 0$ we have that g^{-1} exists.

Since there are no conditions on the matrices in $\mathrm{Gl}(n)$ other than that they be invertible, there are essentially no conditions on the components. Hence the dimension of $\mathrm{Gl}(n, \mathbb{R})$ is n^2. We write the real dimension as $\dim_{\mathbb{R}} \mathrm{Gl}(n, \mathbb{R}) = n^2$ and $\dim_{\mathbb{R}} \mathrm{Gl}(n, \mathbb{C}) = 2n^2$ while the complex dimension is $\dim_{\mathbb{C}} \mathrm{Gl}(n, \mathbb{C}) = n^2$.

The infinitesimal generators of $\mathrm{Gl}(n)$ are obtained in the usual way. We expand a group element about the identity to first order, so we write for $g \in \mathrm{Gl}(n)$

$$g = \mathbb{I}_n + \epsilon M + \dots \quad .$$

The requirement that $\mathrm{Det}\, g \neq 0$ translates into

$$\mathrm{Det}\, g = 1 + \epsilon \operatorname{Tr} M + \cdots \neq 0$$

which is satisfied reagardless of M for infinitesimal ϵ. Consequently, M can be any $n \times n$ matrix. The space of $n\times$ matrices is again n^2 dimensional (real or complex depending on whether we are interested in $\mathrm{Gl}(n, \mathbb{R})$ or $\mathrm{Gl}(n, \mathbb{C})$).

The Lie algebra of $\mathrm{Gl}(n)$ is denoted by $\mathfrak{gl}(n)$ and is a vector space with dimension n^2. Indeed $\dim_{\mathbb{R}} \mathfrak{gl}(n, \mathbb{R}) = n^2$ and $\dim_{\mathbb{C}} \mathfrak{gl}(n, \mathbb{C}) = n^2$.

It is convenient to introduce the following matrices,

$$E_{ij} \tag{10.1.1}$$

with $i, j = 1, \dots, n$, so that there are n^2 matrices. The only non-zero entry of the matrix E_{ij} is its (ij)'th component (i'th row and j'th column) which is taken to be 1. In formulae this means that

$$(E_{ij})_k^{\ l} = \delta_{ik}\, \delta_{jl}.$$

These n^2 matrices form a basis for $\mathfrak{gl}(n)$. Consequently any $X \in \mathfrak{gl}(n)$ can be written as

$$X = \sum_{i\,j} X^{ij}\, E_{ij} \tag{10.1.2}$$

and $X \in \mathfrak{gl}(n, \mathbb{C})$ if all the $X^{ij} \in \mathbb{C}$ or $X \in \mathfrak{gl}(n, \mathbb{R})$ if all the $X^{ij} \in \mathbb{R}$.

Exercise 10.1. Show that the matrices E_{ij} satisfy

$$E_{ij}\, E_{kl} = E_{il}\, \delta_{jk},$$

and consequently the commutation relations,

$$[E_{ij}, E_{kl}] = E_{il}\, \delta_{jk} - E_{kj}\, \delta_{il}.$$

The general linear group, $\mathrm{Gl}(n)$ has the following property. As matrices, group elements act on an n-dimensional vector space W. If W is spanned by n basis vectors, and $g \in \mathrm{Gl}(n)$, then acting with g on each of the basis vectors gives n-new vectors which also span W.

10.2 The Special Linear Group Sl$(n,*)$

Definition 10.2. The Special Linear Group $\mathrm{Sl}(n,\mathbb{R})$ ($\mathrm{Sl}(n,\mathbb{C})$), is the group of all $n\times n$ matrices with real (complex) entries and whose determinant is one.

The dimension of this group is n^2-1. The reason for this is that the determinant condition is one condition on the $n\times n$ matrix. Explicitly $\dim_{\mathbb{R}}(\mathrm{Sl}(n,\mathbb{R}))=n^2-1$ and $\dim_{\mathbb{C}}(\mathrm{Sl}(n,\mathbb{C}))=n^2-1$.

Obviously any $g\in\mathrm{Sl}(n)$ is also in $\mathrm{Gl}(n)$. So acting with M on basis vectors of W gives again a new basis of W. But not only is this true. If you think of W as being $\mathbb{R}^n$ ($\mathbb{C}^n$) and you think of integrating on $\mathbb{R}^n$ ($\mathbb{C}^n$) there is a natural (holomorphic) volume element that we use to integrate with, namely

$$d^nx=dx^1\,dx^2\,\ldots\,dx^n,\quad (d^nz=dz^1\,dz^2\,\ldots\,dz^n)$$

If $g\in\mathrm{Sl}(n)$ then g also preserves this volume form. This means that if we start with coordinates x^i and go to new coordinates y^i by

$$y^i=g^i{}_j\,x^j$$

then

$$d^ny=\mathrm{Det}\,g.\,d^nx=d^nx.$$

(and a similar argument in the complex case).

Lets move on to its Lie algebra. Again near the identity the generators of the group are, for $g\in\mathrm{Sl}(n)$,

$$g=\mathbb{I}_n+\epsilon\,M+\ldots$$

but the fact that $\mathrm{Det}\,g=1$ implies that $\mathrm{Tr}\,M=0$. Hence the Lie algebra $\mathfrak{sl}(n)$ is made up of all traceless $n\times n$ matrices. A reasonable basis for this Lie algebra is made up of the matrices, E_{ij} with $i\neq j$ and say, $E_{ii}-E_{i+1\,i+1}$ for $i=1,\ldots,n-1$. These matrices are all traceless and span a n^2-1 dimensional space,

$$X=\sum_{i=1}^{n-1}X^i\,(E_{ii}-E_{i+1\,i+1})+\sum_{i\neq j}X^{ij}\,E_{ij} \tag{10.2.1}$$

if the entries are real X^i, $X^{ij}\in\mathbb{R}$ then we have $\mathfrak{sl}(n,\mathbb{R})$ while if they are complex X^i, $X^{ij}\in\mathbb{C}$ then we have $\mathfrak{sl}(n,\mathbb{C})$.

In Cartan's classification the Lie algebras $sl(n+1)$ are called the A_n series.

10.3 The Orthogonal Group $O(n,*)$

We have already seen an example of groups of this type, namely $O(2)$. In general we have the

Definition 10.3. The Orthogonal Group $O(n, \mathbb{R})$ ($O(n, \mathbb{C})$), is the group of all orthogonal $n \times n$ matrices with real (complex) entries.

Orthogonality means that for $g \in O(n)$

$$g^{\mathrm{T}}.g = \mathbb{I}_n.$$

The orthogonality condition is invariant under transposition so it really represents n conditions on the diagonal and $n(n-1)/2$ on the upper triangular part. So altogether we have n^2 variables and $n(n+1)/2$ condition leaving us with $n(n-1)/2$ parameters. Consequently, the dimension of $O(n)$ is $n(n-1)/2$.

Orthogonal matrices are clearly also $\mathrm{Gl}(n)$ matrices so that they share their geometric properties. However, $O(n, \mathbb{R})$ matrices also preserve the length of vectors in $\mathbb{R}^n$ where, for $X \in W$ the length^2 is

$$< X, X > \equiv X^{\mathrm{T}}.X$$

and this is invariant under $X \to g.X$ for $g \in O(n, \mathbb{R})$. Infact the inner product of any 2 vectors in W

$$< X, Y > \equiv X^{\mathrm{T}}.Y$$

is invariant providing both vectors are transformed in the same way, i.e. $X \to g.X$ and $Y \to g.Y$.

Now we turn to the Lie algebra $o(n)$. Writing

$$g = \mathbb{I}_n + \epsilon M + \dots$$

we see that the orthogonality condition implies that

$$M^{\mathrm{T}} = -M$$

that is M is anti-symmetric. The number of components that an $n \times n$ anti-symmetric matrix has is $n(n-1)/2$, and this is the dimension of the Lie algebra.

The choice of basis depends on wether n is even or not. We will make a choice later.

10.4 The Special Orthogonal Group $SO(n, *)$

Our prime example of this type of group have been $SO(2)$ and $SO(3)$.

Definition 10.4. The Special Orthogonal Group $SO(n, \mathbb{R})$ ($SO(n, \mathbb{C})$), is the group of all orthogonal $n \times n$ matrices with real (complex) entries whose determinant is one.

If $g \in O(n)$, then we automatically have that

$$\mathrm{Det}\, g = \pm 1.$$

So the group elements which are in the special orthogonal group are all those for which the sign is positive. The dimension of $SO(n)$ is the same as that of $O(n)$. Also the Lie algebra $\mathfrak{so}(n)$ is the same as that $\omega(n)$ since to work out the Lie algebra we are looking close to the identity, so our elements $g \in O(n)$ must have $\mathrm{Det}\, g = 1$.

As a group $SO(n)$, unlike $O(n)$, preserves the volume.

Here is a hint on how to show that the bracket of two elements of $\mathfrak{so}(n)$ are again in $\mathfrak{so}(n)$. $M \in \mathfrak{so}(n)$ iff $M^{\mathrm{T}} = -M$. Let $M_1,\ M_2 \in \mathfrak{so}(n)$, then set $M = [M_1, M_2]$, the hint is: calculate M^{T}.

In Cartans classification the special orthogonal Lie algebras are denoted $\mathfrak{so}(2n) \equiv D_n$ and $\mathfrak{so}(2n+1) \equiv B_n$.

10.5 The Unitary group $U(n)$

We have already seen $U(1)$.

Definition 10.5. The Unitary Group $U(n)$ is the group of all unitary $n \times n$ matrices with complex entries.

The unitarity condition is

$$U^{\dagger}.U = \mathbb{I}_n.$$

Again this equation is invariant under hermitian conjugation and so it represents n real equations on the diagonal, and $n(n-1)/2$ complex equations on the upper triangular part. All in all we have n^2 real equations and $2n^2$ real components, leaving us with n^2 real parameters. Thus the dimension of $U(n)$ is n^2 $\dim_{\mathbb{R}}(U(n)) = n^2$.

Since the matrices are complex they act on complex vector spaces. So, in this case, the vector space W is $\mathbb{C}^n$. (You should think of vectors in this space as being n-component column vectors with complex coefficients.) On a complex vector space there is also a notion of length of a vector, namely for $Z \in W$

$$< Z, Z > \equiv \overline{Z}^{\mathrm{T}}.Z$$

and this is preserved by $Z \to g.Z$ for any $g \in U(n)$.

Write a group element near the identity as

$$g = \mathbb{I}_n + i\epsilon M + \dots\ ,$$

For the Lie algebra $\mathfrak{u}(n)$, the unitarity condition implies that the generators satisfy the equation

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

so they are Hermitian. The dimension of the space of Hermitian $n \times n$ matrices is again n^2, agreeing with the dimension of the group.

10.6 The Special Unitary Group $SU(n)$

The examples that we have seen of these are $SU(2)$ and $SU(3)$. These two groups form a corner stone of our understanding of the fundamental interactions. Other groups of this type are also important.

Definition 10.6. The Special Unitary Group $SU(n)$ is the group of all unitary $n \times n$ matrices with complex entries whose determinant is one.

If $g \in U(n)$ then we know that

$$\operatorname{Det} g = e^{i\phi}$$

for some real ϕ. For $g \in U(n)$ to be in $SU(n)$ we require $\phi = 0$, which is one real equation. Consequently the dimension of $SU(n)$ is the dimension of $U(n)$ minus 1, that is, $\dim_{\mathbb{R}}(SU(n)) = n^2 - 1$.

Not only does $SU(n)$ preserve the length of vectors but it also preserves a volume form on W. The coordinates of $\mathbb{C}^n$ can be taken to be z^i for $i = 1, \ldots, n$. The holomorphic Volume Form is

$$d^n z = dz^1 \ldots dz^n.$$

For $g \in SU(n)$ if we change variables to $w = g.z$ we find that

$$d^n w = \operatorname{Det} g \, d^n z = d^n z.$$

The Lie algebra elements $M \in \mathfrak{su}(n)$ satisfy the Hermiticity condition of the $\mathfrak{u}(n)$ generators as well as having to be traceless, which comes from the fact that the determinant of the group elements are one.

The Lie algebra $\mathfrak{su}(n+1)$ is also in the A_n series in Cartan's classification system.

10.7 The Symplectic Groups $Sp(n, *)$ and $\mathrm{Sp}(n)$

Definition 10.7. The Non-Compact Symplectic Group $Sp(n, \mathbb{R})$ ($Sp(n, \mathbb{C})$) is the group of $2n \times 2n$ matrices, g, with real (complex) entries which satisfy

$$g^{\mathrm{T}}.E.\,g = E$$

where E is the $2n \times 2n$ matrix

$$\begin{pmatrix} 0_n & \mathbb{I}_n \\ -\mathbb{I}_n & 0_n \end{pmatrix}.$$

Notice that the defining condition changes sign under transposition. This means the matrix of equations is anti-symmetric so there are only $2n(2n-1)/2$ independent equations. Consequently the dimension of the group is $4n^2 - 2n(2n-1)/2 = n(2n+1)$.

At first sight this looks like a rather bizarre group, however you probably know it quite well. $Sp(n,\mathbb{R})$ appears in the canonical approach to classical mechanics. Let (q_i, p_i) be our phase space with $i = 1, \ldots, n$. Consider any functions F and G on the phase space and take their Poisson bracket

$$\{F, G\} = \sum_i \left(\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i} \right) = \hat{F}^{\mathrm{T}}.E.\hat{G}$$

where

$$\hat{F} = \left(\frac{\partial F}{\partial q_1} \cdots \frac{\partial F}{\partial q_n} \,|\, \frac{\partial F}{\partial p_1} \cdots \frac{\partial F}{\partial p_n} \right)^T, \quad \hat{G} = \left(\frac{\partial G}{\partial q_1} \cdots \frac{\partial G}{\partial q_n} \,|\, \frac{\partial G}{\partial p_1} \cdots \frac{\partial G}{\partial p_n} \right)^T,$$

So the group of Symplectomorphisms preserves the Poisson Bracket. This means that we send

$$\hat{F} \to g.\hat{F}, \quad \hat{G} \to g.\hat{G}$$

and the Poisson bracket between F and G does not change. In turn this comes from the following; let

$$D = \left(\frac{\partial}{\partial q_1} \cdots \frac{\partial}{\partial q_n} \,|\, \frac{\partial}{\partial p_1} \cdots \frac{\partial}{\partial p_n} \right)^T$$

then $\hat{F} = D.F$ and $\hat{G} = D.G$. The matrix g acts on D, which is equivalent to having acted with g^{-1} on the coordinates of the phase space, i.e.,

$$(q_1 \ldots q_n \,|\, p_1 \ldots p_n)^T \longrightarrow g^{-1}.\ (q_1 \ldots q_n \,|\, p_1 \ldots p_n)^T$$

Let , as usual,

$$g = \mathbb{I}_{2n} + \epsilon M + \ldots,$$

The symplectic condition tells us that

$$M^{\mathrm{T}}.E + E.M = 0,$$

We can re-write this expression as follows

$$(E.M)^{\mathrm{T}} = E.M$$

which means that we can pass to generators

$$M' = E.M$$

and these generators are symmetric matrices.

In fact the Jacobi Identity was invented by Jacobi precisely for the Poisson bracket, so in a sense all Lie algebras stem from this one.

Finally this Lie algebra is equivalently called $C_n \equiv \mathfrak{sp}(n,\,\mathbb{C})$ by Cartan.

10.7.1 The Compact Symplectic Group

The definition of a symplectic group that we have used includes the fact that the group is non compact (we will see a concrete example of this below). The reason we specify this is that there is also a Compact Symplectic Group which has a slightly different definition. That definition requires the concept of the quaternions $\mathbb{H}$ but I would like to avoid that here. Fortunately, there is another way to define the compact symplectic group $\mathrm{Sp}(n)$ (note here we do not specify if it is over $\mathbb{R}$ or $\mathbb{C}$), namely

$$\mathrm{Sp}(n) = U(2n) \cap Sp(n, \mathbb{C}) = SU(2n) \cap Sp(n, \mathbb{C}) \tag{10.7.1}$$

that is the compact symplectic group $\mathrm{Sp}(n)$ is the group of $2n \times 2n$ matrices that are simultaneously unitary and symplectic.

Note that $\dim_{\mathbb{R}}(\mathrm{Sp}(n)) = 2n^2 + n$ as the unitarity (or reality) condition essentially halves the dimension of $Sp(n, \mathbb{C})$.

10.8 Some Accidental Correspondences

There are some striking similarities amongst these groups. Firstly the dimensions of $SU(2)$, $SO(3)$ and $Sp(1, \mathbb{R})$ are the same. We know that $SO(3)$ and $SU(2)$ are very closely related but are not the same. Indeed $Sp(1, \mathbb{R})$ is quite different again in that it is a non-compact group. To see this we use the defining equation. Let

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in Sp(1, \mathbb{R})$$

then

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

which can be satisfied iff

$$ad - bc = 1. \tag{10.8.1}$$

Notice that a, b, c, and d can be as large as we like. This is not the case for $SU(2)$ group elements for which the matrix elements are bounded. In the case of $Sp(1, \mathbb{R})$ since the entries are not bounded we see that the group is non-compact in fact since the condition (10.8.1) simply says the determiant of the matrix is unity we see that $Sp(1, \mathbb{R}) = Sl(2, \mathbb{R})$ and likewise $Sp(1, \mathbb{C}) = Sl(2, \mathbb{C})$. We now deduce that the compact symplectic group

$$\mathrm{Sp}(1) = SU(2) \cap Sl(2, \mathbb{C}) = SU(2) \tag{10.8.2}$$

Notice also that the dimensions of $Sp(2, \mathbb{R})$, $\mathrm{Sp}(2)$ and $SO(5)$ match. There are a number of such matchings of dimensions which do imply relationships between groups (most do not). Infact we have that:

- $SU(2)$ is a double cover of $SO(3,\mathbb{R})$.
- $SU(2) \times SU(2)$ is a double cover of $SO(4,\mathbb{R})$.
- $SL(4,\mathbb{C})$ is a double cover of $SO(6,\mathbb{C})$
- $\mathrm{Sp}(2)$ is a double cover of $SO(5,\mathbb{R})$.
- $SU(4)$ is a double cover of $SO(6,\mathbb{R})$

CHAPTER 11

THE CARTAN-KILLING FORM AND ROOT SPACES

We are now going to begin a rather more systematic study of Lie algebras. We begin by introducing a symmetric and bilinear form on the Lie algebra. This form is degenrate in general, however, our main interest will be in situations where it is not degenerate.

11.1 Cartan-Killing Form

In this section we will define an invariant form or metric on the Lie algebra V.

Consider $X, Y \in V$ then

Definition 11.1. The Cartan-Killing form is defined to be the symmetric and bilinear form

$$B(X,Y) = \mathrm{Tr}\,(\mathrm{ad}_X\ \mathrm{ad}_Y)$$

The way you should think about this is as follows. Since

$$\mathrm{ad}_Z : V \to V, \quad \forall Z \in V$$

is a linear map one can think of it as a matrix. The trace in the definition is a trace over matrices which act on V.

Let T_i be a basis for V and define the structure constants by

$$[T_i, T_j] = f_{ij}^k\, T_k.$$

Let $X = X^i T_i$ and $Y = Y^i T_i$. Then

$$\mathrm{ad}_X\, T_i = [X, T_i] = X^j\, [T_j, T_i] = -f_{ij}^k X^j\, T_k$$

which means that as a matrix ad_X has components

$$(\mathrm{ad}_X)^k{}_i = f_{ij}^k\, X^j.$$

Now we can calculate that

$$\mathrm{ad}_X\,\mathrm{ad}_Y\,T_i = f_{ij}^k\,Y^j\;\mathrm{ad}_X\;T_k = f_{ij}^k f_{kl}^m\,Y^j\,X^l\,T_m$$

so that the matrix which represents $\mathrm{ad}_X\,\mathrm{ad}_Y$ is

$$(\mathrm{ad}_X\,\mathrm{ad}_Y)_i^m = f_{ij}^k f_{kl}^m\,Y^j\,X^l.$$

Consequently, we that,

$$B(X,Y) = f_{ij}^k\,f_{kl}^i\,Y^j\,X^l.$$

You should check that this formula is symmetric in X and Y. We can write this as

$$B(X,Y) = B_{ij}\,X^i\,Y^j$$

where

$$B_{ij} = f_{ik}^l\,f_{jl}^k = \mathrm{Tr}\left(\mathrm{ad}_{T_i}\;\mathrm{ad}_{T_j}\right).$$

Here is a simple example. Let $V = U(1)$ then, as the structure constants are all zero, $B = 0$. Consider now $V = su(2)$ and set $T_i = \frac{i}{2}\sigma_i$, then $f_{ij}^k = \epsilon_{ijk}$ and

$$B_{ij} = \sum_{k,l}\epsilon_{ikl}\epsilon_{jlk} = -2\delta_{ij}.$$

Exercise 11.1. Use the commutation relations

$$[h, X_{\alpha_i}] = \alpha_i(h)\,X_{\alpha_i}$$

for $h \in H$ and X_{α_i} the root vectors of a Lie algebra to deduce that

$$B(h_1,h_2) = 2\sum_i \alpha_i(h_1)\,\alpha_i(h_2).$$

Exercise 11.2. Use the commutation relations of $SU(3)$ (7.6.2) to determine $B(I_3,I_3)$, $B(Y,Y)$ and $B(I_3,Y)$ directly and compare this with the results of the previous exercise.

In the case of the Abelian theory the 'metric' is zero and one says it is degenerate. More generally we will say a metric is degenerate if $\det g_{ij} = 0$ so that the inverse matrix is not defined. In the $SU(2)$ case it is quite a respectable metric. These properties of the Killing-Cartan form play an essential role in the structure theory of Lie algebras. More on this below.

One very important feature of the Cartan-Killing form is that it is invariant under automorphisms of the Lie algebra.

Definition 11.2. An automorphism of a Lie algebra V is a map, ϕ, from V to itself which preserves the bracket,

$$\phi\left([X,Y]\right) = [\phi(X),\phi(Y)], \quad \forall X,Y \in V.$$

We also have

$$\mathrm{ad}_{\phi(X)}(Z) = [\phi(X), Z] = [\phi(X), \phi(.\phi^{-1}(Z))] = \phi([X, \phi^{-1}(Z)]) = \phi(\mathrm{ad}_X(\phi^{-1}(Z))).$$

Now notice that we can get rid of some brackets and write this as

$$\mathrm{ad}_{\phi(X)}(Z) = \phi \circ \mathrm{ad}_X \circ \phi^{-1}(Z).$$

Since ϕ is an automorphism we can think of it as a matrix, ad_X can be thought of as a matrix and so can ϕ^{-1}.

Exercise 11.3. Show that

$$B(\phi(X), \phi(Y)) = B(X, Y).$$

Note: The Cartan-Killing form does not depend on the basis of the algebra that one uses, however, g_{ij} certainly does. For example we saw that in the anti-hermitian basis the Cartan-Killing form for $su(2)$ is

$$B_{ij} = -2\delta_{ij}$$

and then, for example for $h = \sigma_3/2 = -iT_3 = h^i\, T_i$, so in this basis $h^i = -i\delta^{3i}$, we had

$$B(h, h) = -2\, \delta_{ij} h^i\, h^j = 2.$$

If we used the Hermitian basis $T_i = \sigma_i/2$, then the structure constants are $i\epsilon_{ijk}$ and we get

$$B_{ij} = 2\, \delta_{ij}.$$

But in the Hermitian basis we have that $h = \sigma_3/2 = T_3 = \hat{h}^i T_i$ with $\hat{h}^i = \delta^{i3}$ and once again

$$B(h, h) = 2\, \delta_{ij}\, \hat{h}^i\, \hat{h}^j = 2.$$

11.2 The Root Space

There is an honest scalar product defined on the dual vector space of the Cartan subalgebra of a semi-simple Lie algebra. Remember the Cartan subalgebra is also a vector space and taking the dual means having a vector space of 'bra' vectors. The formal definition of semi-simple is given later, but for now we note that it is a Lie algebra for which the Cartan-Killing form is non-degenerate. If we denote the Cartan subalgebra by $\mathfrak{h}$ then its dual is denoted by $\mathfrak{h}^*$. The roots live in $\mathfrak{h}^*$ as we will see.

Let $\alpha \in \mathfrak{h}^*$ then there exists a unique element in $\mathfrak{h}$ which is denoted by h_α such that for every $k \in \mathfrak{h}$

$$\langle \alpha,\, k\rangle = \alpha(k) \simeq B(h_\alpha, k). \tag{11.2.1}$$

(this only makes sense if B is non-degenerate). There is a constant which makes the proportionality $\simeq$ in (11.2.1) an equality. That constant $c(\mathfrak{g})$ depends on the Lie algebra and is chosen for our conveniance- which we will do later.

Definition 11.3. For α a root, that is it is in the space of all roots Δ of $\mathfrak{g}$, the co-root $h_\alpha \in \mathfrak{h}$ is defined to be that element such that $\forall k \in \mathfrak{h}$

$$\alpha(k) = c(\mathfrak{g})\, B(h_\alpha,\, k)$$

What convenience means is that we can just set $c(\mathfrak{g}) = 1$ when we need to, or take it to some other value. In the end it will be ratios of the Cartan-Killing form that will be important for us and the constant drops out.

Once one has the elements h_α there is a natural inner product $\langle\, ,\, \rangle$ on the root space itself which is given by for $\alpha, \beta \in \mathfrak{h}^*$,

$$\langle \alpha, \beta \rangle = c(\mathfrak{g})\, B(h_\alpha, h_\beta).$$

For the $su(n)$ groups one usually takes $c(su(n)) = (2n)^{-1}$.

We see how this works for $su(2)$. There is only one root which we called α and only one Cartan element h. We take

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad X_\alpha = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad X_{-\alpha} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

and we have that

$$\alpha(h) = 2$$

As $\mathrm{ad}_k\, X_{\pm\alpha} = \alpha(k)\, X_{\pm\alpha}$ for $k \in \mathfrak{h}$ we deduce that for any $k_1,\, k_2 \in \mathfrak{h}$

$$B(k_1,\, k_2) = 2\, \alpha(k_1)\, \alpha(k_2) = \sum_{\alpha \in \Delta} \alpha(k_1)\, \alpha(k_2) \tag{11.2.2}$$

(in the sum $\Delta = \{\pm\alpha\}$), so that $B(h, k) = 4\alpha(k)$. But if you look at Definition 11.3 for the co-root you see that in the case of $su(2)$ the equation is $4\alpha(k) = B(h_\alpha,\, k)$, so we deduce that $h_\alpha = h$.

For $su(3)$ the roots are α_1 and α_2 and α_3 $(= \alpha_1 + \alpha_2)$, while there are two Cartan elements h_1 and h_2 and $c(su(3)) = 1/6$. We determined (see part I Exercise 7.3) that

$$\alpha_1(ah_1 + bh_2) = 2a - b, \quad \alpha_2(ah_1 + bh_2) = 2b - a.$$

Hence we are looking for h_{α_1} and h_{α_2} so that

$$B(h_{\alpha_1}, ah_1 + bh_2) = 6(2a - b), \quad B(h_{\alpha_2}, ah_1 + bh_2) = 6(2b - a). \tag{11.2.3}$$

The Cartan elements that we are looking for can be expanded in the two that we have ($\mathfrak{h}$ is 2-dimensional in this case), so set

$$h_{\alpha_1} = mh_1 + nh_2.$$

By a previous exercise 11.1 one has that

$$\begin{aligned} B(h_{\alpha_1}, h_j) &= mB(h_1, h_j) + nB(h_2, h_j) \\ &= 2m\sum_{n=1}^{3} \alpha_n(h_1)\,\alpha_n(h_j) + 2n\sum_{n=1}^{3} \alpha_n(h_2)\,\alpha_n(h_j). \end{aligned} \tag{11.2.4}$$

From (11.2.3) we get

$$B(h_{\alpha_1}, h_1) = 12, \quad B(h_{\alpha_1}, h_2) = -6$$

while from (11.2.4) we have

$$B(h_{\alpha_1}, h_1) = 12m - 6n, \quad B(h_{\alpha_1}, h_2) = -6m + 12n$$

solving we find that, $m = 1$, $n = 0$

$$h_{\alpha_1} = h_1.$$

Exercise 11.4. Prove that $h_{\alpha_2} = h_2$.

Now we have all the pieces that we need to show that

$$\begin{array}{lll} \langle\alpha_1, \alpha_1\rangle = 2 & \langle\alpha_2, \alpha_2\rangle = 2 & \langle\alpha_3, \alpha_3\rangle = 2 \\ \langle\alpha_1, \alpha_2\rangle = -1 & \langle\alpha_1, \alpha_3\rangle = 1 & \langle\alpha_2, \alpha_3\rangle = 1 \end{array} \tag{11.2.5}$$

These equations have a geometric interpretation. Think of all of the α_i as having length $\sqrt{2}$ as dictated in the first line of (11.2.5). Then since $\langle a, b\rangle = |a|\,|b|\cos\theta$ for the inner product of two vectors we see from the second line that $\cos\theta = \pm 1/2$.

CHAPTER 12

SIMPLE AND SEMI-SIMPLE LIE ALGEBRAS

12.1 Non-Degenerate Cartan-Killing Form

Definition 12.1. A Lie algebra with no non-trivial ideals is said to be simple.

Definition 12.2. A Lie algebra is said to be semi-simple if the Cartan-Killing form B is non-degenerate, ie. $\det B \neq 0$.

This is Cartan's criterion. One usually states a different definition and then establishes this as a theorem. Instead we reverse the roles and we have,

Theorem 12.1. A Lie algebra is semi-simple if and only if it has no abelian ideals.

The Lie algebras of $SU(2)$ and $SU(3)$ are, therefore, simple. The Lie algebra $u(2) \oplus u(1)$ is not simple since it has ideals and is not semi-simple since it has an Abelian ideal.

Theorem 12.2. A semi-simple Lie algebra is the direct sum of simple Lie algebras.

We will not try to prove these statements in this course. But we will use them in the sense that we will only first concentrate on semi-simple Lie algebras, putting the non-degeneracy of the Cartan-Killing form to good use. And then we will concentrate on simple Lie algebras since, by the theorems, we can deduce most of the properties of simple Lie algebras from semi-simple Lie algebras.

12.2 Complex Versus Real Lie Algebras

Throughout the following we will be using roots and root vectors liberally. So it is important to recall what we have seen in the explicit cases of $SU(2)$ and $SU(3)$ namely that the root vectors are not in the real Lie algebra. Recall that for $SU(2)$ we have one Cartan element h and one positive root which we can call α and the root vectors $X_{\pm\alpha}$

$$h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad X_\alpha = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad X_{-\alpha} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{12.2.1}$$

In terms of these any element of $\mathfrak{su}(2)$ is expressable as

$$a\,h + bX_\alpha + cX_{-\alpha} \tag{12.2.2}$$

with $a \in \mathbb{R}$ and b, $c \in \mathbb{C}$ where $c = b^*$. In this way there are 3 real parameters and we have that, as a vector space $\dim_{\mathbb{R}} \mathfrak{su}(2) = 3$. But you should note that if we are going to use root vectors then we need to allow complex co-efficients (b and c in this case) otherwise we will not cover the complete vector space of $\mathfrak{su}(2)$.

What happens if we allow for a, b and c to be arbitrary complex numbers (so no relationship between b and c)? What type of Lie algebra will we have? The Lie algebra of $\mathfrak{sl}(2, \mathbb{C})$, as we saw in section 10.2 is generated by the matrices $E_{11} - E_{22}$, E_{12} and E_{21} with complex coefficients. But $h = E_{11} - E_{22}$, $X_\alpha = E_{12}$ and $X_{-\alpha} = E_{21}$ so by allowing complex coefficients in (12.2.2) we obtain the Lie algebra of $\mathfrak{sl}(2, \mathbb{C})$. In this case our real Lie algebra is $\mathfrak{su}(2)$ and its complexification is $\mathfrak{sl}(2, \mathbb{C})$. Conversely one can obtain $\mathfrak{su}(2)$ from $\mathfrak{sl}(2, \mathbb{C})$ by imposing the 'reality' conditions, $a = a^*$ and $b = c^*$.

We could impose a slightly different reality condition on the coefficients in $\mathfrak{sl}(2, \mathbb{C})$, namely that a, b, $c \in \mathbb{R}$ in which case we end up with the Lie algebra of $\mathfrak{sl}(2, \mathbb{R})$ as a different real form.

These examples teach us that it is convenient to start with a complex Lie algebra and then to impose reality conditions to arrive at the various real forms that are of interest. This is what we will do. Strictly speaking we should distinguish between complex and real Lie algebras by writing something like $\mathfrak{g}_{\mathbb{C}}$ for the complexification of the real Lie algebra $\mathfrak{g}$ but I wont and I trust no confusion will arise.

12.3 On the Structure of Simple Lie Algebras

Now we begin a general study of simple Lie algebras. We have many examples of these coming from some of the classical groups. In this section and in the following chapters we work somewhat generally. From here on in, unless otherwise stated, our Lie algebras are 'complexified' that is we understand the underlying vector spaces to be vector spaces over the complex numbers.

In the simple Lie algebra we pick a subset of commuting generators which form the Cartan subalgebra $\mathfrak{h}$. The rest of the generators, the root vectors, are eigenvectors of the Cartan generators and we have for $h \in \mathfrak{h}$,

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

for each root vector X_α. This prompts a

Definition 12.3. Let $\mathfrak{g}$ be a semi-simple Lie algebra and $\mathfrak{h}$ a Cartan Subalgebra, the root space $\mathfrak{g}_\alpha$ of a root $\alpha \in \mathfrak{h}^*$ is

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} \,|\, [h,\, X] = \alpha(h)\, X \;\forall h \in \mathfrak{h}\}$$

To be able to fix the Cartan subalgebra we need to be a bit fussy. So here is a definition:

Definition 12.4. An element h of a Lie algebra V is said to be a regular element (or just regular) if ad_h has the minimal number of zero eigenvalues possible.

Definition 12.5. The Cartan subalgebra of a Lie algebra V is the maximal commuting subalgebra of V that contains a regular element.

Note that for $SU(3)$ we have that ad_{I_3} has I_3 and Y as zero eigenvalue eigenvectors. Y on the other hand has I_3, Y and $X_{\pm\alpha_1}$ as zero eigenvalue eigenvectors, so that it is not a regular element. Note that X_{α_3} has zero eigenvectors X_{α_3}, X_{α_1} and X_{α_2} so it too is not regular, and so on. We find therefore that the regular element is actually I_3.

Theorem 12.3. Let Δ be the set of roots with respect to the Cartan subalgebra $\mathfrak{h}$ then the following hold:

1. $\mathfrak{g} = \mathfrak{h} \oplus \sum_{\alpha\in\Delta} \mathfrak{g}_\alpha$
2. If $\alpha,\ \beta \in \Delta$ then $[\mathfrak{g}_\alpha,\, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$
3. If $\alpha,\ \beta \in \Delta$ and $\alpha + \beta \neq 0$ then $B(\mathfrak{g}_\alpha,\, \mathfrak{g}_\beta) = 0$.
4. The Cartan-Killing form restricted to $\mathfrak{h}$ is non-degenerate.
5. If $\alpha \in \Delta$ then $-\alpha \in \Delta$.
6. $\{\alpha \in \Delta\}$ span $\mathfrak{h}^*$.

Proof:

1. The elements of $\mathfrak{h}$ are such that for $h_1,\ h_2 \in \mathfrak{h}$ then $[\mathrm{ad}_{h_1},\, \mathrm{ad}_{h_2}] = 0$ so the ad_{h_i} are simultaneously diagonalisable. Thus there exists a common basis of eigenvectors $\{e_1, \ldots, e_k\}$ such that $\mathrm{ad}_h(e_i) = \lambda(h)\, e_i$. The rest of the Lie algebra is spanned by the $\{e_i\}$, that is by the root vectors.
2. We determine the commutation relations between the root vectors themselvesby use of the Jacobi identity, for any $h \in \mathfrak{h}$, $X_\alpha \in \mathfrak{g}_\alpha$ and $X_\beta \in \mathfrak{g}_\beta$

$$\begin{aligned} [h,\, [X_\alpha, X_\beta]] &= -[X_\alpha,\, [X_\beta, h] - [X_\beta,\, [h, X_\alpha]] \\ &= (\beta(h) + \alpha(h))\; [X_\alpha, X_\beta]. \end{aligned}$$

so that $[X_\alpha,\, X_\beta] \in \mathfrak{g}_{\alpha+\beta}$ with $\mathfrak{g}_0 = \mathfrak{h}$. (From this we learn that $[X_\alpha, X_\beta] = 0$, or $[X_\alpha, X_\beta]$ is a root vector with root $\alpha+\beta$ or $\alpha+\beta = 0$ so that $[X_\alpha, X_\beta]$ commutes with all elements in $\mathfrak{h}$ so it too is in $\mathfrak{h}$).

3. This is an

   Exercise 12.1. Show that $B(X_\alpha, X_\beta) = 0$ unless $\alpha + \beta = 0$. **Hint:** Note that $\mathrm{ad}_{X_\alpha} . \mathrm{ad}_{X_\beta} : \mathfrak{g}_\gamma \longrightarrow \mathfrak{g}_{\alpha+\beta+\gamma}$ and also check to see if $\mathrm{ad}_{X_\alpha} . \mathrm{ad}_{X_\beta}$ maps h_i to itself.

4. As $\mathfrak{g}$ is semi-simple B is non-degenerate, however, $B(h, X_\alpha) = 0\ \forall h \in \mathfrak{h}$ and $X_\alpha \in \mathfrak{g}_\alpha$ by the previous exercise (let $\beta = 0$). In order for B to be non-degenerate we therefore must have that $B|_\mathfrak{h}$ is non-degenerate.

5. This follows from the exercise as well in that for $X_\alpha \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}$, $B(X_\alpha,\, Y) = 0$ unless $Y \in \mathfrak{g}_{-\alpha}$ and there must be such a Y otherwise $B(\mathfrak{g}_\alpha,\, \mathfrak{g}) = 0$ and B would be degenerate which is not allowed as $\mathfrak{g}$ is semi-simple.

6. If the roots $\{\alpha\}$ do not span $\mathfrak{h}^*$ then there must exist a $h \in \mathfrak{h}$ st $\forall \alpha\ \alpha(h) = 0$, so that $B(h,k) = 2\sum_{\alpha\in\Delta} \alpha(h)\alpha(k) = 0\ \forall k \in \mathfrak{h}$ but this contradicts the fact that $B|_\mathfrak{h}$ is non-degenerate.

In the third case above we have that $[X_\alpha, X_{-\alpha}] \in \mathfrak{h}$ and we would like to know which element it is equal to. We can use the Cartan-Killing form to work this out.

As $B|_\mathfrak{h}$ is non-degenrate it gives us an identification of $\mathfrak{h}^*$ with $\mathfrak{h}$.

Definition 12.6. Given a root $\alpha \in \Delta$ there exist a unique element $h_\alpha \in \mathfrak{h}$ defined by

$$B(h_\alpha,\, k) = \alpha(k), \;\; \forall k \in \mathfrak{h}$$

We will also make use of the following property:

Exercise 12.2. Show that $B(X, [Y, Z]) = B([X, Y], Z)$.

Theorem 12.4. Let Δ be the set of roots with respect to the Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$. Then, with $\alpha \in \Delta$, the following hold:

1. $[X_\alpha,\, X_{-\alpha}] = B(X_\alpha,\, X_{-\alpha}) \,.\, h_\alpha$ and $[\mathfrak{g}_\alpha,\, \mathfrak{g}_{-\alpha}] = \mathbb{C} \,.\, h_\alpha$

2. $B(h_\alpha,\, h_\alpha) \neq 0$.

3. $\dim \mathfrak{g}_\alpha = 1$ and $\mathfrak{g}_{n\alpha}$ is empty except for $n = 0, \pm 1$.

Proof:

1. By Exercise 12.2 for all $h \in \mathfrak{h}$

$$\begin{aligned} B([X_\alpha, X_{-\alpha}], h) &= B(X_\alpha,\, [X_{-\alpha}, h]) \\ &= \alpha(h)\, B(X_\alpha, X_{-\alpha}). \end{aligned}$$

However, $\alpha(h) = B(h_\alpha, h)$ so combining this with the previous formula and the non-degeneracy of B we get $[X_\alpha, X_{-\alpha}] = B(X_\alpha, X_{-\alpha})\, h_\alpha$. as required for the first part. For the second part, as $B(X_\alpha,\, X_\beta)$ is non-zero only if $\beta = -\alpha$ there must be a non-zero $X_{-\alpha}$ since B is non-degenerate.

2. By non-degeneracy of $B|_{\mathfrak{h}}$ for a given $\alpha \in \Delta$ there must exist a $\beta \in \Delta$ so that $B(h_\alpha,\, h_\beta) \neq 0$. Let $V \subset \mathfrak{g}$ be the subspace of the α string through β
$$V = \sum_{n\in\mathbb{Z}} \mathfrak{g}_{\beta+n\alpha}$$
The vector space V is preserved by ad_{h_α} since it preserves root spaces and it is also preserved by $\mathrm{ad}_{X_{\pm\alpha}}$ since these just move us up and down the summands. ($[\mathfrak{g}_\alpha,\, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$). By 1) we can pick[1] X_α and $X_{-\alpha}$ so that $[X_\alpha,\, X_{-\alpha}] = h_\alpha$. We have that
$$\mathrm{Tr}\,(\mathrm{ad}_{h_\alpha})_{|V} = \mathrm{Tr}\left(\mathrm{ad}_{[X_\alpha,\, X_{-\alpha}]}\right)_{|V} = \mathrm{Tr}\left([\mathrm{ad}_{X_\alpha},\, \mathrm{ad}_{X_{-\alpha}}]\right)_{|V} = 0$$
but we can get an explicit expression for the first trace since we know that on rood spaces $(\mathrm{ad}_{h_\alpha})_{|\mathfrak{g}_\gamma} = \gamma(h_\alpha).\mathrm{Id}$ whence
$$0 = \mathrm{Tr}\,(\mathrm{ad}_{h_\alpha})_{|V} = \sum_{n\in\mathbb{Z}} \mathrm{Tr}\,(\mathrm{ad}_{h_\alpha})_{|\gamma_{\beta+n\alpha}} = \sum_{n\in\mathbb{Z}} (\beta+n\alpha)(h_\alpha)\, d_{\beta+n\alpha}$$
where $d_\gamma = \mathrm{Tr}\,\mathrm{Id}_{|\mathfrak{g}_\gamma}$ is the dimension of $\mathfrak{g}_\gamma$. As $\beta(h_\alpha) = B(h_\beta,\, h_\alpha) \neq 0$ and as $d_\gamma \geq 0$, in particular $d_\beta \geq 1$ the above equation implies that $\alpha(h_\alpha) = B(h_\alpha,\, h_\alpha) \neq 0$ as required.

3. This time one takes
$$V = \mathbb{C}\, X_{-\alpha} + \mathbb{C}\, h_\alpha + \sum_{n\geq 1} \mathfrak{g}_{n\alpha}$$
as before V is preserved by ad_{h_α} and $\mathrm{ad}_{X_{\pm\alpha}}$. $\mathrm{Tr}\,(\mathrm{ad}_{h_\alpha})_{|V}$ is zero again and we find
$$-\alpha(h_\alpha) + \sum_{n\geq 1} d_{n\alpha} B(h_\alpha,\, n h_\alpha) = B(h_\alpha,\, h_\alpha)\,.\,(-1 + d_\alpha + \sum_{n\geq 2} n\, d_{n\alpha}) = 0.$$
If any of the $d_{n\alpha} \neq 0$ with $n \geq 2$ then the sum in the formula is greater than one and the formula cannot be solved. Hence we have $d_{n\alpha} = 0$ with $n \geq 2$ and $d_\alpha = 1$.

So putting the pieces together we learn that for any (semi-) simple Lie algebra the commutation relations are
$$\begin{array}{rcll} [h_i, h_j] & = & 0, & h_i,\, h_j \in \mathfrak{h} \\ [h_i, X_\alpha] & = & \alpha(h_i) X_\alpha, & \alpha \in \Delta,\ h_i \in \mathfrak{h} \\ [X_\alpha, X_\beta] & = & N_{\alpha\beta}\, X_{\alpha+\beta}, & \alpha+\beta \in \Delta \\ & = & B(X_\alpha, X_{-\alpha})\, h_\alpha & \alpha+\beta = 0 \\ & = & 0, & \alpha+\beta \neq 0 \quad \alpha+\beta \notin \Delta. \end{array}$$

[1]Scale $\widetilde{X}_{\pm\alpha} = \lambda X_{\pm\alpha}$ with $\lambda = 1/\sqrt{B(X_\alpha, X_{-\alpha})}$ and by linearity of B we obtain $[\widetilde{X}_\alpha,\, \widetilde{X}_{-\alpha}] = \lambda^2 B(X_\alpha, X_{-\alpha})\, h_\alpha = h_\alpha$.

and we also know that the root spaces are all one dimensional.

We can still get more information by considering in detail the α strings through β,

Definition 12.7. Let α, $\beta \in \Delta$ be roots. The α string through β is the set of roots of the form $\{\beta + n\alpha \,|\, n \in \mathbb{Z}\}$ with root spaces $\mathfrak{g}_{\beta+n\alpha}$.

Theorem 12.5. Let Δ be the set of roots, and α, $\beta \in \Delta$ then:

1. There are integers p, $q \geq 0$ such that the α string through β consists of consecutive roots i.e. $-p \leq n \leq q$ and
$$\frac{2B(h_\alpha,\, h_\beta)}{B(h_\alpha,\, h_\alpha)} = p - q$$
2. If $\beta = c\alpha$ with $c \in \mathbb{C}$ then $c = 0$ or ± 1.
3. If $\alpha + \beta \in \Delta$ then $[X_\alpha,\, X_\beta] \neq 0$.

Before giving a proof of this theorem you should note that the formula in the first statement does not have any root dependence on the right hand side while on the left there is no dependence on p or q.

Proof:

1. Suppose there is a consecutive range of integers $n \in [r,\, s]$ such that $\beta + n\alpha \in \Delta$. We let $V = \sum_{n=r}^{s} \mathfrak{g}_{\beta+n\alpha}$ and use the same trick of taking a zero trace $\mathrm{Tr}\,(\mathrm{ad}_{h_\alpha})_{|V} = 0$ and evaluting as before
$$\sum_{n=r}^{s} (\beta + n\alpha)(h_\alpha = (s - r + 1)\left(\beta(h_\alpha) + \frac{(s+r)}{2}\alpha(h_\alpha)\right) = 0$$
whence we have that
$$\frac{2B(h_\alpha,\, h_\beta)}{B(h_\alpha,\, h_\alpha)} = -(r + s)$$
Now suppose there is another consecutive range of integers $n \in [r',\, s']$ such that $\beta + n\alpha \in \Delta$, then we would find that $r' + s' = r + s$. Which means that one of the intervals is contained in the other. If, $s' > s$ we need $r' < r$ for the equality to hold and we would have $[r,\, s] \subset [r',\, s']$. Consequently there is only one interval and the string is connected. As $n = 0$ is allowed the result follows.

2. For $\beta = c\alpha$ we know that the α string through β implies that $2c = p - q$ or $2c \in \mathbb{Z}$ we may also consider the β string through α so that we have $2/c \in \mathbb{Z}$. The allowed values of c are 0, $\pm 1/2$, ± 1 and ± 2. We already know that if α is a root then $\beta = \pm 2\alpha$ is not a root. On the other hand if $\beta = \pm \alpha/2$ that is the same as $\alpha = \pm 2\beta$ but for β a root it would imply that α is not a root.

3. If $[X_\alpha, X_\beta] = 0$ as the α string through β satisfies $-p \leq n \leq q$, we have that the vector space $V = \sum_{n=-p}^{0} \mathfrak{g}_{\beta+n\alpha}$ is preserved by ad_{h_α} and $\mathrm{ad}_{X_{\pm\alpha}}$. Taking the trace of ad_{h_α} restricted to V we find that

$$\sum_{n=-p}^{0} (\beta + n\alpha)(h_\alpha) = (p+1)\beta(h_\alpha) - \frac{p(p+1)}{2}\alpha(h_\alpha) = 0$$

Thus we have that $2B(h_\alpha, h_\beta)/B(h_\alpha, h_\alpha) = p$ but we know that that should be $2B(h_\alpha, h_\beta)/B(h_\alpha, h_\alpha) = p - q$ hence $q = 0$. But $q \geq 1$ as $\alpha + \beta$ is a root so we get a contradiction. □

We now come to a very important part of the general theory. So far we have been using complex coefficients. We now restrict to a real subspace of the Cartan subalgebra, namely

$$\mathfrak{h}_\mathbb{R} = \sum_\alpha \mathbb{R}.h_\alpha \tag{12.3.1}$$

Theorem 12.6. $B|_{\mathfrak{h}_\mathbb{R}}$ is real and positive definite, ie. $B(h,\, h) \geq 0$ $\forall h \in \mathfrak{h}_\mathbb{R}$ and $B(h,\, h) = 0$ implies $h = 0$.

Proof: We already know that for $h,\, k \in \mathfrak{h}$, $B(h,\, k) = 2\sum_{\alpha\in\Delta} \alpha(h)\,\alpha(k)$. Let

$$2\frac{B(h_\alpha, h_\beta)}{B(h_\alpha,\, h_\alpha)} = n_{\alpha,\beta} \in \mathbb{Z}$$

(that this is integral follows from 1) of the previous theorem). We have

$$B(h_\alpha,\, h_\alpha) = \sum_{\gamma\in\Delta} \gamma(h_\alpha)^2 = \frac{1}{4}B(h_\alpha,\, h_\alpha)^2 \sum_{\gamma\in\Delta} n_{\alpha,\gamma}^2$$

and as $B(h_\alpha,\, h_\alpha) \neq 0$ we deduce that

$$B(h_\alpha,\, h_\alpha) = \frac{2}{\sum_{\gamma\in\Delta} n_{\alpha,\gamma}^2} \in \mathbb{Q}, \quad \text{and} \quad B(h_\alpha,\, h_\beta) = \frac{n_{\alpha,\beta}}{\sum_{\gamma\in\Delta} n_{\alpha,\gamma}^2} \in \mathbb{Q}$$

These formulae tell us that if we take $h = \sum c_\alpha \,.\, h_\alpha$ and $d = \sum d_\alpha \,.\, h_\alpha$ with real coefficients c_α and d_α then $B(h,\, k) \in \mathbb{R}$. If $h \in \mathfrak{h}_\mathbb{R}$ and not zero there exists a root β such that $\beta(h) \neq 0$ and

$$B(h,\, h) = 2\sum_{\gamma\in\Delta} \gamma(h)^2 \geq \beta(h)^2 > 0$$

□

12.4 Uniqueness of the Cartan-Killing Form

One can also prove that up to multiplication by a constant the Cartan-Killing form is the **only** invariant bilinear form on a simple Lie algebra. It is not unique on semi-simple Lie algebras.

Suppose there is another invariant bilinear form and call it $\widehat{B}$. Then

$$\begin{aligned}
\widehat{B}(h_\beta, [X_\alpha, X_{-\alpha}]) &= \hat{B}(h_\beta, B(X_\alpha, X_{-\alpha})\, h_\alpha) \\
&= B(X_\alpha, X_{-\alpha})\, \widehat{B}(h_\beta, h_\alpha) \\
&= \widehat{B}([h_\beta, X_\alpha],\, X_{-\alpha}) \\
&= \alpha(h_\beta)\, \widehat{B}(X_\alpha, X_{-\alpha}) \\
&= B(h_\alpha, h_\beta)\, \widehat{B}(X_\alpha, X_{-\alpha}).
\end{aligned}$$

Now we can compare the second and last equality in the previous equation to find that

$$\frac{\widehat{B}(h_\beta, h_\alpha)}{B(h_\beta, h_\alpha)} = \frac{\widehat{B}(X_\alpha, X_{-\alpha})}{B(X_\alpha, X_{-\alpha})}$$

The right hand side of this equality does not depend on β so the ratio

$$\frac{\widehat{B}(h_\beta, h_\alpha)}{B(h_\beta, h_\alpha)}$$

does not depend on β. However, this ratio is symmetric in α and β, so it cannot depend on α either! Consequently we have that there must exist a constant, c so that

$$\frac{\widehat{B}(h_\beta, h_\alpha)}{B(h_\beta, h_\alpha)} = c = \frac{\widehat{B}(X_\alpha, X_{-\alpha})}{B(X_\alpha, X_{-\alpha})}$$

The only time this argument can break down is when $B(h_\alpha, h_\beta) = 0$. However, in a simple algebra we can start with a root α and then move onto a root β so that $B(h_\alpha, h_\beta) \neq 0$ and then go from β to another root γ so that $B(h_\beta, h_\gamma) \neq 0$ and so on until all the roots are exhausted.

Hence we have shown that for a simple Lie algebra any two invariant bilinear forms are equal up to an overall constant, i.e.

$$\widehat{B}(X, Y) = cB(X, Y)$$

for some constant. The situation for semi-simple Lie algeras is somewhat more complicated. Let V_i be a set of simple Lie algebras and let

$$V = \bigoplus_i V_i,$$

be a semi-simple Lie algebra. On each V_i we have a Cartan-Killing form $\widehat{B}_i$ and so in general any invariant bilinear form is of the form $c_i\, \widehat{B}_i$. Hence, the most general invariant bilinear form on V is

$$\sum_i c_i\, \widehat{B}_i$$

(we understand that the $\widehat{B}_i$ give zero for vectors which do not lie in V_i). So we see that there are many different invariant bilinear forms. This really exemplifies why the simple algebras are simple.

CHAPTER 13

THE CLASSICAL LIE GROUPS

We have noted that we can use the Cartan-Killing form to define an inner product on the root space. It is convenient to normalise that inner product in different ways in different groups. This is a matter of choice, usually, made for convenience. So quite generally we take the inner product $\langle \cdot\,, \cdot\rangle$ on $\mathfrak{h}^*$, the dual of $\mathfrak{h}$, to be proportional to the Cartan-Killing form $B(\cdot\,, \cdot)$ and we define the Cartan elemnts h_α with respect to the inner product. We pick $\langle \cdot\,, \cdot\rangle = c(\mathfrak{g})\,.\,B(\cdot\,, \cdot)$ so that,

$$\langle \alpha,\, \beta\rangle = \alpha(h_\beta) \tag{13.0.1}$$

where $c(\mathfrak{g})$ is some constant depending on the Lie algebra to ensure that (13.0.1) holds.

The reason for our choice is that we want to be able to compare lengths of roots between Lie algebras (as will soon become apparent). With this in mind we normalise so to rid ourselves of extraneous Lie algebra factors.

In this chapter we will also make use of the matrix representation of the Lie algebra of $\mathfrak{gl}(n, \mathbb{C})$. The reason for this is that all of the groups of interest are subgroups of $\mathrm{Gl}(n, \mathbb{C})$ and their Lie algebras are sub algebras of $\mathfrak{gl}(n, \mathbb{C})$. So we start there.

13.1 $\mathfrak{gl}(n, \mathbb{C})$

Recall, therefore, that $\mathfrak{gl}(n, \mathbb{C})$ is spanned by the matrices E_{ij} with commutation relations

$$[E_{ij},\, E_{kl}] = E_{il}\delta_{jk} - E_{kj}\delta_{il} \tag{13.1.1}$$

The Cartan subalgebra is

$$\mathfrak{h} = \{h = \sum_{i=1}^{n} h_i\, E_{ii}\,|\, h_i \in \mathbb{C}\} \tag{13.1.2}$$

13.2 $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C}) = \mathrm{A}_{n-1}$

A Cartan subalgebra is given by the diagoanl elements

$$\mathfrak{h} = \{h = \sum_{i=1}^{n} h_i\, e_i |\, h_i \in \mathbb{C},\ \sum_{i=1}^{n} h_i = 0\} \subset \mathbb{C}^n$$

where $e_i = E_{ii}$. A dual basis for $\mathbb{C}^*$ is ω_i such that $\omega_i(e_j) = \delta_{ij}$. We have that

$$[h,\, E_{ij}] = (h_i - h_j).E_{ij}. \tag{13.2.1}$$

In this case we find, recalling that we set $[h,\, X_\alpha] = \alpha(h)\, X_\alpha$, the following roots and root spaces

$$\begin{aligned} \Delta &= \{\pm(\omega_i - \omega_j)\,|\, i < j\} \\ \mathfrak{g}_{\omega_i - \omega_j} &= \mathbb{C}.E_{ij} \text{ for } i \neq j \end{aligned}$$

This means that the root vectors are proportional to E_{ij} for $i \neq j$.

The Cartan-Killing form is given by

$$B(h,\, h) = \sum_{\alpha \in \Delta} \alpha(h)^2 = 2 \sum_{i<j} (\omega_i(h) - \omega_j(h))^2 = 2 \sum_{i<j} (h_i - h_j)^2 = 2n \sum_i h_i^2$$

the last equality follows from the fact that the element h is traceless $0 = (\sum_i h_i)^2 = \sum_i h_i^2 + 2 \sum_{i<j} h_i h_j$.

Exercise 13.1. Show that $B(h,\, k) = 2n \sum_{i=1}^{n} h_i\, k_i$ where $h = \sum h_i\, e_i,\ k = \sum_i k_i\, e_i \in \mathfrak{h}$.

The inner product on the root space is defined to get rid of the overall factor[1] of $2n$ Hence $\langle \cdot\,, \cdot \rangle = B(\cdot\,, \cdot)/2n$. So we have by the previous exercise that, with the root $\alpha = \omega_i - \omega_j$ and on setting $h_{\omega_i - \omega_j} = \sum_l h^l_{\omega_i - \omega_j}\, e_l$, that

$$(\omega_i - \omega_j)(\sum_l k_l\, e_l) = k_i - k_j = \sum_l h^l_{\omega_i - \omega_j} k_l \tag{13.2.2}$$

from which we deduce that (these are the coefficients **not** the Cartan element) $h^l_{\omega_i - \omega_j} = \delta_{il} - \delta_{jl}$ or (now the Cartan element)

$$h_{\omega_i - \omega_j} = e_i - e_j. \tag{13.2.3}$$

Also you should note that with this definition of the inner product we have that

$$\langle \alpha,\, \beta \rangle = \mathrm{Tr}\,(h_\alpha h_\beta) \tag{13.2.4}$$

[1] This means that we define the dual vector h_α to the root α by $B(h_\alpha,\, k) = 2n\alpha(k)\ \forall k \in \mathfrak{h}$.

and that

$$\langle \alpha,\, \alpha \rangle = 2 \tag{13.2.5}$$

so that the length of the roots is $\sqrt{2}$. Lastly, the real Cartan subalgebra is

$$\mathfrak{h}_{\mathbb{R}} = \{h = \sum_{i=1}^{n} h_i\, e_i |\, h_i \in \mathbb{R},\ \sum_{i=1}^{n} h_i = 0\} \tag{13.2.6}$$

It is important to note that, in this case, the inner product does not make the ω_i an orthonormal basis of $\mathfrak{h}^*$ $\langle \omega_i,\, \omega_j \rangle =?$ as the ω_i themselves are not in $\mathfrak{h}^*$ since the trace free condition needs to be imposed.

In short we have shown that

Proposition 13.1. For $n \geq 2$ the group $\mathrm{Sl}(n, \mathbb{C})$ has

1. Roots Space: $\Delta = \{\pm(\omega_i - \omega_j) \,|\, i < j \leq n\}$
2. Positive Roots: $\Delta_+ = \{(\omega_i - \omega_j) \,|\, i < j \leq n\}$
3. Basis for $\mathfrak{h}^*$: $\{\omega_i - \omega_{i+1} \,|\, 1 \leq i < n\}$

13.3 $\mathfrak{g} = \mathfrak{so}(2n,\, \mathbb{C}) = \mathrm{D}_n$

As the matrices that form $\mathfrak{so}(2n)$ are skew transpose it is useful to pass to a basis of such matrices

$$F_{ij} = E_{ij} - E_{ji} \tag{13.3.1}$$

with commutation relations that are zero unless one label appears twice and in that case

$$[F_{ij},\, F_{jk}] = \begin{cases} F_{ik} & \text{provided the labels } i, j, k \text{ are all distinct} \\ 0 & \text{otherwise} \end{cases} \tag{13.3.2}$$

The Cartan subalgebra is

$$\mathfrak{h} = \{h = \sum_{i=1}^{n} h_i F_{2i-1\,2i} \,|\, h_i \in \mathbb{C}\} \tag{13.3.3}$$

that is the Cartan elements are of the form

$$h = \begin{pmatrix} h_1\boldsymbol{\epsilon} & & \\ & \ddots & \\ & & h_n\boldsymbol{\epsilon} \end{pmatrix} \quad \text{where} \quad \boldsymbol{\epsilon} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{13.3.4}$$

We define two types of root vectors:

$$\begin{aligned} X_{k\,l} &= F_{2k-1,\,2l-1} + F_{2k,\,2l} + i(F_{2k-1,\,2l} - F_{2k,\,2l-1}) \\ Y_{k\,l} &= F_{2k-1,\,2l-1} - F_{2k,\,2l} + i(F_{2k-1,\,2l} + F_{2k,\,2l-1}) \end{aligned} \tag{13.3.5}$$

The for $h \in \mathfrak{h}$

$$[h,\, X_{k\,l}] = -i(h_k - h_l)\, X_{k\,l}, \quad [h,\, Y_{k\,l}] = i(h_k + h_l)\, Y_{k\,l}. \tag{13.3.6}$$

Exercise 13.2. Derive the commutation relations (13.3.6).

Let $e_i = iF_{2i-1,\,2i}$ (these are hermitian and had we bothered to diagonalise them, that is chosen a different basis, we would have had real diagonal entries) be a basis of $\mathfrak{h}_{\mathbb{R}}$ and ω_i be the dual basis to the e_i. The root system is

$$\Delta = \{\pm(\omega_i + \omega_j),\, \pm(\omega_i - \omega_j) \,|\, i < j\} \tag{13.3.7}$$

The Cartan-Killing form is

$$B = 2\sum_{i<j}(\omega_i - \omega_j)^2 + 2\sum_{i<j}(\omega_i + \omega_j)^2 = 4(n-1)\sum_i \omega_i^2 \tag{13.3.8}$$

The normalised inner product is once more taken to eliminate the overall factor

$$\langle \cdot\,,\,\cdot \rangle = \frac{1}{4(n-1)} B(\cdot\,,\,\cdot) = \sum_i \omega_i(\cdot)\,\omega_i(\cdot)$$

This means that we define the dual vector h_α to the root α by $B(h_\alpha,\, k) = 4(n-1)\alpha(k)$ $\forall k \in \mathfrak{h}$. In this way one finds that

$$h_{\pm\omega_i \pm\omega_j} = \pm e_i \pm e_j \tag{13.3.9}$$

Exercise 13.3. Prove this.

Once more the roots satisfy $\langle \alpha,\, \alpha \rangle = 2$ and so they have length $\sqrt{2}$.

This time we are in the happy situation that the inner product does make the ω_i an orthonormal basis of $\mathfrak{h}^*$. As the ω_i are linear combinations of the roots and $\omega_i(e_j) = \delta_{ij}$ we have

$$\langle \omega_i,\, \omega_j \rangle = \sum_k \omega_k(e_i)\,\omega_k(e_j) = \delta_{ij}. \tag{13.3.10}$$

as required.

13.4 $\mathfrak{g} = \mathfrak{so}(2n+1,\,\mathbb{C})$

The Cartan subalgebra of $\mathfrak{so}(2n+1,\,\mathbb{C})$ is taken to be that of $\mathfrak{so}(2n,\,\mathbb{C}) \subset \mathfrak{so}(2n+1,\,\mathbb{C})$. We are saying that Cartan subalgebra elemts are,

$$h = \begin{pmatrix} h_1\boldsymbol{\epsilon} & & & \\ & \ddots & & \\ & & h_n\boldsymbol{\epsilon} & \\ & & & 0 \end{pmatrix} \quad \text{where} \quad \boldsymbol{\epsilon} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{13.4.1}$$

where the last zero is the 1×1 zero matrix.

However, there are not only the roots of $\mathfrak{so}(2n,\,\mathbb{C})$, but there are more roots and, consequently, root vectors

$$[h,\, F_{2k-1,\,2n+1} + iF_{2k,\,2n+1}] = ih_k\ (F_{2k,\,2n+1} + iF_{2k-1,\,2n+1}) \tag{13.4.2}$$

and their Hermitian conjugates. Thus

$$\Delta = \{\pm\omega_i,\, \pm(\omega_i + \omega_j), \pm(\omega_i - \omega_j \,|\, i < j\} \tag{13.4.3}$$

The Killing form is $B(\cdot\,,\,\cdot) = 2(2n-1)\sum_i \omega_i(\cdot)\,\omega_i(\cdot)$ and we strip away the prefactor again to define

$$\langle\cdot\,,\,\cdot\rangle = \frac{1}{2(2n-1)}B(\cdot\,,\,\cdot) = \sum_{i=1}^{n}\omega_i(\cdot)\omega_i(\cdot)$$

The e_i and ω_i are an orthonormal basis of $\mathfrak{h}$ and $\mathfrak{h}^*$ $(\omega_i(e_j) = \delta_{ij})$, we have $h_{\pm\omega_i} = \pm e_i$ and $h_{\pm\omega_i\pm\omega_j} = \pm e_i \pm e_j$.

Now we get a slight change in the lengths in that we can have 'short' roots of length 1 (the $\pm\omega_i$) and 'long' roots of length $\sqrt{2}$ (the rest).

Exercise 13.4. Repeat all the steps in this section for $SO(3)$.

13.5 $\mathfrak{g} = \mathfrak{sp}(n,\,\mathbb{C})$

Recall that M an element of $\mathfrak{sp}(n,\,\mathbb{C})$ is a $2n \times 2n$ matrix which satisfies

$$M^T.E + E.M = 0 \tag{13.5.1}$$

If we write M in $n \times n$ blocks

$$\mathfrak{sp}(n,\mathbb{C}) = \left\{\begin{pmatrix} A & B \\ C & -A^T \end{pmatrix} \,|\, A \in \mathfrak{gl}(n,\,\mathbb{C}),\, B^T = B,\, C^T = C\right\} \tag{13.5.2}$$

We take the Cartan sub algebra to be

$$\mathfrak{h} = \{h = \mathbf{diag}(h_1, \ldots, h_n, -h_1, \ldots, -h_n) \,|\, h_i \in \mathbb{C}\} \tag{13.5.3}$$

We take the root vectors to be

$$\begin{aligned}
X_{k,\,l} &= E_{n+k,\,l} + E_{n+l,\,k} \quad k \le l \le n \\
X^T_{k,\,l} &= E_{l,\,n+k} + E_{k,\,n+l} \quad k \le l \le n \\
Y_{k,\,l} &= E_{k,\,l} - E_{n+l\,n+k} \quad k < l \le n
\end{aligned} \tag{13.5.4}$$

You should note that the $X_{k,\,l}$ form a symmetric basis for the matrix B in the decomposition (13.5.2), the $X^T_{k,\,l}$ are a symmetric basis for C while the $Y_{k,\,l}$ form a basis for the non-diagonal parts of $A \oplus -A^T$ as it appears in (13.5.2).

Exercise 13.5. Show that

$$\begin{aligned}
[h, X_{k,\,l}] &= -(h_k + h_l)\, X_{k,\,l}, \\
[h, X^T_{k,\,l}] &= \ (h_k + h_l)\, X^T_{k,\,l} \\
[h,\, Y_{k,\,l}] &= \ (h_k - h_l)\, Y_{k,\,l}
\end{aligned} \tag{13.5.5}$$

The root vectors X and X^T allow for the case where $k = l$ (the diagonal elements of B and C respectively).

Let $e_i = E_{i,\,i} - E_{n+i,\,n+i} \in \mathfrak{h}_{\mathbb{R}}$ and let ω_i be the dual basis. The roots are

$$\Delta = \{\pm 2\omega_i,\ \pm(\omega_i + \omega_j),\ \pm(\omega_i - \omega_j)\,|\, i < j\} \tag{13.5.6}$$

The Killing form is $B(\cdot\,,\,\cdot) = 4(n+1)\sum_{i=1}^{n} \omega_i(\cdot)\,\omega_i(\cdot)$ and the normalised inner product on the root space is

$$\langle \cdot\,,\,\cdot \rangle = \frac{1}{4(n+1)}\, B(\cdot\,,\,\cdot)$$

Almost all roots have length $\sqrt{2}$, however the roots $\pm\omega_i$ have length 2.

Exercise 13.6. Is $\{\omega_i\}$ an orthonormal basis of $\mathfrak{h}^*$ with respect to the inner product? Justify your answer.

We note that even though $\mathfrak{so}(2n+1,\,\mathbb{C})$ and $\mathfrak{sp}(n,\,\mathbb{C})$ have the same dimension and rank, never the less there is a difference in their root systems which can be used to establish that they are not isomorphic Lie algebras.

13.6 The Real Forms $\mathfrak{su}(n)$ and $\mathfrak{sp}(n)$

This section is mostly a set of exercises.

Exercise 13.7. What combinations of the root vectors that we have defined for $\mathfrak{sl}(n,\,\mathbb{C})$ are elements of $\mathfrak{su}(n)$?

Exercise 13.8. What combinations of the root vectors of $\mathfrak{sp}(n, \mathbb{C})$ are elements of $\mathfrak{sp}(n)$ (the Lie algebra of the compact group $Sp(n)$)?

CHAPTER 14

THE CARTAN MATRIX AND DYNKIN DIAGRAMS

We introduce the concept of a Cartan integer on the root space, which is a precursor to the Cartan matrix. The Cartan matrix is defined on the space of simple roots which we will also come to. By studying these objects one is able to determine all semi-simple Lie algebras. We do not quite do that but we do do enough to see how all the classical and exceptional Lie algebra's arise, without giving a proof of completeness.

14.1 Restricting the Length of Strings

Definition 14.1. A Cartan integer for an α string through β with root $\{\beta+n\alpha \,|\, -p \leq n \leq q\}$, $p,\, q \geq 0$, is

$$n_{\alpha\beta} := 2\,\frac{\langle\alpha,\,\beta\rangle}{\langle\alpha,\,\alpha\rangle} = p - q \tag{14.1.1}$$

Clearly if $n_{\alpha\beta} < 0$ it means that $\beta+\alpha \in \Delta$, while if $n_{\alpha,\,\beta} > 0$ then $\beta - \alpha \in \Delta$. On the other hand if $n_{\alpha\beta} = 0$ the only thing we can deduce is that $p = q$, that is one can add and subtract an equal number of α through β. We do have two useful results which are that if $\alpha,\, \beta \in \Delta$ then

$$\beta - 2\frac{\langle\alpha,\,\beta\rangle}{\langle\alpha,\,\alpha\rangle}\,\alpha \in \Delta \tag{14.1.2}$$

and,

$$n_{\alpha\beta}\,.\,n_{\beta\alpha} = 4.\frac{\langle\alpha,\,\beta\rangle}{\langle\alpha,\,\alpha\rangle}\cdot\frac{\langle\beta,\,\alpha\rangle}{\langle\beta,\,\beta\rangle} = 4\cos^2\theta_{\alpha\beta} \tag{14.1.3}$$

where $\theta_{\alpha\beta}$ is the angle between the α and β roots in the $\alpha-\beta$ plane. The first is $\beta + (q-p)\alpha \in \Delta$ which holds since $(q-p) \in [-p,\, q]$. The second is true since in the $\alpha - \beta$ plane $\langle\cdot\,,\,\cdot\rangle$ is a non-degenerate inner product.

$n_{\alpha\beta}$	$n_{\beta\alpha}$	$\theta_{\alpha\beta}$	Relative Size	α String through β
0	0	$\frac{\pi}{2}$		
1	1	$\frac{\pi}{3}$	$\|\beta\|^2 = \|\alpha\|^2$	$\beta,\, \beta - \alpha$
-1	-1	$\frac{2\pi}{3}$	$\|\beta\|^2 = \|\alpha\|^2$	$\beta,\, \beta + \alpha$
2	1	$\frac{\pi}{4}$	$\|\beta\|^2 = 2\|\alpha\|^2$	$\beta,\, \beta - \alpha,\, \beta - 2\alpha$
-2	-1	$\frac{3\pi}{4}$	$\|\beta\|^2 = 2\|\alpha\|^2$	$\beta,\, \beta + \alpha,\, \beta + 2\alpha$
3	1	$\frac{\pi}{6}$	$\|\beta\|^2 = 3\|\alpha\|^2$	$\beta,\, \beta - \alpha,\, \beta - 2\alpha,\, \beta - 3\alpha$
-3	-1	$\frac{5\pi}{6}$	$\|\beta\|^2 = 3\|\alpha\|^2$	$\beta,\, \beta + \alpha,\, \beta + 2\alpha,\, \beta + 3\alpha$

Table 14.1: Allowed values of $n_{\alpha\beta}$ for $|\beta| \geq |\alpha|$

As the $n_{\alpha\beta}$ are integers and the product is ≤ 4, we find

$$n_{\alpha\beta} \,.\, n_{\beta\alpha} \in \{0,\, 1,\, 2,\, 3,\, 4\} \tag{14.1.4}$$

Furthermore, the product can only be 4 if $\alpha = \pm\beta$, in which case we do not get roots or otherwise

$$n_{\alpha\beta} \in \{0,\, \pm 1,\, \pm 2,\, \pm 3\} \tag{14.1.5}$$

Also 0 appears only if α and β are orthogonal and then both $n_{\alpha\beta}$ and $n_{\beta\alpha}$ are zero. If α and β are neither parallel nor orthogonal it means that one of $n_{\alpha\beta}$ or $n_{\beta\alpha}$ is ± 1 and the other is ± 1, ± 2, or $\pm\, 3$. We tabulate the allowed values in Table 14.1.

To motivate Table 14.1 we establish a simple

Proposition 14.1. An α string through β with root $\{\beta + n\alpha \,|\, -p \leq n \leq q\}$, $p,\, q \geq 0$ has $0 \leq p + q \leq 3$.

Proof: By hypothesis, we have that each term in the α string through β, $\beta - p\alpha$, $\ldots$, $\beta + q\alpha$ is a root. Let $\gamma = \beta + q\alpha$, it is a root and the complete α string through γ is $\gamma - (p+q)\alpha$, $\ldots$, γ. By (14.1.1) we have that $n_{\alpha\gamma} = (p+q)$ while by (14.1.5) this sum is bounded by 3. $\square$

In the proof we have used γ as a highest weight state for the '$SU(2)$' raising operator α (actually with ad_{X_α} on the root vectors), and as it is a highest weight the only direction is down (so acting with $\mathrm{ad}_{X_{-\alpha}}$ on the root vectors).

14.2 Positive, Negative and Simple Roots

One can use what we already know to put restrictions on possible Lie algebras, however, the form in which these conditions appear is not optimal. Part of the problem is that while we know that the root space Δ spans $\mathfrak{h}^*$ it does so in a rather inefficient way - there are far more roots than are required to form a basis of $\mathfrak{h}^*$. Our next task is to correct this and get to a definition where we span exactly $\mathfrak{h}^*$. As a first step we have:

Definition 14.2. h_0 is said to be regular if $\alpha(h_0) \neq 0$ for all $\alpha \in \Delta$ and singular otherwise.

Exercise 14.1. What is the relationship between this definition of a regular element and the one we gave before?

Definition 14.3. We give an ordering so that $\alpha \leq \beta$ for $\alpha,\, \beta \in \Delta$ if $\alpha(h_0) \leq \beta(h_0)$ for some fixed regular element h_0.

Definition 14.4. The set of positive roots Δ^+, with respect to a regular element, is given by

$$\Delta^+ = \{\alpha \,|\, \alpha(h_0) > 0\}$$

The set of negative roots Δ^- is defined so that $\Delta = \Delta^+ \cup \Delta^-$ and $\Delta^+ \cap \Delta^- = \emptyset$.

Definition 14.5. A positive root $\alpha \in \Delta^+$ is said to be a simple root if it cannot be written as the sum of any other two positive roots, that is α is positive if $\alpha \neq \beta + \gamma$ for any $\beta,\, \gamma \in \Delta^+$. The set of simple roots is designated by Σ.

Lets unlock these definitions:

For the Lie algebra $\mathfrak{su}(2)$ we have two roots $\pm\alpha$, so that $\Delta = \{\alpha,\, -\alpha\}$, and a regular element $h = \sigma_3$. We know that $\alpha(h) = 2$ so with $h_0 = h$ we have $\Delta^+ = \{\alpha\}$ and also $\Sigma = \{\alpha\}$.

In the case of $\mathfrak{su}(3)$ we have that $\Delta = \{\pm\alpha_1,\, \pm\alpha_2,\, \pm\alpha_3\}$. A regular element is $h_0 = h_1 + h_2 = \mathbf{diag}(1,\, 0,\, -1)$, since we have $\pm\alpha_1(h_0) = \pm 1$, $\pm\alpha_2(h_0) = \pm 1$ and $\pm\alpha_3(h_0) = \pm 2$ so none vanish. As is clear, with this choice of regular element, the positive roots are $\Delta^+ = \{\alpha_1,\, \alpha_2,\, \alpha_3\}$. The simple roots are $\Sigma = \{\alpha_1,\, \alpha_2\}$ as $\alpha_3 = \alpha_1 + \alpha_2$. Notice that Σ has two elements and that these span $\mathfrak{h}^*$.This is not an accident as we will see presently.

Exercise 14.2. Show that $h_1 = \mathbf{diag}(1,\, -1,\, 0)$ is a regular element for $\mathfrak{su}(3)$. Determine the positive roots and the simple roots for this choice. Is $h_1 - h_2 = \mathbf{diag}(1,\, -2,\, 1)$ a regular element?

Theorem 14.1. Fix a regular element $h_0 \in \mathfrak{h} \subset \mathfrak{g}$.

1. If $\alpha_i,\, \alpha_j \in \Sigma$ then $\alpha_i - \alpha_j \notin \Delta$.

2. If $\alpha_i,\, \alpha_j \in \Sigma$, then $\langle \alpha_i,\, \alpha_j \rangle \leq 0$ for $i \neq j$.

3. The elements of Σ are linearly independent (as vectors).

4. If $\alpha \in \Delta^+$ then α is a linear combination of simple roots with positive coefficients $\alpha = \sum c_i\, \alpha_i$ $\alpha_i \in \Sigma$, $c_i \geq 0$. In particular the elements of the set Σ form a basis of $\mathfrak{h}^*_{\mathbb{R}}$.

Proof:

1. Presume to the contrary that $\alpha_i - \alpha_j$ is a root. Then it must be either positive or negative, so either $(\alpha_i - \alpha_j)$ is positive or $(\alpha_j - \alpha_i)$ is positive. Whence either $\alpha_i = \alpha_j + (\alpha_i - \alpha_j)$ so that α_i is not simple or $\alpha_j = \alpha_i + (\alpha_j - \alpha_i)$ so that α_j is not simple.

2. Suppose rather that $\langle \alpha_i,\, \alpha_j \rangle > 0$. In this case the α_i string through α_j tells us that $2\langle \alpha_i,\, \alpha_j \rangle / \langle \alpha_i,\, \alpha_i \rangle = p - q > 0$ which would imply that $p \geq 1$ or that $\alpha_j - \alpha_i$ is a root contrary to the previous result.

3. Let $\Sigma = \{\alpha_1, \ldots, \alpha_r\}$ and assume the elements of Σ are not linearly independent, that is there exist coefficients n_i such that $\sum_{i=1}^{r} n_i\, \alpha_i = 0$. As the we do not know the signs of the coefficients move the negative ones to the right hand side of the equation so that we may write it as $\sum_i p_i\, \alpha_i = \sum_i q_i\, \alpha_i$ with $p_i,\, q_i > 0$ and each α_i appears at most on one side. Then

$$\langle \sum_i p_i\, \alpha_i,\, \sum_j p_j\, \alpha_j \rangle = \langle \sum_i p_i\, \alpha_i,\, \sum_j q_j\, \alpha_j \rangle = \sum_{i \neq j} p_i\, q_j\, \langle \alpha_i,\, \alpha_j \rangle \leq 0$$

   but the first term in the equation is positive since it is a square while the last sum is a sum of negative terms (the product $p_i q_j \geq 0$) the only way for this to happen is if $p_i = 0$ and like wise for q_i as we could have started with $\langle \sum_i q_i\, \alpha_i,\, \sum_j q_j\, \alpha_j \rangle$ to deduce that $q_i = 0$. Consequently all the n_i are zero.

4. If a positive root α is not simple it can be expressed as the sum of two positive roots now repeat this for each of these roots we get that α can be expressed as a sum of simple roots. □

14.3 The Cartan Matrix

By making use of the regular element of the Cartan subalgebra we have managed to structure, or put some order into, the root spaces.

Definition 14.6. Let $\Sigma = \{\alpha_1, \ldots, \alpha_r\}$ be the set of simple roots of Δ. If $\alpha \in \Delta^+$ we write $\alpha = \sum_i k_i\, \alpha_i$ with $k_i \in \mathbb{Z}_{\geq 0}$ and define the level of α to be $\sum_i k_i$.

We see that at level one we have the simple roots.

Definition 14.7. Let the set of simple roots $\Sigma = \{\alpha_1, \ldots, \alpha_r\}$ be a set of r elements. The Cartan Matrix A is the $r \times r$ matrix whose co-efficients are the Cartan numbers $n_{\alpha_j \alpha_i}$

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

One can now use the Cartan matrix together with the formula for the length of the α string through β to determine all the positive roots (and hence all the roots). Suppose we know the positive roots up to level $n = \sum_i k_i$ and we wish to determine them up to level $n+1$ For each root β at level n we must decide if $\beta + \alpha_i$ is a root. As we know all the roots from level one to level n we know how far back the α_i string through β goes; $\beta,\, \beta - \alpha_i, \ldots, \beta - p\alpha_i$. From this we compute how far forward the string can go

$$p - q = 2\frac{\langle \beta,\, \alpha_i \rangle}{\langle \alpha_i,\, \alpha_i \rangle} = \sum_j 2k_j \frac{\langle \alpha_j,\, \alpha_i \rangle}{\langle \alpha_i,\, \alpha_i \rangle} = \sum_j k_j\, A_{ji}$$

In particular $\beta + \alpha_i \in \Delta^+$ if $q = p - \sum_j k_j\, A_{ji} > 0$.

It is useful to have an algorithm which keeps track of the r quantities $\sum_j k_j\, A_{ji}$ for each root as we determine it. We may achieve this by adding to the r quantities the j-th row of the Cartan matrix whenever α_j is added to a root to form a new root.

One should also recall that any α_i string through α_j has $p = 0$ in the length formula since we have already proved that $\alpha_i - \alpha_j$ is not a root for $\alpha_i,\, \alpha_j \in \Sigma$.

Proposition 14.2. Given a positive root $\alpha \in \Delta^+$ of level n there exists a $\beta \in \Delta^+$ at level $n-1$ and a simple root $\alpha_i \in \Sigma$ such that $\alpha = \beta + \alpha_i$.

Proof: Let $0 \neq \alpha = \sum_i n_i\, \alpha_i$ with $\sum_i n_i = n$. Then $\langle \alpha,\, \alpha \rangle = \sum_i n_i \langle \alpha,\, \alpha_i \rangle > 0$. Now as the $n_i \geq 0$ there exists a simple root, say α_k, with $\langle \alpha,\, \alpha_k \rangle > 0$ and $n_k \geq 1$. Thus, since for the α_k string through α p is positive since $\langle \alpha,\, \alpha_k \rangle > 0$ positive so that $\alpha - \alpha_k$ is a root. It is a positive root as $\alpha - \alpha_k = \sum_{i=1}^{k-1} n_i \alpha_i + (n_k - 1)\alpha_k + \sum_{i=k+1}^{r} n_i \alpha_i$ and all the $n_i \geq 0$ while $n_k \geq 1$ (so that it is a positive sum of simple roots). But $\alpha - \alpha_k$ has level $n-1$ and $\alpha = (\alpha - \alpha_k) + \alpha_k$. □

The proposition tells us that as we build up the positive roots level by level, if we come across an empty level $n \geq 3$ then we may as well stop as the next one will also be empty.

14.4 All Rank 2 Semi-Simple Lie Algebras

We make use of Table 14.1 and the results of the previous section to classify all the rank 2 semi-simple Lie algebras. As the rank is 2 the number of simple roots is also 2 and we can read off everything we need to know from the table. First we recall that as the difference of two simple roots $\alpha_i - \alpha_j$ is not a root that any Cartan matrix with both off diagonal entries positive is not allowed. Here I will use the notation α_1 and α_2, in keeping with the fact that these are simple roots (rather than α and β). The strategy throughout this section will be to start at level 1, where we have α_1 and α_2 and then decide if we are allowed to add another α_i to this to reach level 2 and so on until a level is empty. Furthermore, as we noted above, the α_j strings through the α_i satisfy

$$q_j = -2.\frac{\langle \alpha_j,\, \alpha_i \rangle}{\langle \alpha_j,\, \alpha_j \rangle} \geq 0 \quad i \neq j$$

that is $p = 0$ otherwise would have the difference of two simple roots.

Example 14.1. The simplest Cartan matrix arises when α_1 and α_2 are orthogonal

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

This tells us that at level one we have α_1 and α_2 but there is nothing at level two as

$$2.\frac{\langle \alpha_1, \alpha_2 \rangle}{\langle \alpha_1, \alpha_1 \rangle} = -q_1 = 0, \text{ i.e. } \alpha_i(h_{\alpha_j}) = 2\delta_{ij}$$

Consequently, the roots are $\Delta = \{\pm\alpha_1, \pm\alpha_2\}$ and the Lie algebra, as the generators do not mix, is that of $\mathfrak{sl}(2, \mathbb{C}) \oplus \mathfrak{sl}(2, \mathbb{C})$.

Example 14.2. The next simplest case is where the Cartan matrix is

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

Start with α_1 and ask if adding α_2 produces a root at the second level. That is we ask about the α_2 string through α_1; recall $p = 0$ so we have that

$$q_2 = -2.\frac{\langle \alpha_1, \alpha_2 \rangle}{\langle \alpha_2, \alpha_2 \rangle} = 1$$

so $\alpha_1 + \alpha_2 \in \Delta^+$. Now we check level three. We cannot have $\alpha_1 + 2\alpha_2$ since $q_2 = 1$. Likewise, we could not increase the number of α_2 so there are no roots at level 3. The roots are, therefore, $\Delta = \{\pm\alpha_1, \pm\alpha_2, \pm\alpha_3 = \pm(\alpha_1 + \alpha_2)\}$. This is the Lie algebra of $\mathfrak{sl}(3, \mathbb{C})$.

Example 14.3. The next possibility is to have as a Cartan matrix

$$A = \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}$$

From this we read that

$$q_1 = -2.\frac{\langle \alpha_1, \alpha_2 \rangle}{\langle \alpha_1, \alpha_1 \rangle} = 2, \quad q_2 = -2.\frac{\langle \alpha_1, \alpha_2 \rangle}{\langle \alpha_2, \alpha_2 \rangle} = 1$$

At level 1 we have α_1 and α_2. At level 2 we have $\alpha_1 + \alpha_2$ (we are never allowed 2α for any root). At level 3 we could have $2\alpha_1 + \alpha_2$ and $\alpha_1 + 2\alpha_2$. However, the α_2 string through α_1 only goes to $\alpha_1 + \alpha_2$ and not to $\alpha_1 + 2\alpha_2$, as $q_2 = 1$. Hence at level 3 we only have $2\alpha_1 + \alpha_2$. At level 4 the only possibility is $3\alpha_1 + \alpha_2$ (as usual $2(\alpha_1 + \alpha_2)$ is not a root), but this is not allowed as $q_1 = 2$ and the α_1 string through α_2 does not extend past $2\alpha_1 + \alpha_2$.

The root space is then $\Delta = \{\pm\alpha_1, \pm\alpha_2, \pm(\alpha_1 + \alpha_2), \pm(2\alpha_1 + \alpha_2)\}\}$ and the dimension of this Lie algebra is 10 ($= 8 + 2$, the last 2 for the Cartan subalgebra). This is the Lie algebra $\mathfrak{so}(5, \mathbb{C})$.

Example 14.4. The last possibility for the Cartan matrix is

$$A = \begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix} \tag{14.4.1}$$

At level 1 we have α_1 and α_2. At level 2 we have $\alpha_1 + \alpha_2$. At level 3 $\alpha_1 + 2\alpha_2$ is not

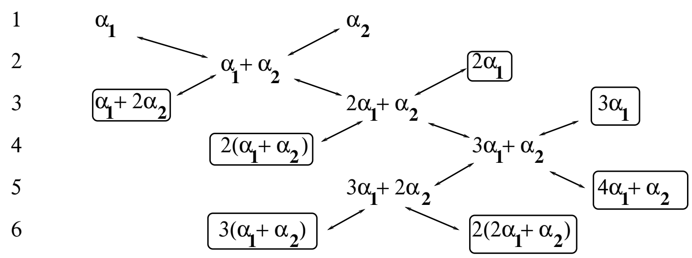


Figure 14.1: The positive root system of the Lie Algebra with Cartan Matrix (14.4.1). The North-East lines are α_1 strings while the North-West lines are α_2 strings.

allowed as $q_2 = 1$ and so we only have $2\alpha_1 + \alpha_2$. At level 4 we can have $3\alpha_1 + \alpha_2$ as $q_1 = 3$. At level 5 there are two possibilities namely

$$4\alpha_1 + \alpha_2 \quad \text{and} \quad 3\alpha_1 + 2\alpha_2$$

We cannot have $4\alpha_1 + \alpha_2$ as the α_1 string through α_2 as the α_1 string does not extend that far($q_1 = 3$) and it cannot be an α_2 string through $4\alpha_1$ as $4\alpha_1$ is not a root. However, $3\alpha_1 + 2\alpha_2$ can be thought of as an α_2 string through $3\alpha_1 + \alpha_2$ and we should check to see if it is allowed. Then

$$p - q = 2\,.\,\frac{\langle \alpha_2,\, 3\alpha_1 + \alpha_2 \rangle}{\langle \alpha_2,\, \alpha_2 \rangle} = -1$$

notice that $p = 0$ as otherwise the string would include $3\alpha_1$ which is not a root, so that $q = 1$ which establishes that $3\alpha_1 + \alpha_2$ is an allowed root. Now we pass to level 6. From level 5 the two possibilities are $4\alpha_1 + 2\alpha_2 = 2(2\alpha_1 + \alpha_2)$ is twice a root so is itself not a root, and $3(\alpha_1 + \alpha_2)$ is 3 times a root so not a root, whence level 6 is empty.

The root space is $\Delta = \{\pm\alpha_1,\, \pm\alpha_2,\, \pm(\alpha_1+\alpha_2),\, \pm(2\alpha_1+\alpha_2),\, \pm(3\alpha_1+\alpha_2), \pm(3\alpha_1+2\alpha_2)\}$. This Lie algebra is 14 dimensional and does not belong to any of the classical groups. It is known as $\mathfrak{g}_2$ and is one of a number of 'exceptional' Lie algebras.

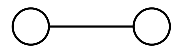
14.5 Dynkin Diagrams

There is yet another way of expressing the information that we have about the allowed Lie algebras and that is through a pictorial device known as a Dynkin diagram.

Definition 14.8. Let Δ be the set of roots of a Lie algebra $\mathfrak{g}$ with simple roots $\Sigma = \{\alpha_1,\, \ldots,\, \alpha_r\}$. The associated Dynkin Diagram is defined as follows

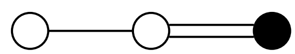
1. Draw a circle for each $\alpha_i \in \Sigma$.
2. Connect the circles representing α_i and α_j by $A_{ij}.A_{ji}$ lines ($= n_{\alpha_i\alpha_j}.n_{\alpha_j\alpha_i}$ and there is no sum over i or j).
3. If $A_{ij}.A_{ji} > 1$ make the circle of the shorter root solid:

Example 14.5. For $\mathfrak{sl}(3, \mathbb{C})$ we only have the two simple roots and $A_{12}.A_{21} = 1$ so we join them with one line and both circles remain clear (not solid).



One of the appealing features of the Dynkin diagrams is that one may reconstruct the Cartan matrix from them. From the Dynkin diagram itself we recover $A_{ij}.A_{ji}$, by Table 14.1 we know that $A_{ij} = 0, \pm1, \pm2, \pm3, \pm4$ furthermore the Dynkin diagram tells us which is the shorter root and so we are able to determine A_{ij}.

Example 14.6. Consider the following Dynkin diagram:



As there are 3 circles, hence 3 simple roots, we have a 3×3 Cartan matrix. We know that all the diagonal entries are 2, $A_{11} = A_{22} = A_{33} = 2$. The first and third simple roots are not connected so that $A_{13} = A_{31} = 0$. The first and second simple roots are connected by 1 line so that we have $A_{12} = A_{21} = -1$. The second and third simple roots are connected by 2 lines so that $A_{23}.A_{32} = 2$ the last circle is filled so that it is the shorter root whence $A_{23} = -2$ and $A_{32} = -1$. Hence, the Cartan matrix is

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

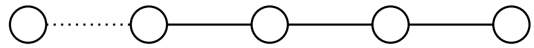
We now determine the Dynkin diagrams for all the classical Lie algebras.

$\mathbf{A}_n = \mathfrak{sl}(n+1, \mathbb{C})$

The roots are $\Delta = \{\pm(\omega_i - \omega_j) | 1 \le i < j \le n+1\}$. The regular element in $\mathfrak{h}$ is chosen to be $h_0 = \mathrm{diag}(h_1, \dots, h_{n+1})$ with $h_1 > \cdots > h_{n+1}$ and $\sum_{i=1}^{n+1} h_i = 0$. The positive roots are, therefore, $\Delta^+ = \{\omega_i - \omega_j | 1 \le i < j \le n+1\}$ and the simple roots are $\Sigma = \{\omega_i - \omega_{i+1} | 1 \le i \le n\}$. We already know that

$$\langle \omega_i - \omega_{i+1}, \omega_j - \omega_{j+1} \rangle = (\omega_i - \omega_{i+1})(e_j - e_{j+1}) = 2\delta_{ij} - \delta_{ij+1} - \delta_{i+1j}$$

This tells us that only adjacent simple roots are connected and that $A_{ii+1}.A_{i+1i} = 1$. Thus the Dynkin diagram is:

A_n 

$$\mathbf{B}_n = \mathfrak{so}(2n+1,\, \mathbb{C})$$

The roots are $\Delta = \{\pm\omega_i;\, \pm(\omega_i+\omega_j),\, \pm(\omega_i-\omega_j)|\, i<j\}$. We take the regular element in $\mathfrak{h}$ to be $h_0 = \mathrm{diag}(h_1 \otimes \boldsymbol{\epsilon}, \ldots, h_n \otimes \boldsymbol{\epsilon}, 0) = \sum_{i=1}^{n} h_i F_{2i-1\,2i}$ (the 2×2 blocks down the diagonal with the last entry being a 1×1 zero) with $h_1 > h_2 > \cdots > h_n$. With respect to this regular element the positive roots are $\Delta^+ = \{\omega_i;\, \omega_i+\omega_j,\, \omega_i-\omega_j|\, i<j\}$ while the simple roots become $\Sigma = \{\omega_1-\omega_2,\, \omega_2-\omega_3, \ldots, \omega_{n-1}-\omega_n,\, \omega_n\}$.

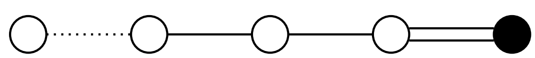
Exercise 14.3. Write $\Sigma = \{\alpha_1, \ldots, \alpha_n\}$ with $\alpha_i = \omega_i - \omega_{i+1}$ for $1 \le i \le n-1$ while $\alpha_n = \omega_n$. Then it is easy to see that

$$\langle \alpha_i, \alpha_i \rangle = 2,\; i = 1, \ldots, n-1, \quad \langle \alpha_i, \alpha_i \rangle = 1$$

so the last circle in a Dynkin diagram should be solid as its associated simple root is shorter than all the others. Furthermore, show that

$$\langle \alpha_i, \alpha_{i+1} \rangle = -1,\; i = 1, \ldots, n$$

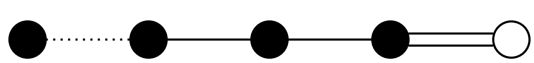
so that $A_{ii+1}.A_{i+1i} = 1$ for $i = 1, \ldots, n-1$ and $A_{nn-1}.A_{n-1n} = 2$. Hence the Dynkin diagram is,

B_n 

$$\mathbf{C}_n = \mathfrak{sp}(n,\, \mathbb{C})$$

The roots are $\Delta = \{\pm 2\omega_i;\, \pm(\omega_i+\omega_j),\, \pm(\omega_i-\omega_j)|\, i<j\}$ and the regular element in $\mathfrak{h}$ is taken to be $h_0 = \mathrm{diag}(h_1, \ldots, h_n, -h_1, \ldots, -h_n)$ with $h_1 > h_2 \cdots > h_n$. In this case the positive roots are $\Delta^+ = \{2\omega_i;\, \omega_i+\omega_j,\, \omega_i-\omega_j|\, i<j\}$ with simple roots $\Sigma = \{\omega_1-\omega_2,\, \omega_2-\omega_3, \ldots, \omega_{n-1}-\omega_n,\, 2\omega_n\}$.

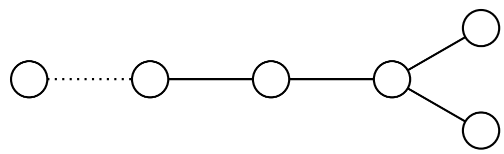
The Dynkin diagram is therefore:

C_n 

$$\mathbf{D}_n = \mathfrak{so}(2n,\, \mathbb{C})$$

The roots are $\Delta = \{\pm(\omega_i + \omega_j),\, \pm(\omega_i - \omega_j) |\, i < j\}$ and the regular element in $\mathfrak{h}$ is taken to be $h_0 = \mathrm{diag}(h_1, \ldots, h_n) \otimes \boldsymbol{\epsilon}$ with $h_1 > h_2 \cdots > h_n$. The positive roots are then $\Delta^+ = \{\omega_i + \omega_j,\, \omega_i - \omega_j |\, i < j\}$ with simple roots $\Sigma = \{\omega_1 - \omega_2,\, \omega_2 - \omega_3, \ldots, \omega_{n-1} - \omega_n,\, \omega_{n-1} + \omega_n\}$. Whence the Dynkin diagram is:

D_n



The general analysis of the conditions on a Dynkin diagram are straightforward though somewhat tedious. So we will skip it. Rather we quote some results.

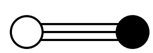
Proposition 14.3. A Lie algebra is simple iff its associated Dynkin diagram is connected.

Corollary 14.1. The Lie algebras $\mathfrak{sl}(n, \mathbb{C})$, $\mathfrak{sp}(n, \mathbb{C})$ and, for $n \neq 2, 4$ $\mathfrak{so}(n, \mathbb{C})$ are simple.

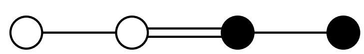
Exercise 14.4. By looking at the Dynkin diagram of $\mathfrak{so}(4, \mathbb{C})$ show that is not simple. As it is semi-simple $\mathfrak{so}(4, \mathbb{C})$ is the direct sum of simple Lie algebras. What are the summands?

The Dynkin diagrams that are allowed by all the constraints imposed on the $n_{\alpha\beta}$ are those of the classical Lie algebras as well as the following **exceptional** Lie algebras (we have not proven that they exist as Lie algebras- but they do). Note that the subscript refers to the rank of the Lie algebra (and hence to the number of circles in the Dynkin diagram)

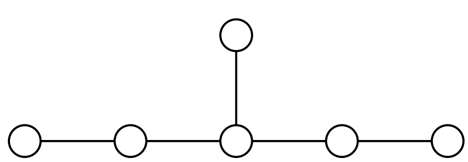
G_2



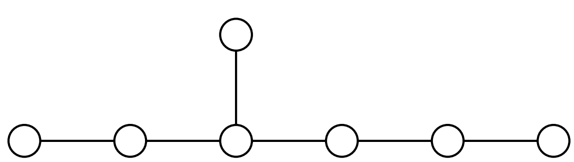
F_4



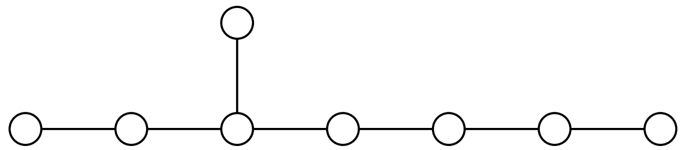
E_6



E_7



E_8



Exercise 14.5. Determine the Cartan matrix of F_4

Proposition 14.4. Two simple complex Lie algebras are isomorphic iff their Dynkin diagrams are the same.

By this last proposition we see that the Lie algebras of $\mathfrak{sl}(2, \mathbb{C})$, $\mathfrak{so}(3, \mathbb{C})$ and $\mathfrak{sp}(1, \mathbb{C})$ are all isomorphic. Likewise $\mathfrak{so}(5, \mathbb{C})$ and $\mathfrak{sp}(2, \mathbb{C})$ are isomorphic as are $\mathfrak{sl}(4, \mathbb{C})$ and $\mathfrak{so}(6, \mathbb{C})$. All of these isomorphisms we had seen before for the real forms of these Lie algebras.

APPENDIX A

YOUNG TABLEAU FOR $SU(N)$ AND $SO(N)$

Here I write down rules for dimensions and the decomposition of tensor products without any proof. There is some motivation though.

Consequently each irreducible tensor with particular symmetry properties is in one to one correspondence with a Young tableau. Instead of thinking of the tensor we may as well just consider the corresponding tableau.

Consider a Young tableau with r_i boxes in the $i-$th row so that the total number of boxes is $r = \sum_i r_i$. In each box we place an index (corresponding to one of the indices of the tensor under consideration). We label the indices by their position in each row so that i_{r_j+k} is the index that belongs in the $(j+1)-$th row and $k-$th column.

Definition A.1. A standard tableau

The rules are

1. One first symmetrizes over all indices in the same row
2. This is followed by anti-symmetrising inices in each column
3.

A.1 Dimension Formulae for $SU(n)$

A.2 Decomposing Tensor products for $SU(n)$

A.3 Dimension Formulae for $SO(n)$

A.4 Decomposing Tensor products for $SO(n)$

APPENDIX B

SPINOR REPRESENTATIONS

B.1 General Relativity versus Special Relativity and Spinors

In general relativity a spinor $\psi_\alpha(x)$ transforms as a scalar under general coordinate transformations,

$$\psi'_\alpha(x') = \psi_\alpha(x) \tag{B.1.1}$$

However, there are also local Lorentz transformations to consider. Under a Local Lorentz transformation the spinor transforms as

$$\psi'_\alpha(x) = \Lambda_\alpha^{\ \beta}(x)\,\psi_\beta(x) \tag{B.1.2}$$

In order to accomodate spinors in General Relativity we also need the concept of an $n-$bein. In particular the $n-$bein $e^a_\mu(x)$ transforms as

$$e'^a_\mu(x) = \Lambda^a_{\ b}(x)\,e^b_\mu(x) \tag{B.1.3}$$

under a local Lorentz transformation while under a general coordinate transformation the $n-$bein transforms as a vector

$$e'^a_\mu(x') = \frac{\partial x^\nu}{\partial x'^\mu}\,e^a_\nu(x) \tag{B.1.4}$$

In order to pass from General Relativity to Special Relativity we fix the $n-$bein and consider it to be a Kronecker delta function

$$e^a_\mu(x) \longrightarrow \delta^a_\mu \tag{B.1.5}$$

so that the metric

$$g_{\mu\nu}(x) = e^a_\mu(x)\,\eta_{ab}\,e^b_\nu(x) \longrightarrow \delta^a_\mu\,\eta_{ab}\,\delta^b_\nu = \eta_{\mu\nu} \tag{B.1.6}$$

becomes the Minkowski metric and, furthermore, we may ignore the $n-$bein altogether. Now everything should be fine, except that the transformation (B.1.1) does not look at all like how a spinor transforms in Special Relativity under a Lorentz boost

$$x'^{\mu} = \Lambda^{\mu}_{\ \nu}\, x^{\nu} \tag{B.1.7}$$

rather, (B.1.2) is closer to what we expect. What is going on?

The answer lies in the transformation (B.1.4) since if we do a Lorentz transformation (B.1.7) the $n-$bein transforms as

$$e'^{a}_{\mu}(x') = \Lambda_{\mu}^{\ \nu}\, e^{a}_{\nu}(x) \longrightarrow \Lambda_{\mu}^{\ \nu}\, \delta^{a}_{\nu} \tag{B.1.8}$$

which no longer agrees with (B.1.5). We are in luck, however, as we may always combine the Lorentz transformation with a Local Lorentz transformation (B.1.3) to have

$$e'^{a}_{\mu}(x') = \Lambda_{\mu}^{\ \nu}\, \Lambda^{a}_{\ b}\, e^{b}_{\nu}(x) \longrightarrow \Lambda_{\mu}^{\ \nu}\, \Lambda^{a}_{\ b}\, \delta^{b}_{\nu} \tag{B.1.9}$$

and we choose $\Lambda^{a}_{\ b}$ so that

$$\Lambda_{\mu}^{\ \nu}\, \Lambda^{a}_{\ b}\, \delta^{b}_{\nu} = \delta^{a}_{\mu} \tag{B.1.10}$$

and the $n-$bein does not change under this combined transformation (we can really ignore it). We see, therefore, that in order to arrive at Special Reltivity without the appearance of an $n-$bein we must 'confuse' Lorentz boosts with Local Lorentz transformations. In particular this means that the spinor transforms as in (B.1.2) with $\Lambda^{a}_{\ b} =$

BIBLIOGRAPHY

[1] R.N. Cahn, **Semi-Simple Lie Algebras and their Representations**, http://phyweb.lbl.gov/~rncahn/www/liealgebras/book.html