Group Theory For Particle Physicists

Course delivered in 2019 by

Dr. Daniele Dorigoni

Durham University

Notes taken by

Richie Dadhley

richie.s.dadhley@durham.ac.uk

Acknowledgements

These are my notes on the 2019 lecture course "Group Theory (For Particle Physicists)" taught by Dr. Daniele Dorigoni at Durham University as part of the Particles, Strings and Cosmology Msc. For reference, the course lasted 3 weeks and was lectured over 12 hours.

I have tried to correct any typos and/or mistakes I think I have noticed over the course. I have also tried to include additional information that I think supports the taught material well, which sometimes has resulted in modifying the order the material was taught. Obviously, any mistakes made because of either of these points are entirely mine and should not reflect on the taught material in any way.

I have also used Dexter Chua's notes on Professor Nick Dorey's 2016 Cambridge Part III course "Symmetries, Fields and Particles". These notes are brilliant and cover some of the material in a more mathematically rigorous way then presented here. They can be found online via

https://dec41.user.srcf.net/notes/III_M/symmetries_fields_and_particles.pdf

I would like to extend a message of thanks to Dr. Dorigoni for teaching this course brilliantly. I really enjoyed this course.

If you have any comments and/or questions please feel free to contact me via the email provided on the title page.

For a list of other notes/works I have available, visit my blog site

www.richiedadhley.com

These notes are not endorsed by Dr. Dorigoni or Durham University.

Richie Dadhley

Contents

1 Groups

1.1 Why Do We Care About Group Theory?

As we will see, symmetries are often related to groups and, as any quantum field theorist knows, symmetries are incredible important and powerful tools in physics. They allow us to massively simplify complex problems and they also reveal *a lot* about the physics. The symmetries can actually be so powerful that it allows us to solve the theory *exactly*. We refer to this as *integrability*. In fact I guess you could argue that, at least at an introductory level, QFT is the study of symmetries in Lagrangians and their corresponding conserved currents.[1]

Example 1.1.1. Electric charge is conserved in particle interactions, and so there is some symmetry in the Lagrangian that corresponds to this.

There are several symmetries that occur in Nature that we may be familiar with, here are some examples:

Symmetry	Group	Continuous Or Discrete
Rotational	SO(3)	Continuous
Lorentz	SO(3, 1)	Continuous
Gauge & Flavour	e.g, SU(3)	Continuous
Parity	$\vec{x} \longrightarrow -\vec{x}$	Discrete
Charge Conjugation	$e^- \longrightarrow e^+$	Discrete
Time Reversal	$t \longrightarrow -t$	Discrete

We have indicated whether the symmetry is a *continuous* or *discrete* symmetry. The names are reasonably self explanatory. In this course we will focus on continuous symmetries as they give rise to Lie algebras (which we will study a lot).

Remark 1.1.2. It actually turns out the neither parity, nor charge conjugation, nor time reversal are proper symmetries of the standard model of particle physics. The combination of charge and parity, known as CP, is a symmetry of the electromagnetism and QCD (strong force), and the full beast charge-parity-time, CPT, is a symmetry of the weak force.

1.2 Group Definitions

The next section is going to contain a lot of definitions, so if you are not used to reading maths notes... enjoy!

[1] If this means nothing to you, look up Noether's Theorem.

1.2.1 Generalities

Definition. [Group] A *group* G is a set $\{g\}$ equipped with a multiplication law

$$\begin{aligned} \bullet : G \times G &\to G \\ (g_1, g_2) &\mapsto g_1 \bullet g_2, \end{aligned}$$

such that:

(i) Closure; $\forall g_1, g_2 \in G$, $g_1 \bullet g_2 \in G$,

(ii) Associativity; $\forall g_1, g_2, g_3 \in G$, $g_1 \bullet (g_2 \bullet g_3) = (g_1 \bullet g_2) \bullet g_3$,

(iii) Identity; there exists a unique $e \in G$ such that $\forall g \in G$ $e \bullet g = g \bullet e = g$, and

(iv) Inverse; $\forall g \in G$ there exists a unique element $g^{-1} \in G$ such that $g^{-1} \bullet g = g \bullet g^{-1} = e$.

Definition. [Order Of A Group] Let $(G, \bullet)$ be a group. Then we call the number of elements in G the *order* of the group.

Remark 1.2.1. We call $\bullet$ a *multiplication*, however it need not multiply two elements by our common understanding of the word. For example $\bullet$ could be addition, as we will see in the examples below.

Exercise

Show that the identity and inverse are unique. *Hint: Suppose that they aren't unique and prove by contradiction.*

Definition. [Subgroup] Let $(G, \bullet)$ be a group and let $H \subset G$ be a subset. Then H is a *subgroup* $(H, \bullet)$ is itself a group.

Remark 1.2.2. Note by the uniqueness of the identity, if $H \subset G$ is to be a subgroup, it must contain e.

Definition. [Abelian Group] Let $(G, \bullet)$ be a group. We say that it is *abelian* if, for all $g_1, g_2 \in G$

$$g_1 \bullet g_2 = g_2 \bullet g_1.$$

Example 1.2.3. The real numbers equipped with addition, $(\mathbb{R}, +)$ form a continuous, abelian group. The identity is simply $0 \in \mathbb{R}$ and the inverse of $a \in \mathbb{R}$ is $-a \in \mathbb{R}$. Associativity and closure should be easy to see from every day use. The order is infinite.

Example 1.2.4. The real numbers, *excluding the origin*, equipped with multiplication, $(\mathbb{R}^*, \times)$, forms a continuous, abelian group. Again closure and associativity should be familiar. The identity is simply $1 \in \mathbb{R}^*$ and the inverse of $a \in \mathbb{R}^*$ is $\frac{1}{a} \in \mathbb{R}^*$. It is because inverse condition that we need to exclude the origin. It might not seem obvious at first that this group is continuous as we have removed the origin. However simply consider redefining all the elements as $a \longrightarrow 1/a$, then the origin is taken all the way to infinity and we're happy. The order is infinite.

Example 1.2.5. The set of integers modulo n equipped with addition, $(\mathbb{Z}_n, +_n)$,[2] form a discrete, abelian group. We can show closure and associativity easily given that we know $(\mathbb{Z}, +)$ is a group. We simply define the addition $+_n$ by

$$[a] +_n [b] := [a+b],$$

where the addition on the right-hand side is the addition on $\mathbb{Z}$. We need to show this is well defined: our equivalence relation is given by

$$a' \sim a \qquad \Longleftrightarrow \qquad a' - a = An$$

where $A \in \mathbb{Z}$ (i.e. a' and a differ by an integer multiple of n). So let's consider $a' = a + An$ and $b' = b + Bn$ where $A, B \in \mathbb{Z}$, so $[a'] = [a]$ and $[b'] = [b]$. Then we have

$$\begin{aligned} [a'] +_n [b'] &:= [a' + b'] \\ &= [(a + An) + (b + Bn)] \\ &= [a + An + b + Bn] \\ &= [(a+b) + (A+B)n] \\ &= [a+b], \end{aligned}$$

where we have used the associativity and abelian nature of $(\mathbb{Z}, +)$ and that $(A+B) \in \mathbb{Z}$ so $[(a+b)+(A+B)n] = [a+b]$. This shows our definition is well defined. We then inherit all the group properties from $(\mathbb{Z}, +)$. In particular, the identity is $[0] \in \mathbb{Z}_n$ and the inverse of $[a] \in \mathbb{Z}_n$ is $[-a] \in \mathbb{Z}_n$. The order is n, as any integers greater than $n-1$ or less than 0 are equivalent to one in the set $\{0, ..., n-1\}$.

Example 1.2.6. The permutation of n elements, denoted S_n, forms a discrete group. I am not going to prove this one here, but set it as an exercise below. The order is $n!$

Exercise

Prove that the above example is true. *Hint: You can prove this by writing a permutation as*

$$\sigma = \begin{pmatrix} a_1 & a_2 & ... & a_n \\ \sigma(a_1) & \sigma(a_2) & ... & \sigma(a_n) \end{pmatrix},$$

and then making a bijective argument.

Any mathematicians reading this will ask the obvious question of "what is the structure preserving map?" That is, what map makes two different groups 'look the same'? The answer is the following definition.

[2] I have put a subscript n on here because technically this addition is different to the additional on integers (it adds equivalence classes). I will use the notation of equivalence relations (square brackets etc) in the proof. If this is not familiar to you, don't worry its not needed to understand the course. However they are useful in maths so I encourage you to read up on them.

Definition. [Group Isomorphism] Let $(G, \bullet)$ and $(H, \circ)$ be two groups. Then if there exists a bijective map $\phi : G \to H$ satisfying

$$\phi(g_1 \bullet g_2) = \phi(g_1) \circ \phi(g_2)$$

for all $g_1, g_2 \in G$, then we call ϕ a *group isomorphism*. We say that the groups $(G, \bullet)$ and $(H, \circ)$ are (group) isomorphic, which we write as $G \cong_{\text{grp}} H$.[3]

In this course we are going to be interested mostly in *matrix* groups, where the multiplication law is simply matrix multiplication. Note that in general these will *not* be abelian. In any proofs that follow, we will assume associativity holds to save time. We will also drop the $\bullet$ basically everywhere and just assume matrix multiplication is implicit.

1.2.2 Matrix Groups

For the non-maths people, unfortunately we are not done with the definitions: we need to define some matrix sets that will appear *a lot*. Even though we are essentially just defining the *sets* below, we will call them the so-and-so group. This is just because they always pop up as matrix groups, so we may as well call them that. The proofs just show that the group conditions meets the restrictions on the set.

Definition. [General Linear Group] The *general linear* over $\mathbb{R}$ group is the matrix group with set[4]

$$GL(n, \mathbb{R}) := \big\{A \in M^{\mathbb{R}}_{n\times n} \mid \det A \neq 0\big\}.$$

We can similarly define $GL(n, \mathbb{C})$. The $\det A \neq 0$ condition is needed so that A is invertible (which is the inverse of A in the group).

Proof. First note that the identity is the $n \times n$ identity matrix $\mathbb{1}_n$ which has $\det \mathbb{1}_n = 1 \neq 0$. Next recall the relation

$$\det(AB) = \det(A)\det(B).$$

So if $\det A, \det B \neq 0$ then $\det(AB) \neq 0$. Using the above relation we also have

$$1 = \det \mathbb{1}_n = \det\big(AA^{-1}\big) = \det(A)\det\big(A^{-1}\big),$$

and so $\det A^{-1} = (\det A)^{-1} \neq 0$. $\square$

Using the fact that $\det \mathbb{1}_n = 1$, and with Remark 1.2.2 in mind, we can define the following matrix group.

Definition. [Special Linear Group] The *special*[5] *linear* group over $\mathbb{R}$ is the matrix group with set

$$SL(n, \mathbb{R}) := \big\{A \in GL(n, \mathbb{R}) \mid \det A = 1\big\}.$$

This is a subgroup of $\mathrm{GL}(n, \mathbb{R})$. Again we can define $\mathrm{SL}(n, \mathbb{C})$ similarly.

Proof. Everything follows through as with the above proof but with $= 1$ everywhere. $\square$

Definition. [Orthogonal Group] The *orthogonal* group is the matrix group with set

$$O(n) := \{A \in GL(n,\mathbb{R}) \,|\, AA^T = A^T A = \mathbb{1}_n\}.$$

Proof. The identity and inverse clearly obey the transpose condition. So just need to show it for closure:

$$(AB)^T(AB) = B^T A^T AB = B^T \mathbb{1}_n B = B^T B = \mathbb{1}_n,$$

where we have used $(AB)^T = B^T A^T$ and the definition of the identity element in a group. Note, using

$$\det A^T = \det A,$$

we have

$$1 = \det \mathbb{1}_n = \det\big(A^T A\big) = \det A^T \det A = (\det A)^2,$$

so $\det A = \pm 1$ for $A \in O(n)$. □

With footnote 4 in mind, we have the following definition.

Definition. [Special Orthogonal Group] The *special orthogonal* group is the matrix group with set

$$SO(n) := \{A \in O(n) \,|\, \det A = 1\}. \tag{1.1}$$

Proof. We've basically done all the work for this. □

Let's give an example of $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ in order to highlight their difference.

Example 1.2.7. Let $n = 2$ then

$$O(2) = \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cup \begin{pmatrix} -\cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \,\Big|\, \theta \in [0, 2\pi) \right\},$$

and

$$SO(2) = \left\{ \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \,\Big|\, \theta \in [0, 2\pi) \right\}.$$

Now consider the actions on these matrices on a general vector in $\mathbb{R}^2$ with $\theta = \pi/2$

$$\begin{pmatrix} \cos\pi/2 & \sin\pi/2 \\ -\sin\pi/2 & \cos\pi/2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ -x \end{pmatrix},$$

which is an anticlockwise rotation by $\pi/2$. We also have

$$\begin{pmatrix} -\cos\pi/2 & \sin\pi/2 \\ \sin\pi/2 & \cos\pi/2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ x \end{pmatrix},$$

which is a reflection about the $x = y$ axis. Indeed $\mathrm{SO}(n)$ is the set of rotations in n-dimensions and $\mathrm{O}(n)$ is both rotations and reflections. The rotations have $\det A = 1$ and reflections $\det A = -1$.

Unlike with the linear groups, we can't extend the definition of the orthogonal groups to the complex numbers. This is because for complex matrices we need to take the *Hermitian conjugate* instead of transpose. These groups are significantly different that we give them separate names.

Definition. [Unitary Group] The *unitary* group is the matrix group with set

$$U(n) := \big\{ A \in GL(n, \mathbb{C}) \,|\, UU^\dagger = U^\dagger U = \mathbb{1}_n, \big\}$$

Proof. The proof for this is basically identical to $\mathrm{O}(n)$ but now we get

$$\det A = e^{i\alpha}, \qquad \alpha \in [0, 2\pi).$$

□

Of course we also have the special case.

Definition. [Special Unitary Group] The *special unitary* group is the matrix group with set

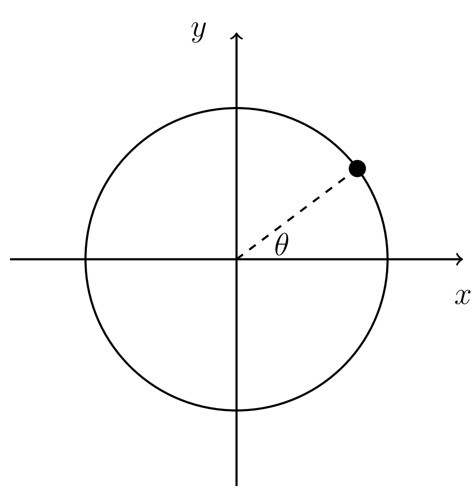
$$SU(n) := \big\{ A \in U(n) \,|\, \det A = 1 \big\}. \tag{1.2}$$

Proof. Again basically done. □

1.3 Lie Groups

All of the above matrix groups are examples of what are known as *Lie groups*. This course is predominantly the study of the Lie groups $\mathrm{SO}(n)$ and $\mathrm{SU}(n)$.

So what is a Lie group? Well we have seen (or at least just said) that the groups we are interested in are continuous. We can therefore think of them as some kind of continuous geometric shape, with each point on the shape corresponding to an element in the group. For example, we could maybe think of $\mathrm{SO}(2)$ as a circle in the xy-plane and identify the elements of $\mathrm{SO}(2)$ by the angle from the x-axis.



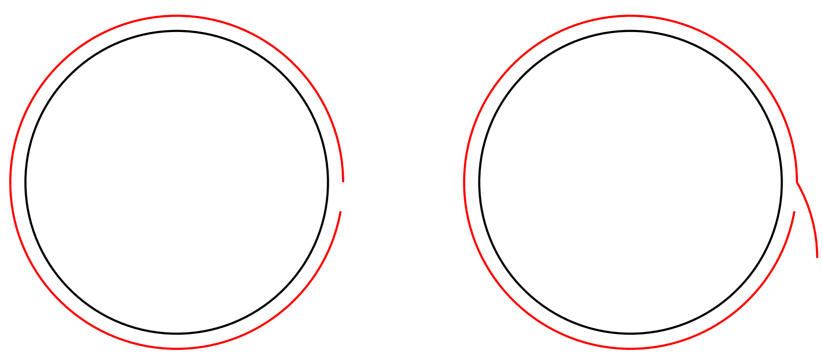
Now we note that in order to do this, we have to 'project' our circle down onto the xy-plane (because the dashed line does not lie on the circle itself). To those familiar with general relativity, this looks a awful lot like a manifold, and indeed that's exactly what it is. For those not familiar, I shall provide a brief explanation of what a manifold is now. If you don't follow this, see any differential geometry textbook, or my notes on Dr. Schuller's GR course.

1.3.1 Manifolds (In a Nutshell)

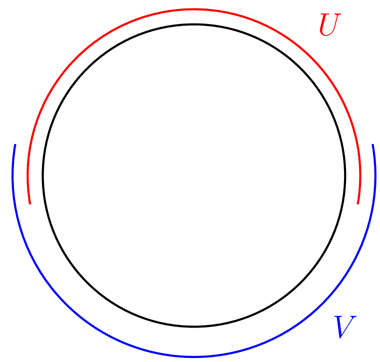
Definition. [Manifold] A *manifold* is the triple $(\mathcal{M}, \mathcal{O}, \mathcal{A})$, where $\mathcal{M}$ is a set, $\mathcal{O}$ is a *topology*,[6] and $\mathcal{A}$ is an *atlas*. The atlas is a collection of doublets (U, φ_U) where $U \in \mathcal{O}$ is an open set in $\mathcal{M}$ and $\varphi_U : U \to \mathbb{R}^n$ is an injective[7] map onto the plane. The dimension of the manifold is n. If $U, V \in \mathcal{O}$ are overlapping sets, i.e. $U \cap V \neq \emptyset$, then we can place conditions on the maps $\varphi_V \circ \varphi_U^{-1} : \varphi(U) \to \varphi(V)$ (i.e. they are maps from open subsets of $\mathbb{R}^n$ to open subsets of $\mathbb{R}^n$) in order to give the manifold itself some properties.

Ok that above definition will mean almost nothing to someone who doesn't know what it already means so let's given an example using the circle above.

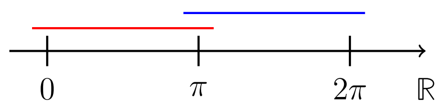
Example 1.3.1. Our set $\mathcal{M}$ is the points on the circle. We now need a topology, $\mathcal{O}$. These need to be open (i.e. we cannot take a 'hard cut' of the circle), and every point of the circle must be in at least one element of the topology. This means we need at least two elements in our topology. Why? Well consider using just one element. We either don't cover the whole circle (left in diagram below) or we cover the same point twice in one element (right below):



The left one is obviously a problem because there's a bit missing. The right one is a problem because we wanted our maps φ_U to be injective, but now we have two points in U that will be mapped to the same point, so its not injective. We therefore need something like the following diagram:



We then define our maps $\varphi_{U/V}$ to the real line $\mathbb{R}$ as shown diagrammatically below:



We obviously require that $\varphi_U(U \cap V) = \varphi_V(U \cap V)$, i.e. we get the same value on the $\mathbb{R}$ line where where the red and blue lines overlap.

We said at the end of the definition above that we can put constraints on the *transition maps* $\varphi_V \circ \varphi_U^{-1}$ and get conditions on the manifold itself. The questions is "what kind of constraints do we use?" Well in the above we have actually already assumed that these transition maps are continuous, but there is a stronger constraint we can imply known as *smoothness*. This is essentially the condition that all the transition maps are infinitely differentiable with continuous result. We call such manifolds *smooth* (or C^∞). By doing this we can talk about maps $f : \mathcal{M} \to \mathcal{M}$ themselves as being smooth by projecting them down into the charts and studying them there.[8] These are hugely important constructions in GR and will be the type of manifold we consider here.

1.3.2 Back To Lie Groups

Ok so we know (or at least know where to learn) what a smooth manifold is, so we can now define a Lie group.

> **Definition.** [Lie Group] A *Lie group* is a continuous group $(G, \bullet)$ whose underlying set is a smooth manifold and where the multiplication map, $\bullet : G \times G \to G$, and inverse map, $i : G \to G$, defined by $i(g) = g^{-1}$, are smooth.

It is clear that we need the groups to be continuous otherwise we wouldn't be able to get a manifold (i.e. our shape wouldn't connect up and so our open sets would be a problem — what is an open set of a single point?)

[8]For a much more detailed, and probably better, explanation see my notes on Dr. Schuller's GR course. Link above.

> **Definition.** [Dimension Of Lie Group] The *dimension* of the Lie group is given by the dimension of the manifold.

> **Definition.** [Lie Subgroup] Let $(G, \bullet)$ be a Lie group and let $(H, \bullet)$ be a subgroup. Then we say $(H, \bullet)$ is a *Lie subgroup* if it is also a Lie group under the restriction of the maps to H.

The mathematicians will now again ask "what are the structure preserving maps?" The answer is again given by the following definition.

> **Definition.** [Lie Group Isomorphism] Let $(G, \bullet)$ and $(H, \circ)$ be two Lie groups that are isomorphic as groups, i.e. $G \cong_{\text{grp}} H$, via the group isomorphism $\phi : G \to H$. If ϕ is also a diffeomorphism (that is it is smooth and its inverse is also smooth) then we say that the Lie groups are isomorphic.

Claim 1.3.2. We claim (without proof) that our lovely matrix groups above are Lie groups. Their dimensions are given by the number of free parameters in the matrix. For example $\text{SO}(n)$ is a $\frac{n(n+1)}{2}$ dimensional Lie group.

Notation. From now on I am very likely to drop the multiplication when writing a group. As in I will call G a (Lie) group without specifying the multiplication.

1.4 Lie Algebras

Lie groups have an associated structure known as a *Lie algebra*. It turns out that a lot of the useful information about a Lie group can be found by studying its associated Lie algebra, it also turns out that the Lie algebra is easier to study. The Lie algebra of a Lie group, G, is the *tangent space to the identity* $e \in G$. To those unfamiliar with GR, a tangent space to a point $p \in \mathcal{M}$ is basically the plane that 'kisses' the manifold at p. We can also think of the Lie algebra as the points infinitesimally close to the identity.

As we just said, we can think of the Lie algebra as the elements infinitesimally close to the identity. Now recall that the Taylor expansion of an exponential of a matrix is

$$e^{\epsilon M} = \mathbb{1} + \epsilon M + \frac{1}{2}\epsilon^2 M^2 + ...,$$

so if we consider ϵ to be some small continuous parameter, we can drop $\mathcal{O}(\epsilon^2)$ terms and obtain

$$e^{\epsilon M} \approx \mathbb{1} + \epsilon M.$$

This is an infinitesimal relation near the identity. So we can get the Lie algebra of a Lie group by taking the exponential of the matrix.[9]

We will give the definition of a Lie algebra in a minute, but first let's study via an example.

[9]Of course not all Lie groups are matrices, but as we said above, in these notes we are basically only considering matrix groups.

1.4.1 Example SO(3)

Consider the Lie group SO(3). As we said before SO(n) are rotations in n-dimensions. Let's consider explicitly a rotation around the z-axis[10] by angle φ_z. The matrix is given explicitly as

$$R_z(\varphi_z) = \begin{pmatrix} \cos\varphi_z & \sin\varphi_z & 0 \\ -\sin\varphi_z & \cos\varphi_z & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Now consider (as if by magic) the matrix

$$T_z := \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then we can show (exercise below) that

$$R_Z(\varphi_z) = e^{\varphi_z T_Z}. \tag{1.3}$$

Exercise

Prove Equation (1.3). *Hint: Find out the general formula for* $(T_z)^n$ *by considering the first few values of* n*. Then Taylor expand the exponential and compare the Taylor expansions of* $\cos\theta$ *and* $\sin\theta$*.*

We can show similar relations for rotations about the x and y axes by angles φ_x and φ_y, respectively, with

$$T_x := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \text{and} \quad T_y := \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

Remark 1.4.1. Note that $T_x, T_y, T_Z \notin SO(3)$! In fact they are all antisymmetric, i.e. they obey $(T_{x/y/z})^T = -T_{x/y/z}$.

Now a general rotation in $\mathbb{R}^3$ can be written as

$$R_x(\varphi_x)R_y(\varphi_y)R_z(\varphi_z) = e^{\varphi_x T_x}e^{\varphi_y T_y}e^{\varphi_z T_z}. \tag{1.4}$$

This is nice, but what we really want is the right-hand side to be a single exponential instead of the product of three! So how do we do this? Well we introduce the Baker-Campbell-Haussdorff (BCH) formula.

Proposition 1.4.2 (BCH Formula)**.** *For any two matrices the following formula holds:*

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]+\dots}, \tag{1.5}$$

where

$$[A,B] = AB - BA$$

is the commutator between matrices. The ... terms on the right hand side are all made up of the commutators.

[10]Note here we're using one of our coordinate systems to define what we mean by z.

Proof. We shall prove that Equation (1.5) holds up the the given terms. Insert a 'book keeping' variable t (we can set $t = 1$ at the end) and consider the Taylor expansion

$$\begin{aligned} e^{tA}e^{tB} &= \left(\mathbb{1} + tA + \frac{1}{2}t^2A^2 + ...\right)\left(\mathbb{1} + tB + \frac{1}{2}t^2B^2 + ...\right) \\ &= \mathbb{1} + t(A+B) + t^2\left(AB + \frac{1}{2}A^2 + \frac{1}{2}B^2\right) + \mathcal{O}(t^3). \end{aligned}$$

Now we want to compare it to something of the form

$$e^{t(A+B)+t^2X} = \mathbb{1} + t(A+B) + t^2X + \frac{1}{2}t^2(A+B)^2 + \mathcal{O}(t^3).$$

Comparing the right-hand sides of these two expressions order by order in t, we have

$$\begin{aligned} X &= \frac{1}{2}A^2 + \frac{1}{2}B^2 + AB - \frac{1}{2}(A+B)^2 \\ &= \frac{1}{2}[A, B], \end{aligned}$$

which is exactly the result we wanted. □

Remark 1.4.3. Note that for Abelian groups we have $[A, B] = 0$ for all $A, B \in G$ and so we get the 'usual' formula

$$e^Ae^B = e^{A+B}.$$

Indeed the reason we are allowed to use this 'identity' when in school is because the real numbers form an Abelian group under multiplication.

So we can now express the right-hand side of Equation (1.4) as a single exponential in terms of T_x, T_y, T_z and their commutators. The question is "what are these commutators?" Insert exercise.

Exercise

Show that

$$[T_x, T_y] = T_z, \qquad [T_z, T_x] = T_y, \quad \text{and} \quad [T_y, T_z] = T_x. \tag{1.6}$$

Remark 1.4.4. As we will see, this will turn out to be an important property in terms of Lie algebras below.

This tells us that all the terms in our single exponential are determined by knowing T_x, T_y and T_z. So we claim that the Lie algebra of our Lie group SO(3) is the vector space spanned by these three matrices. This is *much* easier to study!

1.4.2 Converting Lie Group Properties To Lie Algebra Properties

We have defined our Lie groups as matrices with constrictions imposed, i.e. $A^TA = \mathbb{1}$ etc. The question is "what do these translate to in terms of the Lie algebra?" Well we use the exponential map to find out.

Let's consider the orthogonal condition. Let

$$A = \mathbb{1} + \epsilon a + \mathcal{O}(\epsilon^2)$$

be our infinitesimal expansion around the identity. Now from $A^T A = \mathbb{1}$, we have

$$\begin{aligned}\mathbb{1} &= (\mathbb{1} + \epsilon a)^T(\mathbb{1} + \epsilon a) + \mathcal{O}(\epsilon^2) \\ &= \mathbb{1} + \epsilon(a^T + a) + \mathcal{O}(\epsilon^2),\end{aligned}$$

and so we conclude

$$a^T = -a, \tag{1.7}$$

which says that a is antisymmetric. This is exactly the condition the Ts obeyed.

Exercise

Show that the special condition, $\det A = 1$, translates to

$$\operatorname{Tr} a = 0. \tag{1.8}$$

Remark 1.4.5. Note that antisymmetric matrices are already traceless (it's 0s on diagonal), however for SU(n) we will need this traceless condition to get the dimensions right.

The above two results tell us the the Lie algebra of SO(3) is the 3-dimensional vector space of antisymmetric matrices and T_x, T_y and T_z form a basis for this vector space. In fact we have that the Lie algebra for SO(n) is the vector space of antisymmetric $n \times n$ matrices, which has dimension $\frac{n(n-1)}{2}$.

For SU(n) we are considering complex spaces and so we have a choice to make. We either define the infinitesimal expansion to be

$$H = \mathbb{1} + i\epsilon h + \mathcal{O}(\epsilon^2)$$

or the same without the i. Of course as long as we're consistent it doesn't matter which one we pick. If we take the definition above, then by an analogous calculation to the one that gave Equation (1.7), we have

$$h^\dagger = h, \tag{1.9}$$

so the Lie algebra of SU(n) is the set of $n \times n$ hermitian, traceless matrices. These have *real* dimension $n^2 - 1$. The only thing that changes if we don't use the i is that we get *anti*hermitian matrices, i.e. $h^\dagger = -h$. Of course the dimension is the same in both cases.

Example 1.4.6. The Lie algebra of SU(2) is the set of 2×2 traceless, hermitian matrices. These are just the Pauli matrices!

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.10}$$

These are going to be very useful for us in the future.

1.5 Definitions

Although we have motivated the idea of a Lie algebra via Lie groups, they are actually abstract objects in their own right. That is, we don't need a Lie group in order to define a Lie algebra.

Definition. [Lie Algebra] A *Lie algebra*, $\mathfrak{g}$, is a vector space equipped with a map

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

called the *Lie bracket*, which satisfies the following conditions:

(i) Bilinearity; for all $x,y,z\in\mathfrak{g}$ and $\alpha,\beta\in\mathbb{R}$[11] we require

$$[\alpha x+\beta y,z]=\alpha[x,z]+\beta[y,z],$$

and similarly for the second entry.

(ii) Antisymmetry; for all $x,y\in\mathfrak{g}$

$$[x,y]=-[y,x].$$

(iii) Jacobi identity; for all $x,y,z\in\mathfrak{g}$

$$\big[x,[y,z]\big]+\big[z,[x,y]\big]+\big[y,[z,x]\big]=0.$$

The *dimension* of the Lie algebra is the dimension of the vector space.

Definition. [Lie subalgebra] Let $(\mathfrak{g},[,])$ be a Lie algebra and let $\mathfrak{h}$ be a vector subspace. Then $(\mathfrak{h},[,])$ is a *Lie subalgebra* if it is also a Lie algebra.

Note that the Lie bracket above does *not* need to be the commutator but could be something completely different. However it is relativity easy to show (exercise coming!) that the vector space of matrices forms a Lie algebra when equipped with the commutator.

Exercise

Show the above claim.

Now, because $\mathfrak{g}$ is a vector space, we can express any element in it in terms of a basis. But we have just defined the Lie bracket to be a map to $\mathfrak{g}$, and so the result must be expandable in the basis. This motives the below definition.

Definition. [Structure Constants] Let $\mathfrak{g}$ be a Lie algebra and let $\{T_a\}$ be a basis. Then we define the *structure constants* ${f_{ab}}^c$ via the Lie bracket as

$$[T_a,T_b]=:{f_{ab}}^cT_c. \tag{1.11}$$

It follows from the antisymmetry of the Lie bracket that

$${f_{ab}}^c=-{f_{ba}}^c.$$

Exercise

Show that the Jacobi identity implies

$$f_{ab}{}^{d} f_{cd}{}^{e} + f_{bc}{}^{d} f_{ad}{}^{e} + f_{ca}{}^{d} f_{bd}{}^{e} = 0.$$

Hint: Note that the terms in the Jacobi identity are just cyclic permutations, so you can save some time by just working out $\big[T_a,[T_b,T_c]\big]$ *and then cyclicly permuting the result.*

Remark 1.5.1. As we will see, for the case of a Lie algebra associated to a Lie group, the structure constants capture almost *all* of the properties of the associated Lie group.

As we did in the definition above, we often denote Lie algebras using 'mathfrak' notation. This is especially true for Lie algebras that are associated to Lie groups, we give examples in the table below.

Lie Group	Lie Algebra
$GL(n,\mathbb{R})$	$\mathfrak{gl}(n,\mathbb{R})$
$SL(n,\mathbb{R})$	$\mathfrak{sl}(n,\mathbb{R})$
$SO(n)$	$\mathfrak{so}(n)$
$SU(n)$	$\mathfrak{su}(n)$

For future reference let's write the last two explictly.

$$\mathfrak{so}(n) = \big\{a \in M^{\mathbb{R}}_{n\times n} \,|\, a^T = -a, \, \mathrm{Tr}\, a = 0\big\}. \tag{1.12}$$

$$\mathfrak{su}(n) = \big\{a \in M^{\mathbb{C}}_{n\times n} \,|\, a^\dagger = a, \, \mathrm{Tr}\, a = 0\big\}. \tag{1.13}$$

Example 1.5.2. Using Equation (1.6) we see that the structure constants of $\mathfrak{so}(3)$ are the Levi-Civita symbols, $\epsilon_{ij}{}^{k}$, which are totally antisymmetric in all indices and obey

$$\epsilon_{12}{}^{3} = 1.$$

We don't often write the Levi-Civita tensor (density) this way and so the result is normally written

$$[T_i, T_j] = \epsilon^{ijk} T_k,$$

even though this breaks summation convention.

Once again the mathematicians will ask "*now* what are the structure preserving maps for Lie algebras?" Once again, we define the answer below. First we need to know what it means for two vector spaces to be isomorphic.

Definition. [Vector Space Isomorphism] Let $(A, +_A, \cdot_A)$ and $(B, +_B, \cdot_B)$ be two vector spaces over the same field, say $\mathbb{R}$. Then the bijective map $\phi : A \to B$ is a *vector space isomorphism* iff: for all $a_1, a_2 \in A$ and $\lambda \in \mathbb{R}$

$$\phi(a_1 +_A a_2) = \phi(a_1) +_B \phi(a_2), \quad \text{and} \quad \phi(\lambda \cdot_A a_1) = \lambda \cdot_B \phi(a).$$

We say that the two vector spaces are *isomorphic as vector spaces*, denoted $A \cong_{\text{vec}} B$.

Definition. [Lie Algebra Isomorphism] Let $(\mathfrak{g}, [,]_{\mathfrak{g}})$ and $(\mathfrak{h}, [,]_{\mathfrak{h}})$ be two Lie algebras. Then we call the map $\phi : \mathfrak{g} \to \mathfrak{h}$ a *Lie algebra isomorphism* if it is a vector space isomorphism and, for all $g_1, g_2 \in \mathfrak{g}$

$$\phi([g_1, g_2]_{\mathfrak{g}}) = \big[\phi(g_1), \phi(g_2)\big]_{\mathfrak{h}}$$

holds.

1.5.1 SO(3) & SU(2)

We can now addressed a subtle wording used above: in Remark 1.5.1 we said the structure constants *almost* capture all of the properties of the associated Lie group. Why almost? Well we have already seen that the structure constants for SO(3) are ϵ^{ijk}. Well direct calculation shows that the Pauli matrices, Equation (1.10), also obey

$$[\sigma_i, \sigma_j] = \epsilon^{ijk}\sigma_k.$$

So $\mathfrak{so}(3)$ and $\mathfrak{su}(2)$ have the exact same structure constants! This tells us that the Lie algebras are isomorphic? Indeed we can construct the isomorphism explicitly as

$$\begin{aligned}\phi : \mathfrak{su}(2) &\to \mathfrak{so}(3)\\ \sigma_i &\mapsto T_i.\end{aligned}$$

This doesn't necessarily seem like a bad thing, until we notice that the Lie groups SO(3) and SU(2) are *not* isomorphic! So we see two distinguishable Lie groups have the same Lie algebra.

This particular case is very well known and you can show that $SO(3)$ is actually what is known as a *double cover* of $SU(2)$. I will not explain what this means here, but it is covered in Chau's notes, so the interested reader is directed there.

1.6 What On Earth Is Going On?

Ok, so that was quite a dense lecture with lots of definitions, so let's just have a little recap as to what on Earth we're doing. We want to study the symmetries of physics because they're very powerful and give us important results. We claim that continuous groups are related to symmetries (more on this next lecture). So we define some of our favourite matrix groups. We then see that these are quite hard things to study so we look for an easier structure to study. We claim that Lie algebras associated to Lie groups contain (almost) all the interesting information about the Lie group. We therefore decide to use the Lie algebras, because they are vector spaces and so we can add different elements together and scale them. We also have a basis into which we can decompose elements. These are *very* nice properties to have. It only took us 14 pages to say that.

2 Representations

We started this course saying that groups are important to particle physicists because they are related to symmetries, but we are yet to actually give justification for this claim. We shall hopefully give clarity of this point in this lecture.

2.1 Representations Of Lie Groups

Recall that symmetries in quantum mechanics (QM) correspond to unitary operators acting on the Hilbert space:

$$U : |\psi\rangle \to U|\psi\rangle, \quad \text{and} \quad U : \langle\psi| \to \langle\psi| U^\dagger.$$

We require that the operators are unitary because the physically meaningful thing in QM are probabilities, which always appear as inner products. So if the operator U corresponds to some symmetry of the system the probability shouldn't change and so we have

$$\langle\psi|\psi\rangle = \langle\psi| U^\dagger U |\psi\rangle$$

and so we need $U^\dagger U = \mathbb{1}$. Furthermore, the operator must respect the symmetry group's properties. For example, if we are considering the symmetry of rotations, if U represents this symmetry then we know that the composition of two rotations is a rotation

$$R(\varphi_{x_2}, \varphi_{y_2}, \varphi_{z_2})R(\varphi_{x_1}, \varphi_{y_1}, \varphi_{z_1}) = R(\varphi_{x_3}, \varphi_{y_3}, \varphi_{z_3}),$$

and so the corresponding operator must satisfy

$$U(\varphi_{x_2}, \varphi_{y_2}, \varphi_{z_2})U(\varphi_{x_1}, \varphi_{y_1}, \varphi_{z_1}) = U(\varphi_{x_3}, \varphi_{y_3}, \varphi_{z_3}).$$

In the mathematical lingo, we say that the operators form a *representation* of the symmetry group. Let's be more precise.

Definition. [Representation Of Lie Group] Let G be a Lie group of dimension n and V be a real vector space of the same dimension. Then we obtain a *representation* of G on V by prescribing an invertible, smooth map $D : G \to GL(n, \mathbb{R})$ such that: for all $g_1, g_2 \in G$

$$D(g_1 \bullet g_1) = D(g_1) \cdot D(g_2), \quad \text{and} \quad D(e) = \mathbb{1}_d, \tag{2.1}$$

where the $\cdot$ is matrix multiplication. We call the vector space V the *representation space*. We can extend this definition to $\mathbb{C}$ trivially.

Notation. From now on we will drop the $\cdot$ for matrix multplication and assume it is understood implictly.

Exercise

Show that Equation (2.1) implies

$$D(g^{-1}) = \big[D(g)\big]^{-1}, \tag{2.2}$$

where the left-hand side -1 means the inverse element in the group and the right-hand side -1 means the matrix inverse.

Remark 2.1.1. In fact the above definition of a representation is not as general as possible; all we require is that D be a group homomorphism. That is it maps elements of the Lie group to linear maps acting on the representation space that respect the group structure, i.e. obey Equation (2.1) with $\cdot$ now corresponding to composition of maps. It is true that matrices are such maps, however not all such maps are matrices. For *almost* all of this course, though, we will consider matrix representations and so work off the above definition. I say almost because later we will consider something called the adjoint representation of Lie algebras, which is not a matrix representation but is a linear map. However even in this case we will show how we can write it as a matrix.

Proposition 2.1.2. *If $D(g)$ is a representation of G of dimension n on V then so is*

$$\widetilde{D}(g) := SD(g)S^{-1},$$

where S is a constant, invertible matrix. We say that $D(g)$ and $\widetilde{D}(g)$ are equivalent.

Proof. We just need to show it obeys Equation (2.1). Firstly we have

$$\begin{aligned}\widetilde{D}(g_1 \bullet g_2) &:= SD(g_1 \bullet g_2)S^{-1} \\ &= SD(g_1)D(g_2)S^{-1} \\ &= SD(g_1)S^{-1}SD(g_2)S^{-1} \\ &= \big(SD(g_1)S^{-1}\big)\big(SD(g_2)S^{-1}\big) \\ &= \widetilde{D}(g_1)\widetilde{D}(g_2),\end{aligned}$$

where we have used the fact that $D(g)$ is a representation, inserted $\mathbb{1}_n = S^{-1}S$ and used the associativity of matrix multiplication.

Secondly we have

$$\widetilde{D}(e) := SD(e)S^{-1} = S\mathbb{1}_nS^{-1} = SS^{-1} = \mathbb{1}_n,$$

where again we have used that $D(g)$ is a representation. $\square$

Definition. [Unitary Equivalence] Let $D(g)$ and $\widetilde{D}(g)$ be equivalent representations of G on V. Then if we can choose S to be unitary then $D(g)$ and $\widetilde{D}(g)$ are said to be *unitarily equivalent.*

Remark 2.1.3. It is common to refer to two equivalent representations $D(g)$ and $\widetilde{D}(g)$ that are *not* unitarily equivalent as *unitarily inequivalent.*

Definition. [Unitary Representation] Let G be a Lie group with representation map D. If $D(g)$ is unitary for all $g \in G$ then we say the representation is *unitary.*

2.2 Representations of SU(n)

The main group we are going to be considering in representing is SU(n), Equation (1.2). It is important to note that the idea of a representation holds for a general Lie group. That is we do not need to only consider matrix Lie groups, as we are in these notes. The idea of a representation is to convert the group into a set of matrices, as we know how to calculate the action of a matrix on a vector space (which we just write as a column matrix). However SU(n) is already a matrix group and so we don't really need to do anything to it. But first a comment on notation

Notation. We will denote the elements of a matrix using indices. We will adopt the convention that the contravariant (i.e. upper) index tells us the *row*, and the covariant (lower) index tells us the *column.* For an explicit example, let U be an $n \times n$ matrix, then

$$U = \begin{pmatrix} U^1{}_1 & U^1{}_2 & \dots & U^1{}_n \\ U^2{}_1 & U^2{}_2 & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ U^n{}_1 & U^n{}_2 & \dots & U^n{}_n \end{pmatrix}.$$

2.2.1 Fundamental & Antifundamental Representations

As we have just said, SU(n) already has the properties of a representation, that is it's already a matrix group and obeys the properties Equation (2.1). We can therefore just let D be the identity map. This is known as the fundamental representation.

Definition. [Fundamental Representation Of Lie Group] Let G be a matrix Lie group. Then we define the *fundamental representation* of G over V simply as

$$D(U) = U. \tag{2.3}$$

We can define it via its action on an element $\phi \in V$:

$$D(U) : \phi \mapsto U\phi = U^i{}_j \phi^j. \tag{2.4}$$

There is another important representation of matrix Lie groups related to the fundamental representation. We define it below.

Definition. [Antifundamental Representation Of Lie Group] Let G be a matrix Lie group. Then we define the *antifundamental representation* or *conjugate representation* of G over V as

$$\overline{D}(U) = \overline{U}, \tag{2.5}$$

where the bar denotes complex conjugation. For bookkeeping (i.e. comparison to the fundamental rep.) we denote the vector transforming in the antifundamental representation with a lower index. That is, for $\phi \in V$ we write

$$\overline{D}(U) : \phi \mapsto \overline{U}\phi = {(U^\dagger)_j}^k \phi_k. \tag{2.6}$$

Note that the dimension of the fundamental representation and antifundamental representation agree. This tells us their representation spaces have the same dimension, but we think of them as transposes (i.e. we turn a row index into a column index, $\phi^k \to \phi_k$). For the following we shall denote the latter vector space as $\overline{V}$. This does *not* mean the complex conjugate of V but simply that we lower the index. An example is given below. An interesting question is "are the fundamental and antifundamental representations equivalent?" We will return to this question later.

Remark 2.2.1. Note in Equation (2.6) we used

$${(U^\dagger)_j}^k = {\overline{U}_k}^j$$

as the Hermitian conjugate is both complex conjugation and transpose.

Exercise

Prove that the antifundamental representation is in fact a representation. That is show Equation (2.5) satisfies Equation (2.1).

There is another representation we will use. This one might seem a bit boring, but it will actually prove useful later when discussing Young-Tableauxs, so bear with it.

Definition. [Trivial Representation] Let G be a Lie group. Then we have the *trivial representation* of G over V by the one-point map:

$$D(U) = \mathbb{1} \qquad \forall U \in G. \tag{2.7}$$

It simply acts as

$$D(U) : \phi \mapsto \mathbb{1}\phi = \phi, \tag{2.8}$$

so it 'does nothing'.

Remark 2.2.2. Note that, unlike the fundamental/antifundamental representations, the trivial representation does not require G to be a matrix Lie group.

2.2.2 Tensor Products of D & $\overline{D}$

The way we defined the action of D and $\overline{D}$ looks a lot like the index notation for tensors. So the obvious question is "can we take tensor products of these?" The answer is, of course, yes because the tensor product of two matrices is well defined. We define the representation space of this tensor product construction in the usual manner for the tensor product of vectors. That is we give ϕ one upper index for every D and one lower index for every $\overline{D}$. The dimension of the representation (and therefore also the representation space) is given by $n^{d+\bar{d}}$, where n is the dimension of $D/\overline{D}$, d is the number of Ds and $\overline{d}$ is the number of $\overline{D}$s. To help clarify this let's give a couple examples.

Example 2.2.3. Let D be the n-dimensional fundamental representation of the Lie group G over V. Then define $D_T := D \otimes D \otimes \overline{D}$. Then it acts on vectors in $\phi \in V \otimes V \otimes \overline{V}$. We tend to denote its action in terms of indices as follows:

$$D_T(U) : \phi^{ij}{}_k \mapsto U^i{}_{i'} U^j{}_{j'} \big(U^\dagger\big)^{k'}{}_k \phi^{i'j'}{}_{k'}.$$

The dimension of the tensor product representation, D_T, is n^3. We haven't actually shown that this is in fact a representation. This is the content of the next exercise.

Exercise

Show that D_T defined above forms a representation. That is shows it obeys Equation (2.1). *Hint: Note that* $\mathbb{1}^i{}_j = \delta^i_j$. *The other property is a bit trickier to see. Just write down the action of* $D(UV)$ *on* $\phi^{ij}{}_k$, *expand* $(UV)^i{}_j = U^i{}_k V^k{}_j$ *and then use the fact that you can move around index terms freely.*[a]

[a] This hint might be more cryptic then helpful. If that's the case and you still can't do it, feel free to email me. It's actually not hard to show, but difficult to give much more of a hint without doing the question.

There is a useful trick to notice that can save us a lot of time when we have contracted indices. As with tensors in GR, indices that are contracted (i.e. in $T^{ij}S_j$, j is contracted) are called *dummy* indices and do not transform. This comes from the fact that covariant and contravariant indices transform in exactly the opposite way. We have a similar thing here when we consider the fundamental and antifundamental representations of SU(2). We leave the proof of this as an exercise below.[1]

Exercise

Show that

$$\phi^{jk}{}_k := \phi^{j\ell}{}_k \delta^k_\ell$$

transforms in the fundamental representation. That is

$$D_T(U)\phi^{jk}{}_k = U^j{}_{j'}\phi^{j'k}{}_k = D(U)\phi^{jk}{}_k.$$

Hint: You will need to use the fact that we're considering SU(n) and so $U^\dagger U = \mathbb{1}$.

2.3 Reducible & Irreducible Representations

The last exercise shows that some tensor product constructions don't actually give rise to anything new, and we can essentially consider just the action of one part of it independently.

[1] If you hate these last two exercises, blame Dr. Dorigoni not me... They're exercises in his notes and I don't want to put the answers here for obvious reasons. If you are stuck with either of them, please feel free to email me and I can explain.

In other words, it appears we can 'reduce' complicated constructions into more bite size bits and deal with them one by one. If we can do this, we say the representation if *reducible.*

Definition. [Reducible Representation] A representation of a Lie group,[2] D, of dimension n is called *reducible* if it is equivalent to representation of the form

$$SD(g)S^{-1} = \begin{pmatrix} A(g) & C(g) \\ 0 & B(g) \end{pmatrix}$$

for all $g \in G$.

Remark 2.3.1. In the definition above, the matrix C need not be a square matrix, all we require is that the complete matrix on the right-hand side is $n \times n$ (otherwise it wouldn't be equivalent to D). For example We could have

$$A \in M^{\mathbb{C}}_{d_1\times d_1}, \qquad B \in M^{\mathbb{C}}_{d_2\times d_2}, \quad \text{and} \quad C \in M^{\mathbb{C}}_{d_1\times d_2}.$$

This would give $n = d_1 + d_2$.

Definition. [Completely Reducible] A reducible representation is said to be *completely reducible* if $C(g) = 0$ for all $g \in G$, i.e.

$$SD(g)S^{-1} = \begin{pmatrix} A(g) & 0 \\ 0 & B(g) \end{pmatrix}. \tag{2.9}$$

There is an alternate way we can write the condition of reducible. Note that the representation space of a reducible representation will have an *invariant* subspace. That is if we set the bottom d_2 entries of the column matrix of $\phi \in V$ to 0 we get

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}\begin{pmatrix} \underline{\alpha} \\ 0 \end{pmatrix} = \begin{pmatrix} A\underline{\alpha} \\ 0 \end{pmatrix},$$

where $\underline{\alpha}$ has d_1 entries. We can write this mathematically as follows.

Definition. [Invariant Subspace] Let D be a representation of a Lie group G on V. Then we call the subspace $U \subset V$ an *invariant subspace* if for all $g \in G$ and $u \in U$

$$D(g)u \in U.$$

Now note that if a representation is completely reducible then the representation space consists exactly of 2 separate invariant subspaces. That is,

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} \underline{\alpha} \\ \underline{\beta} \end{pmatrix} = \begin{pmatrix} A\underline{\alpha} \\ B\underline{\beta} \end{pmatrix},$$

and so α and β never talk to each other. We can therefore decompose V into a *direct sum* of its invariant subspaces, in this case

$$V = a \oplus b,$$

where $\underline{\alpha} \in a$ and $\underline{\beta} \in b$ with $\dim V = \dim a + \dim b = d_1 + d_2$.

We can therefore write the condition for completely reducible in a nice mathematical formulation. First we need the definition of an irreducible representation.

Definition. [Irreducible Representation] Let D be a representation of a Lie group[3] G on V. We say that this representation is *irreducible*, or more simply an *irrep*, if V has no non-trivial[4] invariant subspace of any equivalent representation of D.

Definition. [Completely Reducible (Direct Sum)] Let D be a representation of a Lie group on V. Then we call D *completely reducible* if it can be written as a direct sum of irreps:

$$D(g) = A(g) \oplus B(g). \tag{2.10}$$

This acts on $V = a \oplus b$ as

$$D(g)V = A(g)a \oplus B(g)b.$$

Example 2.3.2. In the last exercise (about contracted indices) we have

$$D_T = D \oplus \mathbb{1} \oplus \mathbb{1},$$

where $\mathbb{1}$ is the trivial representation.

Proposition 2.3.3. *Let D be a completely reducible representation. Then the dimension of D is equal to the sum of the dimension of the irreps in the direct sum.*

Theorem 2.3.4 (Maschke)**.** *If a unitary representation is reducible then it is also completely reducible.*

Proof. Let D be our representation of G over V with dimension n. As D is reducible, it has an invariant subspace of V. Let $V_1 \subset V$ be this invariant subspace with dimension d_1. Then, as V is a vector space, we can define a basis

$$\{e_1, ..., e_{d_1}, e_{d_1+1}, ..., e_{d_1+d_2}\},$$

where $n = d_1 + d_2$. We are free to choose this basis such that $\{e_1, ..., e_{d_1}\}$ is a basis for V_1. Define the subspace spanned by $\{e_{d_1+1}, ..., e_{d_1+d_2}\}$ by V_2. We see straight away that V_1 and V_2 are orthogonal.

Now because V_1 is an invariant subspace we know every $v_1 \in V_1$ satisfies

$$D(g)v_1 \in V_1,$$

and so can be decompose it in the basis $\{e_1, ..., e_{d_1}\}$. The inner product with an arbitrary element $v_2 \in V_2$ must vanish by orthogonality, i.e.

$$\big(D(g)v_1, v_2\big) = 0.$$

This holds for all elements $g \in G$ and so in particular holds for $g^{-1} \in G$. Now use the property of inner products:

$$\big(D(g^{-1})v_1, v_2\big) = \Big(v_1, \big(D(g^{-1})\big)^\dagger v_2\Big).$$

Next use the fact that D is unitary and so

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

where the last line comes from Equation (2.2) along with the fact that the inverse in the group is unique (i.e. $(g^{-1})^{-1} = g$). Putting this together we get

$$\big(v_1, D(g)v_2\big) = 0,$$

which tells us

$$D(g)v_2 \in V_2,$$

and so V_2 is an invariant subspace. Finally using the fact that V_1 and V_2 completely span V we have

$$V = V_1 \oplus V_2,$$

and so we can decompose D in a similar manner. □

This theorem is incredibly powerful because it tells us that for $\mathrm{SU}(n)$ the only thing we need to consider is irreducible representations and their direct sums. This *massively* simplifies things.

Remark 2.3.5. Note also for SU(2) Maschke's theorem allows us to stop distinguishing between just reducible and completely reducible. We shall therefore just say reducible (as it's one less word).

Lemma 2.3.6. *Let D be a representation with equivalent representation $\widetilde{D}(g) = SD(g)S^{-1}$. Then D is reducible if, and only if, $\widetilde{D}$ is reducible.*

Exercise

Prove the above Lemma.

Remark 2.3.7. *Basically what you end up showing here is the same as showing that the equivalence of representations forms an equivalence relation (i.e. all the square box notation I was using in my proof that $\mathbb{Z}_n$ is a group). This gives further justification of me saying its worth learning about equivalence classes.*

2.3.1 Symmetric ⊕ Antisymmetric

People familiar with GR will probably know that you can decompose any 2 index tensor into a sum of its symmetric and antisymmetric parts. This subsection aims to show you can do the same thing for the tensor product of two representations.

Let's consider the tensor product of two fundamental representations.

$$D_B(U) = (D \otimes D)(U) : \phi^{ij} \mapsto U^i{}_{i'} U^j{}_{j'} \phi^{i'j'}.$$

What the comment at the start of this subsection is saying is that we want to show that

$$\begin{aligned} \phi^{(ij)} &:= \frac{1}{2}\big(\phi^{ij} + \phi^{ji}\big) && \text{(symmetric)}, \\ \phi^{[ij]} &:= \frac{1}{2}\big(\phi^{ij} - \phi^{ji}\big) && \text{(antisymmetric)}. \end{aligned}$$

are invariant subspaces, and so we can decompose D_B into a direct sum

$$D_B = D_S \oplus D_A,$$

where

$$\begin{aligned} D_S(U) &: \phi^{ij} \mapsto \frac{1}{2}\big(U^i{}_{i'}U^j{}_{j'} + U^j{}_{j'}U^i{}_{i'}\big)\phi^{i'j'} \\ D_A(U) &: \phi^{ij} \mapsto \frac{1}{2}\big(U^i{}_{i'}U^j{}_{j'} - U^j{}_{j'}U^i{}_{i'}\big)\phi^{i'j'}. \end{aligned}$$

It is easy to see that $\phi^{ij} = \phi^{(ij)} + \phi^{[ij]}$, so we just need to show that they are invariant under the action of D. We show this result for the symmetric case and leave the antisymmetric case as an exercise.

Denote the symmetric/antisymmetric parts of V by V_S/V_A respectively. We need to show that

$$D_S(U)\phi_S \in V_S, \quad \text{and} \quad D_A\phi_A = 0$$

for all $\phi_S \in V_S$ and $\phi_A \in V_A$. The general elements of V_S/V_A are given above, so direct calculation gives

$$\begin{aligned} D_S(U)\phi_S &= \frac{1}{4}\big(U^i{}_{i'}U^j{}_{j'} + U^j{}_{j'}U^i{}_{i'}\big)\phi^{i'j'} + \frac{1}{4}\big(U^j{}_{j'}U^i{}_{i'} + U^i{}_{i'}U^j{}_{j'}\big)\phi^{j'i'} \\ &= \frac{1}{4}\big(U^i{}_{i'}U^j{}_{j'} + U^j{}_{j'}U^i{}_{i'}\big)\phi^{i'j'} + \frac{1}{4}\big(U^i{}_{i'}U^j{}_{j'} + U^j{}_{j'}U^i{}_{i'}\big)\phi^{j'i'} \\ &= \frac{1}{4}\big(U^i{}_{i'}U^j{}_{j'} + U^j{}_{j'}U^i{}_{i'}\big)\big(\phi^{i'j'} + \phi^{j'i'}\big) \\ &= \frac{1}{2}U^i{}_{i'}U^j{}_{j'}\big(\phi^{i'j'} + \phi^{j'i'}\big), \end{aligned}$$

which is symmetric in $i \leftrightarrow j$, and so is an element of V_S. Now consider the action on an antisymmetric element: the only thing that will change is the sign between the two terms on the first line and so the same calculation will result in

$$D_S(U)\phi_A = \frac{1}{4}\big(U^i{}_{i'}U^j{}_{j'} - U^j{}_{j'}U^i{}_{i'}\big)\big(\phi^{i'j'} + \phi^{j'i'}\big) = 0,$$

which is the desired result.

Exercise

Show the analogous calculation for $D_A(U)$.

This means we can express $D_B(U)$ as

$$D_B(U) = \begin{pmatrix} D_S(U) & 0 \\ 0 & D_A(U) \end{pmatrix}.$$

Let's just check that dimensions work out. We said that the dimension of the tensor product of representations was the product of the dimensions. This gives (assuming $\dim D = n$)

$$\dim D_B = n^2.$$

Now recall that we said the dimension of a Lie matrix group is equal to the number of free parameters in the matrix. The fundamental representation doesn't do anything to the matrices, and so it's dimension is also given by the number of free parameters. $D_S(U)/D_A(U)$ are symmetric/antisymmetric $n \times n$ matrices, and so have dimensions

$$\dim D_S = \frac{n(n+1)}{2}, \quad \text{and} \quad \dim D_A = \frac{n(n-1)}{2}$$

adding these gives

$$\dim D_S + \dim D_A = n^2 = \dim D_B,$$

which is exactly what we wanted.

Remark 2.3.8. Note there was noting special about us using the tensor product of two fundamental representations. A completely analogous calculation also holds for

$$\overline{D}_B = \overline{D} \otimes \overline{D}, \qquad D_C = D \otimes \overline{D}, \quad \text{and} \quad \overline{D}_C = \overline{D} \otimes D.$$

2.4 Schur's Lemma

We have just invested a considerable amount of time and effort in obtaining a irreps, but any sensible person would ask "why do we care about them?" Well we have already given a reasonable answer above (the idea that, for unitary representations, everything is described in terms of irreps), however there is a physical answer which might be more satisfying to us physicists. It comes in the form of a famous Lemma.

Lemma 2.4.1 (Schur's Lemma). *Let D be an irrep of G over V. Then if there exists a matrix H such that for all $g \in G$*

$$[H, D(g)] = 0 \qquad \Longrightarrow \qquad H = \lambda \cdot \mathbb{1}, \tag{2.11}$$

where $\lambda \in \mathbb{C}$.[5]

Proof. Let $\underline{v} \in V$ be an eigenvector of H with eigenvalue λ,[6] then if H commutes with $D(g)$ then

$$H\big(D(g)\underline{v}\big) = D(g)H\underline{v} = \lambda \cdot \big(D(g)\underline{v}\big).$$

This tells us that $D(g)\underline{v}$ is also a eigenvector of H with the same eigenvalue. This is true for all g and so we conclude that the eigenspace V_λ is an invariant subspace of D (otherwise we would get a different eingenvalue with H). But D is a irrep so it has no non-trivial invariant subspaces, and because the eigenspace is not empty we are forced to conclude that $V_\lambda = V$, so *every* element in V is an eigenvector of H with eigenvalue λ. This is just the statement that $H = \lambda \cdot \mathbb{1}$. □

So why is this a nice physical answer to our question at the start of this section? Well recall that any exact symmetry should commute with the Hamiltonian. So for a group G to be a symmetry, we require

$$[H, D(g)] = 0 \qquad \forall g \in G.$$

[5] Or whatever the field of the vector space is.

[6] Note every matrix has at least one because $\det(B - \lambda \cdot \mathbb{1}) = 0$ has a solution.

Schur's Lemma therefore tells us that the Hamiltonian acts as $E \cdot \mathbb{1}$ on an irrep of a symmetry group G. In fancier language: states in an irrep of an exact symmetry group form a *multiplet* with degenerate energies. This is a very powerful statement, because recalling that small equation of Einstein's, $E = mc^2$, we see that states that are connected by an irrep of an exact symmetry group *have the same mass*! This is the reason why an electron and a positron have the same mass. It is a fact that the dimension of the irrep corresponds to the number of terms in the multiplet,[7] and so if we can find an irrep of dimension n that commutes with H we instantly know that there are n particles with equal mass.

Now if that isn't a physically compelling argument for why irreps are worth studying, then I'm sorry but you're never going to be convinced.

[7]Check you understand why this is the case.

3 Young-Tableaux

In the last lecture we gave arguments for how useful it is to decompose our tensor product of representations into a direct sum of irreps. We gave some explicit examples by finding invariant subspaces. As we saw this took a reasonable amount of work for the two index tensor and we basically only got the answer because we had an idea from GR. It seems like we're doomed when it comes to considering objects with more indices. For example, as we will show soon, the following decomposition is not trivial to see

$$\phi^{(ij)}\varphi^k = \frac{1}{3}\big(\phi^{(ij)}\varphi^k + \phi^{(ik)}\varphi^j + \phi^{(jk)}\varphi^i\big) + \frac{1}{3}\big(2\phi^{(ij)}\varphi^k - \phi^{(jk)}\varphi^i - \phi^{(ik)}\varphi^j\big). \tag{3.1}$$

These two terms are invariant subspaces of $D \otimes D \otimes D$. The first term is maybe not too hard to guess, its just the fully symmetric $\phi^{(ij}\varphi^{k)}$, however the second term doesn't have any nice easy to guess property. Sure it is symmetric in $i \leftrightarrow j$, but the exchange $j \leftrightarrow k$ gives

$$2\phi^{(ik)}\varphi^j - \phi^{(jk)}\varphi^i - \phi^{(ij)}\varphi^k.$$

The middle term hasn't changed at all but the other two have changed sign and factors of 2. A similar thing happens for $i \leftrightarrow k$.

So what are we to do? Well of course we could go via trial and error, but that's not fun. Luckily a brilliant mathematician, called Alfred Young, swoops in and saves the day. He developed a rather remarkable pictorial way to find the decomposition into direct sums in 1990. The pictures even allow us to calculate the dimensions of the irreps. These diagrams are called *Young-Tableaux* and will be the study of this lecture.

Notation. We shall switch to a notation where capital N is the N in $\mathrm{SU}(N)$ and little n is the number of indices. This is just done to make it easier for me to work from Dr. Dorigoni's notes (which do this).

3.1 The Rules

A Young-Tableaux is a pictorial representation to characterise irreps $\mathrm{SU}(N)$[1] and correspond to a particular symmetrisation and antisymmetrisation procedure. It will generate the irreps of a tensor product and it will also give us the dimensions of each irrep. The pictures correspond to drawing boxes. Of course there are rules on how to construct such diagrams, which we lay out below.

[1] You can adjust them for other groups like $\mathrm{SO}(n)$, but in this course we will only be interested in $\mathrm{SU}(N)$.

(i) Each term must have the same number of boxes as there are free indices (i.e. not ones summed over).

(ii) Boxes in the same *row* correspond to *symmetrised* indices.

(iii) Boxes in the same *column* correspond to *antisymmetrised* indices.

(iv) Each row must contain *no more boxes than the one above.*

(v) The rows are aligned to the left.

(vi) The number of rows must not exceed N for SU(N).

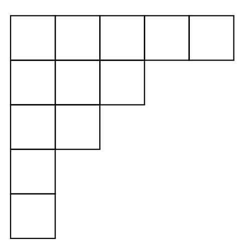
Remark 3.1.1. Note condition (vi) makes perfect sense given condition (iii): rows in the same column are antisymmetrised and for SU(N) the range of the indices is $i = 1, ..., N$. If we have $(N+1)$ indices then at least two of them will have to be the same, and so if we antisymmetrise them all, this term vanishes. For example, for $N = 2$, an object with $n = 3$ indices will vanish if fully antisymmetrised, as we require

$$\phi^{ijk} = -\phi^{ikj} = -\phi^{kji}$$

but if we set $i = 1$ and $j = 2$ then either $k = 1$, and so the second equality gives 0, or $k = 2$ and so the first equality gives 0. A similar argument is made for any other combination for ijk.

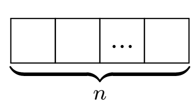
Let's give the pictorial version of the conditions above for clarity, then we'll give some examples.

(i) This one is pretty self explanatory, but here's an example



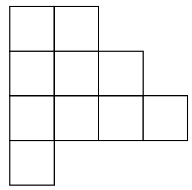
This corresponds to an object with $n = 12$ indices.

(ii) The n index fully symmetrised object $\phi^{(i_1 \dots i_n)}$ corresponds to



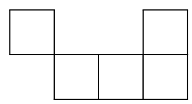
(iii) The n index fully antisymmetrised object $\phi^{[i_1 \dots i_n]}$ is similar to the above one but now the boxes are vertical.

(iv) As we have drawn in (i), each row has no more boxes than the one above it. Note we can have the same number of boxes, as with the last two rows in (i), but a diagram like



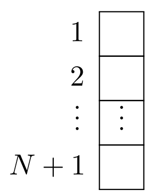
would not be valid.

(v) Again as we have done in (i), all the rows are aligned to the left, so a diagram like



would not be valid, both because the top row has 'gaps' and because the second row starts shifted in.

(vi) We already explained why this is the case in Remark 3.1.1, but pictorially we can write



$$= \quad 0.$$

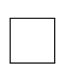
It is worth clarifying what these pictures actually tell us. The number of boxes gives us the number of indices on the elements in our representation space, and the way the boxes are ordered tells us what the invariant subspaces are. So each diagram tells us an invariant subspace, and corresponds to one term in the direct sum of irreps.

We then take the direct sum of all the different diagrams (i.e. all the different invariant subspaces) and so we obtain the full action of the representation on the representation space. So these diagrams tell us both what the vector space V is (i.e. it is the span of the objects whose indices are given by the diagrams) and the decomposition of the representation (as we know the invariant subspaces).

3.2 Fundamental & Antifundamental

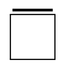
The first, and essentially the building block of all Young-Tableaux, is the fundamental. We give it here as a definition.

Definition. [Fundamental Young-Tableaux] The *fundamental Young-Tableaux* is simply a single box:



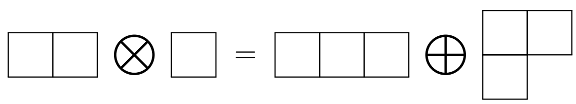
$$= \quad \phi^i \mapsto U^i{}_j \phi^j.$$

We will see later that there is a nice relation between the fundamental and antifundamental Young-Tableaux for SU(N). So now we introduce a nice notation for the antifundamental Young-Tableaux in terms of a definition.

> **Definition.** [Antifundamental Young-Tableaux] The *antifundamental Young-Tableaux* is drawn as a box with a bar over it:
>
>  $= \phi_i \mapsto (U^\dagger)^j{}_i \phi_j.$

3.3 Tensor Products

So how do we write tensor products of Young-Tableaux in terms of direct sums? Well let's give an example here, and then explain why it's correct. We saw (or at least claimed) in Equation (3.1) that we can decompose the tensor product of a 2-index symmetric object with a fundamental object as the direct sum of the fully symmetric object and something that was symmetric in ij but some non-trivial antisymmetry with k. As Young-Tableaux this is

 (3.2)

The first term on the right-hand side, by condition (ii), is just $\phi^{(ij}\varphi^{k)}$, while the other term corresponds to the funny property. Note this terms makes some kind of sense: we have two indices symmetrised and one with some antisymmetry property.

So how do we arrive at this expression? Well the attentive person might realise that all we have done is put the fundamental box in the only two allowed places: on the end of the two-boxes and below it. This is essentially the correct idea, and we will give a more detailed description next lecture, including what to do if you have more than one box to 'distribute'.

How do we see that the last term in the above Young-Tableaux corresponds to the term in Equation (3.1)? Well we need to explain the procedure of symmetrisation/antisymmetrisation in a Young-Tableaux. It goes as follows: for a given Young-Tableaux diagram

1. Assign indices to the boxes, starting at the top left, working along the row and then down to the next column.

2. Apply the permutation operator
$$P = \sum_r p,$$
where r indicates the row number and p permutes the indices in a row.

3. Apply the graded permutation operator
$$Q = \sum_c \mathrm{sgn}(q) q,$$
where c indicates the column number, q permutes the indices in a column, and $\mathrm{sgn}(q)$ is the sign of the permutation.[2]

[2] A permutation q is even (has $\mathrm{sgn}(q) = +1$) is it can be written as the product of an even number of transpositions (something that switches only two indices). Otherwise it is odd (has $\mathrm{sgn}(q) = -1$).

Let's show how this gives the above result.[3]

1. First we have

$$\psi^{ijk} = \begin{array}{|c|c|}\hline i & j \\ \hline k \\ \cline{1-1}\end{array}$$

2. Then we permute along the rows: i.e. symmetrise i and j (k has nothing else in its row so it's left alone)

$$P\big(\psi^{ijk}\big) = \psi^{ijk} + \psi^{jik} = \begin{array}{|c|c|}\hline i & j \\ \hline k \\ \cline{1-1}\end{array} + \begin{array}{|c|c|}\hline j & i \\ \hline k \\ \cline{1-1}\end{array}$$

3. Then we graded permute in the columns: the permutations that do nothing obviously have positive sgn, while both permutations $i \leftrightarrow k$ and $j \leftrightarrow k$ correspond to one transposition and so have negative sgn. This gives

$$Y\big(\psi^{ijk}\big) := (Q \circ P)\big(\psi^{ijk}\big) = \big(\psi^{ijk} - \psi^{kji}\big) + \big(\psi^{ijk} - \psi^{ikj}\big).$$

   This is exactly (apart from the factor 1/3) the second term on the right-hand side of Equation (3.1).

Remark 3.3.1. It turns out that in $\mathrm{SO}(N)$ the contraction of indices will allow for further decomposition into irreps. We will not be concerned with this fact in this course, as we focus on $\mathrm{SU}(N)$.

3.4 Dimensions From Young-Tableaux

As we said at the beginning of this lecture, Young-Tableaux not only give us a way to decompose the tensor product of the representations into a direct sum of irreps, but it also gives us a way to find the dimension of the irreps. We give the prescription of how to do this here. This result is highly non-trivial to see, and we do not provide a proof of it but simply request you believe it's true.

First we define the *Hook* of a box in a Young-Tableaux.

Definition. [Hook In Young-Tableaux] The *Hook* of a box X in a Young-Tableaux is the integer given by summing over the number of boxes directly to the right of X, plus the number of boxes directly below X, plus one for X itself.

Example 3.4.1. Let's give an example of a Hook. Consider the Young-Tableaux

$$\begin{array}{|c|c|c|c|c|}\hline & \mathrm{X} & & & \\ \hline \mathrm{Y} & & & \\ \cline{1-4} & & \\ \cline{1-3} & \\ \cline{1-2}\end{array}$$

[3] We ignore all the factors of 1/2 etc that comes from symmetrisation etc.

We have

$$\text{Hook}(X) = \underset{\text{right}}{3} + \underset{\text{below}}{3} + \underset{\text{self}}{1} = 7, \quad \text{and} \quad \text{Hook}(Y) = \underset{\text{right}}{3} + \underset{\text{below}}{2} + \underset{\text{self}}{1} = 6.$$

Claim 3.4.2. The dimension of a Young-Tableaux of SU(N) is given by the following procedure.

1. Put N in the top left box.
2. Add 1 as you move along the row (so second box has $N+1$, third has $N+2$ etc).
3. Minus 1 as you move down a column (so second row, first column has $N-1$, but second row second column has N — as you add one as you move across row)
4. Multiply all these numbers together and divide by the product of all the Hooks.
5. The result is the dimension.

For clarity, we give a pictorial representation of how to associate the numbers to boxes using the Young-Tableaux given in Example 3.4.1 for the case SU(5):

$$\begin{array}{ccccc} 5 & 6 & 7 & 8 & 9 \\ 4 & 5 & 6 & 7 & \\ 3 & 4 & 5 & & \\ 2 & 3 & & & \end{array}$$

Remark 3.4.3. Note condition (vi) for Young-Tableaux ensures that the dimension is positive definite. That is you will never get 0 or a negative number in a box, as you would need $N+1$ rows to get 0 and more rows to get a negative number.

Example 3.4.4. Let's find the dimension of the following Young-Tableaux for SU(6):

$$\begin{array}{cccc} \square & \square & \square & \square \\ \square & \square & & \\ \square & \square & & \\ \square & & & \\ \square & & & \end{array}$$

Writing the value of the Hook as a number in the box, we have

$$\begin{array}{cccc} 6 & 7 & 8 & 9 \\ 5 & 6 & & \\ 4 & 5 & & \\ 3 & & & \\ 2 & & & \end{array} \Bigg/ \begin{array}{cccc} 8 & 5 & 2 & 1 \\ 5 & 2 & & \\ 4 & 1 & & \\ 2 & & & \\ 1 & & & \end{array} = \frac{6\cdot 7\cdot 8\cdot 9\cdot 5\cdot 6\cdot 4\cdot 5\cdot 3\cdot 2}{8\cdot 5\cdot 2\cdot 1\cdot 5\cdot 2\cdot 4\cdot 1\cdot 2\cdot 1} = 1701,$$

where the slash is meant to indicate a divide.[4] This example highlights the power of Young-Tableaux (imagine trying to find the dimension of a 10 index object with the above symmetrisation/antisymmetrisation).

Exercise

Use the above procedure to show that the dimensions of the decomposition Equation (3.1) works out. That is show that both sides of Equation (3.2) have the same dimension. *Hint: Recall that* $\dim(A\otimes B) = (\dim A)\cdot(\dim b)$ *and* $\dim(A\oplus B) = (\dim A) + (\dim B)$.

3.5 Antifundamental Young-Tableaux From $(N-1)$-Rows

3.5.1 Invariant Tensor

Let's consider the specific case of a Young-Tableaux of SU(N) with exactly N rows. This corresponds to a fully antisymmetrised object and has dimension

$$\begin{array}{|c|}\hline N \\ \hline \vdots \\ \hline 1 \\ \hline \end{array} \Big/ \begin{array}{|c|}\hline N \\ \hline \vdots \\ \hline 1 \\ \hline \end{array} = 1.$$

This looks a lot like the Levi-Civita tensor. So what we're looking at is something like

$$\phi^{[i_1\dots i_N]} = \varphi\epsilon^{i_1\dots i_N},$$

where φ is just some scalar (it doesn't transform under the representation).

Proposition 3.5.1. *The Levi-Civita tensor is an invariant tensor under SU(N). That is,*

$$(\underbrace{D\otimes\dots\otimes D}_{N\text{-times}})(U) : \epsilon^{i_1\dots,i_N} \mapsto \epsilon^{i_1\dots i_N}.$$

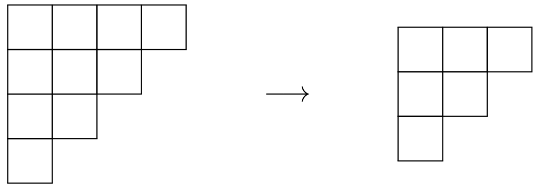
Proof. Just compute the action:

$$(D\otimes\dots\otimes D)(U) : \epsilon^{i_1\dots i_N} \mapsto U^{i_1}{}_{j_1}\dots U^{i_N}{}_{j_N}\epsilon^{j_1\dots j_N} = \det U\epsilon^{i_1\dots i_N},$$

where the second equality is a well known fact (see a linear algebra textbook). But $\det U = 1$ for SU(N) and so we get the result. $\square$

This result tells us that the N-row Young-Tableaux corresponds to the trivial representation, and so in all future Young-Tableaux we can always 'strip off' this part of a diagram. For example, if $N=4$ we would replace

[4] Apologies for how pathetic it looks, I'm new to the Young-Tableaux package and don't know how to make a proper big slash yet.

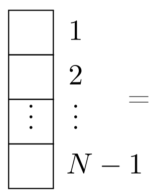


Note that by doing this we will actually violate condition (i) in our Young-Tableaux procedure. It is important to note that we are *not* saying that you remove the indices, but simply that these indices will not transform, and simply come along for the ride. So in order to save ourselves writing it down in every step, we simply 'forget about it for now'.

Remark 3.5.2. It turns out that SO(N) has more than one invariant tensor, and so you can 'forget about' more of the diagram. However, as with other SO(N) remarks, this won't concern us in this course.

3.5.2 Antifundamental

What about if we have $(N-1)$ rows? A similar calculation to the one above tells us that

dim  $= N.$

What representation do we know that has dimension N? Well the fundamental of course (it's $N/1 = N$). The above Young-Tableaux isn't the fundamental though, so what is it? A bit of thought suggests the antifundamental. Let's show this more concretely.

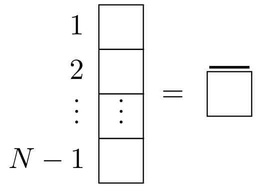
If we write the vector as

$$\phi^{[i_1...i_{N-1}]} = \epsilon^{i_1...i_{N-1}j}\Phi_j,$$

then the transformation is as follows

$$\begin{aligned}(\underbrace{D\otimes ...\otimes D}_{N-1})(U) : \phi^{[i_1...i_{N-1}]} &\mapsto (U^{i_1}{}_{j_1}...U^{i_{N-1}}{}_{j_{N-1}})\epsilon^{j_1...j_{N-1}j}\Phi_j \\ &= (U^{i_1}{}_{j_1}...U^{i_{N-1}}{}_{j_{N-1}})\cdot\delta^j_i\cdot\epsilon^{j_1...j_{N-1}i}\Phi_j \\ &= (U^{i_1}{}_{j_1}...U^{i_{N-1}}{}_{j_{N-1}})\cdot(U^\dagger)^j{}_kU^k{}_i\cdot\epsilon^{j_1...j_{N-1}i}\Phi_j \\ &= (\det U)\epsilon^{i_1...i_{N-1}k}(U^\dagger)^j{}_k\Phi_j,\end{aligned}$$

but the first part is just the invariant tensor from the previous subsection, and so we just get a transformation in the antifundamental representation. So we have

 (3.3)

This tells us that, for SU(N), we don't actually need to consider the antifundamental representation in terms of Young-Tableaux, and we can get all irreps using just the fundamental.

This is good because as we have defined the Young-Tableaux, we only ever used the fundamental representation! To emphasise, our Young-Tableaux construction gives us *all* the irreps for SU(N).

3.6 List Of All Irreps

Now that we have a pictorial tool to list all of the irreps for a given SU(N), let's list some examples.

3.6.1 SU(2)

For SU(2) our invariant tensor is the Young-Tableaux

$$\ydiagram{1,1}$$

so we only need to consider one row. We therefore can list all the irreps as

Young-Tableaux	Tensor	Dimension
$\ydiagram{1}$	ϕ^i	2
$\ydiagram{2}$	$\phi^{(ij)}$	3
$\ydiagram{3}$	$\phi^{(ijk)}$	4
$\vdots$	$\vdots$	$\vdots$

We can therefore characterise the Young-Tableaux by a single number, namely the number of boxes.

Note that for SU(2) the fundamental and the antifundamental are the equivalent:

$$\ydiagram{1} = \overline{\ydiagram{1}}.$$

This is obviously not true for any other SU(N).

3.6.2 SU(3)

For SU(3) we now have at most 2 rows, and have

$$\ydiagram{1,1} = \overline{\ydiagram{1}}.$$

A general Young-Tableaux is of the form

$$\begin{array}{cccccccc} 1 & \dots & \dots & q & 1 & \dots & \dots & p \end{array}$$
$$\begin{ytableau} \ & \dots & \dots & \ & \ & \dots & \dots & \ \\ \ & \dots & \dots & \ \end{ytableau},$$

and so we can simply characterise an arbitrary Young-Tableaux for SU(3) by the double (p, q), which tells us the number of fundamental and antifundamental, respectively, indices.

Remark 3.6.1. These objects can be written as (p, q) tensors that are fully symmetric in all p contravaiarnt indices, fully symmetric in the q covariant indices are are completely traceless, i.e.

$$\phi^{(i_1 \dots i_p)}_{(j_1 \dots j_q)} \qquad \text{with} \qquad \phi^{k i_2 \dots i_p}_{k j_2 \dots j_q} = 0.$$

We do not explain why, but just state that this is true.

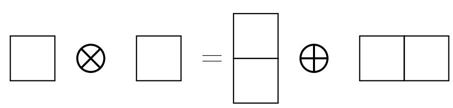
3.6.3 SU(4) & Higher

It is not so easy to characterise a general Young-Tableaux for SU(4) and higher. Simply drawing the Young-Tableaux is most compact way to write down a general tensor.

3.7 Bold Face Dimension Notation

There is a short hand notation to writing the irreps of a Young-Tableaux by its dimension. We simply use a bold font number, and place a bar over it if it's antifundamental. For example the fundamental and antifundamental representations of SU(N) are written as $\mathbf{N}$ and $\overline{\mathbf{N}}$, respectively.

Example 3.7.1. For SU(5), we can write the Young-Tableaux



as

$$\mathbf{5} \otimes \mathbf{5} = \mathbf{15} \oplus \mathbf{10}.$$

Note that $5 \times 5 = 25 = 15 + 10$, so you can always check to see if your answer at least adds up correctly. It is standard convention to list the numbers in decreasing value as we have done.

Exercise

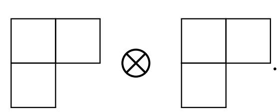
Write the above Young-Tableaux in bold face notation for SU(3). *Hint: Notice something special about the above diagram for SU(3).*

Exercise

Write the Young-Tableaux Equation (3.2) in the bold face notation for SU(6). *Hint: You should get a total dimension of 126.*

Exercise

Verify that $\mathbf{8} \otimes \mathbf{8}$ for SU(3) corresponds to



Comment: We will use this result next lecture, so please actually do this.

4 Decomposing Tensor Products of SU(N)

Ok so we have seen (or at least argued) the importance of irreps to particle physics, and have become comfortable with drawing Young-Tableaux diagrams. We now need to return to the question asked after Equation (3.2). Reworded the question is "given the Young-Tableaux of two irreps, D_{r_1} and D_{r_2}, how do we write their tensor product as a direct sum of irreps?" In other words, how do we take tensor products of Young-Tableaux diagrams? The answer is a procedure known as the *Littlewood-Richardson* (or *Clebsch-Gordan*) rules.

4.1 Littlewood-Richardson Rules

As we did above, let's label our two irreps by D_{r_1} and D_{r_2}, then we find the tensor product $D_{r_1} \otimes D_{r_2}$ via the following procedure.

(i) Draw the Young-Tableaux for D_{r_1} and D_{r_2}.

(ii) Label the rows of D_{r_2} with letters, as indicated in the following example

$$\begin{array}{|c|c|c|c|}\hline a & a & a & a \\ \hline b & b \\ \cline{1-2} c \\ \cline{1-1}\end{array}$$

(iii) Add the boxes of D_{r_2} to D_{r_1} one at a time starting with the first row acording to these following rules:

(a) Augmented Young-Tableaux must be a valid Young-Tableaux.

(b) Boxes with the same label (a, b etc) *cannot* be in the same column (as they are symmetrised in D_{r_2} so if we antisymmetrise the result vanishes).

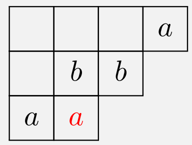
(c) Two or more Young-Tableaux with the same shape *and* the same labels count as one diagram.

(d) Cancel columns with N rows (i.e. remove the invariant parts).

(e) At any given box position, define

$$n_a = \text{number of } a\text{s in the upper-right quadrant from this box},$$

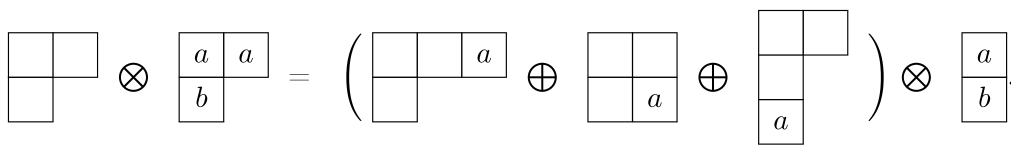
and similarly for n_b etc. Then require $n_a \geq n_b \geq n_c$, and so remove any diagrams that don't obey this, i.e. the following diagram is not valid because the red a has $n_b = 2$ but $n_a = 1$.



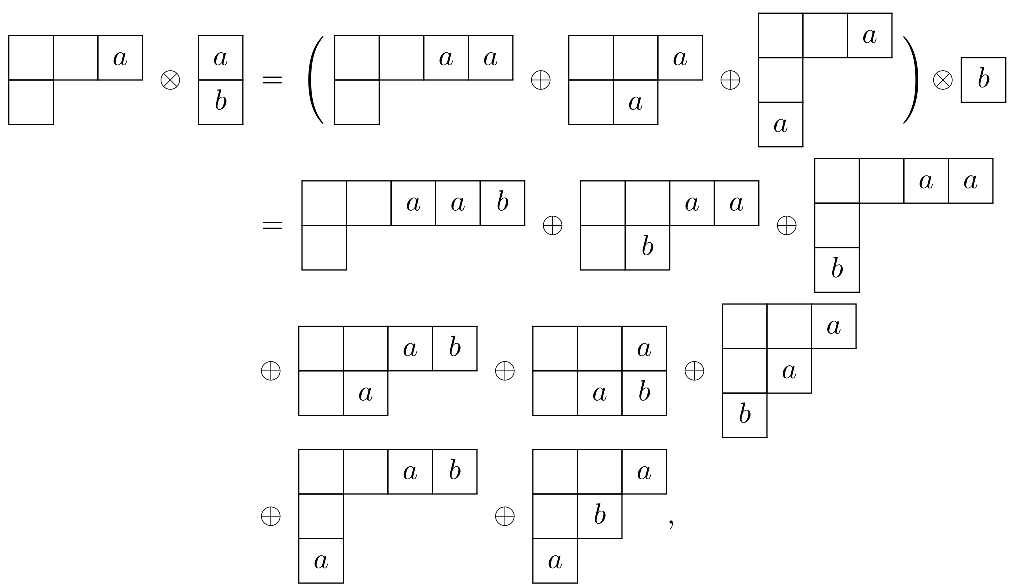
This will probably seem highly cryptic, but it's rather straight forward. Let's give an example.

Example 4.1.1. Let's consider the tensor product $\mathbf{8} \otimes \mathbf{8}$ in SU(3). You showed what this corresponds to as Young-Tableaux at the very end of last lecture, so let's just apply the Littlewood-Richardson rules to find the decomposition.

The first a box gets distributed as

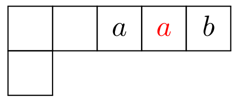
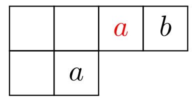
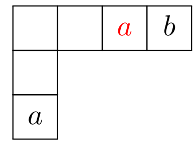


Now we can consider each term on the right-hand side in turn. I will just finish off the first term on the right-hand side and leave the other two as exercises. The only terms we'll cancel (i.e. won't draw) are the ones that don't make valid Young-Tableaux (e.g. 4 columns or more columns then previous row). The rest we'll explain at the end.

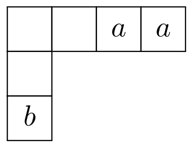
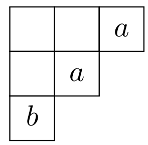
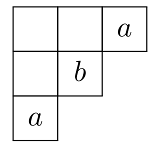


where each row of the calculation corresponds to one term in the brackets.

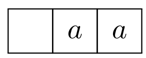
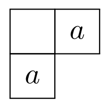
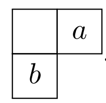
Ok so how do we cancel/simplify this? The following terms go because the red as violate condition (e)

  and 

Then we can simplify these terms using condition (d)

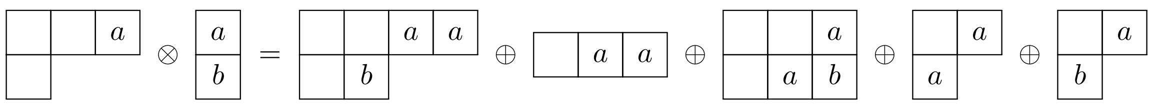
  and 

to become

  and .

Note that the final two diagrams here are different diagrams because, although they have the same shape, they don't have the same label distribution.

So we're left with

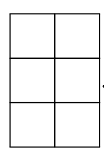


Exercise

Finish the rest of the example and convert it into bold font notation to obtain

$$\mathbf{8} \otimes \mathbf{8} = \mathbf{27} \oplus \mathbf{10} \oplus \overline{\mathbf{10}} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1}$$

Hint: If you have a Young-Tableaux where every column has N rows, then you write it in bold font notation as **1**. *So in this question the* **1** *comes from a diagram of the form*

.

4.2 Final Comment on Young-Tableaux

As we have seen, Young-Tableaux are a very neat and useful trick for the study of SU(N). Other Lie groups, however, are more complicated. For example, for SO(N) the fundamental representation acts as

$$\varphi^i \mapsto M^i{}_j \varphi^j,$$

where $M \in SO(N)$. The decomposition of the tensor product with two indices is not just the symmetric plus antisymmetric. Indeed it turns out that the trace forms an invariant subspace, and so our decomposition is

$$\varphi^i \psi^j = \underbrace{\frac{1}{2}\big(\varphi^i\psi^j + \varphi^j\psi^i\big) - \frac{1}{N}\delta^{ij}\varphi^k\psi^k}_{\text{Symmetric Traceless}} + \underbrace{\frac{1}{N}\delta^{ij}\varphi^k\psi^k}_{\text{Trace}} + \underbrace{\frac{1}{2}\big(\varphi^i\psi^j - \varphi^j\psi^i\big)}_{\text{Antisymmetric}}.$$

We can see that this is the case by showing that the δ^{ij} is an invariant tensor under SO(N):[1]

$$\begin{aligned}\delta^{ij} &\mapsto M^i{}_{i'}M^j{}_{j'}\delta^{i'j'}\\ &= M^i{}_{i'}M^j{}_{i'}\\ &= M^i{}_{i'}(M^T)^{i'}{}_j\\ &= \delta^{ij}\end{aligned}$$

where we have used $MM^T = \mathbb{1}$. This result would not have held for SU(N) because we would need the Hermitian conjugate, not just the transpose.

4.3 Systematic Approach To Irreps

Recall that we can relate a Lie group to a Lie algebra via the exponential map. That is if U is an element of the Lie group, then we can write it as $U = e^X$, where X is an element of the appropriate Lie algebra. The question is can we relate the representation of a Lie group to the representation of a Lie algebra? Well first we need the definition of the representation of a Lie algebra.

Definition. [Representation Of Lie Algebra] Let $(\mathfrak{g}, [,])$ be a Lie algebra of dimension n. Then we obtain a *representation of the Lie algebra* on V, by prescribing a Lie algebra homomorphism, d. That is a map d satisfying: for all $X, Y \in \mathfrak{g}$ and $\alpha, \beta \in \mathbb{C}$[2]

(i) Linearity; $d(\alpha X + \beta Y) = \alpha d(X) + \beta d(Y)$, and

(ii) $d\big([X,Y]\big) = [d(X), d(Y)] = d(X)\circ d(Y) - d(Y)\circ d(X)$, where $\circ$ is the composition as maps.

For the time being we shall assume our representations are matrices, in order to compare to the stuff we've been saying for representations of Lie groups. We will actually deter from this when we introduce the adjoint representation later.

Proposition 4.3.1. *We can obtain a representation on the Lie algebra given one on the Lie group, via the exponential map, defined via*

$$D(e^X) = e^{d(X)}, \tag{4.1}$$

provided d is linear.

Proof. Let D be the representation of our Lie group. Then let

$$U = e^X, \quad \text{and} \quad V = e^Y$$

be two arbitrary elements in the Lie group, with X and Y being elements of the corresponding Lie algebra. Then we have

$$D\big(e^X\big)D\big(e^Y\big) = e^{d(X)}e^{d(Y)} = e^{d(X)+d(Y)+\frac{1}{2}[d(X),d(Y)]+\dots},$$

[1]In SO(N) upper and lower indices aren't different.

where we have used the BCH formula. Then use the fact that D is a representation,

$$D\big(e^X\big)D\big(e^Y\big) = D\big(e^X e^Y\big) = D\big(e^{X+Y+\frac{1}{2}[X,Y]+\dots}\big) = e^{d(X+Y+\frac{1}{2}[X,Y]+\dots)}.$$

Finally use the fact that d is a linear map to obtain

$$e^{d(X)+d(Y)+\frac{1}{2}[d(X),d(Y)]+\dots} = e^{d(X)+d(Y)+\frac{1}{2}d([X,Y])+\dots},$$

which gives us condition (ii). □

Definition. [Equivalent Representations Of Lie Algebras] Let $(\mathfrak{g}, [,])$ be a Lie algebra and let d_1 and d_2 be two representations. Then we say d_1 and d_2 are *equivalent* if there exists a constant matrix S such that

$$d_2(X) = Sd_1(X)S^{-1}, \qquad \forall X \in \mathfrak{g}.$$

Definition. [Reducible Representations Of Lie Algebras] We say a representation d of a Lie algebra is *reducible* if it is equivalent to a block diagonal matrix[3]. We can also define it as the condition that it can be written as the direct sum of irreps:[4] e.g.

$$d = d_a \oplus d_b.$$

Proposition 4.3.2. *Let D and $\widetilde{D}$ be two equivalent representations of a Lie group, i.e.*

$$\widetilde{D}(g) = SD(g)S^{-1} \qquad \forall g \in G.$$

Then their associated Lie algebra representations, d and $\widetilde{d}$, are also equivalent.

Proof. This proof relies on the fact that

$$(SAS^{-1})^n = SA^nS^{-1}$$

for any matrix A. Simply consider the definitions:

$$\begin{aligned}
\widetilde{D}\big(e^X\big) &= SD\big(e^X\big)S^{-1} \\
e^{\widetilde{d}(X)} &= Se^{d(X)}S^{-1} \\
&= S\Big(\sum_{n=0}^{\infty}\frac{(dX)^n}{n!}\Big)S^{-1} \\
&= \sum_{n=0}^{\infty}\frac{(Sd(X)S^{-1})^n}{n!} \\
&= e^{Sd(X)S^{-1}}, \\
\implies \widetilde{d}(X) &= Sd(X)S^{-1} \qquad \forall X \in \mathfrak{g}.
\end{aligned}$$

□

Exercise

Using the block diagonal matrix version for reducible representations of Lie groups, Equation (2.9), show that reducible representations of Lie groups correspond to reducible representations of Lie algebras. In other words, show

$$\exp\left(\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\right) = \begin{pmatrix} e^A & 0 \\ 0 & e^B \end{pmatrix}$$

Exercise

Show that unitary representations of Lie groups correspond to antihermitian representations of Lie algebras. That is

$$\big[D(g)\big]^\dagger = [D(g)]^{-1} \qquad \Longleftrightarrow \qquad \big[d(X)\big]^\dagger = -d(X).$$

So how do the representations of Lie algebras act on the representation space? The answer is

$$d(X) : \psi^{i_1\ldots i_n} \mapsto X^{i_1}{}_j\psi^{ji_2\ldots i_n} + X^{i_2}{}_j\psi^{i_1 j\ldots i_n} + \ldots + X^{i_n}{}_j\psi^{i_1\ldots i_{n-1}j} \tag{4.2}$$

We set the proof (for $n = 2, 3$) as an exercise here[5]

Exercise

If $D(U) : \phi^{ij} \mapsto U^i{}_r U^j{}_s \phi^{rs}$ find the action of the corresponding Lie algebra by putting $U^i{}_j = \delta^i{}_j + \epsilon u^i{}_j$ and considering $\mathcal{O}(\epsilon)$ terms. Similarly write down the action of the Lie algebra on ϕ^{ijk}.

The important point about Lie algebras to understand is that, unlike Lie groups, they are vector spaces and so they have a basis. Putting this together with the fact that the representation map d is linear, we see that for Lie algebras we can find the entire representation by simply knowing it for a basis! This is an extremely useful property. For example it makes dealing with Schur's Lemma much easier. Note, however, that because we do not require d to be invertible it is not generally true that the representation algebra and Lie algebra have the same dimension. That is, its possible that d maps two basis vectors to the same element in the representation, which would give the representation a lower dimension then the Lie algebra itself. There is always a privileged representation which is the Lie algebra itself, this is the topic of the next section.

4.4 The Adjoint Representation

Definition. [Adjoint Representation] Let $[\mathfrak{g}, [,]]$ be a Lie algebra. Then the linear map

$$\begin{aligned} ad : \mathfrak{g} &\to \mathfrak{g} \\ X &\mapsto ad(X), \end{aligned}$$

[5] As it is set as one on Dr. Dorigini's course, and on the off change someone is reading this while doing his course I don't want to just give the answer.

defined via its action:

$$ad(X) : Y \mapsto [X, Y], \tag{4.3}$$

is a representation, called the *adjoint representation*.

Exercise

Prove that the adjoint representation is indeed a representation. That is show that $ad(X)$

(i) Is linear:

$$ad(X)(\alpha Y + \beta Z) = \alpha ad(X)(Y) + \beta ad(X)(Z).$$

(ii) Preserves the commutator:

$$ad_X([Y, Z]) = [ad_X(Y), ad_X(Z)].$$

Remark 4.4.1. Note the linearity condition above is linearity in the argument of $ad(X)$, i.e. in Y not in X itself. It is true that $ad(X)$ is also linear in X (as the commutator is bilinear). For this reason we could define the bilinear map

$$ad : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

as the commutator. This is *not* a representation though as the representation only maps from one copy of $\mathfrak{g}$. To be totally clear it is the *whole* $ad(X)$ that is the representation, not just ad.

Note that unlike the other representations we have considered so far, the adjoint representation does not give a matrix. It is just a linear map, which is all we need for a representation, as stated way back in Remark 2.1.1. We can, though, extract a matrix form for the adjoint representation as follows. We know that the representation is a vector space and has a basis, $\{X_a\}$, and we know that the Lie bracket of two elements is an element itself. This just gives us the structure constants, Equation (1.11). So we can use these structure constants to construct a matrix. To be more clear, we have

$$ad(X_a) : X_b \mapsto [X_a, X_b] = f_{ab}{}^c X_c,$$

so as think of the adjoint representation in terms of the matrices

$$(T_a)_b{}^c := \big(ad(X_a)\big)_b{}^c = f_{ab}{}^c.$$

For a given a this is a $\dim \mathfrak{g} \times \dim \mathfrak{g}$ matrix.

Exercise

Suppose that the structure constants of a Lie algebra $\mathfrak{g}$ in a basis $\{X_a\}$ are $f_{ab}{}^c$. Now switch to a new basis $\{X'_a\}$, related to the old one by $X'_a = S^b{}_a X_b$, where $S^b{}_a$ is a nonsingular matrix. Show that in the new basis the structure constants are

$$f'_{ab}{}^c = S^p{}_a S^q{}_b (S^{-1})^c{}_r f_{pq}{}^r.$$

Comment: Again this is something taken straight from the problem sheets to Dr. Dorigoni's course. However I think it's a useful exercise so have included it here.

4.4.1 Killing form

As we have said many times a Lie algebra is a vector space and this has given us many nice results. However a Lie algebra is even nicer than this: it also comes with a natural *inner product.*

Definition. [Killing Form] Let $(\mathfrak{g}, [,])$ be a Lie algebra. Then we can define an inner product, called the *Killing form* (or *Cartan metric*) by

$$B(X,Y) := \text{Tr}\left(ad(X)\cdot ad(Y)\right), \tag{4.4}$$

where $ad(X)$ is the matrix representing X.

We can write the Killing form in components as

$$\begin{aligned} B(X_a, X_b) &= \text{Tr}\left([ad(X_a)ad(X_b)]_c{}^e\right) \\ &= \text{Tr}\left([ad(X_a)]_d{}^e[ad(X_b)]_c{}^d\right) \\ &= f_{ad}{}^c f_{bc}{}^d \\ &=: g_{ab}. \end{aligned}$$

Note that $g_{ab} = g_{ba}$, which we expect from an inner product.

4.4.2 Casimir Operator

An important application of the Killing form is what is known as the *Casimir operator.*

Definition. [Casimir Operator] Let g_{ab} be the components of the Killing form for a Lie algebra $(\mathfrak{g}, [,])$. Then **if** the Killing form is invertible we can define

$$g^{ab} := (g^{-1})_{ab}.$$

In **any** representation, d, we can then define the *Casimir operator*

$$C_d = \sum_{a,b=1}^{\dim\mathfrak{g}} g^{ab}\cdot d(X_a)\cdot d(X_b), \tag{4.5}$$

where $\{X_a\}$ is a basis for the Lie algebra.

It is a fact that the Casimir operator commutes with all elements in the representation. As the Lie bracket is linear, we can write this as

$$[C_d, d(X_a)] = 0 \qquad \forall a \in \{1, ..., \dim \mathfrak{g}\}.$$

This is a very powerful result; if d is an irrep then we know, from Schur's Lemma, that

$$C_d = \lambda \cdot \mathbb{1}.$$

Example 4.4.2. As an example, consider $\mathfrak{su}(2)$. Here the Killing form is

$$g_{ab} = -2\delta_{ab}. \tag{4.6}$$

This is invertible, and we obtain

$$g^{ab} = -\frac{1}{2}\delta^{ab}.$$

So here the Casimir is given by

$$C = -\frac{1}{2}\Big(\big[d(X_1)\big]^2 + \big[d(X_2)\big]^2 + \big[d(X_3)\big]^2 \Big).$$

This tells you that whenever d is an irrep of SU(2) the Casimir is a multiple of the identity. This is often written as

$$J^2 = J_1^2 + J_2^2 + J_3^2 = \lambda \mathbb{1},$$

to make the connection with the angular momentum of a particle. We will see this more in detail soon.

Exercise

Prove Equation (4.6). *Hint: Use*

$$[X_a, X_b] = \sum_c \epsilon_{abc} X_c$$

for $\mathfrak{su}(2)$*, where* ϵ_{abc} *is the Levi-Civita tensor.*

5 | Systematic Approach To Finite Dimensional Irreps

We now want to give some systematic approach to getting the irreps of finite dimensional $\mathfrak{su}(N)$. We will consider $\mathfrak{su}(2)$ and $\mathfrak{su}(3)$.

5.1 $\mathfrak{su}(2)$

Recall that the Lie algebra $(\mathfrak{su}(2), [,])$ is the set of 2×2, antihermitian,[1] traceless matrices, and the Lie bracket is the commutator. A basis for such matrices are the *Pauli matrices*, Equation (1.10). We shall actually scale the matrices slightly, and use the basis

$$\begin{aligned} \tau_1 &= -i\frac{\sigma_1}{2} = \begin{pmatrix} 0 & -i/2 \\ -i/2 & 0 \end{pmatrix}, \\ \tau_2 &= -i\frac{\sigma_2}{2} = \begin{pmatrix} 0 & -1/2 \\ 1/2 & 0 \end{pmatrix}, \\ \tau_3 &= -i\frac{\sigma_3}{2} = \begin{pmatrix} -i/2 & 0 \\ 0 & i/2 \end{pmatrix}. \end{aligned} \tag{5.1}$$

The reason we use these is because now the commutation relation becomes

$$[\tau_i, \tau_j] = \epsilon_{ijk}\tau_k. \tag{5.2}$$

This is a useful basis but there is a different one which makes connection with QFT easier, which is something we ultimately want to do (as this is a course for particle physicists). Recall that in QM and QFT we have raising and lowering (or ladder) operators which increase the eigenvalues of a given operator. The question is, can we do something similar with $\mathfrak{su}(2)$? The answer is yes,[2] and is given by

$$\begin{aligned} E_+ &:= i\tau_1 - \tau_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\ E_- &:= i\tau_1 + \tau_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \\ H &:= 2i\tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned} \tag{5.3}$$

[1] Note we have actually changed convention here compared to Lecture 1, where we had Hermitian matrices.
[2] Otherwise I wouldn't be saying all this

Exercise

Show that
$$[H, E_\pm] = \pm 2E_\pm, \quad \text{and} \quad [E_+, E_-] = H. \tag{5.4}$$

Remark 5.1.1. Note that Equation (5.1) are complex matrices but Equation (5.3) are real matrices. This might seem like a problem for the latter to be a basis, but we have to remember that our underlying field is the complex numbers, so we can still span the whole space.

Ok, so now let's consider a representation d. What form do our new basis elements take? Well we note that H is unitary, so we can expect the representation $d(H)$ be unitary too. Now it's a fact that any unitary matrix is diagonalisable.[3] That is we can always find an equivalent representation $\widetilde{d}(H) = Sd(H)S^{-1}$ such that we get a diagonal matrix. We shall therefore always do this.

As $d(H)$ is diagonal, we can construct the representation space such that each element is an eigenvector of $d(H)$. That is $d(H)$ is a $\dim d \times \dim d$ matrix, so we can construct our representation space as a $\dim d$ column matrix. This obviously smells a lot like QM, and so we use bra-ket notation. As $d(H)$ is unitary, it is Hermitian, and so we know the eigenvalues are real. We shall also assume that these eigenvalues are unique, so we can label the eigenstates by their eigenvalues. That is, we write

$$d(H) |k\rangle = k |k\rangle . \tag{5.5}$$

So what about the action of $E_\pm$ on our states? Well, that's an exercise.

Exercise

Show that
$$d(H)d(E_\pm) |k\rangle = (k \pm 2)d(E_\pm) |k\rangle .$$

Hint: Using Equation (5.4).

The result of this exercise tells us that

$$d(E_\pm) \propto |k \pm 2\rangle ,$$

the question is "what are the proportionality constants?" Well we rescale our system such that
$$d(E_-) |k\rangle = |k - 2\rangle . \tag{5.6}$$

So we just need to find the coefficient for $d(E_+)$. As the title of this lecture says, we want to consider *finite* dimensional representations, and so we require there to be some bound on the values of k. In particular we require that there is some maximum value, $k = j$, such that

$$d(E_+) |j\rangle = 0. \tag{5.7}$$

[3] This is a consequence of something called the *spectral theorem*. For more details see Simon and my notes on Dr. Schuller's QM course, available on my blog site.

We call this state the *highest weight state*. We can use this to find the action on a general state. We get a recursion relation as follows: let $d(E_+)\,|k\rangle = r_{k+2}\,|k+2\rangle$, then we have

$$\begin{aligned} d(E_+)\,|k-2\rangle &= d(E_+)d(E_-)\,|k\rangle \\ &= \big(d(H) + d(E_-)d(E_+)\big)\,|k\rangle \\ &= (k + r_{k+2})\,|k\rangle\,, \end{aligned}$$

giving the relation

$$r_k = \begin{cases} k + r_{k+2} & k \neq j+2 \\ 0 & k = j+2. \end{cases}$$

This is solved by

$$r_{j-2k} = (k+1)(j-k). \tag{5.8}$$

To clarify,[4] get the highest weight state by setting $k = -1$:

$$r_{j+2} = r_{j-2(-1)} = (-1+1)(j+1) = 0.$$

We see that we also have $r_{-j} = 0$, which corresponds to the fact we must also put a lower bound on the values of k. In terms of the lowering operator this is the statement that

$$d(E_-)\,|{-j}\rangle = 0. \tag{5.9}$$

This gives us a weight 'lattice':

$$\begin{tikzcd} -j \arrow[r, bend left] & -j+2 \arrow[l, bend left] \arrow[r, bend left] & -j+4 \arrow[l, bend left] & \overset{d(E_+)}{\underset{d(E_-)}{\cdots}} & j-4 \arrow[r, bend left] & j-2 \arrow[l, bend left] \arrow[r, bend left] & j \arrow[l, bend left] \end{tikzcd}$$

The conclusion we draw from this result is that for each value of j we have a $(j+1)$-dimensional irrep with basis elements

$$|j\rangle\,, |j-2\rangle\,, ..., |{-j}+2\rangle\,, |{-j}\rangle\,,$$

which we can write as a column matrix explicitly as

$$|j\rangle = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad ... \qquad |{-j}\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Our matrices $H, E_\pm$ take the form[5]

$$d(H) = \begin{pmatrix} j & & 0 \\ & \ddots & \\ 0 & & -j \end{pmatrix}, \qquad d(E_-) = \begin{pmatrix} 0 & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{pmatrix} \qquad d(E_+) = \begin{pmatrix} 0 & r_j & & 0 \\ & \ddots & r_{j-2} & \\ & & \ddots & \ddots \\ 0 & & & 0 \end{pmatrix}.$$

[4] As this took me a few minutes to see.

[5] I wasn't sure how to make the 0s big, but basically everything blank is a 0.

Remark 5.1.2. There is a nice way to convert these irreps into Young-Tableaux. We're considering $SU(2)$, so, as we described in Section 3.6.1, we can categorise any Young-Tableaux by its dimension. An irrep with dimension n has $(n-1)$ boxes. So, from the fact that the dimension of our irreps are $(j+1)$, our Young-Tableaux are just j horizontal boxes.

Example 5.1.3. Let's consider the example of $j=1$, then we have

$$d(H) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad d(E_-) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad d(E_+) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

which are exactly Equation (5.3), so this is the fundamental representation. This agrees with the remark above as we expect the Young-Tableaux to just be a single box, which is the fundamental representation. Our states are $|\pm 1\rangle$.

Example 5.1.4. Now let's consider $j=2$. Here we have

$$d(H) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \qquad d(E_-) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad d(E_+) = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

We have three states $|\pm 2\rangle$ and $|0\rangle$. Our Young-Tableaux here is simply

$$\begin{array}{|c|c|}\hline & \\ \hline \end{array}\,.$$

Exercise

Check that the Casimir Equation (4.5) is indeed a multiple of the identity for the irreps of $j=1,2$. *Hints: 1) We have already found the Killing form in Example 4.4.2. 2) Be careful: the Equation (4.5) is expressed in terms of $d(\tau)$s, you need to convert this into $d(H)/d(E_\pm)$ first.*

The generalisation of the above exercise for general j is

$$C = \frac{j}{2}\left(\frac{j}{2}+1\right)\frac{\mathbb{1}}{2}.$$

To a quantum physicist this whole lecture will have looked *very* familiar, and this last result in particular; recall that the spin operator acts as

$$S^2 = s(s+1)\frac{\mathbb{1}}{2},$$

so we see that j is twice the spin. Equally $d(H)$ is playing the role of $2S_z$.

Remark 5.1.5. Of course we could have divided j by two everywhere and obtained exactly the spin, however we have been using a mathematician's convention and they prefer to carry 2s around then 1/2s.

As a final comment before moving on to SU(3), let's just make a comment on how you relate the index notation $\phi^{(ij\dots)}$ to kets. We do it for $j=1$ and set $j=2$ as an exercise.[6]

[6] Again this is because it's set as an exercise on the course and I don't want to put the answers on here. If you don't understand what I did below please feel free to email me for clarity.

Example 5.1.6. For $j = 1$ we have a single index, which transforms as

$$d(X) : \phi^i \mapsto X^i{}_j \phi^j.$$

Now consider the action of $d(H)$: we have $H = \text{diag}(1, -1)$, so we get

$$\begin{aligned} d(H) : \phi^1 &\mapsto H^1{}_j \phi^j = \phi^1 \\ d(H) : \phi^2 &\mapsto H^2{}_j \phi^j = -\phi^2, \end{aligned}$$

so we relate

$$\phi^1 \sim |1\rangle, \quad \text{and} \quad \phi^2 \sim |-1\rangle,$$

and obtain

$$d(H) = \text{diag}(1, -1),$$

which is just the fundamental representation, in agreement with the previous comments.

Exercise

Repeat the calculation above but now for $j = 2$ to obtain

$$\phi^{11} \sim |2\rangle, \qquad \phi^{(12)} \sim |0\rangle, \quad \text{and} \quad \phi^{22} \sim |-2\rangle.$$

This tells us that

$$d(H) = \text{diag}(2, 0, -2),$$

which we agrees with what we wrote before.

5.2 $\mathfrak{su}(3)$

Let's now consider the Lie algebra of SU(3). We have seen that this is the set of 3×3, traceless, hermitian matrices. The commonly used basis for this space are the so-called *Gell-Mann* matrices:

$$\begin{aligned} &\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \qquad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\ &\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \text{and} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \end{aligned} \tag{5.10}$$

An observant person might realise that these matrices have the Pauli matrices (i.e. the basis elements of $\mathfrak{su}(2)$) embedded in them. For example λ_1, λ_2 and λ_3 contain exactly the Pauli matrices in the top left corner.

Claim 5.2.1. We can group the Gell-Mann matrices into three groups, each of which obeys an $\mathfrak{su}(2)$ algebra (i.e. the structure constants is the Levi-Civita tensor). The groups are

(i) λ_1, λ_2 and λ_3,

(ii) λ_4, λ_5 and $\frac{1}{2}(\sqrt{3}\lambda_8 + \lambda_3)$, and

(iii) λ_6, λ_7 and $\frac{1}{2}(\sqrt{3}\lambda_8 - \lambda_3)$.

Proof. This can easily be checked just by calculating all the commutation relations, but we gain little insight by doing this, so just state its true here.[7] □

We now want to do a similar thing to the $\mathfrak{su}(2)$ case and use a smart basis that corresponds to raising and lowering operators. We have a bit more of a challenge here though, as we have 3 $\mathfrak{su}(2)$s to consider. Luckily the result is known so, as if by magic, we just state it. We label each $\mathfrak{su}(2)$ by α, β and $(\alpha + \beta)$ with

$$H_{\alpha+\beta} = H_\alpha + H_\beta, \quad \text{and} \quad E_{\pm(\alpha+\beta)} = [E_{\pm\alpha}, E_{\pm\beta}].$$

Explicitly we get

$$H_\alpha = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad E_\alpha = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad E_{-\alpha} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$H_\beta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad E_\beta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad E_{-\beta} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$H_{\alpha+\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad E_{\alpha+\beta} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad E_{-\alpha-\beta} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

As the notation suggests, the idea is that each label forms one $\mathfrak{su}(2)$ group, and the $E_\pm$s are the raising and lowering operators within each group. We call α and β *simple roots*, whereas $(\alpha + \beta)$ is a *non-simple root*.

We can just focus on the simple roots (as the non-simple ones are obtainable from simple ones). First note that

$$[H_\alpha, H_\beta] = 0,$$

and so, from the definition of a representation of a Lie algebra,

$$[d(H_\alpha), d(H_\beta)] = 0.$$

This tells us that we can simultaneously diagonalise both $d(H_\alpha)$ and $d(H_\beta)$, as they are both unitary. We then proceed as before to label the states of the representation space, but now we have two weights (i.e. eigenvalues) to keep track of. We define the states by

$$d(H_\alpha)\,|m,n\rangle = m\,|m,n\rangle\,, \quad \text{and} \quad d(H_\beta)\,|m,n\rangle = n\,|m,n\rangle\,. \tag{5.11}$$

[7] By all means feel free to check yourself.

Now we want to define the action of the lowering operators, $d(E_{-\alpha})$ and $d(E_{-\beta})$, on these states as before. First consider $d(E_{-\alpha})$: we know

$$[d(H_\alpha), d(E_{-\alpha})] = -2d(E_{-\alpha}),$$

which, following the calculation from the previous section, tells us that $d(E_{-\alpha})$ lowers the value of m by 2. The question is "does it effect n?"

Exercise

Verify that
$$[d(H_\beta), d(E_{-\alpha})] = d(E_{-\alpha}).$$
Use this to show that $d(E_{-\alpha})$ increases the value of n by 1. *Hint: You can just find the commutator of H_β and $E_{-\alpha}$ and then use the definition of d to obtain the above result.*

Putting the result of the above exercise together with the comment just before it, and making a similar argument for $d(E_{-\beta})$, we rescale our states so that they satisfy

$$d(E_{-\alpha})\,|m,n\rangle = |m-2,n+1\rangle\,, \quad \text{and} \quad d(E_{-\beta})\,|m,n\rangle = |m+1,n-2\rangle\,. \tag{5.12}$$

We define our highest weight state by the condition

$$d(E_\alpha)\,|m,n\rangle = 0 = d(E_\beta)\,|m,n\rangle\,.$$

From this condition and Equation (5.12), we can again produce the whole weight lattice with lowest weight state
$$d(E_{-\alpha})\,|\widetilde{m},\widetilde{n}\rangle = 0 = d(E_{-\beta})\,|\widetilde{m},\widetilde{n}\rangle\,.$$

Definition. [Root Lattice] We define the *root lattice* to the be the lattice of all the states of simple roots.

Example 5.2.2. Consider the fundamental representation. We can find the states by considering the index expressions. We have

$$d(H_\alpha) : \phi^i \mapsto (H_\alpha)^i{}_j\phi^j,$$

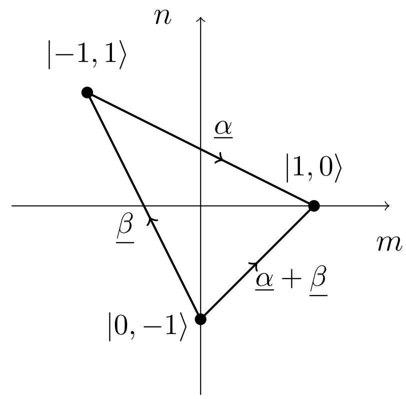
and similarly for $d(H_\beta)$. Therefore, using the matrix expressions above, we get

$$\begin{aligned} d(H_\alpha) &: \phi^1 \mapsto \phi^1 \\ d(H_\alpha) &: \phi^2 \mapsto -\phi^2 \\ d(H_\alpha) &: \phi^3 \mapsto 0. \end{aligned}$$

Doing the same thing for $d(H_\beta)$ gives the states

$$|1,0\rangle\,, \qquad |-1,1\rangle\,, \quad \text{and} \quad |0,-1\rangle\,. \tag{5.13}$$

The first/last is the highest/lowest weight state, respectively. The root lattice is depicted below.



We have indicated the states on the diagram. The raising/lowering operators move you from point to point on the root lattice, going with/against the vectors

$$\underline{\alpha} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad \text{and} \quad \underline{\beta} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}. \tag{5.14}$$

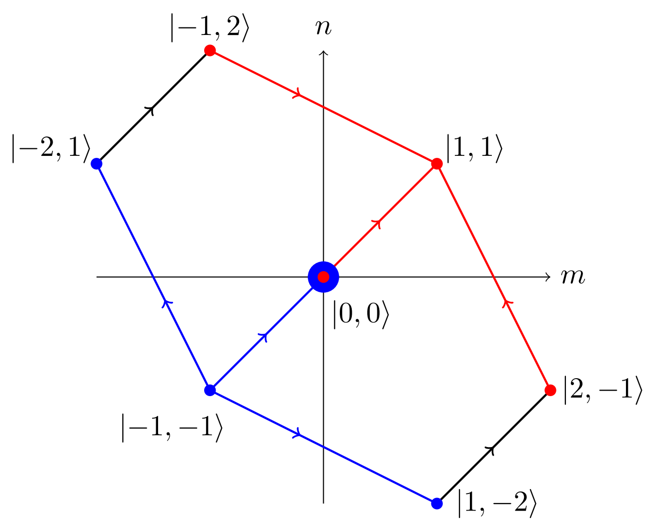
That is, for example,

$$d(E_\alpha) : |-1,1\rangle \mapsto |1,0\rangle\,, \quad \text{and} \quad d(E_{-\beta}) : |-1,1\rangle \mapsto |0,-1\rangle\,.$$

Exercise

Finish obtaining Equation (5.13), i.e. do the $d(H_\beta)$ part.

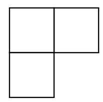
Example 5.2.3. Now let's consider the representation which has the highest/lowest weight states $|1,1\rangle\,/\,|-1,-1\rangle$, respectively. We don't know the expressions for $d(H_{\alpha/\beta})$ here,[8] but we can obtain the root lattice by plotting these two states and applying the raising and lowering operators, i.e. use Equation (5.14). We get the following diagram.



[8]Note we could obtain them using the index aproach, but I think you'd agree the method used here is a lot faster.

I have tried to make it clear how you obtain the points: the red points/arrows are the highest weight state and the lowering operators acting on it; the blue points/arrows are the lowest weight state and the raising operators acting on it. The black arrows are included to show you can 'close' the diagram using $d(E_{\pm(\alpha+\beta)})$.

Note that we can get to the origin in two different ways. We count these as two separate states, so in total we have 8 states. This tells us the the dimension of the representation space is 8, which we can use to obtain the Young-Tableaux:

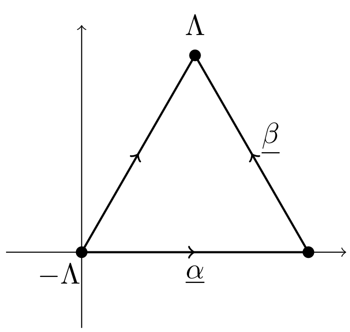


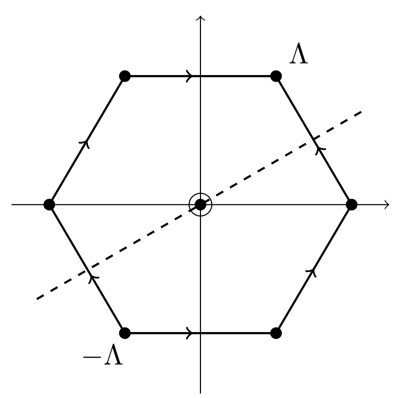
which for SU(3) does indeed have dimension 8. We can do a similar thing for the previous example (with dimension 3) to get the single box Young-Tableaux, which is the fundamental representation, as required.

Remark 5.2.4. Note that in the root lattice diagrams we can identify the highest/lowest weight states by looking where the arrows point to/away from. This is because the arrows representing raising, so they all point towards the highest weight state and away from the lowest weight state. Combining this with the Young-Tableaux argument given at the end of the last example, we see how much information is really packed into these diagrams!

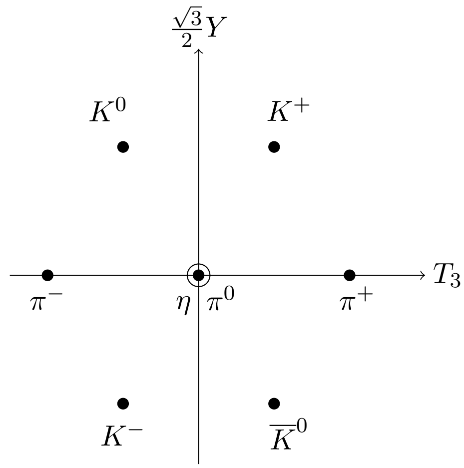
5.2.1 The Eightfold Way

The above remark just made a point about how much information is contained in these diagrams, however it seems a shame that they're not very nice shapes. By which I mean, both of them are squashed versions of nice shapes (i.e. an equilateral triangle and a hexagon). The question is: "can we make them look nicer?" The answer is yes, and we will do this next lecture, but here's the basic idea. There root space comes with a metric, and as we've drawn them the metric is not in some nice form. We make the diagrams look nicer by considering a change of basis, making the metric into the Euclidean metric. This will make the above two diagrams look like the following. In both diagrams, Λ labels the highest weight state and $-\Lambda$ the lowest weight state. The dashed line is explained in a minute.

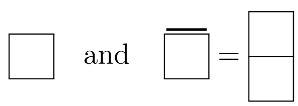




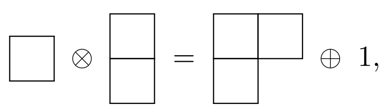
The hexagon diagram has direct relation to particle physics. The story goes (roughly) as follows: in the 1950s particle physicists were trying to work out the symmetries of the strong force. After a lot of work they realised that the combination of *isospin* and *strangeness* were (almost) conserved by the strong interactions. They also found that certain hadrons with the same spin had (almost) degenerate masses, which suggested a symmetry. If you plot the third component of isospin, T_3, against the so-called *hyper charge*,[9] Y, you got the following diagram:



Hmm... this looks awful familiar. This lead them to the idea that the fundamental objects are quarks/antiquarks, which transform in the fundamental/antifundamental representations of SU(3), i.e.



are the quark and antiquark respectively. Mesons (a quark-antiquark pair) are therefore given by



or

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

This gives us exactly the hexagon diagram above. This result is often referred to as the *eightfold way*. As the diagram corresponds to an irrep, by the argument made at the end of lecture 2, we see that these things have the same mass![10] So these root diagrams have very physical importance for us.

Remark 5.2.5. The dashed line on the hexagon diagram represents a physical symmetry known as Weyl symmetry.

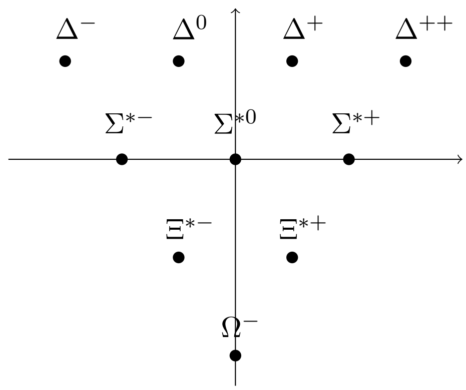
Similarly for baryons (which are 3 quarks) we get the decomposition

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1}.$$

[9] Baryon number + strangeness.

[10] It turns out these things do actually have different mass. This comes from the fact that the different generations of quarks form separate invariant subspaces of the $SU(2)_L$ part of the standard model. Therefore the different generations of quarks have different masses which ruins the result here.

The two **8**s correspond to hexagons as above, while the **10** corresponds to the following big triangle, known as the *baryon decuplet*.



6 Lorentz Group & Cartan Classification

Remark 6.0.1. Unfortunately this lecture is the one that is must heavily bashed by the lack of time on the course, so a lot of the material is sort of brushed over or set as exercises. I shall try and flush out this lecture with additional information to help clarify things, however as with previous exercises, I won't type the answers to any of the exercises set in Dr. Dorigoni's notes/problem sheets.

We saw last lecture that $\mathfrak{su}(2)$ appeared to be some kind of 'building block' for other Lie algebras. In this lecture we are going to show that this is actually more powerful than just the case of $\mathfrak{su}(3)$ discussed in the previous lecture, by first considering the Lorentz group and then touching on Cartan's classification.

6.1 Lorentz Group

The Lorentz group, denoted $\mathrm{SO}(3,1)$,[1] is the group whose elements are 4×4, invertible matrices which we denote by Λ. We write their action on elements in $\mathbb{R}^4$ as

$$X'^{\mu} = \Lambda^{\mu}{}_{\nu} X^{\nu}.$$

They preserve the pseudo-inner product on $\mathbb{R}^4$, i.e.

$$X'^{\mu}\eta_{\mu\nu}X'^{\nu} = X^{\mu}\eta_{\mu\nu}X^{\nu}, \tag{6.1}$$

where we use signature

$$\eta_{\mu\nu} = \mathrm{diag}(1,-1,-1,-1).$$

Physically the Lorentz group corresponds to spatial rotations and Lorentz boosts.

Remark 6.1.1. The Lorentz group is an example of what are known as *non-compact groups*. We will not discuss technically what this means here but simply say that it corresponds to the range of the parameters being open intervals.[2] This is the case for the Lorentz group because we can only boost asymptotically to the speed of light. That is the parameter $\beta := v/c$ has range $\beta \in (-1,1)$, which is open. We will return to this fact shortly.

[1] Some people write $\mathrm{SO}(1,3)$, it doesn't matter, the numbers just indicate the number of $+$s and $-$s in the metric.

[2] See a book on topology for more details on compact spaces.

Exercise

Use Equation (6.1) to show that

$$\Lambda^{\rho}{}_{\mu}\eta_{\rho\tau}\Lambda^{\tau}{}_{\nu} = \eta_{\mu\nu}.$$

This is often given as a defining property of the Lorentz group. Use this result to show that

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \epsilon\omega^{\mu}{}_{\nu}$$

is the infinitesimal deviation from the identity element in SO(3,1), provided $\omega_{\mu\nu} = -\omega_{\nu\mu}$.

The previous result tells us about the Lie algebra. The generators are the $\omega^{\mu}{}_{\nu}$s, and the antisymmetry condition tells us that the dimension is $d = \frac{4(4-1)}{2} = 6$. These are the three spatial rotations and the three boosts, and you can show they obey the commutation relations

$$\begin{aligned}[J_i, J_j] &= \epsilon_{ijk}J_k \\ [J_i, K_j] &= \epsilon_{ijk}K_k \\ [K_i, K_j] &= -\epsilon_{ijk}J_k,\end{aligned} \tag{6.2}$$

where the Js are the generators of spatial rotations and the Ks are generators of boosts.

6.1.1 Smart Basis

The first relation in Equation (6.2) looks just like a $\mathfrak{su}(2)$, but the second two mess it all up. The question is "can we change coordinates in such a way as to produce two[3] sets of $\mathfrak{su}(2)$?" The answer is yes, and is accomplished by defining

$$N_i := \frac{1}{2}(J_i - iK_i) \quad \text{and} \quad \overline{N}_i := \frac{1}{2}(J_i + iK_i) \tag{6.3}$$

Claim 6.1.2. The expressions Equation (6.3) form a basis for $\mathfrak{so}(3,1)$ and obey the commutation relations

$$\begin{aligned}[N_i, N_j] &= \epsilon_{ijk}N_k \\ [\overline{N}_i, \overline{N}_j] &= \epsilon_{ijk}\overline{N}_k \\ [N_i, \overline{N}_j] &= 0.\end{aligned} \tag{6.4}$$

Exercise

Prove the above claim.

The above claim gives us exactly what we wanted, two separate copies of $\mathfrak{su}(2)$ embedded in $\mathfrak{so}(3,1)$. We write this as

$$\mathfrak{so}(3,1) = \underbrace{\mathfrak{su}(2)_L}_{N} \times \underbrace{\mathfrak{su}(2)_R}_{\overline{N}},$$

[3]Note we know it's two because the dimension is 6 and each $\mathfrak{su}(2)$ has dimension 3.

where the L/R stand for left/right, respectively. The reason for this will become clear in just a moment. Therefore the irreps of the Lorentz algebra are completely specified once we specify the irreps of the two $\mathfrak{su}(2)$s. This is brilliant because in lecture 3 we classified all the representations of SU(2) (which we can convert into representations of the Lie algebra). They were specified by a single integer, j, related to the dimension $d = j + 1$. So we can categorise all of the representations of $\text{SO}(3,1)$ using two integers, (j_1, j_2), with dimension $d = (j_1 + 1)(j_2 + 1)$. We give some examples below.[4]

(j_1, j_2)	Name	Symbol	Dimension	Young-Tableaux
$(0,0)$	Scalar	ϕ	1	$1_L \times 1_R$
$(1,0)$	Left-handed Weyl Spinor	ψ_α	2	$\square_L \times 1$
$(0,1)$	Right-handed Weyl Spinor	$\overline{\psi}_{\dot{\alpha}}$	2	$1_L \times \square_R$
$(1,1)$	Vector	$A_{\alpha\dot{\alpha}}$	4	$\square_L \times \square_R$
$(2,0)$	Self-dual 2-form	$F_{\alpha\beta}$	3	$\square\square_L \times 1_R$
$(0,2)$	Antiself-dual 2-form	$B_{\dot{\alpha}\dot{\beta}}$	3	$1_L \times \square\square_R$

<u>*Remark 6.1.3*</u>. Now things are a little subtle because we're treading the line between mixing the Lie group, which have the matrices $\Lambda^\mu{}_\nu$, and the Lie algebra, which have the basis elements $\{N_i, \overline{N}_i\}$. The objects in the table above are in the representation space of the Lie group (that's why we can draw Young-Tableaux), but we want to use the nice properties of the Lie algebra to study things. What we have to remember is that the two structures are related by the exponential map, and we can relate their representations this way. I shall try to be as explicit as possible in the following but it's likely I'll make a couple errors.

Notation. As we have done in the table above, we will denote elements of $SU(2)_L$ with α, β etc., and we will denote elements of $SU(2)_R$ with $\dot{\alpha}, \dot{\beta}$ etc. The reason is that this is the usual notation used in places like supersymmetry. Note that both indices take values in $\{1, 2\}$.

We should stop a second a make a few comments on the table above. The first three entries are fine, but we call the $(1,1)$ entry a vector. As is required, it has an α and a $\dot{\alpha}$ index, but we're used to writing vectors with a single spacetime index, μ. So what's going on? Well, as we will see shortly, it turns out that something of the form $A_{\alpha\dot{\alpha}}$ does indeed transform as a vector in $\text{SO}(3,1)$, i.e. we can 'repackage' the information such that

$$A^\mu \mapsto \Lambda^\mu{}_\nu A^\nu.$$

A similar thing holds for the $(2,0)$ and $(0,2)$ entries, but we won't discuss that here.

[4]The Young-Tableaux here might look a little strange. The important thing to note is the we are *not* taking the tensor product of two Young-Tableaux, but the Cartesian product. This just corresponds to categorising the representations by the double (j_1, j_2), see a linear algebra book if this doesn't make sense.

Remark 6.1.4. As we just said, the last two are not going to be important to us here but for completeness, basically they obey

$$F = \star F, \quad \text{and} \quad B = -\star B,$$

where $\star$ is the *Hodge dual*.[5] In spacetime components this can be written

$$F_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}F_{\rho\sigma}, \quad \text{and} \quad B_{\mu\nu} = -\frac{1}{2}\epsilon_{\mu\nu\rho\sigma}B_{\rho\sigma}.$$

These structures play crucial roles in the study of so-called *Yang-Mills instantons*.

There is another important representation to consider, but this one is *reducible*. It is known as a *Dirac spinor* and is given by $(1,0)\oplus(0,1)$, which we write in matrix form as[6]

$$\psi_D = \begin{pmatrix} \psi_\alpha \\ \overline{\psi}_{\dot{\alpha}} \end{pmatrix}.$$

It has dimension $\dim(1,0)+\dim(0,1) = 2+2 = 4$, however it is *not* a vector as

$$(1,0)\oplus(0,1) \neq (1,1).$$

Remark 6.1.5. As was the case with the js in the previous lecture, we are using the mathematician's notation with integers not half integers. A physicist would write the Dirac spinor as $(\frac{1}{2},0)\oplus(0,\frac{1}{2})$. Same for the other terms in the table above.

6.1.2 Left-Handed vs Right-Handed Spinors

Let's consider the left-handed Weyl spinors first. As per the table above, they transform like the fundamental representation in $SU(2)_L$ and via the trivial representation in $SU(2)_R$. Keeping Remark 6.1.3 in mind, we can convert this into a statement about the representations of the Lie algebras $\mathfrak{su}(2)_{L/R}$. In terms of our basis $\{N_i, \overline{N}_i\}_{i\in\{1,2,3\}}$, we have[7]

$$d(N_i) = \tau_i = -\frac{i}{2}\sigma_i, \quad \text{and} \quad d(\overline{N}_i) = 0, \tag{6.5}$$

where we have used the basis Equation (5.1) (so the commutators work nicely). So what does this representation look like in terms of the Lie group? That is we want to find an expression for

$$D(\Lambda) : \psi_\alpha \mapsto \Lambda^\beta{}_\alpha \psi_\beta$$

in terms of the representation of the Lie algebra. How do we do this? Well we use the exponential map to obtain

$$\Lambda^\mu{}_\nu = \exp\big(\omega^\mu{}_\nu\big).$$

Note this is a finite transformation as we don't have the small parameter ϵ, as we did in the exercise above. As we said after this exercise, $\omega^\mu{}_\nu$ has 6 free parameters, 3 of which are the

[5] See a book on differential geometry.
[6] It is also often written with ψ_L and ψ_R as entries.
[7] Note the representation $d(X) = 0$ in the Lie algebra corresponds to $D(e^0) = \mathbb{1}$ in the Lie group, which is exactly the trivial representation.

rotations and the other 3 the boosts. We shall denote these by r_i and b_i with $i \in \{1,2,3\}$. Now the $d(N_i)/d(\overline{N}_i)$ span the representation, and so we obtain

$$D(\Lambda) : \psi_\alpha \mapsto \exp\big[n_i d(N_i)\big]^\beta{}_\alpha \psi_\beta,$$

with $n_i \in \mathbb{C}$ and where we don't get any $d(\overline{N}_i)$ terms by Equation (6.5).

The question is "what are the n_is?" Well they are linear combinations of the r_i and b_i mentioned above, and we can decide how by using our required interpretation. We want the r_is to be the rotations, and we have seen previously that the rotations are given by $\exp(a_i A_i)$ where a_i are the rotation angles and A_i are the rotation matrices in the Lie algebra. We therefore want the r_i term to come in the form

$$\exp(r_i \tau_i) = \exp\bigg(-\frac{i}{2} r_i \sigma_i\bigg).$$

Similarly we want the boost parts *not* to look like a rotation, and so we don't want the i factor, i.e. we want something of the form

$$\exp(-ib_i\tau_i) = \exp\bigg(-\frac{b_i}{2}\sigma_i\bigg).$$

Using Equation (6.5), we therefore take

$$n_i = r_i - ib_i. \tag{6.6}$$

Note these two sign conventions have been chosen so that they line up with Equation (6.3). Putting this together we get

$$D(\Lambda) : \psi_\alpha \mapsto M^\beta{}_\alpha \psi_\beta, \qquad \text{with} \qquad M^\beta{}_\alpha := \exp\bigg(-\frac{ir_i}{2}\sigma_i - \frac{b_i}{2}\sigma_i\bigg)^\beta{}_\alpha. \tag{6.7}$$

<u>*Remark 6.1.6*</u>. Note that $M^\beta{}_\alpha \in SL(2,\mathbb{C})$ and *not* $SU(2)$. That is it is not unitary. This is a result of a theorem which says that non-compact groups cannot have unitary representations, and in Remark 6.1.1 we said that the Lorentz group is non-compact. Note that this result stems from the fact that we have complexified the n_is. If we had not (e.g. if we'd set $n_i = r_i + b_i$) then we could have got two types of rotation in Equation (6.7), and we would have been studying SO(4), which is compact.

We can now redo the whole game for the right-handed spinors. In this case we have the representation opposite to Equation (6.5), namely

$$\widetilde{d}(N_i) = 0, \quad \text{and} \quad \widetilde{d}(\overline{N}_i) = \tau_i = -\frac{i}{2}\sigma_i.$$

If we consider the same Lorentz transformation as above, everything follows through the same, apart from now we use

$$\overline{n}_i = r_i + ib_i,$$

and obtain

$$\widetilde{D}(\Lambda) : \psi_{\dot\alpha} \mapsto (M^*)^{\dot\beta}{}_{\dot\alpha} \psi_{\dot\beta}, \qquad \text{with} \qquad (M^*)^{\dot\beta}{}_{\dot\alpha} := \exp\bigg(-\frac{ir_i}{2}\sigma_i + \frac{b_i}{2}\sigma_i\bigg)^{\dot\beta}{}_{\dot\alpha}. \tag{6.8}$$

Note that basically the only difference between the representations on left-handed and right-handed Weyl spinors is the sign before the boost part. This corresponds physically to a property known as *helicity*. Helicity basically tells you the projection of the spin of a massless particle (which Weyl spinors are) onto its momentum, we call the two options left- and right-handed (hence the names we've been using). These names come from our hands: make a thumbs up but don't curl your fingers all the way in, now imagine your thumb points in the direction of momentum, then your fingers tell you about the spin direction. A right-handed spinor has spin-momentum projection like your right hand looks, and similarly for a left-handed spinor.

Left-Handed Right-Handed

6.1.3 Vectors

We can now return to the comment we made about about the fact that $A_{\alpha\dot{\alpha}}$ is a vector. Let's looks how it transforms:

$$D(\Lambda) \times \widetilde{D}(\Lambda) : A_{\alpha\dot{\alpha}} \mapsto M^{\beta}{}_{\alpha} (M^*)^{\dot{\beta}}{}_{\dot{\alpha}} A_{\beta\dot{\beta}} = (MAM^{\dagger})_{\alpha\dot{\alpha}}.$$

This still doesn't look anything like the transformation of a vector. We recover our usual vector type transformation by introducing the following vector of matrices

$$\sigma^{\mu} := (\mathbb{1}_{2\times 2}, -\sigma_1, -\sigma_2, -\sigma_3).$$

We can use this to repackage the information of a vector X^{μ} as a 2×2 matrix. We define

$$X_{\alpha\dot{\alpha}} := X^{\mu}\eta_{\mu\nu}\sigma^{\nu} = \begin{pmatrix} X^0 + X^3 & X^1 - iX^2 \\ X^1 + iX^2 & X^0 - X^3 \end{pmatrix}.$$

Now we have just chosen to label the entries of this 2×2 matrix with an $\alpha\dot{\alpha}$, but haven't shown it actually relates to the $\alpha\dot{\alpha}$ notation of left-handed/right-handed representations. Well it turns out that if you consider the Lorentz transformation

$$X'^{\mu} = \Lambda^{\mu}{}_{\nu} X^{\nu},$$

where this Λ is the same as the one in Equations (6.7) and (6.8), it translates to

$$X'_{\alpha\dot{\alpha}} = (MXM^{\dagger})_{\alpha\dot{\alpha}}. \tag{6.9}$$

The proof of this is the content of the next exercise.[8]

[8]This is something that was set as a problem sheet question on the course, so I don't want to type the answer. If the question is unclear at all, please feel free to email me for further clarity.

Exercise

Given the matrices

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

$$K_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad K_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad K_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

show that the vector representation of SO$(3,1)$ corresponds to the $(1,1)$ representation of $SU(2)_L \times SU(2)_R$. That is prove that Equation (6.9) holds. *Hint: Construct the explicit representation of the generators* $N_i/\overline{N}_i$ *of* $SU(2)_L \times SU(2)_R$ *starting from the* J_i/K_i *matrices given above.*

6.2 Cartan Classification

So we have done a lot of work regarding representations of groups, the final question we want to ask is "can we classify all Lie algebras and their representations?" The answer is yes and no. The no part just means that their are too many Lie algebras, and so we need to restrict ourselves to a smaller set. The yes will take some time to get to. First we need to introduce some definitions.[9]

6.2.1 Some More Definitions/Theorems

Definition. [Abelian Lie Algebra] A Lie algebra $(\mathfrak{g}, [,])$ is said to be *Abelian* if the Lie bracket of any two elements vanishes. That is, for all $g_1, g_2 \in \mathfrak{g}$

$$[g_1, g_2] = 0.$$

Definition. [Lie Subalgebra] Let $(\mathfrak{g}, [,])$ be a Lie algebra. Then then we call $\mathfrak{h} \subset \mathfrak{g}$ a *Lie subalgebra* if it is a subspace and it is closed under the Lie bracket, i.e.

$$[h_1, h_2] \in \mathfrak{h}, \qquad \forall h_1, h_2 \in \mathfrak{h}.$$

[9]Some of these may have appeared above. I have decided to present them here again anyways just so this section is easier to read.

Definition. [Invariant Lie Subalgebra/Ideals] An *invariant Lie subalgebra* is a Lie subalgebra that goes into itself under commutation with *any* element of the full Lie algebra. That is, for all $g \in \mathfrak{g}$ and $h \in \mathfrak{h}$

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

We also refer to these as *ideals*.

Corollary 6.2.1. *Every Lie algebra possesses two ideals, namely* $\mathfrak{h} = \{0\}$ *and* $\mathfrak{h} = \mathfrak{g}$. *We refer to these as trivial ideals.*

Exercise

Verify the above Corollary. *Hint: This is not a trick question, it is very straight forward.*

Definition. [Simple Lie Algebra] A Lie algebra is called *simple* if it has $\dim \mathfrak{g} > 1$[10] and it has no non-trivial ideals.

Remark 6.2.2. Ideals is a similar concept to an invariant subspace of a representation. Similarly, simple Lie algebras are akin to irreps.

It is the *Abelian ideals* that mess up our classification process. This is just because once you hit one element of an Abelian ideal, everything commutes and so you loose all the information (i.e. the structure constants are all vanishing). This motives the next definition.

Definition. [Semisimple Lie Algebras] A Lie algebra is said to be *semisimple* if it has no Abelian ideals.

Theorem 6.2.3. *A semisimple Lie algebra can be written as a direct sum of simple Lie algebras.*

Theorem 6.2.4 (Cartan)**.** *A Lie algebra is semisimple if, and only if, its Killing form is non-degenerate. That is* g^{ab} *is well defined.*

Proof. See page 41 of Dexter Chua's notes for part of the proof. □

Definition. [Ad-Diagonalisable] Let $(\mathfrak{g}, [,])$ be a Lie algebra. We say an element $X \in \mathfrak{g}$ is *ad-diagonalisable* if the adjoint representation $\mathrm{ad}(X)$ is diagonalisable.

6.2.2 Standard Form Of Semisimple Lie Algebras

Essentially what we're going to try and do is use some smart basis such that our Lie algebra becomes a bunch of $\mathfrak{su}(2)$s. This basis is known as the *Chevalley basis*. We are going to use our discussion of $\mathfrak{su}(3)$ from last lecture as a guiding example.

Step 1

Find a maximal set of independent, commuting, ad-diagonalisable elements, $\{H_1, ..., H_r\}$. The value r is known as the *rank* of the Lie algebra, and the subalgebra

$$\mathfrak{h} := \text{span}_{\mathbb{C}}\{H_1, ..., H_r\} \tag{6.10}$$

is known as the *Cartan subalgebra*, it is *not* unique. The idea is going to be to simultaneously diagonalise the (adjoint) representation of all of these, as we did last lecture. This is why we require $\{H_i\}$ to be ad-diagonalisable.

Example 6.2.5. We saw last lecture that for $\mathfrak{su}(2)$, $r = 1$ and we can chose $H = \sigma_3$. We also saw that $r = 2$ for $\mathfrak{su}(3)$ and had $H_1 \sim \lambda_3$ and $H_2 \sim \lambda_8$. This result generalised for $\mathfrak{su}(N)$, where $r = N - 1$.

Step 2

Consider the algebra as a representation space on its own, i.e. use the adjoint representation. From the definition of a representation, and the fact that $[H_i, H_j] = 0$, we have

$$[ad(H_i), ad(H_j)] = 0 \qquad \forall i, j \in \{1, ..., r\}.$$

Proposition 6.2.6. *Let $\mathfrak{h}$ be a rank r Cartan subalgebra of $\mathfrak{g}$. Then any $X \in \mathfrak{g}$ that satisfies $[X, H_i] = 0$ for all $i \in \{1, ..., r\}$, then $X \in \mathfrak{h}$.*

This proposition basically tells us that every diagonal element is non-zero in at least one of the H_is, as if one wasn't then the diagonal matrix with only that one entry in it would commute with all other H_is. Combining this result with the fact that the matrices of the adjoint representation are $\dim \mathfrak{g} \times \dim \mathfrak{g}$ in size, we get the following important result.

Corollary 6.2.7. *The simultaneous eigenvectors*[11] *of the $ad(H_i)$s form a basis for the whole Lie algebra $\mathfrak{g}$.*

We call these simultaneous eigenvectors *root vectors* and denote them by $E_{\underline{\alpha}}$, where $\underline{\alpha} = (\alpha^1, ..., \alpha^r)$ are the simultaneous eigenvalues.[12] We call $\underline{\alpha}$ the *root* and it lives in an r-dimensional vector space, called the *root space*. The set of all roots is called the *root system*, and it corresponds to the spectrum of the Cartan subalgebra.

The eigenvector condition tells us

$$ad(H_i)E_{\underline{\alpha}} := [H_i, E_{\underline{\alpha}}] = \alpha_i E_{\underline{\alpha}}, \tag{6.11}$$

where the middle expression is just the definition of the adjoint representation. Comparing to the previous lecture, we see that the root vectors are just the generalisation of the step operators. This is nice, but last lecture the commutators were $[H, E_\pm] = \pm 2E_\pm$, we want to recover this.

[11] We allow the eigenvalue to be zero here.

[12] We have actually used the fact that the eigenvectors are non degenerate. That is, each $\underline{\alpha}$ has a unique eigenvector $E_{\underline{\alpha}}$

The Killing form induces a metric on $\mathfrak{h}$:

$$g_{ij} = f_{i\ d}^{\ c} f_{j\ c}^{\ d}, \qquad \text{with} \qquad i,j \in \{1,...,r\} \text{ and } c,d \in \{1,...,\dim\mathfrak{g}\}.$$

By Cartan's theorem this is invertible, and so we have

$$g^{ij} := (g^{-1})_{ij}.$$

We define the inner product as

$$\langle X, Y\rangle := X_k g^{ks} Y_s \tag{6.12}$$

for $X, Y \in \mathfrak{h}$. Then we define, for every root $\underline{\alpha}$,

$$H_{\underline{\alpha}^\checkmark} := \frac{2g^{ij}\alpha_i H_j}{\langle\underline{\alpha},\underline{\alpha}\rangle}, \tag{6.13}$$

which is just a linear combination of Cartan elements. We also define

$$\alpha_i^\checkmark := \frac{\alpha_i}{\langle\underline{\alpha},\underline{\alpha}\rangle}, \tag{6.14}$$

which we call a *coroot*. This gives us

$$H_{\underline{\alpha}^\checkmark} = \langle\underline{\alpha}^\checkmark, H\rangle = \alpha_i^\checkmark g^{ij} H_j,$$

which in turn gives us

$$[H_{\underline{\alpha}^\checkmark}, E_{\underline{\alpha}}] = \frac{2g^{ij}\alpha_i}{\langle\underline{\alpha},\underline{\alpha}\rangle}[H_j, E_{\underline{\alpha}}] = \frac{2g^{ij}\alpha_i}{\langle\underline{\alpha},\underline{\alpha}\rangle}\alpha_j E_{\underline{\alpha}} = 2E_{\underline{\alpha}},$$

which is what we wanted to obtain.

Remark 6.2.8. Note we can think of $\langle\underline{\alpha},\underline{\alpha}\rangle$ as an inner product on the root space, telling us the length the root $\underline{\alpha}$ w.r.t. g^{ij}. This is exactly what we were talking about last lecture with the root diagrams not having nice shapes but being squashed. We can make them nicer by taking a change of basis such that this metric becomes the Euclidean one.

Exercise

Consider the weight lattice given last lecture for the fundamental represntation of SU(3), i.e. the squashed triangle with states $|1,0\rangle, |-1,1\rangle$ and $|0,-1\rangle$. The Killing metric on this root space is

$$g^{ij} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

so that the inner product

$$\langle\underline{\alpha},\underline{\beta}\rangle = \begin{pmatrix} 2 & -1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \end{pmatrix} = -3.$$

By considering the inner product of all the weights, convince yourself that geometrically the weight lattice form the vertices of an equilateral triangle.

Comment: Again this one is taken straight from the problem sheets on the course, but I've included it to illustrate the fact that we can make nice shapes. As always, feel free to email me if you want further explanation.

Step 3

The last thing we have to do is find the commutators between the different root vectors, i.e.

$$[E_{\underline{\alpha}}, E_{\underline{\beta}}] = ?$$

Well recall last lecture we set

$$E_{\alpha+\beta} = [E_\alpha, E_\beta].$$

Exercise

Use Equation (6.11) to show

$$ad(H_i)\big([E_{\underline{\alpha}}, E_{\underline{\beta}}]\big) = (\alpha_i + \beta_i)[E_{\underline{\alpha}}, E_{\underline{\beta}}],$$

thereby justifying what we defined last lecture.

This result also explains why last lecture we only needed to consider the simple roots α and β and not the other root $(\alpha + \beta)$. This result generalises to the following definition.

Definition. [Simple Root] A simple root is a positive root $\underline{\alpha}$ that can not be written as a sum of two positive roots.

Remark 6.2.9. We say *positive* root, because obviously if we have the root $\underline{\alpha}$, then $-\underline{\alpha}$ is also a root, so we decide to split our root space in two and define simple roots only using the positive ones. Note *we decide* which roots are positive, it is not something that is given to us. In perhaps more technical language, we take an $(r-1)$-dimensional hyperplane of our root space and say "everything above this plane is positive, and everything below it is negative".

Corollary 6.2.10. *Our root space has exactly r simple roots.*

Proof. This just follows from the fact that our root space is a r-dimensional lattice space, and so we have r linearly independent roots that we can use to span the space. □

This Corollary allows us to categorise all other roots, simply: we call a non-simple root *positive* if it can be written as a sum of simple roots with all coefficients being positive. Likewise we have a negative root. Note that a root is either positive or negative, as it either lies above the hyperplane or below it.

This basis on the root space induces a nice basis on the Cartan subalgebra given by $\{H_{\underline{\alpha}_i^\checkmark}\}$ such that

$$[H_{\underline{\alpha}_i^\checkmark}, E_{\underline{\alpha}_j}] = C_{ij} E_{\underline{\alpha}_j}, \tag{6.15}$$

where C_{ij} is a $r \times r$ matrix, known as the *Cartan matrix.*

Example 6.2.11. For $\mathfrak{su}(2)$ the rank is $r = 1$ and we just have $C_{11} = 2$. For $\mathfrak{su}(3)$ the rank is $r = 2$ and we have

$$C_{ij} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

The $\mathfrak{su}(3)$ example above shows us that although we have a bunch of embedded $\mathfrak{su}(2)$s, they do talk to each other as the off diagonal elements are non-vanishing. However it is a nice surprise that for a general $\mathfrak{su}(N)$ the Cartan matrix is the $(N-1)\times(N-1)$ matrix of the form

$$C_{ij} = \begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix},$$

with all the missing elements being 0. This tells us that although the embedded $\mathfrak{su}(2)$s speak to each other, they only speak with their 'neighbours' and not *everyone*, i.e. it is only the just off diagonal elements that are non-zero.

Proposition 6.2.12. *All the information about a simple Lie algebra can be extracted from the Cartan matrix and the simple roots.*

The 'proof' of this proposition is the following idea: you start with the highest weight state and use the Cartan metric to work downwards and obtain all other states. For clarity, the states are given by

$$d(H_{\underline{\alpha}_i^{\checkmark}})\,|\lambda_1,...,\lambda_r\rangle = \lambda_i\,|\lambda_1,...,\lambda_r\rangle\,,$$

where d is some representation. The highest weight state is defined via

$$d(E_{\underline{\alpha}})\,|\Lambda_1,...,\Lambda_r\rangle = 0,$$

for all simple roots $\underline{\alpha}$.

Cartan managed to classify *all* (not just $\mathfrak{su}(N)$) simple Lie algebras by examining the possible Cartan matrices and possible allowed root lattices. He developed a system to indicate these pictorially, known as *Dynkin diagrams.* Dr. Dorigoni did not have time to discuss these in this course, but details about them can be found in Dexter Chua's notes or on Dr. Schuller's "Lectures on the Geometric Anatomy of Theoretical Physics".

6.3 Lie Groups Relevant In Physics

We end this course with a brief mention of some of the Lie groups relevant in physics.

(i) In the standard model, the gauge group is $SU(3)\times SU(2)\times U(1)$.

(ii) Grand unified theories, need a group that can contains the standard model group as a subgroup. Two possibilities are $SU(5)$ and $SO(10)$.

(iii) In superstring theory we have multiple groups, including $SO(10)$[13] and something called E_8 (classified as a Dynkin diagram).

6.4 Dykin Diagrams

To come when I get time to type this up. These were not part of the course, but just something I think worth including.

[13]See my notes on Dr. Shiraz Minwalla's string theory course for why we need $SO(10)$.

Useful Texts & Further Readings

Similar Courses

- Dexter Chua's Notes On Professor Nick Dorey's 2016 Cambridge Part III course: *Symmetries, Fields and Particles*. Available Online.
- Dr Frederic Schuller's *Lectures on the Geometric Anatomy of Theoretical Physics*. Available on YouTube or via Simon Rea's notes. The most relevant lectures are 13-18, however knowledge from the previous lectures is obviously assumed in the teaching.
- Hans Samelson's *Notes on Lie Algebras*. Available online.

Books

- Jones, Hugh F. *Groups, Representations & Physics.* CRC Press, 1998.
- Fuchs, Jürgen, and Christoph Schweigert. *Symmetries, Lie Algebras & Representations: A Graduate Course For Physicists.* Cambridge University Press, 2003.
- Georgi, Howard, and Richard Slansky. *Lie Algebras In Particle Physics.* Physics Today 36 (1983): 62.
- Nakahara, Mikio. *Geometry, Topology & Physics.* CRC Press, 2003.